An application of Bayesian analysis and Markov chain Monte Carlo methods to the estimation of a regional trend in annual maxima

Benjamin Renard, V. Garreta, M. Lang

Author-produced version of the article published in Water Resour. Res., 42, W12422, doi:10.1029/2005WR004591.
The original publication is available at http://www.agu.org

An application of Bayesian analysis and MCMC methods to the estimation of a regional trend in annual maxima.

B. RENARD[1], V. GARRETA[1], M. LANG[1]

(1) *Cemagref Centre de Lyon, U.R. Hydrologie-Hydraulique, 3 bis Quai Chauveau, CP 220, 69336 Lyon cedex 09, France.*

telephone:33 4 72 20 87 72

fax:33 4 78 47 78 75

e-mail : renard@lyon.cemagref.fr

Index terms:

0520 Data analysis: algorithms and implementation

0560 Numerical solutions

3275 Uncertainty quantification

1817 Extreme events

1860 Streamflow

Keywords : Regional analysis; Trend detection; MCMC ; Bayesian analysis ; Gibbs ; Metropolis ;

Abstract

Bayesian analysis is becoming increasingly popular in a number of fields, including hydrology. It appears to be a convenient framework for deriving complex models in agreement with both physical reality and statistical requirements. The aim of this paper is to present an application to the regional frequency analysis of extremes in a non-stationary context. A non-stationary regional model is thus proposed, together with the related hypotheses. The Bayesian inference of this model is then described. Markov chain Monte Carlo (MCMC) methods are needed for this purpose, because of the dimensionality of the model, and are described in this paper. The usefulness of such a model is then illustrated on a hydrological case study concerning annual maximum discharges of several sites. The advantage of regional analysis compared to at-site estimation is thus highlighted. Moreover, the Bayesian framework allows for a direct and comprehensive inference based on the posterior distribution, and is able to take into account modeling uncertainties, which is particularly useful when the stationarity of a series can neither be ensured nor be totally rejected.

I. Introduction

Bayesian analysis has become increasingly popular in recent years, and has shown its practical benefits in a number of applied problems. As an example, publications dealing with Bayesian inference are now usual in the field of hydrology (Coles and Tawn, 1996; Khodja et al., 1998; Lu and Berliner, 1999; Perreault et al., 1999, 2000a, b, c; Thyer and Kuczera, 2000, 2003a, b; Wang, 2001; Thyer et al., 2002; Coles and Pericchi, 2003; Coles et al., 2003; Parent and Bernier, 2003; Perreault and Fortin, 2003; Tapsoba et al., 2004; Moyeed and Clarke, 2005; Reis and Stedinger, 2005; Renard et al., 2006). This popularity can be primarily explained by the rediscovery of the usefulness of Markov chain Monte Carlo (MCMC) algorithms in the early 1990s (Gelfand and Smith, 1990). The aim of this article will be to describe an application of the Bayesian analysis and the MCMC methods to the estimation of a regional trend in the distribution of annual maximum discharges of several sites.

The basis of the Bayesian paradigm is to use the posterior distribution of parameters for inference. Suppose $\boldsymbol{X}$ is a vector (or a matrix) of observed data, whose probability distribution is parameterized by $\boldsymbol{\theta}$. The parameters are here considered as random quantities. Let $\pi(\boldsymbol{\theta})$ denote the prior distribution of $\boldsymbol{\theta}$, which reflects the knowledge about the parameters before observing the data, and $p(\boldsymbol{X}|\boldsymbol{\theta})$ be the likelihood of the data. The posterior distribution is then defined by the Bayes relationship:

$$p(\boldsymbol{\theta}|\boldsymbol{X}) = \frac{p(\boldsymbol{X}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int p(\boldsymbol{X}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}} \tag{1}$$

Except for a small number of dimensions ($d \leq 3$), the posterior distribution is not easily tractable. First of all, a high dimensional distribution cannot directly be used to simply summarize the parameters' properties. Moreover, the denominator of equation (1), which is called the marginal likelihood of observations, can generally not be computed, which means

that the posterior distribution is only known up to a constant of proportionality. In order to overcome these difficulties, one way is to use a multivariate sample from the posterior distribution. Marginal distributions will then be easily estimated using marginal samples, leading to classical point, interval or distribution estimates.

