WATER RESOURCES RESEARCH, VOL. 43, W08403, doi:10.1029/2006WR005525, 2007

Click Here for Full Article

# Usefulness of the reversible jump Markov chain Monte Carlo model in regional flood frequency analysis

M. Ribatet,[1,2] E. Sauquet,[1] J. M. Grésillon,[1] and T. B. M. J. Ouarda[2]

Received 13 September 2006; revised 3 May 2007; accepted 17 May 2007; published 3 August 2007.

[1] Regional flood frequency analysis is a convenient way to reduce estimation uncertainty when few data are available at the gauging site. In this work, a model that allows a non-null probability to a regional fixed shape parameter is presented. This methodology is integrated within a Bayesian framework and uses reversible jump techniques. The performance on stochastic data of this new estimator is compared to two other models: a conventional Bayesian analysis and the index flood approach. Results show that the proposed estimator is absolutely suited to regional estimation when only a few data are available at the target site. Moreover, unlike the index flood estimator, target site index flood error estimation seems to have less impact on Bayesian estimators. Some suggestions about configurations of the pooling groups are also presented to increase the performance of each estimator.

**Citation:** Ribatet, M., E. Sauquet, J. M. Grésillon, and T. B. M. J. Ouarda (2007), Usefulness of the reversible jump Markov chain Monte Carlo model in regional flood frequency analysis, *Water Resour. Res.*, *43*, W08403, doi:10.1029/2006WR005525.

## 1. Introduction

[2] Extreme value theory is now widely applied when modeling block maxima or exceedences over a threshold are of interest. In particular, the Generalized Pareto Distribution (**GPD**) describes the limiting distribution of normalized excesses of a threshold as the threshold approaches the endpoint of the variable [*Pickands*, 1975]. The GPD has a distribution function defined by:

$$G(x;\mu,\sigma,\xi) = 1 - \left[1 + \frac{\xi(x-\mu)}{\sigma}\right]^{-1/\xi}, x > \mu, 1 + \frac{\xi(x-\mu)}{\sigma} > 0 \tag{1}$$

where $\sigma > 0$, $\xi \in \mathbb{R}$. $\mu$, $\sigma$ and $\xi$ are respectively the location, scale, and shape parameters.

[3] Thus, when extreme values must be estimated, this approximation is frequently used. Most applications based on this result are related to environmental sciences, as extreme wind speed [*Payer and Kuchenhoff*, 2004], extreme sea level [*Bortot and Coles*, 2000; *Pandey et al.*, 2004], or extreme river discharge [*Northrop*, 2004].

[4] However, one must often deal with small samples and large uncertainties on estimation. Several publications point out the problem of the shape parameter estimation. This parameter is of great interest as it determines the tail behaviour of the distribution. Therefore, many authors analyzed the performance of particular estimators given a specified range for the shape parameter: *Rosbjerg et al.* [1992] for the method of moments; *Coles and Dixon* [1999] for the maximum likelihood; *Hosking and Wallis* [1987] for the probability weighted moments; *Juarez and Schucany* [2004] for the minimum density power divergence estimator; and *Martins and Stedinger* [2000] for a proposed generalized maximum likelihood. However, these results provide the most accurate estimator given the shape parameter, which is never the case in practice. Therefore, *Park* [2005] introduced a systematic way of selecting hyperparameters for his proposed generalized maximum likelihood estimator.

[5] All these approaches only deal with information from the target site sample. However, it is frequent in hydrology to perform a Regional Frequency Analysis (**RFA**). Traditional RFA consists of two steps: (a) delineation of homogeneous regions, i.e., a pooling group of stations with similar behaviour; (b) regional estimation, i.e., estimate target site distribution from the regional information.

[6] More recently, Bayesian approaches have been applied with success to incorporate regional information in frequency analysis [*Coles and Tawn*, 1996; *Northrop*, 2004; *Seidou et al.*, 2006; *Ribatet et al.*, 2007]. Empirical Bayesian estimators have also been proposed [*Kuczera*, 1982; *Madsen and Rosbjerg*, 1997]. One of the advantages of these approaches is to distinguish the at-site information from the other site data in the estimation procedure. This is an important point as, no matter how high the homogeneity level may be, the only data which represent perfectly the target site are obviously the target site one. Thus the whole information available is used more efficiently. In addition, according to *Ribatet et al.* [2007], the Bayesian approaches allow to relax the scale invariance property required by the most applied RFA model, i.e., the index flood [*Dalrymple*, 1960].

[7] However, a preliminary study on simulated data showed that the approach developed by *Ribatet et al.* [2007] may lead to unreliable estimates for larger return

[1]Unité de Recherche Hydrologie-Hydraulique, CEMAGREF Lyon, Lyon Cedex, France.
[2]INRS-ETE, University of Quebec, Québec, Canada.

Copyright 2007 by the American Geophysical Union.
0043-1397/07/2006WR005525$09.00

periods ($T > 20$ years) when small samples are involved. This poor performance is mainly due to the large variance on the shape parameter estimation. Consequently, for such cases, attention must be paid to the regional estimation procedure for the shape parameter.

[8] The basis of our new development was formerly proposed by *Stephenson and Tawn* [2004]. They use reversible jump Markov Chain Monte Carlo (MCMC) techniques [*Green*, 1995] to attribute a non-null probability to the Gumbel case. Therefore, realizations are not supposed to be Gumbel distributed but have a non-null probability to be Gumbel distributed. An application to extreme rainfall and sea level is given. In this work, this approach is extended to take into account a regional shape parameter, not only the Gumbel/Exponential case, within a RFA framework. The reversible jump technique allows to focus on a "likely" shape parameter value given by the hydrological relevance of the homogeneous region. Thus this approach may reduce the shape parameter variance estimation while relaxing the scale invariance property.

[9] The main objectives of this article is first to present new developments in the methodology proposed by *Stephenson and Tawn* [2004] required for a RFA context; second is to assess the quality of two Bayesian models based on the index flood hypothesis: the regional Bayesian model proposed by *Ribatet et al.* [2007] (**BAY**) and the new proposed Bayesian approach applying reversible jumps Markov chains (**REV**). They are compared to the classical index flood approach of *Dalrymple* [1960] (**IFL**). The assessment is developed through a stochastic generation of regional data performed in order to obtain realistic features of homogeneous regions. Detailing the index flood concept is out of the scope of this article. Estimation procedure can be found in the study of *Hosking and Wallis* [1997].

[10] The paper is organized as follows. The next two sections concentrate on methodological aspects. Section 2 describes the Bayesian framework including the specific MCMC algorithm, required to extend the work by *Stephenson and Tawn* [2004]. Section 3 presents the simple and efficient algorithm to generate stochastically hydrological homogeneous regions. A sensitivity analysis is performed in section 4 to assess how quantile estimates and related uncertainties are influenced by the values of two parameters of the reversible jump Markov chains. Section 5 compares the performance of each estimator on six representative case studies. The impact of the bias in the target site index flood estimation is analyzed in section 6, while suggestions for building efficient pooling groups are presented in section 7. Finally, some conclusions are drawn in section 8.

## 2. Methodology

[11] In the Bayesian framework, the posterior distribution of parameters must be known to derive quantile estimates. The posterior distribution $\pi(\boldsymbol{\theta}|x)$ is given by the Bayes Theorem [*Bayes*, 1763]:

$$\pi(\boldsymbol{\theta}|x) = \frac{\pi(\boldsymbol{\theta})\pi(x;\boldsymbol{\theta})}{\int_\Theta \pi(\boldsymbol{\theta})\pi(x;\boldsymbol{\theta})d\boldsymbol{\theta}} \propto \pi(\boldsymbol{\theta})\pi(x;\boldsymbol{\theta}) \tag{2}$$

where $\boldsymbol{\theta}$ is the vector of parameters of the distribution to be fitted, $\Theta$ is the parameter space. $\pi(x;\boldsymbol{\theta})$ is the likelihood function, $\boldsymbol{x}$ is the vector of observations, and $\pi(\boldsymbol{\theta})$ is the prior distribution.

[12] In this study, as excesses over a high threshold are of interest, the likelihood function $\pi(\boldsymbol{x};\boldsymbol{\theta})$ is related to the GPD (see equation (1)).

### 2.1. Prior Distribution

[13] In this section, the methodology to elicit the prior distribution is presented. In this study, regional information is used to define the prior distribution. Furthermore, the prior is specific as it must account for a fixed shape parameter $\xi_{\text{Fix}}$ with a non-null probability $p_\xi$. Let $\Theta_0$ be a sub-space of the parameter space $\Theta$ of $\boldsymbol{\theta}$. More precisely, $\Theta_0 = \{\boldsymbol{\theta} \in \Theta: \xi = \xi_{\text{Fix}}\}$. $p_\xi$ is a hyper-parameter of the prior distribution. The approach is to construct a suitable prior distribution on $\Theta$; then, for $p_\xi$ fixed, to modify this prior to account for the probability of $\Theta_0$.

[14] For clarity purposes, the prior distribution is defined in two steps. First, an initial prior distribution $\pi_{\text{in}}(\boldsymbol{\theta})$ defined on $\Theta$ is introduced. Second, a revised prior distribution $\pi(\boldsymbol{\theta})$ is derived from $\pi_{\text{in}}(\boldsymbol{\theta})$ to attribute a non-null probability to the $\Theta_0$ sub-sample.

#### 2.1.1. Initial Prior Distribution

[15] As the proposed model is fully parametric, the initial prior distribution $\pi_{\text{in}}(\boldsymbol{\theta})$ is a multivariate distribution entirely defined by its hyper-parameters. In our case study, the initial prior distribution corresponds to the one introduced by *Ribatet et al.* [2007]. Consequently, the marginal prior distributions were supposed to be independent lognormal for both location and scale parameters and normal for the shape parameter. Thus,

$$\pi_{\text{in}}(\boldsymbol{\theta}) \propto J \exp\left[(\boldsymbol{\theta}' - \gamma)^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}' - \gamma)\right] \tag{3}$$

where $\boldsymbol{\gamma}$ and $\boldsymbol{\Sigma}$ are hyper-parameters, $\boldsymbol{\theta}' = (\log \mu, \log \sigma, \xi)$, and $J$ is the Jacobian of the transformation from $\boldsymbol{\theta}'$ to $\boldsymbol{\theta}$, namely $J = 1/\mu\sigma$. $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \gamma_3)$ is the mean vector, $\boldsymbol{\Sigma}$ is the covariance matrix. As marginal priors are supposed to be independent, $\boldsymbol{\Sigma}$ is a $3 \times 3$ diagonal matrix with diagonal elements $d_1$, $d_2$, $d_3$.

[16] Hyper-parameters are defined through the index flood concept, that is, all distributions are identical up to an at-site-dependent constant. Consider all sites of a region except the target site, say the $j$th site. A set of pseudo target site parameters can be computed:

$$\tilde{\mu}_i = C^{(j)}\mu_*^{(i)} \tag{4}$$

$$\tilde{\sigma}_i = C^{(j)}\sigma_*^{(i)} \tag{5}$$

$$\tilde{\xi}_i = \xi_*^{(i)} \tag{6}$$

for $i \neq j$, where $C^{(j)}$ is the target site index flood, and $\mu_*^{(i)}$, $\sigma_*^{(i)}$, $\xi_*^{(i)}$ are respectively the location, scale and shape at-site parameter estimates from the rescaled sample, for example, normalized by its respective index flood estimate. Under the hypothesis of the index flood concept, pseudo-

parameters are expected to be distributed as parameters of the target site.

