Regional flood frequency analyses involving extraordinary flood events at ungauged sites: further developments and validations

Chi Cong Nguyen, Eric Gaume, Olivier Payrastre

Regional flood frequency analyses involving extraordinary flood events at ungauged sites: further developments and validations

C.C. Nguyen[a,b], E. Gaume[a], O. Payrastre[a]

[a] *LUNAM Université, IFSTTAR, GERS, F-44344 Bouguenais, France*
[b] *Danang university, Vietnam*

Abstract

Flood frequency analyses are often based on recorded series at gauging stations. However, the length of the available data sets is usually too short to provide reliable estimates of extreme design floods. Hence, hydrologists have tried to make use of alternative sources of information to enrich the datasets used for the statistical inferences. Two main approaches were therefore proposed. The first consists in extending the information in time, making use of historical and paleoflood data. The second, spatial extension, consists in merging statistically homogeneous data to build large regional data samples. Recently, a combination of the two techniques aiming at including estimated extreme discharges at ungauged sites of a region in the regional flood frequency analyses has been proposed. This paper presents a consolidation of this approach and its comparison with the standard regional flood frequency approach proposed by Hosking & Wallis. A modification of the likelihood function is introduced to enable the simultaneous calibration of a regional index flood relation and of the parameters of the regional growth curve. Moreover, the efficiency of the proposed method is evaluated based

*Corresponding author. Tel.:+33(0)2 40 84 58 84, eric.gaume@ifsttar.fr

on a large number of Monte Carlo simulated data sets. This work confirms that extreme peak discharges estimated at ungauged sites may be of great value for the evaluation of large return period (typically over 100 years) flood quantiles. They should therefore not be neglected despite the uncertainties associated to these estimates.

Keywords:

Floods, regional analysis, statistics, Extremes , GEV

---

1. Introduction

Although a large number of statistical inference methods have been progressively developed, the question of estimating extreme design floods is still problematic due to the generally limited amount of available data. Continuous discharge series at gauged sites are generally too short to provide reliable estimates of extreme quantiles - typically the 100-year or higher return period quantiles (NERC, 2000). To cope with this difficulty, hydrologists have tried to complement the available data sets, either through a "temporal extension", incorporating data on historical and paleofloods (Hosking and Wallis, 1986a,b; Stedinger and Cohn, 1986; Cohn and Stedinger, 1987; Gary and Stedinger, 1987; Sutcliffe, 1987; Minghui and Stedinger, 1989; Sheffer et al., 2003; Reis et al., 2005; Neppel et al., 2010; Payrastre et al., 2011), or through a "spatial extension", merging data sets in regions considered as statistically homogeneous, "trading space for time" according to the words of Hosking & Wallis (Hosking and Wallis, 1997; Heinz and Stedinger, 1998; Charles and Stedinger, 1999; Ouarda et al., 2001; Kjeldsen et al., 2002; Merz and Blöschl, 2003; Seidou et al., 2006; Ribatet et al., 2007; Norbiato et al., 2007; Wallis

et al., 2007; Kjeldsen and Jones, 2009; Haddad and Rahman, 2011).

Recently, Gaume et al. (2010) observed that estimated extreme peak discharges at ungauged sites are often available, but never really used in flood frequency studies and proposed a method to incorporate such information in regional flood frequency analyses.

The proposed approach is based on the index flood principle (Dalrymple, 1960), assuming that, within a statistically homogeneous region, all local statistical distributions are identical apart from a site-specific scaling factor: the index flood. Usually, the index flood corresponds to the mean of the local series (Hosking and Wallis, 1997). The approach proposed by Gaume et al. (2010) is based on the calibration of an index flood relation linking the characteristics of the watersheds and the index flood value. Although this relationship represents an additional homogeneity requirement that may limit the extent of the region used for the statistical analysis, it also enables to estimate the index flood at ungauged sites, and thus to incorporate the corresponding ungauged extremes in the regional sample.

Based on several case studies, Gaume et al. (2010) showed the possible great value of such an approach, depending on the characteristics of available extreme flood inventories. The index flood relationship proposed, of the form S^{β} (S being the area of the watershed and β a parameter to be calibrated), appeared satisfactory in the test regions. The presented inference results were based on a Bayesian MCMC framework (Castellarin, 2005; Reis et al., 2005; Seidou et al., 2006; Ribatet et al., 2007; Castellarin et al., 2007; Payrastre et al., 2005, 2011) to adjust the regional growth curve with associated 90% credibility intervals. The results showed that the incorporation of ungauged

extremes could lead to a significant reduction of the width of the computed credibility intervals.

In the initial version of the method (Gaume et al., 2010), the index flood relation was adjusted prior to the calibration of the regional growth curve, and the uncertainties associated with its calibration were not taken into account. This led certainly to underestimate the credibility intervals and over-rate the added value of the ungauged extremes and of the proposed method. The effects of possible variations (heterogeneities) in the average relation calibrated in a given region should also be considered for a fair comparison with other statistical methods. This paper proposes an extension of the initial method to account for uncertainties in the calibrated index flood relation. It also tests the effect of possible regional variations in the average relation on the efficiency of the proposed statistical inference approach.

The performances (i.e. widths and correctness of computed credibility intervals) of the proposed approach and of the standard regional frequency approach proposed by Hosking & Wallis (1997) are first compared in the case where gauged data only are considered. The comparison is based on samples generated through Monte Carlo simulations in order to be able to verify the accuracy of the calculated credibility intervals and to introduce controlled heterogeneities in the samples. In a second step, both approaches are applied to the statistical analysis of a data set from the Ardèche region in France composed of 168 records at 5 gauging stations and 18 estimated ungauged extremes.

The paper is organised as follows: section 2 presents the basics and adaptations of the two regional flood frequency methods: Hosking & Wallis and

the proposed approach. The performances of the two approaches are compared based on simulated samples of random variables in section 3. In section 4, the methods are applied to the real-world case study. Conclusions are drawn in section 5.

2. Tested regional flood frequency analysis methods

2.1. The index flood hypothesis

The two approaches considered in this paper are based on the same fundamental simple scaling hypothesis or index-flood principle (Dalrymple, 1960): in a statistically homogeneous region, all the local annual maximum peak discharge distributions are supposed to be identical apart from a site-specific scaling factor. This hypothesis is summarized in equation 1:

$$Q_i(F) = \mu_i q(F) \tag{1}$$

Where F is the probability of non-exceedance, i is the index of the site $(i = 1, ..., s)$, s the total number of sites in the homogeneous region, $Q_i(F)$ is the discharge quantile, $q(F)$ is the regional dimensionless (i.e. reduced) quantile and μ_i is the index flood (or scaling factor).

The index flood may be any constant value proportional to the expectancy of the local distribution. Usually, when only data from gauged sites are considered, the index flood is estimated by the at-site sample mean (Hosking and Wallis, 1997; Castellarin, 2005; Castellarin et al., 2007). A regional flood frequency method where the index flood is computed as the average of the local series of annual maxima will be called hereafter method of Hosking & Wallis. This, even if a likelihood based Bayesian MCMC procedure rather

than a L-moment based procedure, as suggested by Hosking & Wallis (1997), is used to calibrate the parameters of the regional growth curve. Gaume et al. (2010) suggested an alternative approach to account for extreme discharge estimates that may be available at ungauged sites. An inventory of ungauged extremes may includes h extreme peak discharges Q_k $(k = 1, .., h)$, each Q_k corresponding to the largest flood at site k during a period of length n_k. In order to include this additional information in the regional dataset, Gaume et al. (2010) proposed to use an index flood relation linking the index flood value to the catchment area S, since an average annual peak discharge can obviously not be computed at ungauged sites:

$$\mu_i = S_i^{\beta} \; and \; \mu_k = S_k^{\beta} \tag{2}$$

Where S_i and S_k are the catchment areas at the corresponding sites, and β a coefficient to be calibrated. More complex relations based on various climatic and physio-geographic characteristics may be tested in the future, but at the price of an increased number of parameters to be calibrated. In the initial version of the method, the value of β was adjusted through a regression between the log tranform of the average annual peak discharges and the watershed areas at gauged sites.

It is proposed here to calibrate β along with the parameters of the regional growth curves using a modified version of likelihood as described below.

2.2. Likelihood of the observed sample

The inference approach applied herein is directly derived from Gaume et al. (2010) and inspired by numerous previous works (Reis et al., 2005;

Renard et al., 2006; Payrastre et al., 2011): i.e. based on the likelihood of the available data sets and a Bayesian MCMC algorithm for the estimation of the parameters of the growth curve and of their posterior distribution according to the observed data set.