MCMC methods are tools for simulating data from an arbitrary distribution. The principle is to create a Markov chain $(\boldsymbol{\theta}^{(t)})$ which converges to a stationary distribution that is the desired distribution. Some general schemes exist to achieve this simulation, namely the Gibbs (Geman and Geman, 1984) and the Metropolis (Metropolis and Ulam, 1949; Metropolis et al., 1953) samplers. A more general version of the latter is known as the Metropolis-Hastings algorithm (Hastings, 1970). The interest of MCMC methods lies in the fact that the preceding samplers are usable for sampling from a very wide range of distributions. Despite this quasi-universality, some problems can arise during the algorithm implementation, in particular numerical overflows, slowness of the runs, or difficulties in convergence assessment. The books by Gelman et al. (1995) and Tanner (1996) provide some guidelines to improve the algorithms' performances. Recently, El Adlouni et al. (2006) compared different methods to check the convergence of the chains.

An interesting application of the Bayesian framework lies in the estimation of changes in hydrological series. For at-site Gaussian data, such an application has been developed by Lee and Heghinian (1977) and Bernier (1994). Perreault et al. (2000a; b) generalized this approach by dealing with both change detection and frequency analysis in a non-stationary context. Renard et al. (2006) adapted the latter approach in the case of extreme data. In some situations, it can be anticipated that the cause of change has an impact at a wide spatial scale (e.g., climate-related change). In such cases, it would be relevant to search for consistent changes within a sample of several sites in a homogeneous region. Perreault et al. (2000c) therefore proposed a model for estimating a simultaneous single change in the mean of

several Gaussian series. Unfortunately, the adaptation of this approach to the extreme values context is far from obvious, both for practical and theoretical reasons.

The aim of this article is to describe a Bayesian regional model for annual maximum discharges in a non-stationary context. It is organized as follows. Firstly, the Bayesian framework is described (section 2). Models used at local and regional scales are presented, together with the hypotheses we made. The use of the posterior distribution for parameter estimation, change detection and frequency analysis is also explained. In section 3, the strategy we adopted to sample from the posterior distribution is described, based on the combined use of two MCMC algorithms. A hydrological case study involving six gauging stations in France is finally presented in section 4, before drawing some conclusions and discussing perspectives (section 5).

All the algorithms used in this study are available at the following address: http://www.lyon.cemagref.fr/hh/PNRH-NS/.

II. Bayesian inference

II.1. Non-stationary regional model

Let $X_t^{(i)}$ denote the annual maximum discharge of a station *i (i=1,…,p)* for a year *t*. Extreme value theory provides asymptotic arguments to the use of the generalized extreme value (GEV) distribution (Fisher and Tippett, 1928; Jenkinson, 1955), whose probability density function (pdf) can be written as follows:

$$GEV(x;\alpha,\beta,\xi) = \frac{1}{\alpha}\left[1-\xi\left(\frac{x-\beta}{\alpha}\right)\right]^{\frac{1}{\xi}-1} \exp\left\{-\left[1-\xi\left(\frac{x-\beta}{\alpha}\right)\right]^{\frac{1}{\xi}}\right\}, \quad (2)$$

$$\text{with } \alpha>0 \text{ and } 1-\xi\left(\frac{x-\beta}{\alpha}\right)>0.$$

α, β and ξ are the scale, position and shape parameters. In the stationary case, the annual maxima series can thus be modeled with the GEV distribution, $X_t^{(i)} \sim GEV(\alpha_i, \beta_i, \xi_i)$ ($M_0^{(i)}$ model). If the stationarity of the series can not be ensured, a pragmatic approach suggested by Coles (2001) consists in assuming that for a given year, the annual maximum still arises from a GEV distribution, but that the parameters of this distribution will be a function of time. In this paper, it will be assumed that the location parameter is a function of time, whereas the scale and shape parameters remain constant. This leads to the $M_1^{(i)}$ model, $X_t^{(i)} \sim GEV(\alpha_i, \beta_i (1+\delta_i t), \xi_i)$. Alternative models may be used, for instance polynomial trends, step changes, trends on the scale or the shape parameter, etc. The methodology presented in this paper can be used for the inference of such models. However, because extreme discharges series are known to be affected by a strong natural variability together with large measurement uncertainties, we favored a parsimonious description of the temporal changes in the annual maxima distribution.