[17] Information from the target site sample cannot be used to elicit the prior distribution. Thus $C^{(j)}$ in equations (4) and (5) must be estimated without use of the $j$th sample site.

[18] In this case study, $C^{(j)}$ is estimated through a Generalized Linear Model (GLM) defined by:

$$\begin{cases} \mathbb{E}\left[\log C^{(j)}\right] & = \nu, \qquad \nu = X\beta \\ \mathrm{Var}\left[\log C^{(j)}\right] & = \phi V(\nu) \end{cases} \tag{7}$$

where $X$ are basin characteristics (possibly log transformed), $\phi$ is the dispersion parameter, $V$ is the variance function, and $\nu$ is the linear predictor. *McCullagh and Nelder* [1989] give a comprehensive introduction to GLM. Other alternatives for modeling the target site index flood can be considered such as Generalized Additive Models [*Wood and Augustin*, 2002], Neural Networks [*Shu and Burn*, 2004], or Kriging [*Merz and Blöschl*, 2005]. However, the variance of $C^{(j)}$ should be estimated. Indeed, as $C^{(j)}$ is estimated without use of the target site data, uncertainties due to this estimation must be incorporated in the prior distribution.

[19] From these pseudo parameters, hyper-parameters can be computed:

$$\gamma_1 = \frac{1}{N-1}\sum_{i\neq j}\log\tilde{\mu}^{(i)}, \qquad d_1 = \frac{1}{N-1}\sum_{i\neq j}\mathrm{Var}\left[\log\tilde{\mu}^{(i)}\right] \tag{8}$$

$$\gamma_2 = \frac{1}{N-1}\sum_{i\neq j}\log\tilde{\sigma}^{(i)}, \qquad d_2 = \frac{1}{N-1}\sum_{i\neq j}\mathrm{Var}\left[\log\tilde{\sigma}^{(i)}\right] \tag{9}$$

$$\gamma_3 = \frac{1}{N-1}\sum_{i\neq j}\tilde{\xi}^{(i)}, \qquad d_3 = \frac{1}{N-2}\sum_{i\neq j}\left(\tilde{\xi}^{(i)} - \gamma_3\right)^2 \tag{10}$$

Under the independence assumption between $C^{(j)}$ and $\mu_*^{(i)}$, $\sigma_*^{(i)}$, the following relations hold:

$$\mathrm{Var}\left[\log\tilde{\mu}^{(i)}\right] = \mathrm{Var}\left[\log C^{(j)}\right] + \mathrm{Var}\left[\log\mu_*^{(i)}\right] \tag{11}$$

$$\mathrm{Var}\left[\log\tilde{\sigma}^{(i)}\right] = \mathrm{Var}\left[\log C^{(j)}\right] + \mathrm{Var}\left[\log\sigma_*^{(i)}\right] \tag{12}$$

The independence assumption is not too restrictive as the target site index flood is estimated independently from $\mu_*^{(i)}$, $\sigma_*^{(i)}$.

[20] Note that $\mathrm{Var}[\log \cdot_*^{(i)}]$ are estimated thanks to Fisher information and the delta method. Estimation of $\mathrm{Var}[\log C^{(j)}]$ is a special case and depends on the method for estimating the at-site index flood. Nevertheless, it is always possible to carry out an estimation of this variance, at least through standard errors.

#### 2.1.2. Revised Prior Distribution

[21] The initial prior distribution $\pi_{\mathrm{in}}(\boldsymbol{\theta})$ gives a null probability to the sub-sample $\Theta_0$. Thus, from this initial prior $\pi_{\mathrm{in}}(\boldsymbol{\theta})$, a revised prior $\pi(\boldsymbol{\theta})$ is constructed to attribute a non-null probability to the $\Theta_0$ sub-sample. According to *Stephenson and Tawn* [2004], $\pi(\boldsymbol{\theta})$ is defined as:

$$\pi(\boldsymbol{\theta}) = \begin{cases} (1-p_\xi)\pi_{\mathrm{in}}(\boldsymbol{\theta}) & \text{for } \boldsymbol{\theta} \in \Theta\backslash\Theta_0 \\ p_\xi \pi_{\xi_{\mathrm{Fix}}}(\boldsymbol{\theta}) & \text{for } \boldsymbol{\theta} \in \Theta_0. \end{cases} \tag{13}$$

where $p_\xi \in [0, 1]$ and with

$$\pi_{\xi_{\mathrm{Fix}}}(\boldsymbol{\theta}) = \frac{\pi_{\mathrm{in}}(\mu, \sigma, \xi_{\mathrm{Fix}})}{\int_{\mu,\sigma}\pi_{\mathrm{in}}(\mu, \sigma, \xi_{\mathrm{Fix}})\mathrm{d}\mu\mathrm{d}\sigma} \tag{14}$$

for $\boldsymbol{\theta} \in \Theta_0$. The integral in equation (2) can be easily evaluated by standard numerical integration methods.

[22] By construction, the new prior distribution $\pi(\boldsymbol{\theta})$ gives the required probability to the sub-space $\Theta_0$. *Stephenson and Tawn* [2004] have already applied formulations (13) and (14) with success for sea level maxima and rainfall threshold exceedences.

### 2.2. Posterior Estimation

[23] As it is often the case in Bayesian analysis, the integral in equation (2) is insolvable analytically. MCMC techniques are used to overcome this problem. Yet, due to the duality of $\pi(\boldsymbol{\theta})$ distribution, standard Metropolis-Hastings [*Hastings*, 1970] within Gibbs [*Geman and Geman*, 1984] methods are not sufficient. Reversible jump techniques [*Green*, 1995] are used to allow moves from the two-dimensional space $\Theta_0$ to the three-dimensional space $\Theta\backslash\Theta_0$ and vice versa.

[24] The classical Bayesian analysis, on $\Theta\backslash\Theta_0$, is performed with Gibbs cycle over each component of $\boldsymbol{\theta}$ using Metropolis-Hastings updates, with random walk proposals [*Coles and Tawn*, 1996].

[25] *Stephenson and Tawn* [2004] extended this algorithm to incorporate the mass on the Gumbel/Exponential case. However, as our approach does not only focus on the $\xi_{\mathrm{Fix}} = 0$ case, a new algorithm must be implemented. To help understand the algorithmic developments, some details about the classical Metropolis-Hastings algorithm and the reversible jump case are reported in Appendix A.

[26] The proposed algorithm must deal with two-dimensional changes: a change to $\Theta_0$ from $\Theta\backslash\Theta_0$ space and vice versa. These two types of special moves must be defined cautiously. As inspired by *Stephenson and Tawn* [2004], quantiles associated to a non-exceedence probability $p$ are set to be equal at current state $\boldsymbol{\theta}_t$ and proposal $\boldsymbol{\theta}_{\mathrm{prop}}$, $p$ being fixed.

[27] For a proposal move to $\Theta\backslash\Theta_0$ from $\Theta_0$, i.e., $\xi_t = \xi_{\mathrm{Fix}}$ and a proposal shape $\xi_{\mathrm{prop}} \neq \xi_{\mathrm{Fix}}$, the candidate move is to change $\boldsymbol{\theta}_t = (\mu_t, \sigma_t, \xi_t)$ to $\boldsymbol{\theta}_{\mathrm{prop}} = (\mu_{\mathrm{prop}}, \sigma_{\mathrm{prop}}, \xi_{\mathrm{prop}})$ where

$$\mu_{\mathrm{prop}} = \mu_t \tag{15a}$$

$$\sigma_{\mathrm{prop}} = \sigma_t \frac{\xi_{\mathrm{prop}}\left(y^{-\xi_t} - 1\right)}{\xi_t\left(y^{-\xi_{\mathrm{prop}}} - 1\right)} \tag{15b}$$

$$\xi_{\mathrm{prop}} \sim \mathcal{N}\left(\tilde{\xi}, s_\xi^2\right) \tag{15c}$$

where $y = 1 - p$, $p$ being fixed, $\tilde{\xi}$ is taken to be the mode of the marginal distribution for $\xi$ when there is no mass on

$\Theta_0$ [*Stephenson and Tawn*, 2004], and $s_\xi$ is the standard deviation selected to give good mixing properties to the chain. As it is usually the case with Metropolis-Hastings updates, this move is accepted with probability min(1, $\Delta$) with

$$\Delta = \frac{\pi\left(\mu_{\text{prop}}, \sigma_{\text{prop}}, \xi_{\text{prop}}|x\right)}{\pi(\mu_t, \sigma_t, \xi_{\text{Fix}}|x)} \frac{p_\xi}{1-p_\xi} \left[\phi\left(\xi_{\text{prop}}; \tilde{\xi}, s_\xi^2\right) J_{\xi_{\text{Fix}}}\left(\xi_{\text{prop}}\right)\right]^{-1} \tag{16}$$

where $\phi(\cdot; m, s^2)$ denotes the density function of the normal distribution with mean $m$ and variance $s^2$, and $J_{\xi_{\text{Fix}}}$ is the Jacobian of the parameter transformation for quantile matching, that is:

$$J_{\xi_{\text{Fix}}}(\xi) = \frac{\xi_{\text{Fix}}}{\xi} \frac{y^{-\xi} - 1}{y^{-\xi_{\text{Fix}}} - 1} \tag{17}$$

If the move is accepted, then $\boldsymbol{\theta}_{t+1} = (\mu_{\text{prop}}, \sigma_{\text{prop}}, \xi_{\text{prop}})$, else $\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t$.

[28] For a proposal move to $\Theta_0$ from $\Theta \backslash \Theta_0$, i.e., $\xi_t \neq \xi_{\text{Fix}}$ and a proposal shape $\xi_{\text{prop}} = \xi_{\text{Fix}}$, the proposal is to change $\boldsymbol{\theta}_t = (\mu_t, \sigma_t, \xi_t)$ to $\boldsymbol{\theta}_{\text{prop}} = (\mu_{\text{prop}}, \sigma_{\text{prop}}, \xi_{\text{prop}})$ where

$$\mu_{\text{prop}} = \mu_t \tag{18a}$$

$$\sigma_{\text{prop}} = \sigma_t \frac{\xi_{\text{prop}}\left(y^{-\xi_t} - 1\right)}{\xi_t\left(y^{-\xi_{\text{prop}}} - 1\right)} \tag{18b}$$

$$\xi_{\text{prop}} = \xi_{\text{Fix}} \tag{18c}$$

This move is accepted with probability min (1, $\Delta$) where

$$\Delta = \frac{\pi\left(\mu_{\text{prop}}, \sigma_{\text{prop}}, \xi_{\text{Fix}}|x\right)}{\pi(\mu_t, \sigma_t, \xi_t|x)} \frac{1-p_\xi}{p_\xi} \phi\left(\xi_t; \tilde{\xi}, s_\xi^2\right) J_{\xi_{\text{Fix}}}(\xi_t) \tag{19}$$

If the move is accepted, then $\boldsymbol{\theta}_{t+1} = (\mu_{\text{prop}}, \sigma_{\text{prop}}, \xi_{\text{prop}})$ else $\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t$.

[29] Obviously, special moves introduced in this study are not the only conceivable ones. Other reversible jumps can be explored, see for example the study of *Stephenson and Tawn* [2004]. However, for this application, the proposed moves seem to be particularly well suited. Indeed, a preliminary study shows that the location parameter was well estimated by a regional Bayesian approach. Thus a special move which only affects the shape and scale parameters should be consistent.