Considering the regional sample **D** described above, including both (i) s series of gauged annual maximum discharges $Q_{i,j}$, $j = 1, \cdots, n_i$ the index of the year, and (ii) the h estimated largest peak discharges Q_k over n_k years at h ungauged sites, the standard expression of the likelihood of the regional sample **D** would be the following:

$$\ell\left(\mathbf{D} \mid \theta\right) = \prod_{i=1}^{s}\left[\prod_{j=1}^{n_i} f_\theta\left(\frac{Q_{i,j}}{\mu_i}\right)\right] \prod_{k=1}^{h}\left[f_\theta\left(\frac{Q_k}{\mu_k}\right)\right] \prod_{k=1}^{h}\left[F_\theta\left(\frac{Q_k}{\mu_k}\right)\right]^{(n_k-1)} \quad (3)$$

Where f_θ and F_θ are respectively the probability density function and the cumulative probability function of the selected statistical distribution for the regional growth curve, and θ corresponds to the vector of parameters to be estimated. The GEV distribution, often used to describe peak discharge growth curves (Lu and Stedinger, 1992; Stedinger and Lu, 1995; Coles and Powell, 1996; Coles and Tawn, 1996; Heinz and Stedinger, 1998; Seidou et al., 2006), was selected here (Eq. 4 and 5). The vector θ comprises the position, scale and shape parameters (ξ, α, κ) of the GEV distribution.

$$F_\theta\left(Q\right) = \exp\left[-\left(1 - \frac{\kappa\left(Q-\xi\right)}{\alpha}\right)^{1/\kappa}\right]_{\alpha>0} \quad (4)$$

$$f_\theta(Q) = \frac{1}{\alpha}\left(1 - \frac{\kappa(Q-\xi)}{\alpha}\right)^{1/\kappa-1} \exp\left[-\left(1 - \frac{\kappa(Q-\xi)}{\alpha}\right)^{1/\kappa}\right]_{\alpha>0} \quad (5)$$

In equation 3, the first term corresponds to the probability of the gauged series. It is the only necessary term if continuous series of measured annual maximum discharges are used. The second term is the probability of the ungauged extremes. The third complementary term is the probability associated to the fact that the ungauged extreme value has not been exceeded during the remaining $(n_k - 1)$ years at each ungauged site. The index values μ can be estimated before the calibrating the regional growth curve : method of Hosking & Wallis and initial version of the proposed approach. Equation 2 can be also directly introduced in the formulation of the likelihood (eq. 3) adding one parameter, β, to be calibrated. This replacement necessitates a slight but determining modification of the expression of likelihood (section 2.4). Depending on the option, the set of parameters to be calibrated will count either 3 or 4 parameters.

2.3. Bayesian Monte Carlo Markov Chain algorithm

The Bayesian Monte Carlo Markov Chain procedure is now relatively common for hydrological applications (Reis et al., 2005; Renard et al., 2006; Seidou et al., 2006; Ribatet et al., 2007; Neppel et al., 2010; Gaume et al., 2010; Payrastre et al., 2011; Viglione et al., 2013) and will only be briefly presented here. Part of the algorithms used is included in the R software library nsRFA (Viglione, 2013). Recalling the Bayes' theorem, the likelihood

of the sample given the parameters of the statistical model $\ell(\mathbf{D} \mid \theta)$ can be related to the likelihood or density of probability of the parameters given the sample $p(\theta \mid \mathbf{D})$ (posterior distribution):

$$p(\theta \mid \mathbf{D}) = \frac{\ell(\mathbf{D} \mid \theta)\, p(\theta)}{p(\mathbf{D})} \tag{6}$$

Where $p(\theta)$ is the so called prior distribution of θ, which summarizes any prior or alternative knowledge on the distribution of θ, and $p(\mathbf{D})$ is the probability of the sample $\mathbf{D}$ which is unknown. When prior information on the distribution of θ does not exist, then $p(\theta)$ is often taken as uniform. It is the case here, which implies that $p(\theta \mid \mathbf{D})$ is proportional to $\ell(\mathbf{D} \mid \theta)$. The statistical model being chosen, it is possible to compute the probability density function of its parameters according to the observed data sample.

$p(\theta \mid \mathbf{D})$ being known, or more precisely $\ell(\mathbf{D} \mid \theta)$ which is proportional to it according to our hypotheses (uniform prior), parameter sets θ will be sampled according to $p(\theta \mid \mathbf{D})$ to build their posterior distributions and compute the corresponding credibility limits for the discharge quantiles. The MCMC algorithms, combining random walk Monte Carlo methods with Markov chains, are a class of algorithms for the efficient sampling from multivariate random distributions (Tanner, 1996; Robert and Casella.G, 2004).

Figure 1 illustrates the Bayes MCMC inference procedure: posterior distributions of two of the estimated parameters for 3 different Markov chains run in parallel.

2.4. Modification of the likelihood formulation

A modification of the values of the index floods μ_i and μ_k, corresponding to a change of scale of the analyzed discharges, will also affect the values

of the optimal parameters ξ and α, the variable $(Q/\mu - \xi)/\alpha$ remaining unchanged. The absolute value of the cumulative density function (eq. 4) will be unaffected by such a change of variable, but not the absolute value of the density function, due to the factor $1/\alpha$ in front of the expression of the density (eq. 5). Any method based on the maximization of likelihood expression 3, will therefore tend to minimize the value of α, and hence to maximize the values of the indices $\mu_i = S_i^\beta$ if they are calibrated. Therefore, the use of likelihood expression 3 in an inference procedure where the parameter β is adjusted will inevitably lead to biased estimates of β (over-estimations) and of the parameters of the regional growth curve which are linked to β. This is confirmed by numerical tests as illustrated in Figure 1. To avoid this disruptive effect, one possibility consists in including only cumulative densities in the likelihood expression (eq. 7).

$$\ell\left(\mathbf{D} \mid \theta\right) = \prod_{i=1}^{s} \left[\prod_{j=1}^{n_i} \left[F_\theta\left(\frac{Q_{i,j}^U}{S_i^\beta}\right) - F_\theta\left(\frac{Q_{i,j}^L}{S_i^\beta}\right) \right] \right] \prod_{k=1}^{h} \left[F_\theta\left(\frac{Q_k^U}{S_k^\beta}\right) - F_\theta\left(\frac{Q_k^L}{S_k^\beta}\right) \right] \prod_{k=1}^{h} \left[F_\theta\left(\frac{Q_k^U}{S_k^\beta}\right) \right]^{(n_k-1)} \quad (7)$$

Where the exponents U and L indicate the upper and lower estimate for the considered discharge values (upper and lower guesses for an estimated ungauged extreme or known uncertainty range for measured discharges). It has been verified that equations 7 and 3 provide very close results if the difference between $Q_{i,j}^U$ and $Q_{i,j}^L$ is small. In the rest of the manuscript $Q_{i,j}^U$ will be set equal to 1.01 $Q_{i,j}$ and $Q_{i,j}^L$ to 0.99 $Q_{i,j}$.

Figure 1 illustrates the type of results obtained through the application of

the MCMC algorithm using either expressions 3 or 7 of the likelihood, when the index flood relation is directly introduced in the likelihood formulation and the parameter β is calibrated along with the parameters of the regional growth curve. The simulated regional data sample is described in section 3. These results show the tremendous under-estimation of the parameter α of the regional growth curve when equation 3, containing density functions, is used. The assessment seems correct with equation 7, even if affected by uncertainties, due to the finite number of records (i.e. to the limited information content of the data set).

2.5. MCMC convergence diagnosis

Like for any other optimization method, convergence of the MCMC algorithm towards the posterior distribution of the model parameters may sometimes be difficult to reach, especially when the number of parameters to be calibrated is increased. For the inference on a single data set, convergence problems may be solved by tunning the MCMC algorithm: better selection of the parameter starting values, adjustment of the Monte Carlo Markov chain random search controling factors. The tests presented hereafter implied repeated applications of the MCMC algorithm on a large number of data samples. It was therefore necessary to strengthen the algorithm to ensure rapid convergence in most cases, to verify automatically the convergence of the MCMC algorithm for each run and to avoid introducing results corresponding to insufficient convergence or convergence failures in the analysis. Among the large number of methods developed for convergence monitoring of MCMC algorithms (Cowles and Carlin, 1996; Salaheddine et al., 2006), the popular Gelman & Rubin test has been selected (Gelman and Rubin, 1992).

It consists in running several Markov chains in parallel and in computing the evolution during the iterations of the square root of the ratio between average within-chain and average between-chains variance of the likelihood. This ratio should ideally be equal to 1 if convergence is achieved.

The test was applied herein with two chains and a threshold value of 1.05. Figure 2 shows the evolution of the Gelman & Rubin criterion with the number of iterations when MCMC algorithms are run on synthetic data samples presented in section 3. To achieve fast convergence, a first MCMC chain was systematically run with a limited number of iterations (30000) to explore the shape of likelihood function in the vicinity of the parameter values maximizing the likelihood and evaluate the covariance matrix of the parameters (important controling factor of the MCMC chain). Figure 2 shows that the convergence is rapidly achieved with the reference approach, when it appears to be more difficult with the proposed approach. According to the conducted tests, the maximum iteration number was set to 100 000 and the various MCMC outputs computed on the basis on the last 30 000 iterations. This number of iterations was sufficient to ensure the convergence of the MCMC algorithm for all the simulations presented in this paper.

3. First test of the two methods

The two regionalisation approaches (standard and proposed) have been first tested for the statistical analysis of simulated random data sets to evaluate their performances, before their application to a real-world case study (statistical analysis of data samples available in the Ardèche region, South of France). A special emphasis has been put on (i) the validation the credibility

intervals computed with the MCMC procedure and (ii) the analysis of the sensitivity to regional variability in the index flood relation.

3.1. Method

(i) Monte Carlo simulation procedure.

The test set is composed of 1000 simulated random samples drawn from a regional GEV distribution (ξ=3.34, α=2.24, κ=-0.16) and with an index flood relation (β=0.76) adjusted on the Ardèche set analysed in section 4. The samples have the same characteristics as the Ardèche regional gauged sample: catchment areas and record lengths, see table 1.

Both regionalisation approaches are then applied to each simulated dataset, including the estimation of the posterior distribution of parameters thanks to the Bayesian MCMC procedure. The cases for which convergence of the MCMC algorithm is not achieved, according to the Gelman & Rubin test, are discarded. The posterior distribution of the 100-year quantile is then built, and the corresponding maximum likelihood value $\hat{Q}_i^{ML}(0.99)$ (mode of this posterior distribution since no prior information is considered) and 90% centered credibility interval $\left[\hat{Q}_i^5(0.99), \hat{Q}_i^{95}(0.99)\right]$ are obtained for each site and each dataset. The final result is a set of 1000 (or less if some MCMC chains did not converge which did not happen here) of maximum likelihood quantiles $\hat{Q}_i^{ML}(0.99)$ and corresponding 90% credibility intervals. They will be shown as box plots.