The approach applied at the regional scale consists in deriving a regional model comprising local parameters, which are related to a given site, and regional parameters, which are assumed to be equal for all sites within the region, as suggested by Katz et al. (2002). In the stationary case, it was therefore assumed that the shape parameter is identical for all sites, which leads to the M_0 regional model $X_t^{(i)} \sim GEV(\alpha_i, \beta_i, \xi)$. A non-stationary model was derived by applying a regional trend to the location parameter, which leads to the M_1 regional model $X_t^{(i)} \sim GEV(\alpha_i, \beta_i (1+\delta t), \xi)$. It should be noticed that the assumptions we made about the data distribution, notably the hypothesis of identical shape and trend parameters for all sites within the region, may be irrelevant in heterogeneous regions. Methods for creating homogeneous regions have been extensively studied (e.g. Hosking and Wallis, 1997, or

Ouarda et al., 1999), and will not be described in this paper. However, the adequacy between the data and the model will have to be properly checked.

The estimation of the parameters of the M_1 model is the main objective of this paper. For a given year, data thus consist of a vector of annual maxima. Consequently, the multivariate distribution of this vector will be needed for estimation. Multivariate extreme value distributions constitute an area of active research for statisticians, but remain difficult to handle in practice, especially with more than two or three dimensions. Although some promising approaches have been proposed (Coles and Tawn, 1990; Bortot et al., 2000; Favre et al., 2004), this aspect of the problem is beyond the scope of this paper. We thus make the hypothesis of independence between stations. As noted by Katz et al. (2002), this should have little effect on point estimates, but could result in an underestimation of the parameter variances. The effect of ignoring spatial dependence will be discussed in more detail in a later section.

II.2. Bayesian inference

The first step in performing the Bayesian inference of the M_1 model is to compute the likelihood. Let $\boldsymbol{X}^{(i)} = (x_t^{(i)})_{t=1,\ldots,n}$ be the column vector of annual maxima for a particular site i, and $\boldsymbol{X} = (\boldsymbol{X}^{(1)},\ldots,\boldsymbol{X}^{(p)})$ be the data matrix. Let $\boldsymbol{\theta} = (\alpha_1, \beta_1, \ldots, \alpha_p, \beta_p, \xi, \delta)$ denote the parameters vector of size $2p+2$ arising from the M_1 model with p sites. Under the hypothesis of spatial independence between sites, the likelihood is computed as follows:

$$\begin{aligned} p(\boldsymbol{X} \mid \boldsymbol{\theta}) &= p\left(\boldsymbol{X}^{(1)},\ldots,\boldsymbol{X}^{(p)} \mid (\alpha_1, \beta_1, \ldots, \alpha_p, \beta_p, \xi, \delta)\right) \\ &= \prod_{t=1}^{n} p\left(x_t^{(1)},\ldots,x_t^{(p)} \mid (\alpha_1, \beta_1, \ldots, \alpha_p, \beta_p, \xi, \delta)\right) \\ &= \prod_{t=1}^{n}\prod_{i=1}^{p} GEV\left(x_t^{(i)}; (\alpha_i, \beta_i\,(1+\delta t), \xi)\right) \end{aligned} \tag{3}$$

The next step is to specify a prior distribution $\pi(\boldsymbol{\theta})$ for the parameters, without using the data. Several approaches can be considered for this purpose: prior knowledge of an expert (Coles and Tawn, 1996), additional data (Perreault et al., 2000c), empirical relationships with explanatory variables (Renard et al., 2006), regional information (Ribatet et al., 2006), etc. The lack of information about the phenomenon studied sometimes leads to the use of improper priors. However, such distributions can be problematic in the context of Bayesian model choice which will be described later. Consequently, it will be supposed in this paper that prior information is available and can be translated into a proper prior distribution.