## 3. Generation Procedure

[30] In this section, the procedure implemented to generate stochastic homogeneous regions is described. The idea consists in generating sample points in a neighborhood of the L-moment space (Mean, L-CV, L-Skewness). The generation procedure can be summarized as follows:

[31] (1) Set the center of the neighborhood $(l_{1,R}, \tau_R, \tau_{3,R})$ or equivalently parameters of the regional distribution $(\mu_R, \sigma_R, \xi_R)$;

[32] (2) Generate $N$ points $(l_{1,i}, \tau_i, \tau_{3,i})$ uniformly in the sphere $\mathcal{B}((l_{1,R}, \tau_R, \tau_{3,R}); \varepsilon)$;

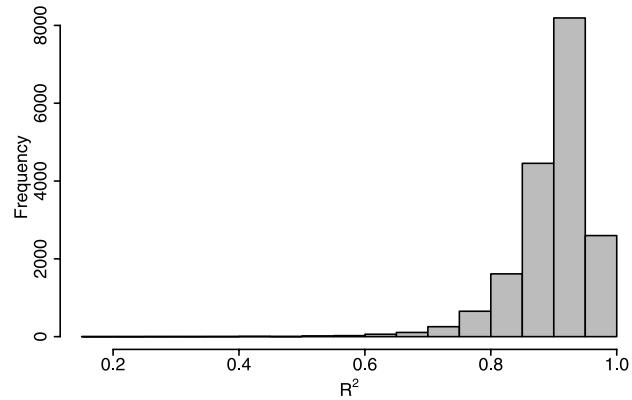


**Figure 1.** Histogram of the coefficient of determination for the regressive model (7). Application of section 5.

[33] (3) Generate $N$ index floods $C$ using the scaling model parametrization:

$$C = \alpha \text{Area}^\beta \tag{20}$$

Catchment areas are defined as realizations of a lognormal random variable.

[34] (4) For each $(l_{1,i}, \tau_i, \tau_{3,i})$, compute adimensional parameters by:

$$\xi_i^* = \frac{3\tau_{3,i} - 1}{1 + \tau_{3,i}} \tag{21a}$$

$$\sigma_i^* = \left(\xi_i^* - 1\right)\left(\xi_i^* - 2\right) l_{1,i}\tau_i \tag{21b}$$

$$\mu_i^* = l_{1,i} - \frac{\sigma_i^*}{1 - \xi_i^*} \tag{21c}$$

[35] (5) Then, compute at-site parameters from:

$$\xi_i = \xi_i^* \tag{22a}$$

$$\sigma_i = C_i \sigma_i^* \tag{22b}$$

$$\mu_i = C_i \mu_i^* \tag{22c}$$

[36] (6) Simulate samples from a GPD with parameters $(\mu_i, \sigma_i, \xi_i)$.

[37] As a GLM is used to elicit the prior distribution, the scaling model (20) must be altered to avoid giving an advantage to the Bayesian approaches over the index flood model. For this purpose, a noise in relation (20) at step 3 is introduced. Thus areas are altered by adding uniform random variables varying in ($-0.5 \times$ Area, $0.5 \times$ Area).

[38] This distortion is necessary to ensure that the regressive model is not too competitive and is consistent with observations. Indeed, large deviations to the area-index flood relationship are often encountered in practice. In the following applications, $\alpha = 0.12$, $\beta = 1.01$ and Area $\sim$

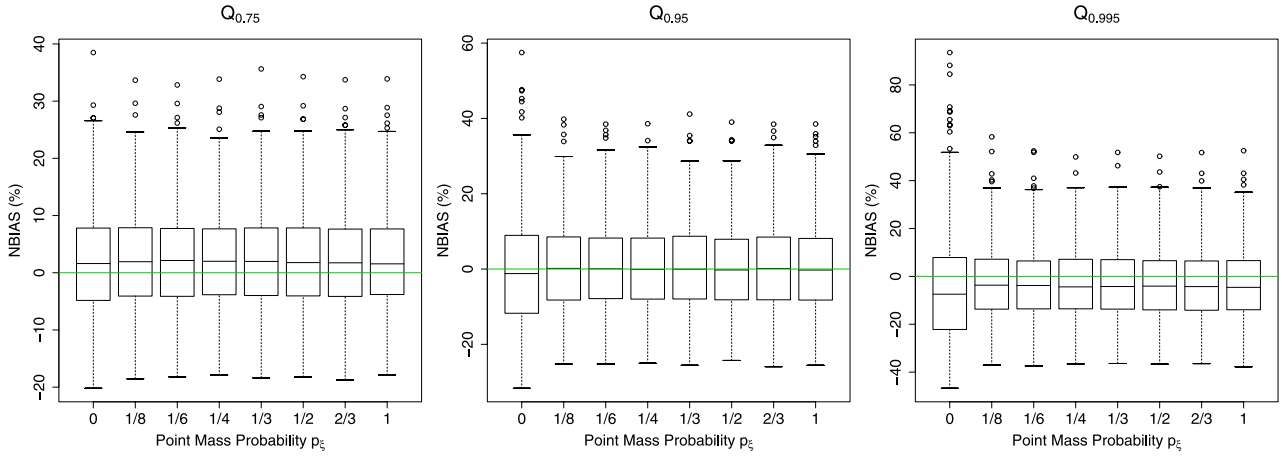


**Figure 2.** Effect of $p_\xi$ value on quantile estimation with non-exceedence probabilities 0.75, 0.95 and 0.995. Sample size 10. $\xi_{\text{Fix}} = 0.26$.

LN(4.8, 1). These values arise from a previous study on a French data set [*Ribatet et al.*, 2007] and ensure realistic magnitudes. For the application of section 5, the coefficients of determination for the regressive model (7) vary from 0.20 to 0.99, with a mean value of 0.89. The histogram of these coefficients of determination is presented in Figure 1. The radius $\varepsilon$ in the generation algorithm is set to 0.04. This value is chosen to reflect variability met in practice while preserving a low dispersion around the regional distribution. The $\varepsilon$ value primarily impacts the proportions of regions satisfying $H_1 < 1$. For specific applications, regions with a heterogeneity statistic $H_1$ such as $H_1 > 1$ may be discarded.

## 4. Sensitivity Analysis

[39] In this section, a sensitivity analysis for the algorithm introduced in section 2.2 is carried out. The primary goal is to check if results are not too impacted by the choice of the two user-selectable parameters $p_\xi$ and $\xi_{\text{Fix}}$. For this purpose, the effect of both $p_\xi$ and $\xi_{\text{Fix}}$ values on estimates and

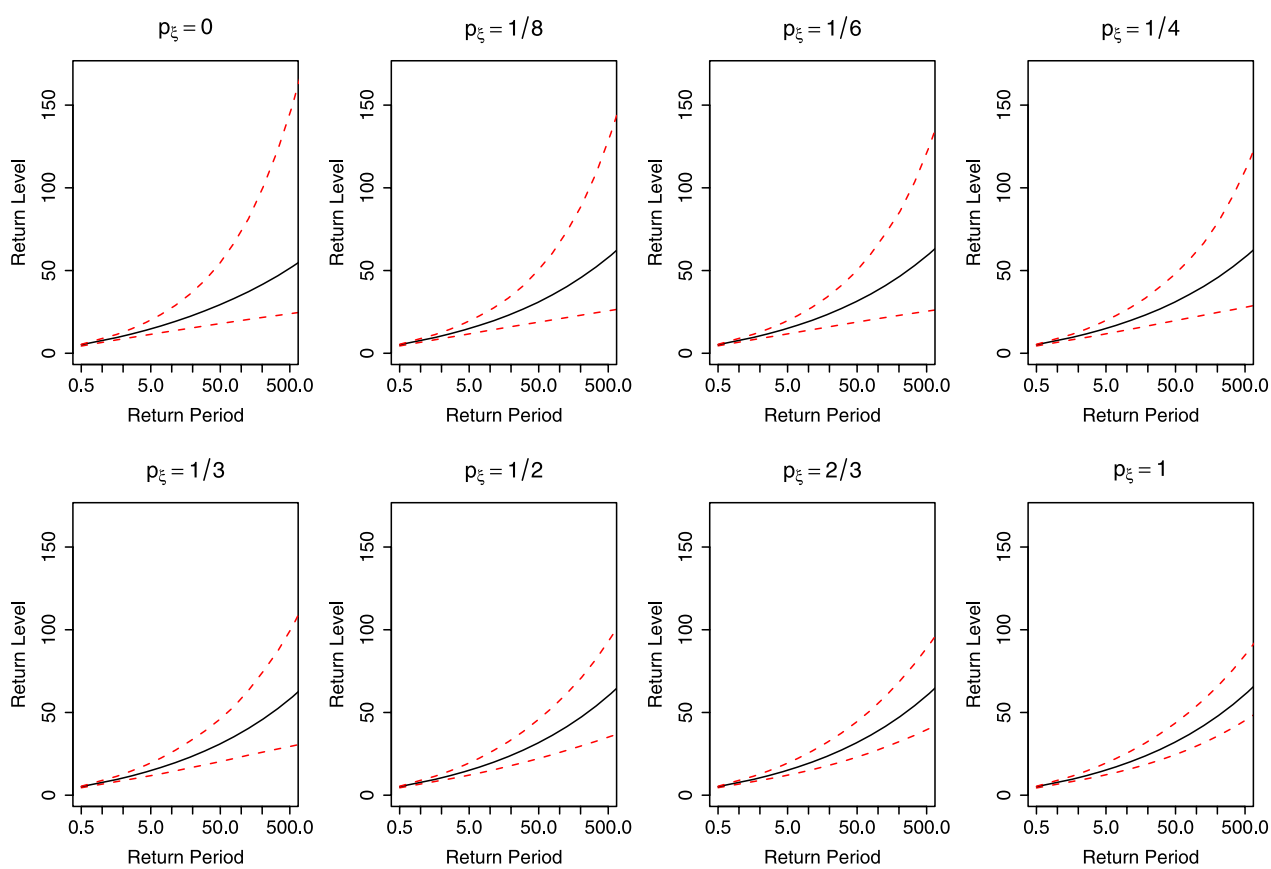


**Figure 3.** Effect of $p_\xi$ value on 90% posterior credibility interval. Sample size 10.

**Table 1.** Posterior Proportions (in Percent) of Events $\{\boldsymbol{\theta} \in \Theta_0\}$ for Different Values of $p_\xi$ and $\xi_{\text{Fix}}$[a]

| $\xi_{\text{Fix}}$ Features | | $p\xi$ Values | | | | | |
|---|---|---|---|---|---|---|---|
| $R_{\text{Shape}}$ | $D_{\text{Shape}}$ | 1/8 | 1/6 | 1/4 | 1/3 | 1/2 | 2/3 |
| −0.50 | 2e−5 | 0.00 | 0.03 | 0.00 | 0.00 | 0.05 | 0.00 |
| 0.00 | 0.06 | 10.07 | 14.55 | 17.27 | 21.99 | 41.53 | 61.84 |
| 0.50 | 0.70 | 38.88 | 46.94 | 59.96 | 67.42 | 81.88 | 92.17 |
| 0.83 | 1.00 | 46.21 | 57.33 | 67.53 | 76.08 | 85.33 | 92.20 |
| 1.00 | 0.87 | 48.24 | 55.14 | 68.90 | 76.16 | 86.05 | 91.85 |
| 1.50 | 0.41 | 32.72 | 45.61 | 54.62 | 66.18 | 82.11 | 89.90 |
| 2.00 | 0.10 | 22.95 | 22.83 | 35.06 | 49.82 | 57.86 | 81.92 |
| 2.50 | 0.01 | 13.93 | 7.04 | 9.86 | 36.21 | 38.89 | 42.28 |

[a]Target sample size 60.

credibility intervals is examined. For this sensitivity analysis, the parameter vector of the regional distribution is set to (0.64, 0.48, 0.26). The regions have 20 sites with a sample size of 70. For the whole sensitivity analysis, 10,000 regions were generated. The target site has a sample size of 10. We concentrate on estimates at sites with very few data, to exhibit the main differences in the most restricting configuration. Other configurations were found to demonstrate features similar to Figure 2 and Figure 4.