(ii) Reliability of the computed credibility intervals.

Monte Carlo simulations enable to test the relevance of the posterior distributions of the discharge quantile $\hat{Q}_i(0.99)$ obtained through the Bayesian

MCMC procedure. The real quantile value $Q_i(0.99)$ is known, and its probability of non-exceedance $\hat{F}(Q_i(0.99))$ can be computed according to the computed posterior distribution of the discharge quantile $\hat{Q}_i(0.99)$ for each simulated sample. If these posterior distributions of the estimated discharge quantiles $\hat{Q}_i(0.99)$ are unbiased - i.e. reflect the correct density of probability of the real quantile $Q_i(0.99)$, the values of $\hat{F}(Q_i(0.99))$ estimated for the 1000 samples should be uniformly distributed over $[0, 1]$. Figure 3 illustrates how the distributions of $\hat{F}(Q_i(0.99))$ should be interpreted. For instance in case (c), the real quantile value has an estimated probability of non-exceedance close to 0 or 1 for too many cases : the real quantile lies too often at the margins of the computed posterior distribution for the estimated quantiles, sign that the variance of the computed distribution is too low and that the computed intervals are too narrow. Likewise, case (d) illustrates the case of an overestimation of the variance of the posterior distributions and resulting too large credibility intervals. Case (b) reveals an underestimation of the quantiles.

(iii) Introduction of regional variability in the index flood relation.
The Monte Carlo simulations have been repeated with the introduction of a regional variability in the index flood relation to test the robustness of the new proposed method. For each generated sample, the index flood value at site i, $\mu_i = S_i^{\beta}$ was replaced by a random value drawn from a lognormal distribution with mean S_i^{β} and standard deviation $\delta * S_i^{\beta}$, δ being successively set to 0.1 and 0.3. Statistical tests indicate that values of δ equal to 0 or 0.1 are compatible with the gauged sample available in the Ardèche region.

3.2. Results

(i) Case of homogeneous data sets

Figure 4 presents the inference results based on 1000 simulated samples, without introduction of regional variability in the index flood relation. The fluctuations of the maximum likelihood estimates $\hat{Q}_i^{ML}(0.99)$ (respectively $\hat{q}^{ML}(0.99)$ for the regional growth curve), and associated 90% credibility upper and lower bounds $\left[\hat{Q}_i^5(0.99), \hat{Q}_i^{95}(0.99)\right]$, are presented both, for the estimates corresponding to each gauged watershed area of the region and for the regional growth curve. In order to facilitate comparisons, all estimations have been divided by the real quantile value $Q_i(0.99)$ (or $q(0.99)$). Therefore the position of the real quantile is 1 for all the box plots.

Figure 4.a and b appear overall very similar except for the regional growth curves: the maximum likelihood estimates and the width of computed credibility intervals fluctuate in similar ranges, particularly for the larger catchment areas. The uncertainties increase logically in the proposed method for smaller catchment areas, because the sensitivity of the index flood to the value of the parameter β is increasing with the watershed area according to equation 2. Conversely, the smallest areas have a lower weight in the determination of the index flood relation and as a consequence the calibrated index flood values are less accurate for lower surfaces.

The regional growth curves have different significations in both methods: distribution of the reduced discharges (values divided by the expectancy of the distribution) in the method of Hosking & Wallis and distribution of the discharges for a 1 km^2 watershed in the proposed approach. This is largely in the extrapolation range of the index flood relation for low values of watershed

areas and explains the large uncertainties for the estimated quantiles for the regional growth curve in the proposed approach.

An in depth analysis of both figures reveals nevertheless some differences of the two regionalisation approaches: the median values of the maximum likelihood and 5% and 95% credibility bounds for the 100-year quantile are slightly lower when the standard Hosking & Wallis method is used, but the uncertainties affecting these estimates (widths of the box plots) appear larger except for the smallest considered watershed area (63 km^2). The median maximum likelihood estimated value lower than 1 for the standard method suggests also that it is slightly biased.

The reliability tests conducted on the computed credibility intervals indicate furthermore that, while the intervals seem correctly estimated with the proposed method (Fig. 5.e and f) with the exception of the smallest watershed area where a tendency to the overestimation of the quantile appears. The credibility intervals computed with the standard method appear too narrow (Fig. 5.a and b). A complementary test presented in appendix Appendix A indicates that a 20% to 40% underestimation affects the computed widths of credibility intervals in this case study: a correction factor m equal to 1.2 to 1.4 should be applied to the standard-deviation of the posterior distributions of $\hat{Q}_i(0.99)$ to obtain a uniform distribution of $\hat{F}(Q_i(0.99))$ (Fig. 5.c and d). This under-estimation of the variance of the posterior distributions of $\hat{Q}_i(0.99)$ for the standard method is logical since the uncertainty associated with the estimation of the local sample means is not accounted for in the expression of the likelihood. Its effect on the quantile estimations can therefore not be evaluated by the Bayesian MCMC procedure.

The origins of the observed limited biases - tendency to underestimate the quantile on average for the standard method and to overestimate the quantile for the smaller areas for the proposed approach - are difficult to depict.

These results are overall extremely satisfactory. When the regional sample corresponds to the proposed statistical model (regional growth curve of the GEV type and power-law linking the index flood and the area of the watershed), the proposed approach appears consistent and accurate. The parameters of the regional statistical model seem correctly retrieved from samples with only a slight bias in the estimation of the parameter β affecting the estimates of the discharge quantiles of the watersheds with the smallest areas. In the standard method, an additional source of uncertainty seems to be introduced by the computation of local sample means, estimates of the local sample expectancies, leading to an unexplained minor but detectable underestimation of the quantiles and to an increase of the estimation uncertainties of these quantiles if compared to the proposed method. The credibility intervals computed through the MCMC algorithm underestimate in this case the uncertainties associated to the calculated quantiles.

The additional parameter β appears to be a factor of stabilization rather than a source of further uncertainties in a flood frequency analysis. Fixing an index flood relation seems to help to filter out the variability of the local sample means due to sampling fluctuation. This is an unexpected result. The proposed approach, initially designed to introduce ungauged extremes in flood frequency analysis, is as efficient, if not more efficient than the standard approach when applied to gauged series. To be really able to draw such a

conclusion, it is necessary to consider that the regional data set may not follow perfectly the simple suggested index flood relation. How does the method resist to some regional variability in the index flood relation? It is the focus of the next section.

(ii) Impact of some regional variability in the index flood relation

Without surprise, the introduction of variability in the index flood relation increases both, the estimated credibility intervals and the uncertainties affecting the estimates of the maximum likelihood, 5% and 95% credibility bound estimates (Fig. 6).

The effect remains nevertheless limited when the variability is relatively low (Fig. 6.a). The results obtained for $\delta = 0.1$ do not differ much from the results obtained in the ideal case ($\delta = 0$ in Fig. 4.b). The maximum likelihood estimate is still unbiased and it could be verified that the credibility intervals were also only slightly biased. Uncertainties increase drastically when the regional variability is increased (Fig. 6.b), but still with a limited bias, on average, on the maximum likelihood estimate of the quantile (overestimation) and some bias appearing in the estimated credibility intervals. Of course, the biases on the individual local estimates can be significant as shown by the widths of the box plots.

These results finally suggest that the proposed approach may resist to a certain level of heterogeneity of the proposed index flood relationship in the considered region. Important heterogeneities should nevertheless be avoided through a careful selection of the regional data set based on the plot of the local sample means versus watershed areas (Gaume et al., 2010). This poses an additional constraint to the delineation of homogeneous regions if compared to the method of Hosking & Wallis, limiting the size of the regional gauged sets, but opening the possibility to valuate the ungauged extremes. This constraint should be kept in mind for a fair comparison of both methods.

4. Application to a real-world example with ungauged extremes: the Ardèche area in France

4.1. The Ardèche data set

The Ardèche case study had already been selected by Gaume et al. (2010) for the initial tests of the regionalisation method involving ungauged extremes. It has been again used here to illustrate the changes introduced by the new inference procedure including the calibration of the index flood relationship.

The Ardèche region is located in the southeast of Massif Central in southern France (Fig. 7). It is one of the areas in Europe exposed to the most frequent and severe flash floods (Gaume et al., 2009). The region includes six main rivers: the Ardèche River 2380 km^2, Ouvèze River 120 km^2, Eyrieux River 860 km^2, Doux River 630 km^2, Ay River 110 km^2 and Cance River 380 km^2. Five stream gauges provide accurate, continuous and long series of discharges in the region. The gauged dataset of annual maximum peak

discharges counts 168 records. The situation of the 5 gauging stations is reported in figure **??** and the characteristics of corresponding samples are presented in table 1.

Using the large set of documentary sources available in the region, ungauged largest peak discharge estimates over the last 50 years could be retrieved at 14 additional sites (Fig. 8). Similarly, information on the 50 years preceding the gauging period could be found at 4 of the gauged sites. The set of ungauged extremes is finally composed of 18 maximum values over 50 years, that is equivalent to 900 censored records, six times more than the regional gauged data set (Table 2). Note that the selected set of ungauged extremes has been reduced if compared to the set used by Gaume et al. (2010). A minimum distance between sites on the same stream has for instance been imposed to avoid redundancy (see appendix Appendix B for a short discussion about the sampling of extremes).

The homogeneity of the growth curves within the region has been tested using the Hosking & Wallis heterogeneity measure (Hosking and Wallis, 1997). The H_1 value based on the 5 gauged series is equal to 2.7, which suggests that the region is possibly heterogeneous. Homogeneity has nevertheless been assumed to proceed with the computations. The influence of regional heterogeneities in the growth curves on the inference results will be studied in a future publication.