Given the likelihood of the data and the prior distribution of the parameters, the posterior distribution of the parameters can be computed from the Bayes' theorem (equation 1). Because of the limitations explained in the introduction, a sample simulated from this distribution using a MCMC algorithm is needed in practice. This sample is used to estimate the marginal posterior distribution of each parameter and related characteristics (mean, variance, probability intervals, etc.).

An attractive characteristic of Bayesian analysis is the possibility of computing model probabilities. Suppose that $\Lambda = \{M_1, ..., M_q\}$ is a collection of models possibly suitable for describing the data. Assume that $\pi(.)$ is a prior probability distribution on Λ, which means that a prior probability can be assigned to each model without using the data. The posterior probability of a model M_j can then be computed with the Bayes theorem:

$$p(M_j \mid \boldsymbol{X}) = \frac{p(\boldsymbol{X} \mid M_j)\pi(M_j)}{\sum_{i=1}^{q} p(\boldsymbol{X} \mid M_i)\pi(M_i)} \tag{4}$$

This equation requires the computation of the marginal likelihood of observations for each model, which is equal to the normalizing factor of equation (1):

$$p(\boldsymbol{X} \mid M) = \int p(\boldsymbol{X}, \boldsymbol{\theta} \mid M) d\boldsymbol{\theta} = \int p(\boldsymbol{X} \mid \boldsymbol{\theta}, M) \pi(\boldsymbol{\theta} \mid M) d\boldsymbol{\theta} \quad (5)$$

As stated previously, this integration is generally intractable with a large number of dimensions. Alternative methods therefore have to be used. The approach proposed by Chib (1995) was used in this paper, and is described in Appendix 1.

The posterior probabilities of the models considered can be helpful in a decision-making process. Bayes factors (Kass and Raftery, 1995) can be used for this purpose, as they provide a measurement of the pertinence of one model compared to another. The Bayes factor between model M_i and model M_j is defined as follows:

$$B_{i,j} = \left(\frac{p(M_i \mid \boldsymbol{X})}{p(M_j \mid \boldsymbol{X})} \right) \Bigg/ \left(\frac{\pi(M_i)}{\pi(M_j)} \right) = \frac{p(\boldsymbol{X} \mid M_i)}{p(\boldsymbol{X} \mid M_j)} \quad (6)$$

This quantity is then compared to 1, with stronger confidence in model M_i for high values. Kass and Raftery (1995) furnished guidelines for interpreting Bayes factors, together with additional developments. For instance, composite Bayes factors can be used to compare two sets of models (e.g. several stationary models and several trend models).

Hydrologists involved in flood risk assessment are interested in p-quantiles of the annual maxima distribution, that is the value which is exceeded with a probability of 1-p. For a GEV distribution, the quantile q_p can be computed as:

$$q_p = \beta + \frac{\alpha}{\xi} \left(1 - (-\log(p))^{\xi} \right) \quad (7)$$

Given a model M_k, the posterior distribution of the quantile can thus easily be obtained from the posterior distribution of the parameters. Of course, in the case of a non-stationary model, the quantile posterior distribution will also depend on time. For instance, in the case of the M_1 model, the p-quantile for a given site i at year t will be:

$$q_p(t) = \beta_i \, (1 + \delta t) + \frac{\alpha_i}{\xi} \left(1 - (-\log(p))^{\xi} \right) \quad (8)$$

An alternative to the selection of a model using Bayes factors or any other criterion is to average over all the models considered. This can be done using the following equation:

$$p(q_p(t) \mid \boldsymbol{X}) = \sum_{k=1}^{q} p(q_p(t) \mid \boldsymbol{X}, M_k) p(M_k \mid \boldsymbol{X}) \qquad (9)$$

This approach is known as Bayesian model averaging. It allows the uncertainty related to the choice of a model to be taken into account. An underlying assumption is that the data arise from one of the models considered. Of course, as a model is a simplification of reality, this hypothesis is unlikely to be strictly fulfilled. However, by considering contrasting models, a more realistic quantification of the uncertainties can be obtained, compared to the approach consisting in choosing a single model on the basis of a given criterion. The paper by Hoeting et al. (1999) provides a comprehensive review of Bayesian model averaging.