### 4.1. Effect of $P_\xi$

[40] The evolution of the normalized biases (expressed in percent) for return levels with non-exceedence probabilities 0.75, 0.95, and 0.995 associated to several $p_\xi$ values are depicted in Figure 2. Each boxplot is obtained from at-site estimates computed on more than 365 stochastic homogeneous regions. The case $p_\xi = 0$ corresponds to a classical Bayesian approach free from any point mass. In addition, to analyze only the effect of the parameter $p_\xi$, $\xi_{\text{Fix}}$ is temporarily fixed to the theoretical regional shape parameter.

[41] From Figure 2, the quantile estimates distribution seems to be stationary, provided that $p_\xi > 0$. Introducing a point mass does not impact $Q_{0.75}$ estimates, whereas significant reduction in median biases and scatter of estimates is noticeable for more extremal quantiles.

[42] Figure 3 shows the posterior distributions of return levels and 90% posterior credibility intervals for several $p_\xi$ values.

[43] It is clear that credibility intervals are sensitive to the $p_\xi$ value. This result is consistent as more and more proposals in the MCMC simulation belong to $\Theta_0$ as $p_\xi$ increases. Thus, by construction, the Markov chain is less variable. As denoted by *Stephenson and Tawn* [2004], the special case $p_\xi = 1$ is particular as uncertainty in the shape parameter is not considered. In that case, credibility intervals could be falsely narrow.

### 4.2. Effect of $\xi_{\text{Fix}}$

[44] It is important to analyze the influence of the choice of $\xi_{\text{Fix}}$ on the simulated Markov chains and thus its impact on estimations. Indeed, when specifying an unreasonable $\xi_{\text{Fix}}$ value, the estimations must not differ significantly from the conventional Bayesian ones. For this purpose, Table 1 displays the posterior proportions of events $\{\boldsymbol{\theta} \in \Theta_0\}$ for several $\xi_{\text{Fix}}$ and $p\xi$ values. This table is obtained with a target site sample size of 60. For each specified $\xi_{\text{Fix}}$ value, two features are computed to measure the relevance of the $\xi_{\text{Fix}}$ value: (a) $R_{\text{Shape}}$ the ratio of $\xi_{\text{Fix}}$ to the true shape parameter and (b) $D_{\text{Shape}}$ the ratio of the marginal posterior density from a conventional Bayesian analysis evaluated in $\xi_{\text{Fix}}$ and $\tilde{\xi}$.

[45] $R_{\text{Shape}}$ characterizes how much the point mass differs from the true value. $D_{\text{Shape}}$ quantifies the distance of $\xi_{\text{Fix}}$ from the estimator of the shape parameter proposed by *Ribatet et al.* [2007]. Thus, from these two statistics, consistency of the posterior proportions with deviations from theoretical and empirical values can be analyzed.

[46] The results in Table 1 show that values of $\xi_{\text{Fix}}$ that are not consistent with the data imply low proportions of state in $\Theta_0$. Thus, for such values, the proposed model is quite similar to a conventional Bayesian analysis. However, for two different values of $\xi_{\text{Fix}}$ ($R_{\text{Shape}}$ equal to 0.83 and 1), the posterior proportions are quite equivalent. This empha-

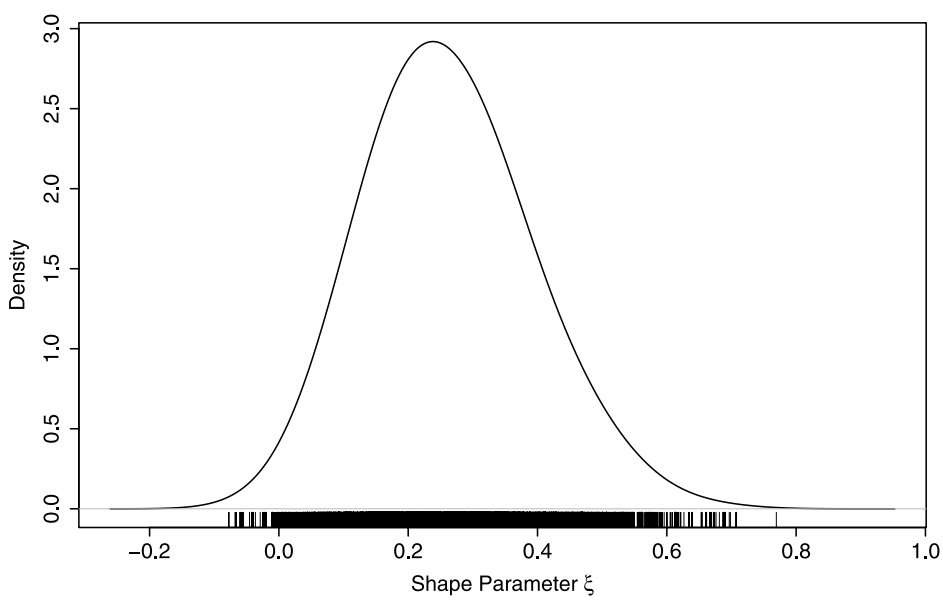


**Figure 4.** Posterior marginal density for the shape parameter.

**Table 2.** Characteristics of the Sixth Case Studies[a]

| | $(\mu_R, \sigma_R, \xi_R)$ | $N_{\text{Site}}$ | $(n_{\text{Site}}, n_{\text{Size}})$ | $N_{\text{Events}}$ |
|---|---|---|---|---|
| Conf1 | (0.64, 0.48, 0.26) | 10 | (9, 50) | 450 |
| Conf2 | (0.64, 0.48, 0.26) | 20 | (9, 30) × (10, 18) | 450 |
| Conf3 | (0.64, 0.48, 0.26) | 15 | (14, 50) | 700 |
| Conf4 | (0.66, 0.48, 0.08) | 10 | (9, 50) | 450 |
| Conf5 | (0.66, 0.48, 0.08) | 20 | (9, 30) × (10, 18) | 450 |
| Conf6 | (0.66, 0.48, 0.08) | 15 | (14, 50) | 700 |

[a]The target site is omitted in the couple $(n_{\text{Site}}, n_{\text{Size}})$ and has a sample size of 10, 25, and 40.

sizes the large uncertainty on the shape parameter estimation for small sample sizes. Uncertainty on the shape parameter estimation is also corroborated by the posterior marginal distribution of a conventional Bayesian analysis (see Figure 4).

[47] As noticed above, these results are obtained with a target site sample size of 60. This particular sample size was selected as it is the most illustrative case. However, the posterior proportions are quite similar when dealing with other target site sample sizes, even for very small sample sizes, this is less noticeable.

## 5. Simulation Study

[48] In this section, performance of three different estimators is analyzed: a conventional Bayesian estimator (BAY) introduced by *Ribatet et al.* [2007], the proposed estimator based on reversible jumps (REV), and the index flood estimator (IFL). In particular, the BAY estimator is related to the initial prior distribution defined in section 2.1.1. Thus the BAY estimator is identical to the REV approach with $p_\xi = 0$.

[49] For the proposed estimator, the point Mass probability $p_\xi$ was set to be a function of the $H_1$ statistic of *Hosking and Wallis* [1997]; that is:

$$p_\xi = \frac{\exp(-H_1)}{1 + \exp(-H_1)} \tag{23}$$

[50] For this parametrization, necessary requirements are satisfied; that is, $p_\xi \to 0$ when $H_1 \to +\infty$ and $p_\xi \to 1$ when $H_1 = -\infty$. Moreover, for $H_1 = 0$, $p_\xi = 0.5$ which corresponds to the estimator introduced by *Stephenson and Tawn* [2004]. Note that $p_\xi$ in equation (23) is defined with the negative inverse of the so-called so called *logit* function.

[51] Thus, for this choice, as underlined by the sensitivity analysis, credibility intervals are related to the degree of confidence of the point mass $\xi_{\text{Fix}}$ to be the true shape parameter and implicitly to the level of homogeneity of the regions.

[52] In addition, the non-exceedence probability $p$ used for quantiles matching in our algorithm (see section 2.2) is equal to $1 - 1/2n$, where $n$ is the target site sample size. This last point guarantees that quantiles associated with non-exceedence probability $1 - 1/2n$ for both proposal and current state of the Markov chain are identical. Other choices for $p$ are arguable. Here we introduce a quantile matching equation for a value closely related to the scale parameter and for which uncertainties are not too large.

[53] The analysis was performed on six different case studies summarized in Table 2. The configurations differ by the way information is distributed in space; that is, (a) "small regions" with well-instrumented but few sites (Conf1 and Conf4); (b) "large regions" with less instrumented and numerous sites (Conf2 and Conf5); and (c) "medium regions" with well-instrumented sites and an intermediate number of gauging stations. Conf1 (resp. Conf2, Conf3) correspond to Conf4 (resp. Conf5, Conf6) apart from the $(\mu_R, \sigma_R, \xi_R)$ values. The target site sample size takes the values in 10, 25, and 40. 1000 Regions were generated for each configuration. Markov chains with a

**Table 3.** Performance of BAY and IFL Estimators for Quantile $Q_{0.75}$, $Q_{0.95}$, and $Q_{0.995}$[a]

| | $Q_{0.75}$ | | | $Q_{0.95}$ | | | $Q_{0.995}$ | | |
|---|---|---|---|---|---|---|---|---|---|
| Model | NBIAS | SD | NMSE | NBIAS | SD | NMSE | NBIAS | SD | NMSE |
| | | | | *Conf1* | | | | | |
| BAY | 0.015 | 0.123 | 0.015 | 0.001 | 0.187 | 0.035 | −0.006 | 0.318 | 0.101 |
| IFL | 0.037 | 0.189 | 0.037 | 0.025 | 0.195 | 0.038 | −0.004 | 0.230 | 0.053 |
| | | | | *Conf2* | | | | | |
| BAY | 0.019 | 0.122 | 0.015 | 0.030 | 0.249 | 0.063 | 0.110 | 0.561 | 0.326 |
| IFL | 0.041 | 0.183 | 0.035 | 0.025 | 0.191 | 0.037 | −0.022 | 0.221 | 0.049 |
| | | | | *Conf3* | | | | | |
| BAY | 0.019 | 0.110 | 0.012 | 0.006 | 0.174 | 0.030 | −0.003 | 0.292 | 0.085 |
| IFL | 0.035 | 0.188 | 0.037 | 0.025 | 0.195 | 0.039 | −0.002 | 0.222 | 0.049 |
| | | | | *Conf4* | | | | | |
| BAY | 0.009 | 0.104 | 0.011 | −0.007 | 0.149 | 0.022 | −0.021 | 0.233 | 0.054 |
| IFL | 0.023 | 0.157 | 0.025 | 0.022 | 0.163 | 0.027 | 0.022 | 0.192 | 0.037 |
| | | | | *Conf5* | | | | | |
| BAY | 0.018 | 0.109 | 0.012 | 0.012 | 0.193 | 0.037 | 0.033 | 0.378 | 0.144 |
| IFL | 0.036 | 0.168 | 0.029 | 0.033 | 0.173 | 0.031 | 0.024 | 0.197 | 0.039 |
| | | | | *Conf6* | | | | | |
| BAY | 0.024 | 0.103 | 0.011 | 0.001 | 0.151 | 0.023 | −0.038 | 0.222 | 0.050 |
| IFL | 0.028 | 0.168 | 0.029 | 0.028 | 0.177 | 0.032 | 0.028 | 0.202 | 0.042 |

[a]Target site sample size: 10.