Provided the preceding assumption, this case study finally offers an interesting opportunity to apply and compare the approaches described above, in two different contexts: (i) regional dataset limited to gauged sites, and (ii) availability of information on the largest discharge at ungauged sites.

4.2. Results and discussion

Figure 8 and table 3 summarize the results obtained for the Ardèche case study. They should be interpreted at the light of the conclusions drawn in the previous section. The presented results correspond to the two extreme gauged catchment surfaces in the region: 63 km^2 (Borne watershed at Puylaurens), and 2240 km^2 (close to the outlet of the entire Ardèche watershed at Saint Martin).

The two methods - Hosking & Wallis and the proposed approach - lead to close results when applied to the gauged data set (Table 3, cases a and b). The widths of the estimated credibility intervals are slightly larger in the case of the proposed approach, but these intervals are underestimated by about 20% in the reference approach according the previous results. The underestimation of the quantiles is not systematic when the reference approach is used as illustrated by table 3. The inference results are depending on the characteristics of the sample.

The main advantage of the proposed approach lies in the possibility to incorporate information on ungauged extremes in the analysis. In the Ardèche case, the ungauged extremes appear consistent with the gauge series and the GEV distribution (Fig. 8). Their incorporation in the statistical inference leads to a clear reduction of the credibility intervals that are divided by a factor of about 2. This confirms the possible great value of the proposed approach in regions for which information on extremes floods at ungauged sites is available. The assumption of homogeneity of the index flood relationship limits the number of gauged sites that can be included in the homogeneous region. But, the extremes at ungauged sites represent a potentially large ad-

ditional source of information (equivalent to 900 records in the present case), that may reduce significantly the uncertainties attached to estimated high return period quantiles.

Figure 9 compares the modelled relation between average discharges and watershed areas and the observed locally estimated mean discharges at the light of the proposed Lognormal model to reproduce regional variability in the index flood relationship. This illustrates the consistency of the proposed index flood relation with the observed data. The second smallest watershed, the Eyrieux, appears to have relatively smaller empirical mean discharge and could be possibly regarded as non-consistent with the other sites as far as the index flood relation is considered. This low mean value may nevertheless be attributed to sampling fluctuation but may likewise be explained by the fact that this site is located further North and less influenced by the Mediterranean climate. This last comment illustrates that any regional flood frequency method is based on homogeneity hypotheses that can hardly be verified. It is therefore certainly sound to test the robustness of the obtained inference results to different hypotheses and region delineations. In this perspective, the proposed approach enriches the available panel of methods and possible hypotheses.

Finally, the same Monte Carlo simulations as the ones used in section 3 were conducted with generated samples including 18 ungauged sites where only the highest discharge over 50 years is known to verify if the conclusions drawn on the real-world example can be generalized (Fig. 10.d). In fact, the results of an anlysis conducted on one single data set may be attributed to sampling variability. On average, the credibility intervals appear

to be divided by two on a sample of 1000 simulated datasets when the "ungauged" extremes are taken into account, which is the order of magnitude obtained for the Ardèche data set. 100-year quantile estimates based on the L-moments (standard method suggested by Hosking & Wallis) have also been computed for sake of verification (Fig. 10.a). L-moment and maximum likelihood based estimates appear to have similar performances on this data sample (figures 10.a and b). The L-moment estimates exhibit a slightly lower dispersion (L-moment are less sensitive to sampling variability) but a higher bias.

5. Conclusions

The approach initially proposed by Gaume et al. (2010), aiming at incorporating information on ungauged extremes in regional flood frequency analyses, has been consolidated and thoroughly tested. The proposed method and algorithm, after some necessary adjustments, in particular of the formulation of the likelihood, appear to be able to estimate accurately quantiles and their uncertainty bounds on the basis of a regional sample combining continuous gauged series and ungauged censored data. The introduction of an index flood relationship and of an additional parameter to be calibrated, proved to be a stabilization factor rather than a factor of complexity and a source of uncertainty. The method shows a relative robustness to some regional variability in the index flood relation. The additional homogeneity constraint limits the possible extent of homogeneous regions if compared to the reference approach, but this is compensated by the additional information brought by the estimated ungauged extremes. In the case of the Ardèche

region the ungauged data set represents 900 censored recording station-years: six times the amount of the available regional gauged series. Its incorporation in the regional flood frequency analysis led to a division by two of the estimated 100-year quantile 90% credibility intervals.

Uncertainties on the extreme peak discharge estimates were not considered herein and the ungauged extreme dataset that could be retrieved in the Ardèche region may appear as particularly rich if compared to what could reasonably be obtained in other regions. Previous works, conducted on the usefulness of historic records in flood frequency studies, indicated that the accuracy level of the estimated extreme discharges and the number of extremes for which an estimated discharge is available, have little influence on the inference results when the historic period of record if fixed Payrastre et al. (2011). It remains to be confirmed that the same conclusion holds for ungauged extremes when included in regional flood frequency analyses.

In the same line of thoughts, the possible variability in the growth curves in a region considered as homogeneous has not been considered yet. Its influence could be evaluated in the same manner as the influence of variability in the index flood relation has been considered in section 3. It may modify partly the judgment on the relative efficiency of the various tested regional flood frequency analysis methods.

Finally, it is important to keep the conclusion of section 4 in mind. All the regional flood frequency methods are based on homogeneity hypotheses that can hardly be verified. It is therefore certainly sound to test the robustness of the obtained inference results to different hypotheses and region delineations: to combine various methods. In this perspective, the proposed

approach does not replace but enriches the available panel of methods and possible hypotheses. It may lead, as shown herein, to quantile estimates that are theoretically equally or even more accurate than the estimates obtained with standard methods. Its main constraint may also be considered as an advantage: the two necessary homogeneity hypotheses limit the extent of homogeneous regions that may be, on the other hand, hydrologically more consistent and less questionable than in standard regional flood frequency analyses.

Acknowledgements

This work has benefited from the support of the French Ministry of foreign affairs.

Appendix A. Method for the evaluation of the underestimation factor for the computed credibility intervals

In many cases, the proposed reliability test for the computed credibility intervals presented in section 3.1 indicates that the intervals are too narrow (Fig. 5). This is due the fact that some sources of uncertainties are not accounted for in the bayesian MCMC framework used to compute these intervals. The following procedure is proposed to evaluate the underestimation factor m of the computed intervals. Based on the hypothesis that the shapes of the posterior distribution of the discharge quantiles $\hat{Q}_i(0.99)$ are correct and that only their dispersion is underestimated, the procedure to evaluate the underestimation factor m is the following.

The distance between the real quantile value $Q_i(0.99)$ and the median value of the computed posterior distribution $Me\left(\hat{Q}_i(0.99)\right)$ is artificialy reduced by the factor m : the probability of non-exceedance $\hat{F_m}\left(Q_i(0.99)\right)$ (eq. A.1) is computed rather than the probability $\hat{F}\left(Q_i(0.99)\right)$ based on the MCMC results for each simulated sample.

$$\hat{F_m}\left(Q_i(0.99)\right) = \hat{F}\left(Me\left(\hat{Q}_i(0.99)\right) + 1/m\left[Q_i(0.99) - Me\left(\hat{Q}_i(0.99)\right)\right]\right) \tag{A.1}$$

The factor m leading to the distribution of $\hat{F_m}\left(Q_i(0.99)\right)$ closest to a uniform distribution for the 1000 simulated samples, gives an estimate of the underestimation factor of the credibility intervals.

Appendix B. Sampling ungauged extremes in a region

The regional samples of ungauged extremes must be carrefully built to avoid introducing bias into the statistical analysis. At least, the extreme data collation method must seek to meet three conditions: (i) be based on an explicit sampling strategy that can be translated into the likelihood formulation, (ii) be comprehensive fo each selected location, (iii) limit possible dependencies between records and samples. The next three sections explain how these three conditions were accounted for to build the Ardèche regional sample of ungauged extreme floods.

Appendix B.1. Explicit and systematic sampling method for regional extremes

Isolated extreme discharge values are often reported but can not be directly integrated in regional flood frequency analyses if their representativeness is unknown. Over which period of time can it be asserted that a larger

flood event did not occur at the same site ? And, even more crucial, these values have been reported because they appear as extreme, i.e. this means implicitly that they seem large if compared the known flood events observed in neighbouring comparable catchments. Over how many years and for how many similar catchments (station) could it be finally asserted that the considered extreme discharge value - or rescalled discharge value - has not been exceeded ? Large uncertainties concerning the record length (station.years) corresponding to each extreme may drastically reduce their added value in regional flood frequency analyses as illustrated by the Slovakian test case in Gaume et al. (2010). An inventory of regional extremes should therefore be based, if possible, on a systematic data mining procedure. The various sites in a region where information on extremes may be available should be determined a priori: typically significant towns along rivers where large flloods may have induced damages and may have been reported. An inventory of past large floods should be realized at each of these sites based on local memory and archives and it must at least be verified that some given threshold values have or have not been exceeded at the considered sites in the recent past: non-exceedance is a statistically valuable information.

Appendix B.2. Comprehensiveness

Collation of extreme events over a given period of time provide censored data sets: only discharges exceeding a given threshold, that depends on the site and the considered period, are documented. Flood events exceeding the threshold should be exhaustively listed at each site over the considered past period for a correct statitical inference. To be on the safe side, it is generally preferable to limit the considered past period of time and to raise the thresh-

old value to ensure exhaustiveness, even if this limits the information content of the inventories of large or extreme floods. For this reason, the collation period has been limited in the Ardèche to the last 50 years (typical depth of local memory). Likewise, the inventory has been limited to the largest known flood event at each ungauged site. This last choice facilitates also the computation of the plotting positions for the ungauged extremes, which can become tricky: the return periods of the ungauged extremes reported on the cumulated distribution plots are estimated through Monte Carlo simulations (Gaume et al., 2010).