III. MCMC methods

The principle of MCMC methods is to create a random walk in the parameters space which converges to a stationary distribution that is the joint posterior distribution (Gelman et al., 1995). Several algorithms have been proposed for this purpose, and have been described and compared by Tierney (1994a; b). Whatever the method used, an important feature is to check that the algorithm has been run for long enough to ensure the simulated sample to be representative of the target distribution. Consequently, the first iterations are usually not used for inference, and are considered as burning time. Moreover, the simulated values are not independent: a large number of iterations may thus be needed to obtain an acceptable accuracy.

In the following section, the MCMC strategy used in this paper is exposed. $p(\boldsymbol{\theta} \mid \boldsymbol{X})$ denotes the target posterior distribution of a d-dimensional parameter vector $\boldsymbol{\theta} = \left(\theta_1, ..., \theta_d\right)$.

III.1. Use of a combined algorithm

Metropolis sampler

The Metropolis algorithm (Metropolis and Ulam, 1949; Metropolis et al., 1953) has been widely used in Bayesian applications, because of its simplicity and its efficiency. Its principle can be summarized as follows: starting from a vector generated at iteration *i*-1, a new candidate vector is generated thanks to a symmetric jump distribution. If this new vector leads to an increasing value of the target distribution, it is accepted as the generated value at iteration *i*. Otherwise, the ratio between the new and the previous value of the target distribution is computed, and used as the acceptance probability of the candidate vector. In case of rejection, the generated vector at iteration *i* remains equal to that of iteration *i*-1.

The Metropolis algorithm was used in this paper with a Gaussian jump distribution with variance matrix Σ. Appendix 2 describes the way the algorithm was implemented. To start a Metropolis simulation, a starting point and a jump variance are therefore needed. The choice of a suitable starting point can be made in different ways: as an example, if informative priors are used, the prior mode can be chosen. Other methods for deriving a rough estimate of the posterior mode are available, and are described by Gelman et al. (1995). The choice of a jump variance is more delicate and can be problematic in practice, whereas it is the most important task to derive an efficient algorithm. More accurately, if the jump variance is too high, most of the candidate vectors will be generated far from the high-density area of the posterior distribution, thus leading to rejection of the candidate vector. The algorithm will then generate a great number of tied values (*i.e.* successive identical values), and the number of iterations necessary to visit the whole parameters space will be unacceptable. Alternatively, with a too small jump variance, almost all candidate vectors will be accepted, because each new vector will be generated very close to the precedent, leading to a ratio of target distributions close to one. Once again, the parameters space will be visited too slowly.

Several solutions can be considered for deriving an adequate variance matrix. As an example, it may be possible to use the observed Fisher information matrix as a first approximation. This may be relevant for large samples, because of the predominant influence of the likelihood compared to the prior distribution. Alternatively, if a proper prior distribution is specified, it can also be used to approximate the variance matrix. The most efficient solution may be to use a Taylor series expansion of the posterior density. However, an analytical computation of the derivatives will often be intractable, and numerical approximates will have to be computed.

Gibbs/Metropolis sampler

In order to overcome these difficulties, a possibility is to modify the characteristics of the algorithm during the iterations, by using the previously simulated values. For instance, it may be possible to periodically update the variance matrix of the jump distribution. This leads to the construction of adaptive algorithms. Gelman et al. (1995) provide some guidelines for this purpose. Some examples of adaptive Metropolis samplers can be found in the papers by Haario et al. (2001), Thyer et al. (2002) or Marshall et al. (2004).