**Table 4.** Performance of BAY and REV Estimators for Quantile $Q_{0.75}$, $Q_{0.95}$, and $Q_{0.995}$[a]

| | $Q_{0.75}$ | | | $Q_{0.95}$ | | | $Q_{0.995}$ | | |
|---|---|---|---|---|---|---|---|---|---|
| Model | NBIAS | SD | NMSE | NBIAS | SD | NMSE | NBIAS | SD | NMSE |
| | | | | *Conf1* | | | | | |
| BAY | 0.015 | 0.123 | 0.015 | 0.001 | 0.187 | 0.035 | −0.006 | 0.318 | 0.101 |
| REV | 0.011 | 0.119 | 0.014 | −0.012 | 0.159 | 0.026 | −0.046 | 0.213 | 0.047 |
| | | | | *Conf2* | | | | | |
| BAY | 0.019 | 0.122 | 0.015 | 0.030 | 0.249 | 0.063 | 0.110 | 0.561 | 0.326 |
| REV | 0.005 | 0.105 | 0.011 | −0.026 | 0.154 | 0.024 | −0.066 | 0.269 | 0.077 |
| | | | | *Conf3* | | | | | |
| BAY | 0.019 | 0.110 | 0.012 | 0.006 | 0.174 | 0.030 | −0.003 | 0.292 | 0.085 |
| REV | 0.014 | 0.103 | 0.011 | −0.008 | 0.139 | 0.019 | −0.042 | 0.185 | 0.036 |
| | | | | *Conf4* | | | | | |
| BAY | 0.009 | 0.104 | 0.011 | −0.007 | 0.149 | 0.022 | −0.021 | 0.233 | 0.054 |
| REV | 0.010 | 0.102 | 0.011 | 0.002 | 0.136 | 0.018 | −0.001 | 0.182 | 0.033 |
| | | | | *Conf5* | | | | | |
| BAY | 0.018 | 0.109 | 0.012 | 0.012 | 0.193 | 0.037 | 0.033 | 0.378 | 0.144 |
| REV | 0.013 | 0.097 | 0.010 | 0.000 | 0.126 | 0.016 | −0.014 | 0.171 | 0.030 |
| | | | | *Conf6* | | | | | |
| BAY | 0.024 | 0.103 | 0.011 | 0.001 | 0.151 | 0.023 | −0.038 | 0.222 | 0.050 |
| REV | 0.031 | 0.099 | 0.011 | 0.033 | 0.133 | 0.019 | 0.034 | 0.174 | 0.032 |

[a]Target site sample size: 10.

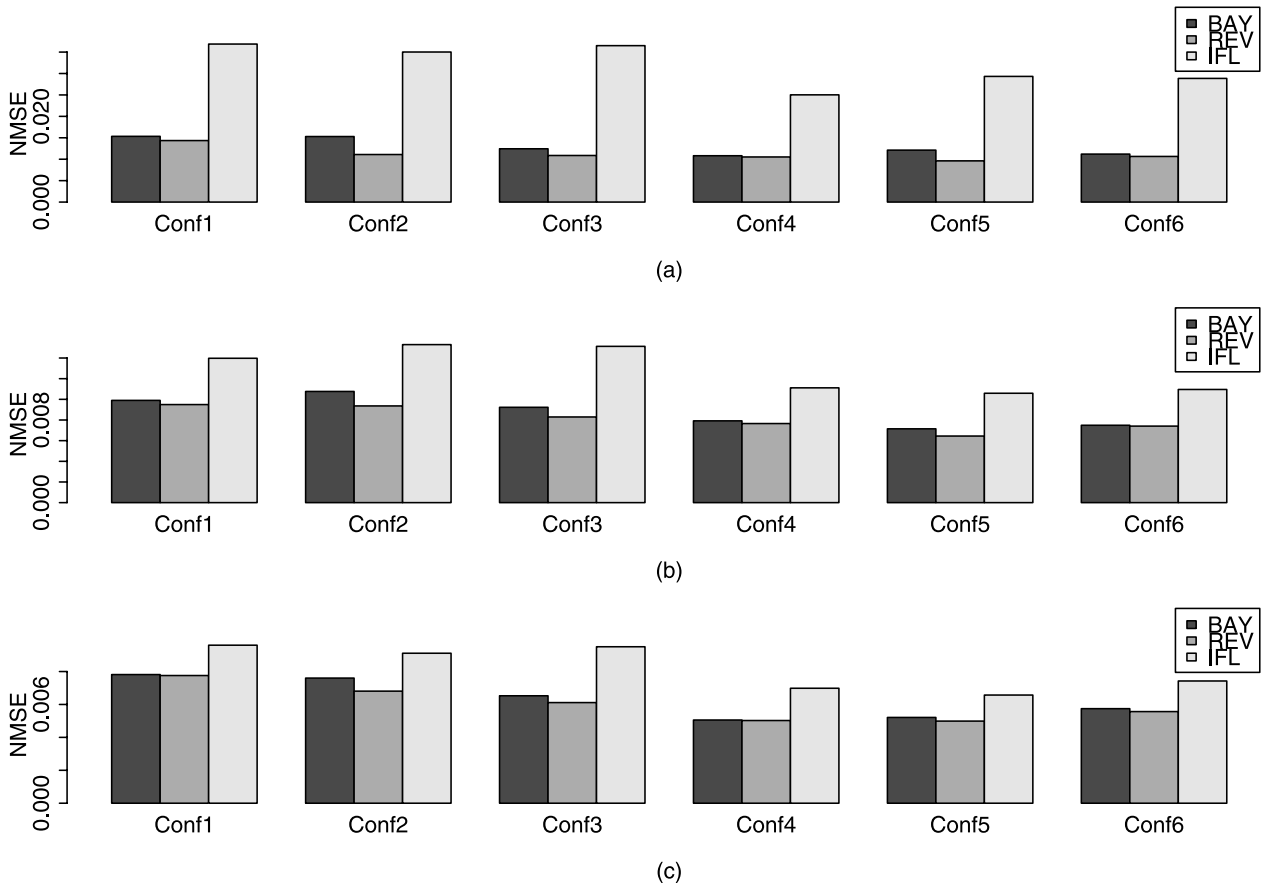


**Figure 5.** Evolution of the NMSE for quantile $Q_{0.75}$ in function of the region configuration. Target site sample size: (a) 10, (b) 25, and (c) 40.

length of 15,000 were generated. To ensure good mixing properties for all simulated Markov chains, an automated trial and error process was used to define proposal standard deviations of the MCMC algorithm. Furthermore, the first 2000 iterations were discarded to ensure that the equilibrium was reached.

[54] The performance of each estimator is assessed through the three following statistics:

$$\mathrm{NBIAS} = \frac{1}{k}\sum_{i=1}^{k}\frac{\hat{Q}_i - Q_i}{Q_i} \tag{24}$$

$$\mathrm{SD} = \sqrt{\frac{1}{k-1}\sum_{i=1}^{k}\left(\frac{\hat{Q}_i - Q_i}{Q_i} - \mathrm{NBIAS}\right)^2} \tag{25}$$

$$\mathrm{NMSE} = \frac{1}{k}\sum_{i=1}^{k}\left(\frac{\hat{Q}_i - Q_i}{Q_i}\right)^2 \tag{26}$$

where $\hat{Q}_i$ is the estimate of the theoretical value $Q_i$, and $k$ is the total number of theoretical values.

### 5.1. BAY vs. IFL Approach

[55] Table 3 shows that, for a small target site sample size and quantiles $Q_{0.75}$ and $Q_{0.95}$, the BAY approach is more competitive than the IFL one. Indeed, the three BAY statistics (NBIAS, SD, NMSE) are smaller than the ones related to IFL. However, for Conf2 and Conf5, IFL $Q_{0.95}$ estimates are more competitive. These two case studies correspond to the same configuration, i.e., numerous sites with short records. IFL estimates for $Q_{0.995}$ are always more accurate than BAY for all configurations.

[56] These results indicate that the relative performance of BAY compared to IFL depends on the pooling group. Thus, for the BAY approach and quantiles $Q_{0.75}$ and $Q_{0.95}$, it seems preferable to work with less gauging stations but which have larger data series, independently of the target site sample size. The sensitivity to the configuration of the sites and the availability of long time series is a drawback for the application of this Bayesian approach.

[57] These conclusions obtained on stochastic regions are in line with a previous analysis on a French data set [*Ribatet et al.*, 2007]. The BAY approach is suited to work with "small" or "medium" regions and well-instrumented gauging stations. In addition, this approach is accurate for "reasonable" quantile estimation (see the bad performance of BAY for $Q_{0.995}$ in Table 3).

[58] However, the white noise introduced in the generation procedure is independent of the target site sample size. It only regards both Bayesian approaches. Thus the performance of the BAY estimator for large sample sizes may be too impacted. Indeed, while the IFL estimation procedure is

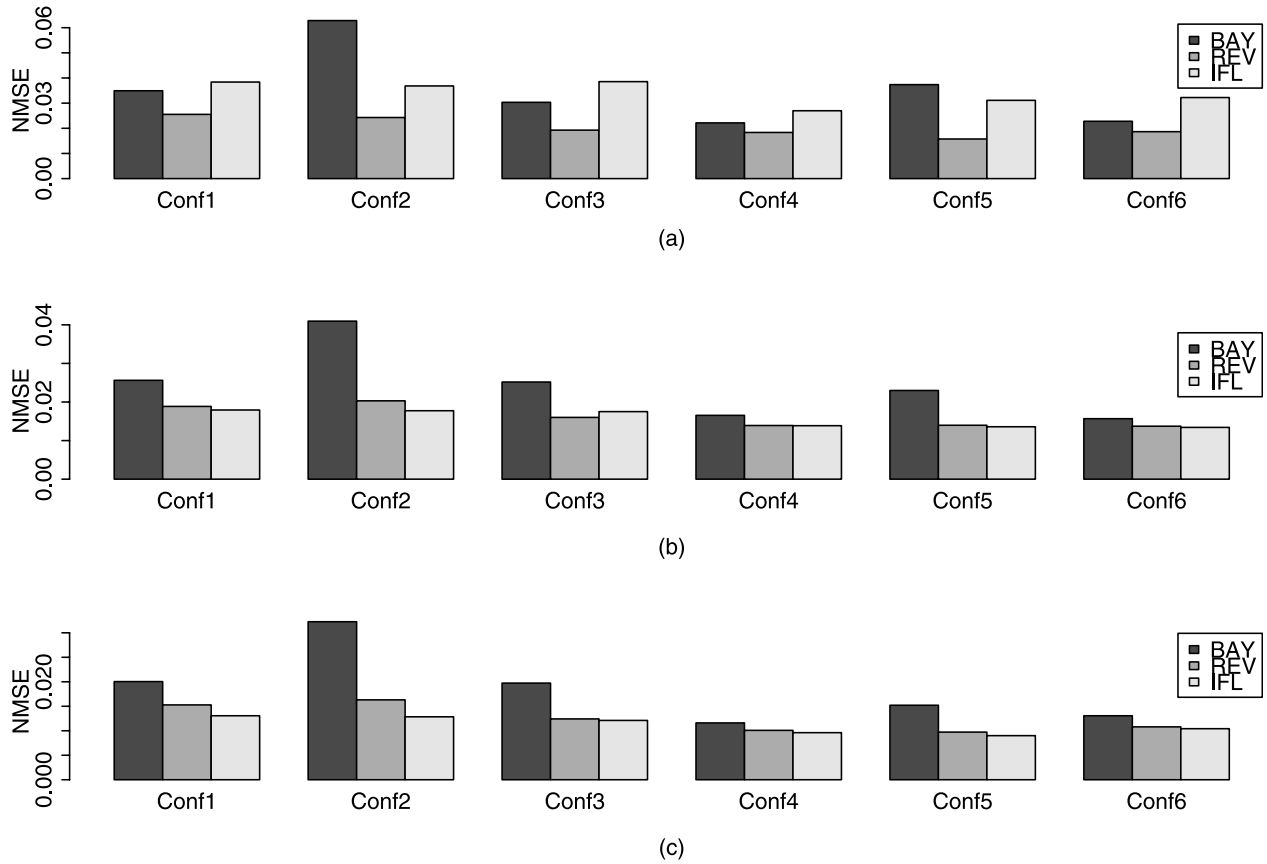


**Figure 6.** Evolution of the NMSE for quantile $Q_{0.95}$ in function of the region configuration. Target site sample size: (a) 10, (b) 25, and (c) 40.

not altered, both Bayesian approaches must deal with artificially generated biases.