Appendix B.3. Indepedency

Likelihood formulations and the computation of the credibility intervals are based on the assumption of independence between the various recorded maximum annual peak discharges. But, extreme peak discharges induced by the same rainfall event sometimes occur at the same date at nearby ungauged sites, suggesting at least weak depedencies in the inventory of ungauged extremes. To prevent any effect of these dependencies on the results and especially to limit the risk of underestimation of the credibility intervals due to an implicite overrating of the information content of the ungauged extreme inventory neglecting depedencies, the data set had to be refined. When the largest reported flood occurred at the same date at close ungauged sites, one randomly selected site has been kept in the final ungauged data set. Keeping the largest value would have introduced a sampling bias. The refinement of the Ardèche ungauged extreme set led to eliminate more than half of the records initially used by Gaume et al. (2010), hence resulting in a slight increase only of the computed credibility intervals.

Castellarin, A., 2005. Probabilistic behavior of a regional envelope curve. Water Resources Research 41, W06018.

Castellarin, A., Vogel, R., Matalas, N., 2007. Probabilistic envelope curves for design flood estimation at ungauged sites. Water Resources Research 43, W04406.

Charles, N., Stedinger, J., 1999. Development of regional regression relationship with censored data. Water Resources Research 35(4), 775–784.

Cohn, T., Stedinger, J., 1987. Use of historical information in a maximum likelihood framwork. Journal of Hydrology 96 (1-4), 215–223.

Coles, S., Powell, E., 1996. Bayesian methods in extreme value modelling. A review and new developments. Int. Stat. Rev 43, 1–48.

Coles, S., Tawn, J., 1996. A bayesian analysis of extreme rainfall data. Appl. Stat 45, 463–478.

Cowles, M., Carlin, B., 1996. Markov Chain Monte Carlo convergence diagnostics: a comparative review. Statist-Assoc 91, 883–904.

Dalrymple, T., 1960. Flood frequency analyses. Water Supply: Geological Survey, Reston,Virginia, USA , 1543–A.

Gary, D., Stedinger, J., 1987. Regional regression of flood characteristics employing historical information. Journal of Hydrology 96, 255–264.

Gaume, E., Bain, V., Bernardara, P., Newinger, O., Barbuc, M., Bateman, A., Blaskovicova, L., Blöschl, G., Borga, M., Dumitrescu, A., Daliakopoulos, I., Garcia, J., Irimescu, A., Kohnova, S., Koutroulis, A., Marchi, L.,

Matreata, S., Medina, V., Preciso, E., Sempere-Torres, D., Stancalie, G., Szolgay, J., Tsanis, J., Velasco, D., Viglione, A., 2009. A compilation of data on European flash floods. Journal of Hydrology 367, 70–80.

Gaume, E., Gaal, L., Viglione, A., Szolgay, J., Kohnova, S., Blöschl, G., 2010. Bayesian MCMC approach to regional flood frequency analyses involving extraordinary flood events at ungauged sites. Journal of Hydrology doi:10.1016/j.jhydrol.2010.01.008.

Gelman, A., Rubin, D., 1992. Inference from iterative simulation using multiple sequences (with discussion). Statistical Science 7, 457–511.

Haddad, K., Rahman, A., 2011. Regional flood frequency analysis in eastern Australia: Bayesian GLS regression-based methods within fixed region and ROI framework: Quantile Regression vs. Parameter Regression Technique. Journal of Hydrologic Engineering 920, doi.org/10.1061/(ASCE)HE.1943-5584.0000395, 1–6.

Heinz, D., Stedinger, J., 1998. Using regional regression within index flood procedures and an empirical Bayesian estimator. Journal of Hydrology 210, 128–145.

Hosking, J., Wallis, J., 1986a. Paleoflood hydrology and flood frequency analysis. Water Resources Research 22, 543–550.

Hosking, J., Wallis, J., 1986b. The value of historical data in flood frequency analysis. Water Resources Research 22, 1606–1612.

Hosking, J., Wallis, J., 1997. Regional frequency analysis: An approach Based on L-Moments. Cambridge University Press, London, UK.

Kjeldsen, T., Jones, D., 2009. A formal statistical model for pooled analysis of extreme. Hydrology Research 40(5), 465–480.

Kjeldsen, T., Smithers, J., Schulze, R., 2002. Regional flood frequency analysis in the KwaZulu-Natal province, South Africa, using the index-flood method. Journal of Hydrology 255, 194–211.

Lu, L., Stedinger, J., 1992. Sampling variance of normalized GEV/PWM quantile estimators and regional homogeneity test. Journal of Hydrology 138, 223–245.

Merz, R., Blöschl, G., 2003. A process typology of regional floods. Water Resources Research 39(12), 1340.

Minghui, J., Stedinger, J., 1989. Flood frequency analysis with regional and historical information. Water Resources Research 25, 925–936.

Natural Environment Research Council, 2000. Flood estimation Handbook. NERC - Center for Ecology & Hydrology. Wallingford, UK.

Neppel, L., Renard, B., Lang, M., Ayral, P., Coeur, D., Gaume, E., Jacob, N., Payrastre, O., Pobanz, K., Vinet, F., 2010. Flood frequency analysis using historical data: accounting for random and systematic errors. Hydrological Sciences Journal 55, 192–208.

Norbiato, D., Borga, M., Sangati, M., Zanon, F., 2007. Regional frequency analysis of extreme precipitation in the eastern italian Alps and the August 29, 2003 flash flood. Journal of Hydrology 345, 149–166.

Ouarda, T., Girard, C., Cavadias, G., Bobee, B., 2001. Regional flood frequency estimation with canonical correlation analysis. Journal of Hydrology 254, 157–173.

Payrastre, O., Gaume, E., Andrieu, H., 2005. Use of historical data to assess the occurrence of floods in small watersheds in the French Mediterranean area. Advances in Geosciences 2, 1–8.

Payrastre, O., Gaume, E., Andrieu, H., 2011. Usefulness of historical information for flood frequency analyses: Developments based on a case study. Water Resources Research 47, W08511.

Reis, D., Stedinger, J., Martins, E., 2005. Bayesian GLS regression with application to LP3 regional skew estimation. Water Resources Research 41, W10419.

Renard, B., Garreta, V., Lang, M., 2006. An application of Bayesian analysis and Markov Chain Monte Carlo methods to the estimation of a regional trend in annual maxima. Water Resources Research 42(12), W12422.

Ribatet, M., Sauquet, E., Gresillon, J., Ouarda, T., 2007. Usefulness of the reversible jump Markov Chain Monte Carlo in regional flood frequency analysis. Water Resources Research 43, W08403.

Robert, C., Casella.G, 2004. Monte Carlo statistical methods. Springer, New York. 2nd ed.

Salaheddine, A., Anne-Catherine, F., Bernard, B., 2006. Comparison of methodologies to assess the convergence of Markov Chain Monte Carlo methods. Computational statistics and data analysis 50, 2685–2701.

Seidou, O., Ouarda, T., Barbet, M., Bruneau, P., Bobe, B., 2006. A parametric Bayesian combination of local and regional information in flood frequency analysis. Water Resources Research 42, W11408.

Sheffer, N.A., Enzel, Y., Benito, G., Grodek, T., Poart, N., Lang, M., Naulet, R., Coeur, D., 2003. Paleofloods and historical floods of the Ardche river, France. Water Resources Research 39(12), 1379.

Stedinger, J., Cohn, T., 1986. Flood frequency analysis with historical and paleoflood information. Water Resources Research 22, 785–793.

Stedinger, J., Lu, L., 1995. Appraisal of regional and index flood quantile estimators. Stochastic Hydrology and Hydraulics 9, 49–75.

Sutcliffe, J., 1987. The use of historical records in flood frequency analysis. Journal of Hydrology 96, 159–171.

Tanner, M., 1996. Tools for statistical inference: Methods for the Exploration of Posterior, Distributions and Likelihood Functions. Springer, Evanston, USA.

Viglione, A., 2013. Non-supervised regional frequency analysis R library. http://cran.r-project.org/web/packages/nsRFA/.

Viglione, A., Merz, R., Salinas, J.L., Blöschl, G., 2013. Flood frequency hydrology: 3.A Bayesian analysis. Water Resources Research 49, WR010782, doi:10.1029/2011WR010782.

Wallis, J.R., Schaefer, M., Barker, B., Taylor, G., 2007. Regional precipitation-frequency analysis and spatial mapping for 24-hour and 2-

hour durations for Washington State. Hydrology and Earth System Sciences (11), 415–442.