The adaptive MCMC algorithm used in this paper is based on the Gibbs sampler (Geman and Geman, 1984). This sampling scheme is based on the principle of alternating conditional sampling: within an iteration i, each component of the parameters vector is updated successively, by sampling in the full conditional densities. More accurately, at iteration i, the component q is updated by sampling from the distribution $p\left(\theta_q \mid \theta_1^{(i)}, ..., \theta_{q-1}^{(i)}, \theta_{q+1}^{(i-1)}, ..., \theta_d^{(i-1)}, \boldsymbol{X}\right)$. Unfortunately, this full conditional density does generally not have the form of a standard distribution. It is thus necessary to add a new simulation step for each of the d conditional samplings. Several authors have proposed tools for this purpose. As an example, the Griddy-Gibbs sampling algorithm introduced by Ritter and Tanner (1992) is based on the inversion of the cumulative distribution function (cdf) of

the full conditional distribution, this cdf being estimated on a discrete grid of points. Alternatively, Gilks and Wild (1992) proposed using an acceptation-rejection (AR) method to sample from log-concave full conditional distributions, or the joint use of AR and Metropolis-Hastings algorithms if the log-concavity cannot be ensured (Gilks et al., 1995).

The approach applied in this paper is based on the suggestion of Gelman et al. (1996) and Coles and Tawn (1996) to use a Metropolis algorithm on the one-dimensional distribution $p\left(\theta_q \mid \theta_1^{(i)}, \ldots, \theta_{q-1}^{(i)}, \theta_{q+1}^{(i-1)}, \ldots, \theta_d^{(i-1)}, \boldsymbol{X}\right)$. If a univariate Gaussian distribution is used as the jump distribution, a jump variance is needed at each iteration of the Gibbs algorithm, and for each of the d conditional samplings. Choosing a constant jump variance for each conditional sampling may lead to a poor performance of the algorithm. Moreover, given the large number of Gibbs loops, it is not possible to manually check the convergence of the one-dimensional Metropolis runs. An automated procedure is therefore needed to compute an adequate jump variance for each conditional distribution.

To this end, it is possible to use the assumption that the distribution to sample from is roughly normal. This is not a drastic restriction in the Bayesian context, because the posterior distribution is asymptotically Gaussian under some regularity conditions. It is then possible to use the following property of the multivariate Gaussian distribution: the full conditional distribution is still Gaussian, and its variance does not depend on the values used for conditioning. This implies that, for each parameter, the variance can be transmitted between two iterations of the Gibbs algorithm. Moreover, this variance can be updated after each Metropolis run, in order to comply with an adequate acceptance rate. Gelman et al. (1995) recommended keeping the acceptance rate between 0.23 with a large number of dimensions (d>5) and 0.44 (d=1). Consequently, at the end of each one-dimensional Metropolis run, the acceptance rate τ is computed on the generated values. If $\tau > 0.44$, the jump variance is increased in order to decrease the acceptance rate. Similarly, if $\tau < 0.23$, the jump variance is

decreased. This leads to the implementation of an adaptive Gibbs/Metropolis sampler, which is described in Appendix 3.

Compared to the Metropolis sampler, this adaptive algorithm has a larger complexity. In order to simulate a sample of length N, the Metropolis sampler will thus execute N loops including computations of the target distribution, while this number will be $dN_{Metro}N$ for the Gibbs/Metropolis sampler, where N_{Metro} is the length of the one-dimensional Metropolis runs. However, it is anticipated that the adaptation rules will improve the efficiency of the algorithm, by decreasing the sensitivity to starting point or jump variance misspecification.

Combining samplers

Although very useful from a practical point of view, adaptive algorithms suffer from an important theoretical drawback: because the Markov chain properties of the simulated values are changed during the iterations, the stationarity of the target distribution is no longer ensured and has to be proven. Examples where an adaptive algorithm fails to generate values from the desired distribution have been provided by Tierney and Mira (1999) or Haario et al. (2001). At the opposite, the latter authors succeed in establishing the correct ergodic properties of their AM adaptive sampler. The adaptive Gibbs/Metropolis used in this paper was empirically tested on a number of cases, and gave so far acceptable results. However, we are not able to prove its ergodic properties.