[59] The main idea for the REV approach is to combine the good performance of the BAY estimator for "reasonable" quantiles and the efficiency of the IFL approach for larger quantiles.

### 5.2. BAY vs. REV Approach

[60] The comparison of the two Bayesian estimators is summarized in Table 4. REV leads to more accurate estimated quantiles, in particular for $Q_{0.95}$ and $Q_{0.995}$. This last point confirms the benefits of using a regional shape parameter through a reversible jump approach.

[61] By construction of the algorithm described in section 2.2, Markov chains generated from the REV approach are less variable than the ones generated from the BAY model. Thus REV is associated to smaller standard deviation than BAY whatever the configuration is (Table 4). Moreover, if the regional fixed shape parameter $\xi_{\mathrm{Fix}}$ is suited, REV should have the same biases than BAY. Thereby, the REV estimator always leads to a smaller NMSE.

### 5.3. Global Comparison

[62] Figures 5, 6 and 7 illustrate the results for different target site sample sizes and regions. We concentrate on the NMSE criteria since it measures variation of the estimator around the true parameter value.

[63] From Figure 5, it is clear that Bayesian estimations, i.e., BAY and REV, of $Q_{0.75}$ are more accurate, specially for a target site sample size of 10. For larger target site sample sizes, Bayesian approaches are always more competitive than the IFL estimator, even if this is less clear-cut on the graphs. Furthermore, BAY and REV estimators often have the same performance. This result is logical as the $Q_{0.75}$ value is mostly impacted by the location parameter $\mu$. Thus reversible jumps do not have a significant result on REV $Q_{0.75}$ estimation.

[64] The plots in Figure 6 and those displayed in Figure 5 are quite different. For a target site sample size of 10, both Bayesian approaches are the most accurate (except for BAY applied on Conf2 and Conf5), and the REV estimator leads always to the smallest NMSE. Thus REV is the most competitive model. For larger target site sample sizes, REV is at least as accurate as IFL, except for Conf2.

[65] For $Q_{0.995}$ and a target site sample size of 10, REV is the most accurate model, except for Conf2. As the target site sample size increases, the IFL approach becomes more efficient. However, for these cases, NMSE for the REV estimator are often close to the IFL ones. Although the BAY approach performs poorly for $Q_{0.995}$, its NMSE for Conf6 is close to the REV and IFL ones.

[66] In conclusion, these results illustrate the good overall performance of the *REV* model. Indeed, this approach benefits from the efficiency of the BAY estimator for quantiles with small non-exceedence probabilities while being as competitive as the IFL approach for larger non-exceedence probabilities.

[67] However, the Bayesian approaches outperform the index flood model, but differences in accuracy seem to be

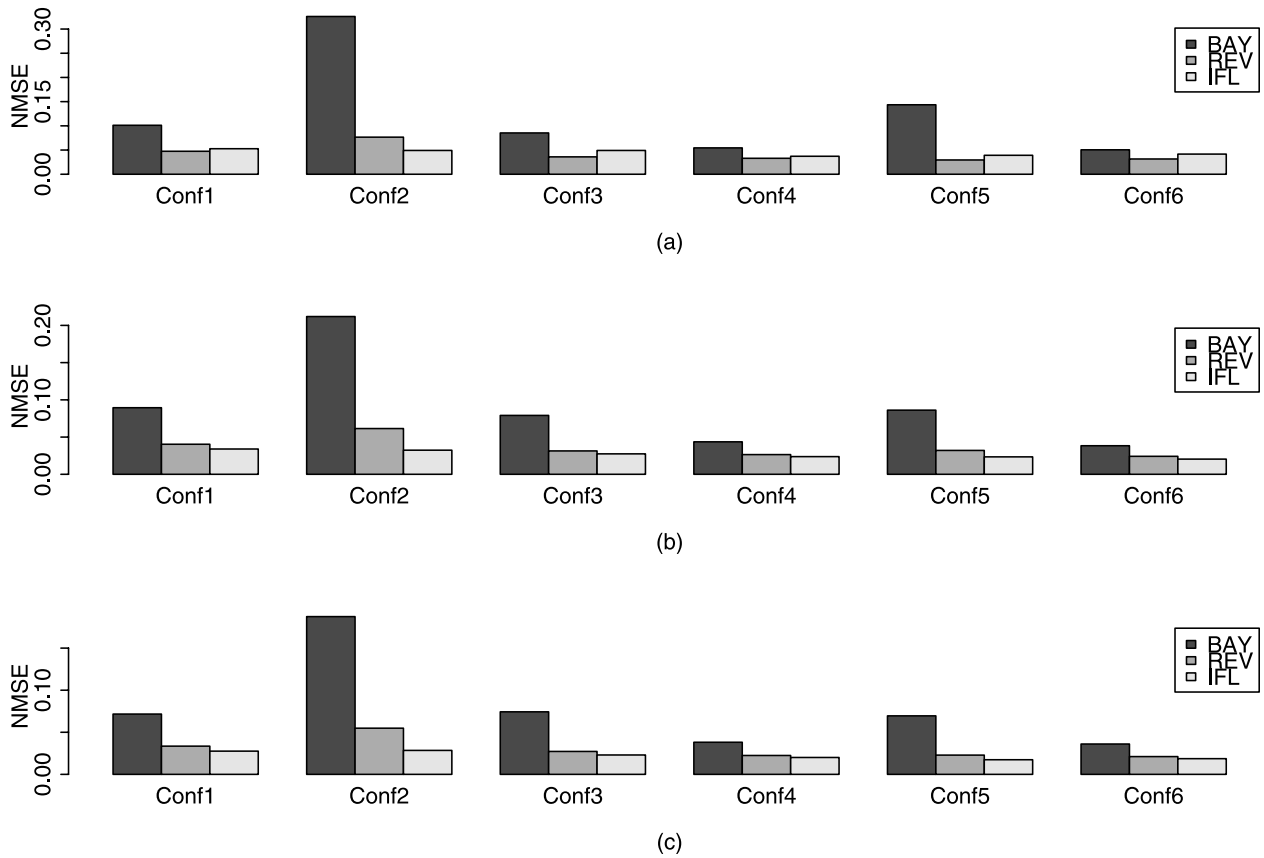


**Figure 7.** Evolution of the NMSE for quantile $Q_{0.995}$ in function of the region configuration. Target site sample size: (a) 10, (b) 25, and (c) 40.

less and less significant as the sample site increases. This may be related to the white noise introduced in the generation procedure. Indeed, this white noise is independent of the target site sample size and may strongly penalize the performances of the both Bayesian approaches. The next section tries to outline the effect of the target site index flood estimation error to the quantile estimates.

## 6. Effect of Bias on the Target Site Index Flood Estimation

[68] According to the model being considered, two types of biases are encountered for the target site index flood estimation. Indeed, on one hand, the index flood for the IFL model is derived from the target site sample. On the other hand, for BAY and REV approaches, the index flood is estimated from a scaling model. Thus biases on index flood estimation are due to the relevance of this scaling model but also to the index flood error estimation for the other sites within the region.

[69] To illustrate these two types of biases, the normalized bias on target site index flood estimation is computed as follows:

$$\mathrm{Bias}(C) = \frac{\hat{C} - C}{C} \tag{27}$$

where $C$ is the target site index flood, and $\hat{C}$ is an estimate of $C$. Figure 8 depicts changes in NBIAS for quantile $Q_{0.95}$ in function of Bias($C$). As normalized biases are considered, statistics for the six configurations are plotted in the same graphic. Solid lines correspond to local polynomial regression fits to help underline trends.

[70] Scatterplots in Figure 8 show clearly these two types of biases. Indeed, on one hand, the range of Bias($C$) is not the same for IFL than for BAY and REV, particularly for a target site sample size of 25 and 40. On the other hand, for the BAY and REV approaches, biases on index flood estimation are independent of the target site sample size; while this is not the case for IFL. This last point is also illustrated as the bias ranges for the Bayesian approaches remain the same for all target site sample size. Thus, for large sample size, efficiency of the Bayesian estimators may be too much impacted as the artificial bias introduced in the generation procedure is too penalizing.

[71] The Bayesian approaches do not have the same behaviour as the IFL model. In particular, BAY and REV seem to be less sensitive to a large bias in target site index flood estimation. NBIAS for the IFL model are clearly linear with a response $y = x$. This last point is an expected result. Indeed, apart from sampling variability, if a unique regional distribution exists, quantile IFL estimate biases are only induced by biases on target site index flood estimates. Thus the relevance of the generation procedure is corroborated.