List of Tables

List of Figures

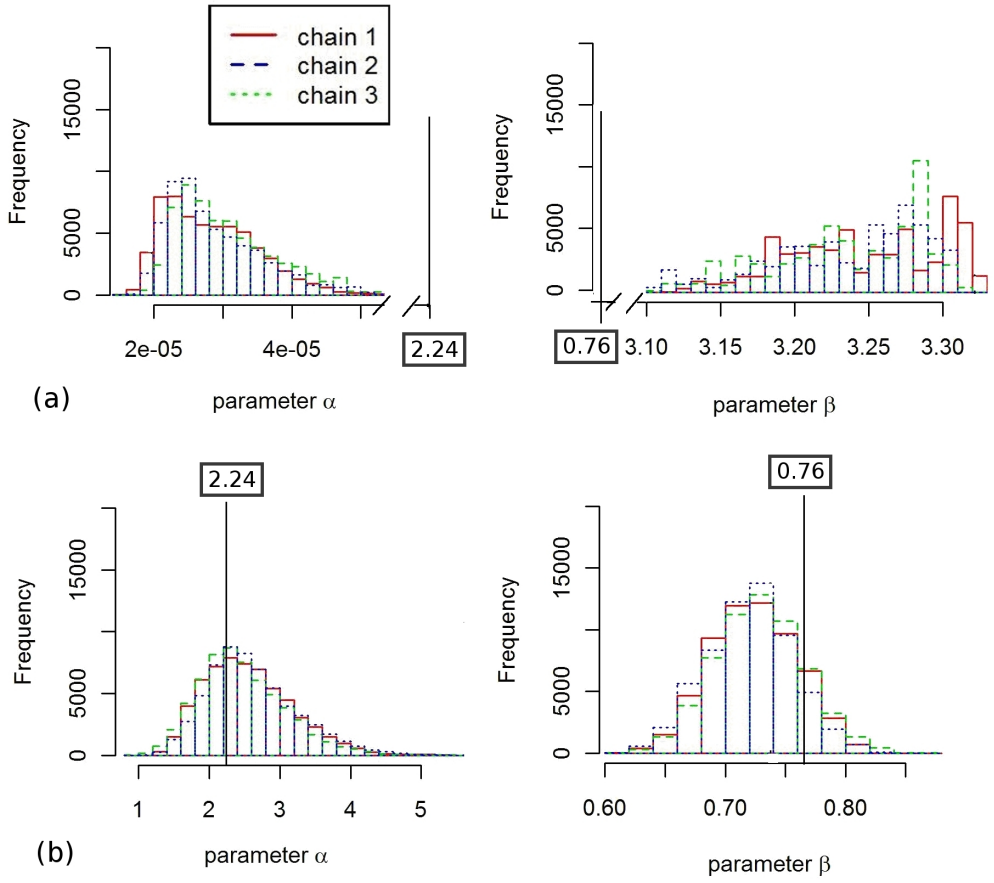

Figure 1: Posterior distributions for the parameters α and β. Calibration of a regional growth curve and an index flood relation, based on a simulated sample of 1000 records: $\alpha = 2.24$, $\xi = 3.34$, $k = -0.16$ and $\beta = 0.76$. (a) totally biased result with the standard likelihood expression 3 and (b) correct result with likelihood expression 7.

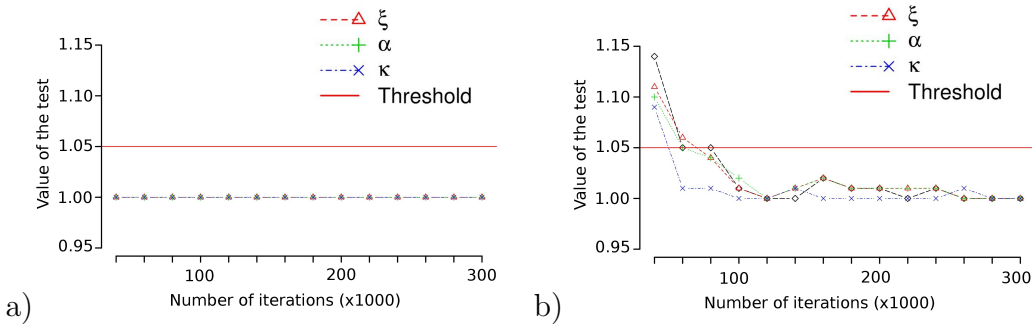

Figure 2: Gelman & Rubin's convergence diagnosis with the regional data sets of the Ardèche: a) reference approach; and b) proposed approach.

Table 1: The sample size and the surface of the gauging stations in the regional Ardèche.

Station	River name	Period of data	N(year)	S (km^2)
Beauvene	Eyrieux	1969-1998	30	392
Vogüe	Ardèche	1966-2003	38	636
Chambonas	Chassezac	1971-1997	27	507
St.Laurent	Borne	1968-1997	30	63
St.Martin	Ardèche	1963-2005	43	2240

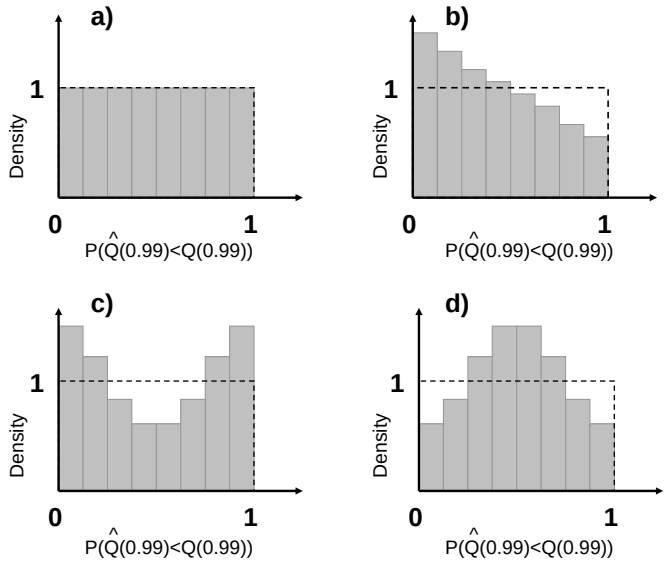

Figure 3: Posterior distributions of $\hat{F}(Q_i(0.99))$: (a) perfect credibility intervals, (b) tendency to over-estimate the quantile, (c) too narrow estimated credibility intervals, (d) too large estimated credibility intervals.

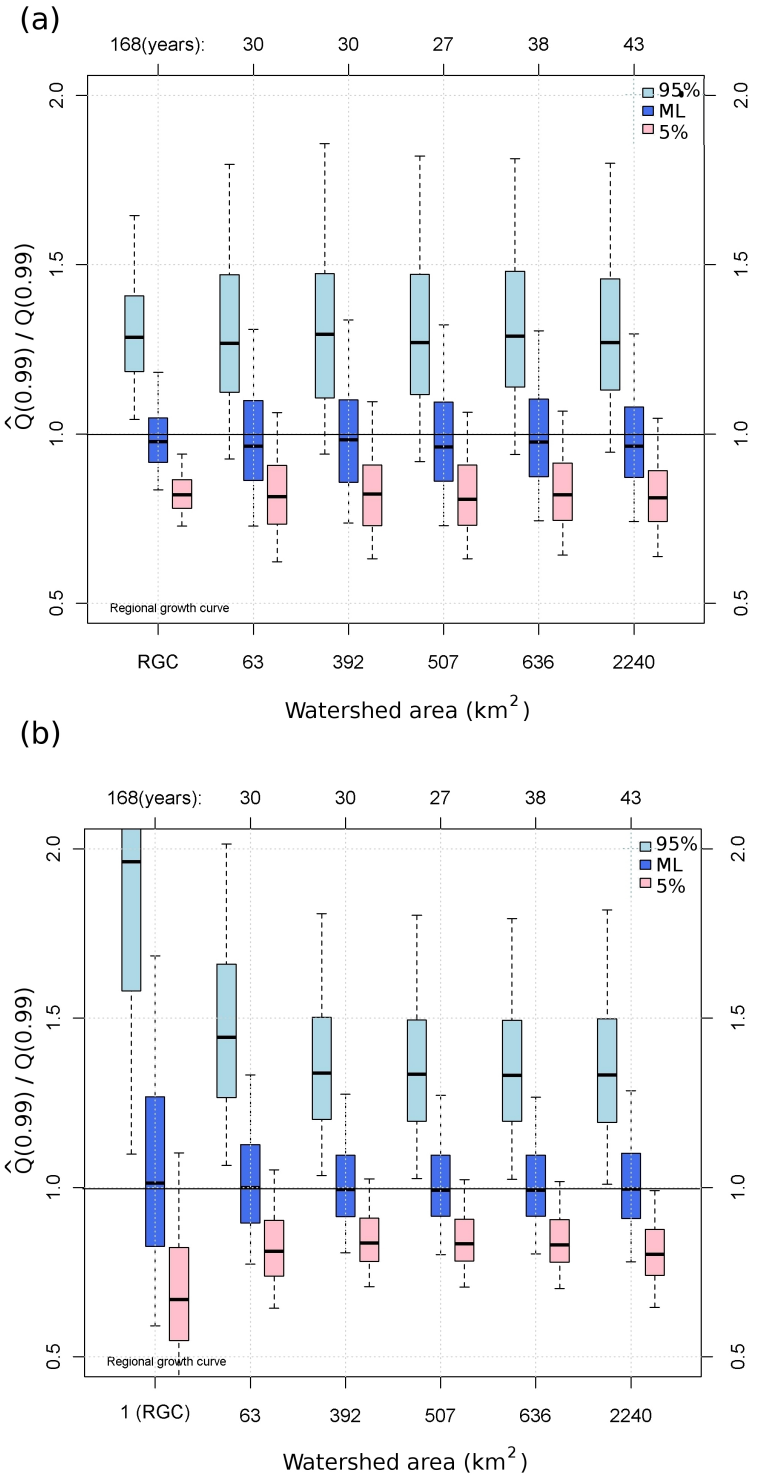

Figure 4: Estimates of the 100-year quantile through Bayesian MCMC regional flood frequency inferences conducted on 1000 different samples: (a) Method of Hosking & Wallis and (b) proposed method. Box plots of the 1000 maximum likelihood and 5% and 95% credibility bounds for each considered watershed area. All the values have been divided