Based on these considerations, should the use of adaptive algorithms be avoided? Actually, a number of such algorithms have been developed (Gilks and Wild, 1992; Roberts and Gilks, 1994; Gilks et al., 1995; Tierney and Mira, 1999; Haario et al., 2001, 2005) and have given acceptable results in most applications. As an illustration, Haario et al. (2001) noted concerning their adaptive AP sampler that "for many practical applications, the error produced by the AP algorithm is, however, ignorable", although a counter-example can be constructed. Nevertheless, in order to avoid errors due to particular target distributions, the

following two-step procedure was adopted: the adaptive Gibbs/Metropolis algorithm is first used for a limited number of iterations. The simulated values are then used to compute a mean vector $\boldsymbol{\mu}$ and a variance matrix $\boldsymbol{\Sigma}$, but will not be used for inference. As a second step, a traditional Metropolis algorithm without adaptation, whose starting point is equal to $\boldsymbol{\mu}$, is used in order to simulate the sample which will be used for inference. The jump distribution is a Gaussian distribution with variance matrix $c^2\boldsymbol{\Sigma}$, with $c = 2.4/\sqrt{d}$, as suggested by Gelman et al (1995). Because no adaptation is made on this latter sampler, the theoretical properties of the simulated Markov chain are ensured.

III.2. Influence of starting characteristics on algorithm performance

The aim of this section is to highlight some properties of the Metropolis, the adaptive Gibbs/Metropolis and the combined samplers previously described, based on simplified target distributions. This comparison does not aim at deriving general conclusions about the sampler performances, but just at illustrating some issues which can be encountered in practice.

Samples were generated from the three following target distributions:

- ✓ $f_1(x_1,...,x_5) = f_1(\boldsymbol{x}) = N(\boldsymbol{x} \mid \boldsymbol{0}; \boldsymbol{\Sigma})$, where the variance matrix has the form $\boldsymbol{\Sigma}(i,j) = 0.8^{|i-j|}$, for $i, j = 1,...,5$.

- ✓ $f_2(x_1,...,x_5) = \prod_{i=1}^{5} [1+0.2x_i]^{-6} \exp\left\{-[1+0.2x_i]^{-5}\right\}$, *i.e.* f_2 is a multivariate distribution with independent GEV marginal distributions, whose location parameters are 0, scale parameters are 1 and shape parameters are –0.2.

- ✓ $f_3(x_1,...,x_5) = f_3(\boldsymbol{x}) = 0.5N(\boldsymbol{x} \mid \boldsymbol{0}; \boldsymbol{I}) + 0.5N(\boldsymbol{x} \mid \boldsymbol{8}; \boldsymbol{I})$, *i.e.* f_3 is a mix of two multivariate normal distributions with independent marginal distributions and mean vectors equal to $\boldsymbol{0}$=(0;0;0;0;0) and $\boldsymbol{8}$=(8;8;8;8;8).

Inadequate starting characteristics were specified in order to evaluate the samplers' robustness. Starting vectors are equal to ***-10*** for the f_1 and f_3 distributions, and ***–4*** for the f_2 distribution (in order to ensure that $1-\xi\left(\frac{x-\beta}{\alpha}\right)>0$). The variance matrix $\boldsymbol{\Omega}$ of the Gaussian jump distribution used in the Metropolis algorithm is equal to:

$$\boldsymbol{\Omega}=\begin{pmatrix} 0.01 & 0 & 0 & 0 & 0 \\ 0 & 0.1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 0 & 100 \end{pmatrix} \qquad (10)$$

In this way, some components will have an underestimated variance, and some others an overestimated one. Similarly, the starting variances used in the Gibbs/Metropolis sampler are equal to (0.01;0.1;1;10;100).

The Metropolis algorithm was used to simulate 50 000 values from the target distribution. The first 20 000 iterations were considered as burning time. The remaining 30 000 values were then used for inference. The Gibbs/Metropolis algorithm was used with N_{Metro}=100, and a sample of length 5000 was simulated, with 2000 iterations of burning time. The combination of these two samplers was made as follows: the Gibbs/Metropolis sampler was used during 1000 iterations, with the starting points and the jump variance previously described. The mean vector and the variance matrix were then computed on the last 500 values, and used to derive the starting point and the variance matrix of the jump distribution of a Metropolis algorithm. 50 000 values were thus simulated, and the last 30 000 were used for inference.