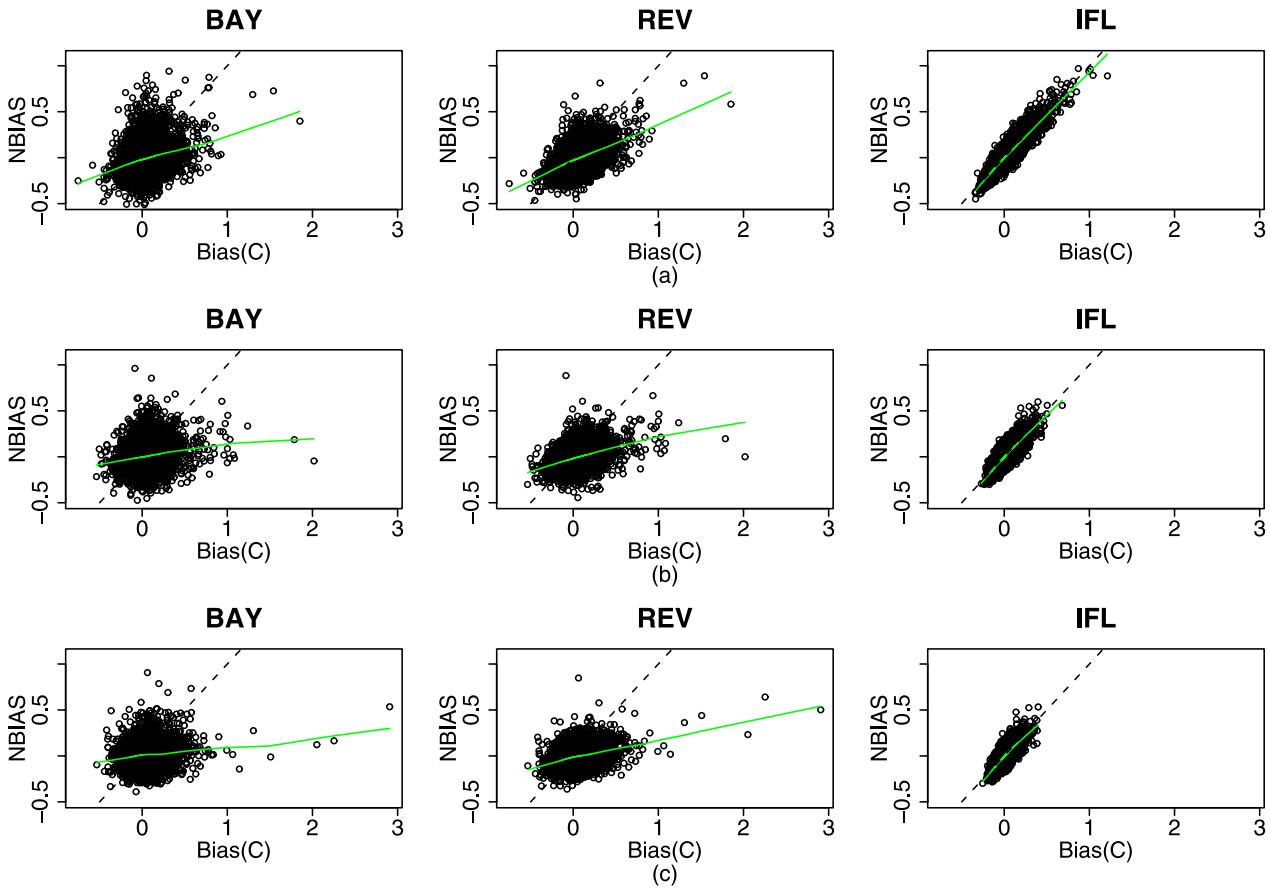


**Figure 8.** Evolution of NBIAS for $Q_{0.95}$ in function of the normalized bias on target site index flood estimation (Bias(C)). Target site sample size: (a) 10, (b) 25, and (c) 40. Solid green lines: local smoothers, black dashed lines: $y = x$.

[72] The main difference between the BAY and REV estimators is the dispersion around local smoothers. Indeed, REV has a smaller range while preserving the same robustness to the bias on target site index flood estimation.

[73] These results and conclusions are independent of the target site index flood estimation procedure. However, the performance of the two Bayesian estimators is related to the bias and variance of the target site index flood estimate. Thus, for similar variance, these results should be identical if general autoregressive modelings Generalized Additive Models (GAMs) or Kriging were used.

## 7. Suggestions for Region Configuration

[74] This section attempts to present some suggestions for building suitable pooling groups according to the considered estimator. *Hosking and Wallis* [1997] already advice not to build regions greater than 20 sites because of the small gain affected with additional stations. However, they only focus on the IFL methodology. We attempt to do the same for the two Bayesian estimators considered in this study. For this purpose, Tables 5, 6, and 7 include the NMSE and the related standard errors for each configuration and target site sample size.

[75] From Table 5, the IFL estimator seems to have the same performance level independently of the configuration. This result points out that the information is not used optimally as regions with the most information (i.e., Conf3 and Conf6) do not always lead to better estimations. This last point corroborates a previous comments of *Ribatet et al.* [2007].

[76] Table 6 shows that the BAY estimator is more accurate with "medium" regions, i.e., Conf3 and Conf6. However, results for "small" regions, i.e., Conf1 and Conf4, are often close to the best ones, especially for a light tail. Thus it is preferable to work with well-instrumented sites, i.e., Conf1, Conf3, Conf4, and Conf6.

[77] Table 7 shows that the REV estimator more efficient with "medium" regions, i.e., Conf3 and Conf6. In addition, it seems to be more accurate with few but well-instrumented gauging stations rather more but less instrumented ones. Nevertheless, for a light tail, all configurations seems to lead to similar performance levels.

[78] Tables 5, 6, and 7 show that the estimation of $Q_{0.75}$ is independent of the region configuration for all estimators. Thus it seems that the regional information is not relevant for quantiles with small non-exceedence probabilities.

## 8. Conclusions

[79] This article introduced a new Bayesian estimator which uses regional information in an innovative way. The proposed model accounts for a fixed regional shape parameter with a non-null probability. Thus, as in the study

**Table 5.** Changes in NMSE for $Q_{0.75}$, $Q_{0.95}$, and $Q_{0.995}$ in Function of the Region Configuration and the Target Site Sample Size for the IFL Estimator[a]

| | Heavy Tail | | | Light Tail | | |
|---|---|---|---|---|---|---|
| Model | Conf1 | Conf2 | Conf3 | Conf4 | Conf5 | Conf6 |
| | | | *Target Site Sample Size 10* | | | |
| $Q_{0.75}$ | 0.037 (3e − 3) | 0.035 (3e − 3) | 0.037 (3e − 3) | 0.025 (2e − 3) | 0.029 (2e − 3) | 0.029 (3e − 3) |
| $Q_{0.95}$ | 0.038 (3e − 3) | 0.037 (3e − 3) | 0.039 (3e − 3) | 0.027 (4e − 3) | 0.031 (2e − 3) | 0.032 (3e − 3) |
| $Q_{0.995}$ | 0.053 (4e − 3) | 0.049 (3e − 3) | 0.049 (4e − 3) | 0.037 (2e − 3) | 0.039 (3e − 3) | 0.042 (4e − 3) |
| | | | *Target Site Sample Size 25* | | | |
| $Q_{0.75}$ | 0.014 (8e − 4) | 0.015 (1e − 3) | 0.015 (1e − 3) | 0.011 (7e − 4) | 0.011 (7e − 4) | 0.011 (7e − 4) |
| $Q_{0.95}$ | 0.018 (1e − 3) | 0.018 (1e − 3) | 0.018 (1e − 3) | 0.014 (9e − 4) | 0.014 (9e − 4) | 0.013 (9e − 4) |
| $Q_{0.995}$ | 0.034 (2e − 3) | 0.032 (2e − 3) | 0.027 (2e − 3) | 0.024 (2e − 3) | 0.023 (2e − 3) | 0.020 (1e − 3) |
| | | | *Target Site Sample Size 40* | | | |
| $Q_{0.75}$ | 0.010 (6e − 4) | 0.009 (6e − 4) | 0.010 (6e − 4) | 0.007 (4e − 4) | 0.007 (4e − 4) | 0.007 (5e − 4) |
| $Q_{0.95}$ | 0.013 (8e − 4) | 0.013 (8e − 4) | 0.012 (8e − 4) | 0.010 (6e − 4) | 0.009 (5e − 4) | 0.010 (6e − 4) |
| $Q_{0.995}$ | 0.028 (2e − 3) | 0.028 (2e − 3) | 0.023 (2e − 3) | 0.020 (1e − 3) | 0.017 (1e − 3) | 0.019 (1e − 3) |

[a]Related standard errors are displayed in brackets.

of *Ribatet et al.* [2007], the regional information is still used to elicit the prior distribution. However, the prior distribution is now a mixture of a GEV/GPD and a GEV/GPD with only two parameters; the remaining one corresponds to the fixed regional shape parameter.

[80] The estimation procedure is achieved using reversible jump Markov chains [*Green*, 1995]; and theoretical details for simulated suited Markov chains were presented. A sensitivity analysis for the proposed algorithm was performed. The results showed that the estimates are consistent provided that the probability attributed to the fixed regional shape parameter is positive. In addition, as noticed by *Stephenson and Tawn* [2004], the credibility intervals are sensitive to this probability value. Thus the proposed estimator relates this probability value to the homogeneity degree of the region, using the heterogeneity statistic of *Hosking and Wallis* [1997]. Therefore, the credibility intervals take into account the belief about the fixed regional shape parameter to be the true value.

[81] A performance analysis was carried out on stochastic data for three different estimators. For this purpose, another algorithm which generates stochastic homogeneous regions was implemented. The good overall performance of the proposed estimator has been demonstrated. Indeed, on one hand, this approach combines the accuracy of the regional Bayesian approach of *Ribatet et al.* [2007] for quantiles associated to small exceedence probabilities. On the other hand, the duality of the prior distribution (and the fixed regional shape parameter) allows the proposed estimator to be at least as efficient as the index flood model. Thus this new estimator seems very suited for regional estimation when the target site is not well instrumented.

[82] Furthermore, the two Bayesian approaches considered here appear to be less sensitive to biases on target site index flood estimation than the index flood estimator. Thus the Bayesian approaches are more readily adaptable which is a major advantage as errors on the index flood estimation are often uncontrollable.

[83] As noticed by *Ribatet et al.* [2007], the index flood model does not use information optimally. This point is corroborated in this study as the model initiated by *Dalrymple* [1960] is not inevitably more accurate as the information within the pooling group increases. This is not the case for the Bayesian approaches. In addition, they seem to be more accurate when dealing with regions with well-instrumented sites, particularly for large quantiles.

[84] All statistical analyses were carried out by the use of the *R Development Core Team* [2006]. For this purpose, the

**Table 6.** Changes in *NMSE* for $Q_{0.75}$, $Q_{0.95}$, and $Q_{0.995}$ in Function of the Region Configuration and the Target Site Sample Size for the BAY Estimator[a]

| | Heavy Tail | | | Light Tail | | |
|---|---|---|---|---|---|---|
| Model | Conf1 | Conf2 | Conf3 | Conf4 | Conf5 | Conf6 |
| | | | *Target Site Sample Size 10* | | | |
| $Q_{0.75}$ | 0.015 (1e − 3) | 0.015 (1e − 3) | 0.012 (8e − 4) | 0.011 (6e − 4) | 0.012 (9e − 4) | 0.011 (9e − 4) |
| $Q_{0.95}$ | 0.035 (2e − 3) | 0.063 (4e − 3) | 0.030 (2e − 4) | 0.022 (1e − 3) | 0.037 (3e − 3) | 0.023 (2e − 3) |
| $Q_{0.995}$ | 0.101 (1e − 2) | 0.326 (3e − 2) | 0.085 (6e − 3) | 0.054 (5e − 3) | 0.144 (1e − 2) | 0.050 (3e − 3) |
| | | | *Target Site Sample Size 25* | | | |
| $Q_{0.75}$ | 0.010 (6e − 4) | 0.011 (7e − 4) | 0.009 (5e − 4) | 0.008 (5e − 4) | 0.007 (4e − 4) | 0.007 (5e − 4) |
| $Q_{0.95}$ | 0.026 (2e − 3) | 0.041 (3e − 3) | 0.025 (1e − 3) | 0.017 (1e − 3) | 0.023 (1e − 3) | 0.016 (9e − 4) |
| $Q_{0.995}$ | 0.089 (8e − 3) | 0.212 (2e − 2) | 0.079 (4e − 3) | 0.044 (3e − 3) | 0.086 (6e − 3) | 0.038 (2e − 3) |
| | | | *Target Site Sample Size 40* | | | |
| $Q_{0.75}$ | 0.008 (5e − 4) | 0.008 (5e − 4) | 0.007 (4e − 4) | 0.005 (3e − 4) | 0.005 (3e − 4) | 0.006 (4e − 4) |
| $Q_{0.95}$ | 0.020 (1e − 3) | 0.032 (2e − 3) | 0.020 (1e − 3) | 0.012 (8e − 4) | 0.015 (9e − 4) | 0.013 (8e − 4) |
| $Q_{0.995}$ | 0.072 (5e − 3) | 0.187 (2e − 2) | 0.074 (5e − 3) | 0.038 (3e − 3) | 0.070 (6e − 3) | 0.036 (2e − 3) |

[a]Related standard errors are displayed in brackets.