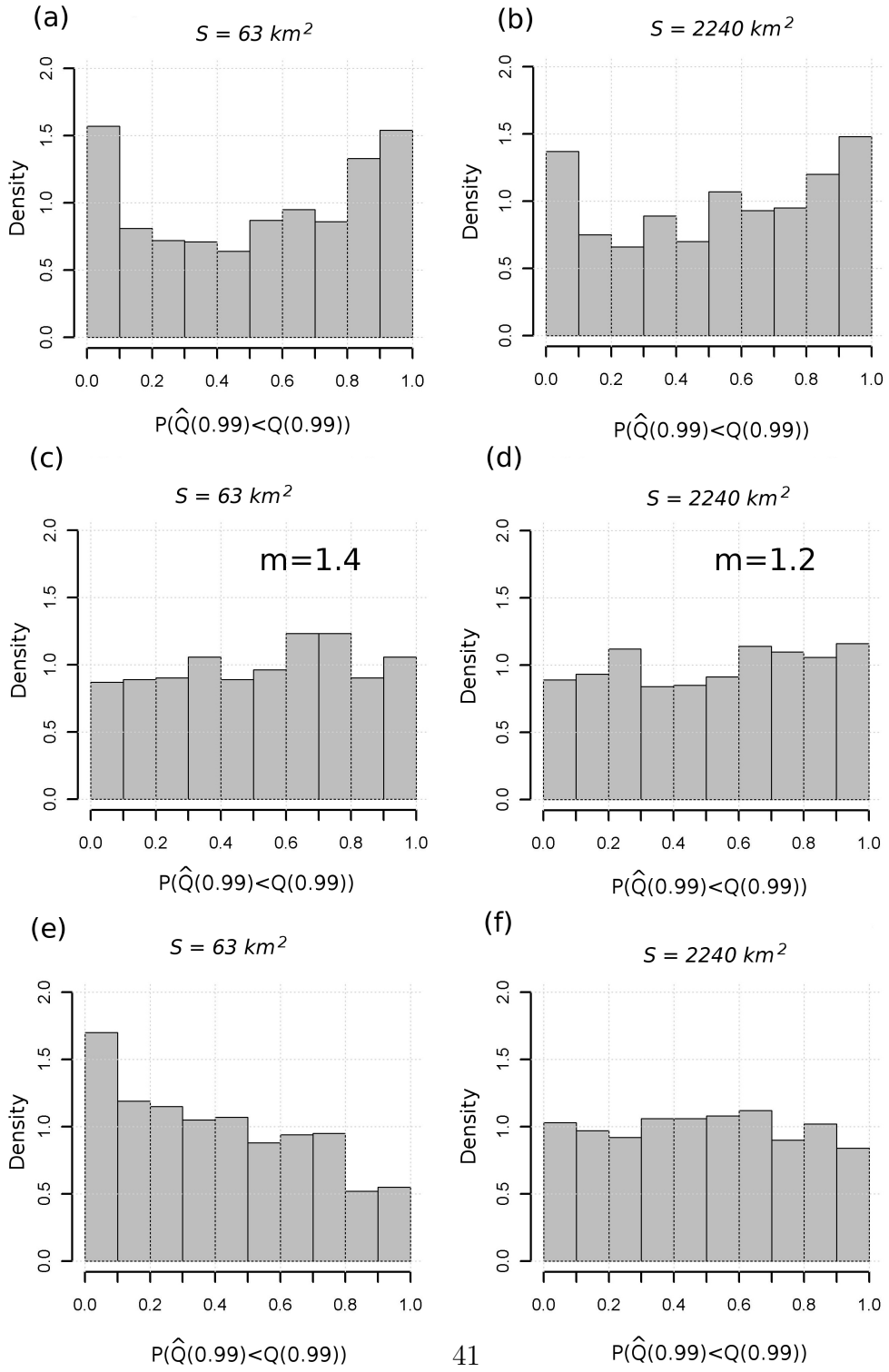

Figure 5: Distributions of the values of $\hat{F}\left(Q_i\left(0.99\right)\right)$ estimated for the 1000 simulated samples: (a,b) Method of Hosking & Wallis, (c,d) Hosking& Wallis with correction factor m for the posterior distributions and (e,f) proposed method.

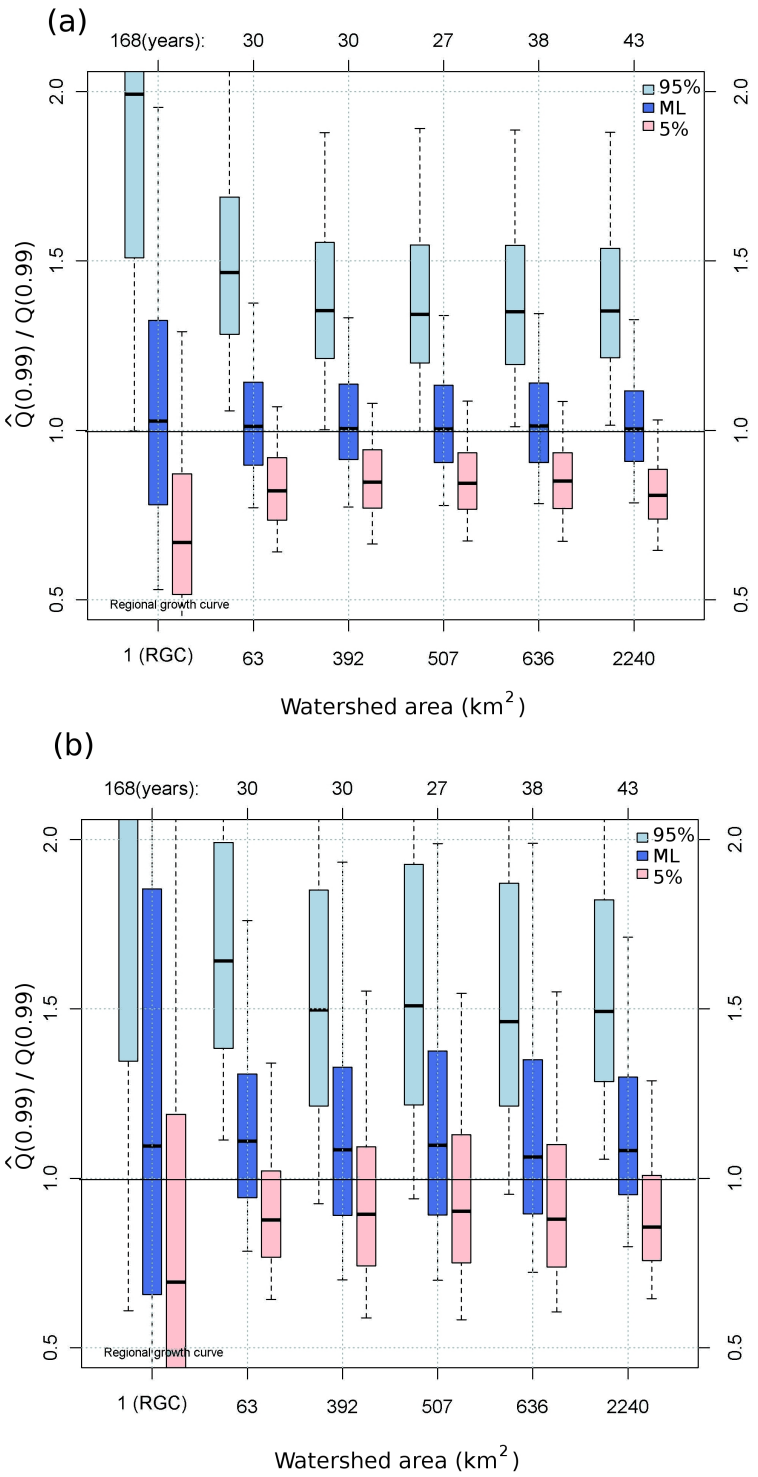

Figure 6: Results of Bayesian MCMC regional flood frequency inferences conducted on 1000 samples with the proposed method when a regional variability in the index flood relation is introduced: (a) $\delta = 0.1$, (b) $\delta = 0.3$.

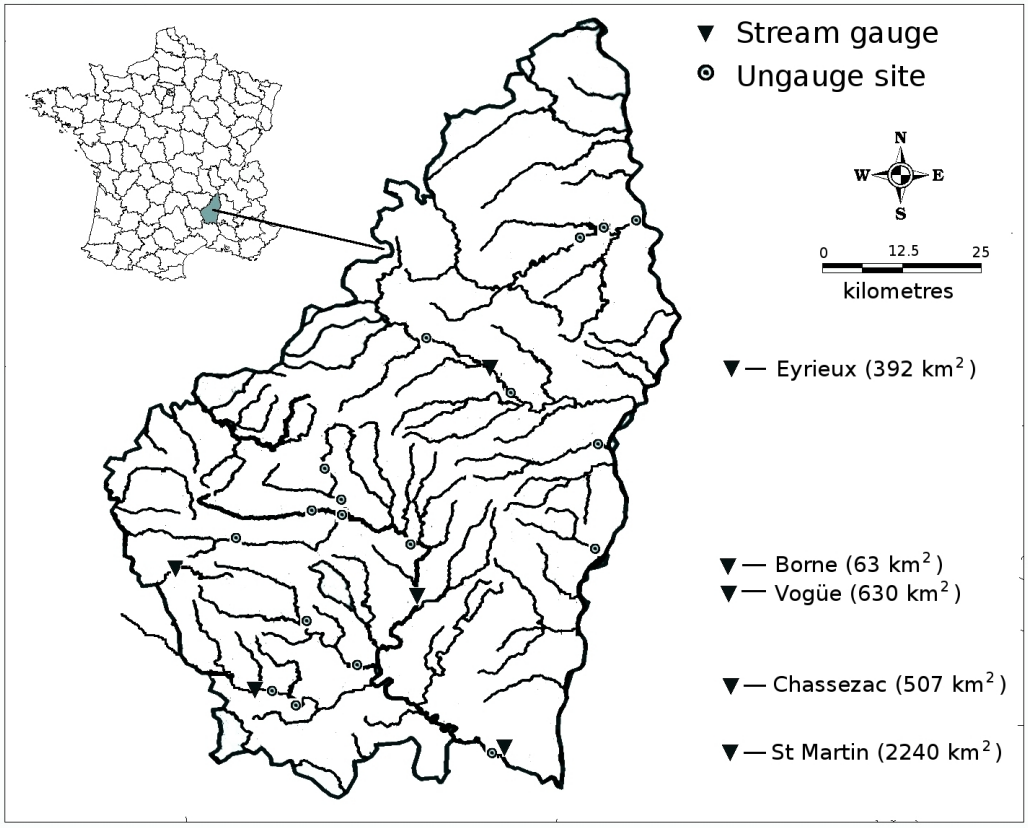

Figure 7: The Ardèche region, location of the 5 gauged sites (triangles) and the 18 sites where ungauged extremes were retrieved (dots).

Table 2: Detail of the extreme floods reported to be the highest ones in the last 50 years (or last 50 years before the gauged period)

Location	River name	Date	Q (m^3/s)	S (km^2)
Pont de Rolandy	Ardèche	22/09/1992	1150	150
Aubenas	Ardèche	26/09/1992	2200	480
Sauze St.Martin	Ardèche	30/09/1958	4500	2240
Rosieres	Beaune	04/10/1958	1820	210
Joyeuse	Beaune	30/09/1958	1000	100
Chambon	Borne	30/09/1958	100	11
Vans	Boudaric	03/11/1989	130	6
Burzet	Bourges	22/09/1992	350	47
Chambonas	Chassezac	21/09/1980	3360	510
Dorne	Dorne	03/08/1963	630	78
Lamastre	Doux	03/08/1963	970	242
Pont de Cesar	Doux	03/08/1963	1500	635
Barrage des Collanges	Eyrieux	03/081963	1685	343
Meyras	Fontoliere	22/09/1992	900	130
Meysee	Lavezon	30/09/1960	500	56
Rieut ord	Loire	01/09/1992	444	62
Pouzin	Ouveze	10/08/1967	700	140
Saliouse	Saliouse	21/09/1980	300	61

Table 3: Comparison the credibility intervals (CI) at St. Laurent and St.Martin gauging stations to three cases corresponding to return periods T=100 years.

Surface (km^2)	N_{cont} (years)	N_{hist} (years)	Case	$\hat{Q}_i^5\,(0.99)$ (m^3/s)	$\hat{Q}_i^{ML}\,(0.99)$ (m^3/s)	$\hat{Q}_i^{95}\,(0.99)$ (m^3/s)
	168	0	(a)	423.8	520.9	712.5
63	168	0	(b)	382.3	479.5	682.1
	166	900	(c)	442.8	500.4	574.2
	168	0	(a)	5538.6	6806.8	9310.8
2240	168	0	(b)	5892.4	7602.2	10747.6
	166	900	(c)	5477.3	6560.0	7847.1

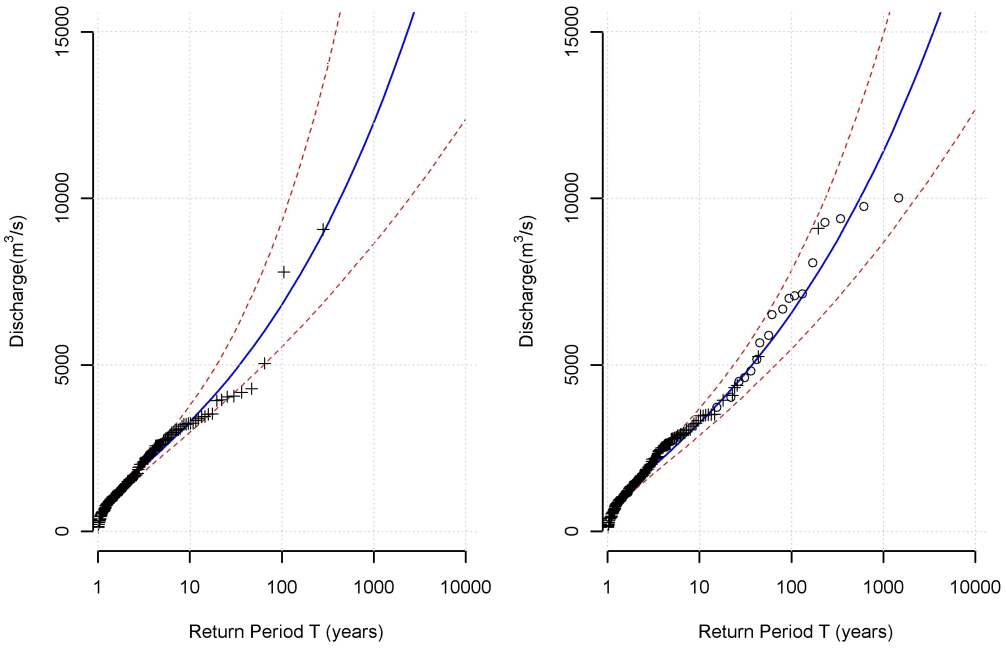

Figure 8: Results of the Bayesian MCMC regional flood frequency inferences conducted on the Ardèche dataset. Results provided for the Ardèche at Saint Martin (2240 km^2): (a) Method of Hosking & Wallis and (b) proposed approach including ungauged extremes. Statistical distributions corresponding to the maximum likelihood (continuous line) and 5% and 95% credibility bounds (dotted lines). Gauged values (crosses) and ungauged extremes (dots).

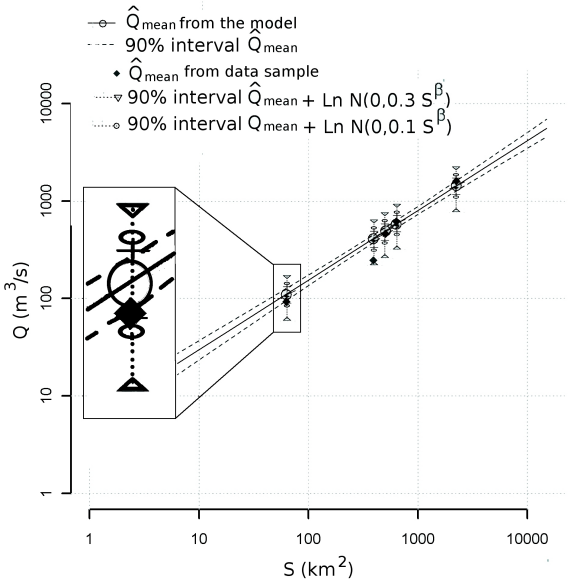

Figure 9: Calibrated relation between average discharges and watershed areas, sample mean discharges and variation ranges corresponding to the tested lognormal random models.

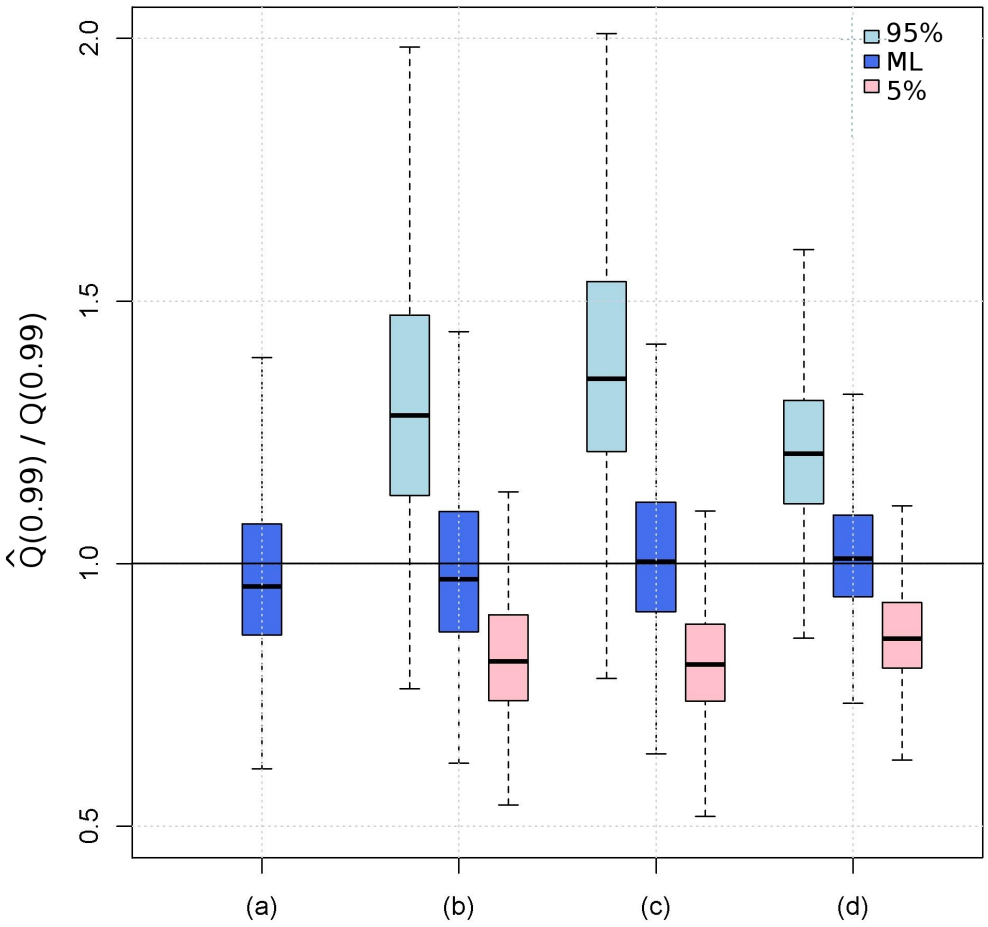

Figure 10: Maximum likelihood estimate and 5% and 95% bounds of the computed credibility intervals for 4 tested methods on 1000 simulated samples having the same characteristics as the Ardèche data sample (St Martin case): (a) method of Hosking and Wallis with L-moments, (b) method of Hosking and Wallis used in this paper, (c) proposed method applied on gauged series only, (d) proposed method with 18 ungauged extremes.