**Table 7.** Changes in NMSE for $Q_{0.75}$, $Q_{0.95}$, and $Q_{0.995}$ in the Function of the Region Configuration and the Target Site Sample Size for the REV Estimator[a]

| | Heavy Tail | | | Light Tail | | |
|---|---|---|---|---|---|---|
| Model | Conf1 | Conf2 | Conf3 | Conf4 | Conf5 | Conf6 |
| | | | *Target Site Sample Size 10* | | | |
| $Q_{0.75}$ | 0.014 (1e − 3) | 0.011 (7e − 4) | 0.011 (7e − 4) | 0.011 (6e − 4) | 0.010 (7e − 4) | 0.011 (9e − 4) |
| $Q_{0.95}$ | 0.026 (2e − 3) | 0.024 (2e − 3) | 0.019 (1e − 3) | 0.018 (1e − 3) | 0.016 (1e − 3) | 0.019 (2e − 3) |
| $Q_{0.995}$ | 0.047 (3e − 3) | 0.077 (2e − 2) | 0.036 (2e − 3) | 0.033 (2e − 3) | 0.030 (2e − 3) | 0.032 (3e − 3) |
| | | | *Target Site Sample Size 25* | | | |
| $Q_{0.75}$ | 0.009 (6e − 4) | 0.009 (6e − 4) | 0.008 (5e − 4) | 0.008 (5e − 4) | 0.006 (4e − 4) | 0.007 (5e − 4) |
| $Q_{0.95}$ | 0.019 (1e − 3) | 0.020 (2e − 3) | 0.016 (9e − 4) | 0.014 (1e − 3) | 0.014 (9e − 4) | 0.014 (9e − 4) |
| $Q_{0.995}$ | 0.040 (3e − 3) | 0.061 (1e − 2) | 0.031 (2e − 3) | 0.026 (2e − 3) | 0.032 (3e − 3) | 0.024 (2e − 3) |
| | | | *Target Site Sample Size 40* | | | |
| $Q_{0.75}$ | 0.008 (5e − 4) | 0.007 (5e − 4) | 0.006 (4e − 4) | 0.005 (3e − 4) | 0.005 (3e − 4) | 0.006 (3e − 4) |
| $Q_{0.95}$ | 0.015 (1e − 3) | 0.016 (1e − 3) | 0.012 (9e − 4) | 0.010 (7e − 4) | 0.010 (5e − 4) | 0.011 (6e − 4) |
| $Q_{0.995}$ | 0.034 (2e − 3) | 0.055 (1e − 2) | 0.027 (2e − 3) | 0.022 (2e − 3) | 0.023 (2e − 3) | 0.021 (1e − 3) |

[a]Related standard errors are displayed in brackets.

algorithm presented in section 2.2 was incorporated in the evdbayes packages [*Stephenson and Ribatet*, 2006]. The algorithm for the generation procedure is available on request from the author.

## Appendix A: The Metropolis-Hastings Algorithm

[85] In this section, the Metropolis-Hastings algorithm is presented. According to the results derived by *Green* [1995], some details will be given to consider the reversible jump case. The basic idea of the Metropolis-Hastings algorithm is to obtain a Markov chain that converges to a known stationary distribution. The strength of the Metropolis-Hasting approach is that the convergence is reached whatever the initial state of the Markov chain is and that the distributions could be known up to a constant.

[86] Let $f$ denote the target distribution of interest. Most often, in Bayesian inference, $\pi$ will be the posterior distribution for the parameters. Let $q(\cdot\,, x)$ be the proposal distribution, that is, the proposal states will be sampled from this proposal distribution given the current state $x_t$. The Metropolis-Hastings algorithm can be summarized as follows:

[87] 1. Generate $u$ from a uniform distribution on [0, 1]
[88] 2. Generate $x_{\text{prop}}$ from $q(\cdot\,, x_t)$
[89] 3. $\Delta_{\text{class}} \leftarrow \frac{f(x_{\text{prop}})}{f(x_t)} \frac{q(x_t|x_{\text{prop}})}{q(x_{\text{prop}}|x_t)}$
[90] 4. if $u < \min(1, \Delta_{\text{class}})$ then
[91] 5. $x_{t+1} \leftarrow x_{\text{prop}}$
[92] 6. else
[93] 7. $x_{t+1} \leftarrow x_t$
[94] 8. endif
[95] 9. Go to 1.

[96] The initial Metropolis-Hastings algorithm cannot account for dimensional switch. For this purpose, the "jumps" between sub-spaces must be defined (see equations (15a), (15b), and (15c) and (18a), (18b), and (18c)), and the quantity $\Delta_{\text{class}}$ must be redefined each time a jump is considered. Here only a simple case of the reversible jumps approach is considered (see section 3.3 of *Green*, 1995). If only two moves $m_1(x_t)$ and $m_2(x_t)$ can occur with probabilities $p_1$ and $p_2$, respectively, then the quantity $\Delta_{\text{class}}$ must be replaced by $\Delta_{\text{rev}}$. Consequently, for a proposal move of type $m_1$:

$$\Delta_{\text{rev}} = \Delta_{\text{class}} \frac{p_1}{p_2} J_1 \tag{A1}$$

where $J_1$ is the Jacobian of the transformation $x_t \mapsto m_1(x_t)$. If the proposal move is of type $m_2$, then

$$\Delta_{\text{rev}} = \Delta_{\text{class}} \frac{p_2}{p_1} J_2 \tag{A2}$$

where $J_2$ is the Jacobian of the transformation $x_t \mapsto m_2(x_t)$.

[97] **Acknowledgments.** The authors wish to thank Alec Stephenson for providing the original codes of his article. The financial support provided by the National Science and Engineering Research Council of Canada (NSERC) is acknowledge. We are also grateful to the editor, the associate editor, and two anonymous referees for their useful criticism of the original version of the paper.

## References

Bayes, T. (1763), An essay towardstowards solving a problem in the doctrine of chance, *Philosophical Transaction of the Royal Society*, *53*.
Bortot, P., and S. Coles (2000), The multivariate Gaussian tail model: An application to oceanographic data, *J. R. Stat. Soc., Ser. C*, *49*(1), 31–49.
Coles, S., and M. Dixon (1999), Likelihood-based inference for extreme value models, *Extremes*, *2*(1), 5–23.
Coles, S., and J. Tawn (1996), A Bayesian analysis of extreme rainfall data, *J. R. Stat. Soc., Ser. C*, *45*(4), 463–478.
Dalrymple, T. (1960), Flood frequency analysis, *U. S. Geol. Surv. Water Supply Pap.*, *1543 A.*
Geman, S., and D. Geman (1984), Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, *IEEE Trans. Pattern Anal. Mach. Intell., PAMI*, *6*(6), 721–741.
Green, P. (1995), Reversible jump Markov chain Monte Carlo computation and Bayesian model determination, *Biometrica*, *82*, 711–732.
Hastings, W. K. (1970), Monte Carlo sampling methods using Markov chains and their applications, *Biometrika*, *57*, 97–109.
Hosking, J., and J. Wallis (1987), Parameter and quantile estimation for the Generalized Pareto Distribution, *Technometrics*, *29*(3), 339–349.
Hosking, J. R. M., and J. R. Wallis (1997), *Regional Frequency Analysis*, Cambridge Univ. Press, New York.
Juárez, S., and W. Schucany (2004), Robust and efficient estimation for the Generalized Pareto Distribution, *Extremes*, *7*(3), 237–251.
Kuczera, G. (1982), Combining at-site and regional information: An empirical Bayes approach, *Water Resour. Res.*, *18*(2), 306–314.

Madsen, H., and D. Rosbjerg (1997), Generalized least squares and empirical Bayes estimation in regional partial duration series index-flood modeling, *Water Resour. Res.*, *33*(4), 771–781.

Martins, E., and J. Stedinger (2000), Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data, *Water Resour. Res.*, *36*(3), 737–744.

McCullagh, P., and J. A. Nelder (1989), *Generalized Linear Models*, CRC Press, Boca Raton, FL.

Merz, R., and G. Blöschl (2005), Flood frequency regionalisation—Spatial proximity vs. catchment attributes, *J. Hydrol.*, *302*(1–4), 283–306.

Northrop, P. (2004), Likelihood-based approaches to flood frequency estimation, *J. Hydrol.*, *292*(1–4), 96–113.

Pandey, M., P. Van Gelder, and J. Vrijling (2004), Dutch case studies of the estimation of extreme quantiles and associated uncertainty by bootstrap simulations, *Environmetrics*, *15*(7), 687–699.

Park, J.-S. (2005), A stimulation-based hyperparameter selection for quantile estimation of the generalized extreme value distribution, *Math. Comput. Simul.*, *70*(4), 227–234.

Payer, T., and H. Kuchenhoff (2004), ModellingModelling extreme wind speeds at a German weather station as basic input for a subsequent risk analysis for high-speed trains, *J. Wind Eng. Ind. Aerodyn.*, *92*(3–4), 241–261.

Pickands, J. I. (1975), Statistical inference using extreme order statistics, *Ann. Stat.*, *3*, 119–131.

R Development Core Team (2006), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.

Ribatet, M., E. Sauquet, J.-M. Grésillon, and T. B.M. J. Ouarda (2007), A regional Bayesian POT model for flood frequency analysis, *Stochastic Environ. Res. Risk Assess. (SERRA)*, *21*(4), 327–339.

Rosbjerg, D., H. Madsen, and P. Rasmussen (1992), Prediction in partial duration series with generalised pareto-distributed exceedances, *Water Resour. Res.*, *28*(11), 3001–3010.

Seidou, O., T. Ouarda, M. Barbet, P. Bruneau, and B. Bobée (2006), A parametric Bayesian combination of local and regional information in flood frequency analysis, *Water Resour. Res.*, *42*(11), W11408, doi:10.1029/2005WR004397.

Shu, C., and D. H. Burn (2004), Artificial neural network ensembles and their application in pooled flood frequency analysis, *Water Resour. Res.*, *40*(9), W09301, doi:10.1029/2003WR002816.

Stephenson, A., and M. Ribatet (2006), *A User's Guide to the evdbayes Package (Version 1.1)*.

Stephenson, A., and J. Tawn (2004), Bayesian inference for extremes: Accounting for the three extremal types, *Extremes*, *7*(4), 291–307.

Wood, S., and N. Augustin (2002), GAMs with integrated model selection using penalized regression splines and applications to environmental modellingmodelling, *Ecol. Modell.*, *157*(2–3), 157–177.

J. M. Grésillon, M. Ribatet, and E. Sauquet, Unité de Recherche HH, Cemagref Groupement de Lyon, 3bis quai Chauveau CP220, 69336 Lyon Cedex 09, France. (ribatet@lyon.cemagref.fr)

T. B. M. J. Ouarda, INRS-ETE, University of Quebec, 490, de la Cournne, Quebec, QC, G1K 9A9, Canada.