The results presented in Figure 1 are related to the first components of the target distributions, whose jump variances were strongly underestimated. In the Metropolis case, it can be observed that the simulated values are a very poor representation of the expected distributions. In fact, both the burning time and the number of iterations are too low to

provide an adequate description of the target distribution, because of the small size of the jumps. Thereby, for the f_1 distribution, the chain is still influenced by the too low starting value. In the f_2 case, the algorithm seems unable to sample a value from the tail of the distribution. In the f_3 case, the algorithm is locked up in the first mode, which can be explained by the fact that the jump variance is too low to be able to reach the other mode with a significant probability. In contrast, the samples arising from the adaptive Gibbs/Metropolis algorithm provide an adequate representation of the three target distributions. Consequently, the combined algorithm appears useful for overcoming the issues observed with the Metropolis sampler for the f_1 and the f_2 distributions. The f_3 case is more mitigated: although the combined algorithm is able to sample in the two modes of the target distributions, the accuracy of the estimates is not satisfactory. This accuracy is however likely to be enhanced by increasing the number of iterations of the Metropolis part of the combined algorithm.

Similarly, results for the last components of each target distribution, whose jump variances were strongly overestimated, are presented in Figure 2. In the Metropolis case, the influence of the starting value is still observable for the f_1 target distribution. The fact that the marginal distributions are strongly positively correlated can explain this phenomenon: while the first components remain in the left tail of the distribution because of a too low jump variance, the other components are unlikely to reach the center or the right tail of the distribution. Conversely, in the f_2 and the f_3 cases, the values simulated with the Metropolis algorithm provide an adequate representation of the target distributions. Once again, the Gibbs/Metropolis sampler behaves satisfactorily, whatever the target distribution considered, while the accuracy of the values simulated by the combined sampler is adequate in the f_1 and the f_2 cases, but is not satisfactory with the f_3 distribution.

These results show that the combined sampler is able to overcome some issues observed with a simple Metropolis algorithm, by deriving acceptable starting characteristics. However,

because the inference will still be conducted using a Metropolis sample, the limitations related to the use of a Gaussian jump distribution may remain. As an illustration, the shape of the strongly multimodal distribution f_3 is too different from that of the jump distribution, which prevents the algorithm from behaving satisfactorily. A more subtle implementation, including a non-Gaussian instrumental distribution, is thus necessary to avoid these problems. Consequently, the Metropolis-Hastings algorithm is likely to be a more suitable tool in this case, because of the possibility of using an asymmetrical jump distribution. Moreover, this example also emphasizes the importance of monitoring convergence of the chain, preferably by several methods, as suggested by El Adlouni et al. (2006). In a real-life problem, as the target distribution would be unknown, a convergence assessment based on a single chain is likely to arrive at a false convergence in the case of a sampler locked up in one of the two modes. Conversely, by using several independent chains, obtained with overdispersed starting points (Gelman et al., 1995), the existence of several modes is more likely to be detected.

IV. Case study

IV.1. Data and prior elicitation

Six undisturbed hydrological stations located in the center of France were used (Figure 3). A description of the general characteristics of the watersheds is provided in Table 1. Daily discharges are available for various periods. Only years for which all series are free from missing values were considered, which leads to 34 years of data between 1961 and 2002, with 1981, 1982, 1983, 1988, 1992, 1996, 1997 and 2001 as missing years. For each station, the annual maximum value is extracted from the daily discharges series.

The M_1 model detailed in section II.1 is used to describe these data, with a minor modification to account for the beginning of the study period: $X_t^{(i)} \sim GEV(\alpha_i, \beta_i (1+\delta(t-1960)), \xi)$. A prior distribution is specified as follows. Firstly,

empirical knowledge is used for the regional parameters. The shape parameter is assumed to follow a Gaussian distribution with zero mean and standard deviation 0.3, which implies that the interval [-0.6;0.6] encompasses more than 95% of the density. This prior distribution can be compared with the Martins and Stedinger (2000) geophysical prior, which is less variable and is entirely included in the interval [-0.5;0.5]. We also assume that the regional trend parameter follows a Gaussian distribution. Because no evidence of climate-related changes has been reported for high flows in France, the mean is set to zero. Moreover, we consider that it is very unlikely that a change of more than 100% could have affected the location parameter $\beta_{j,t} = \beta_j (1 + \delta(t - 1960))$ between 1960 and 2000. More accurately, the prior variance is computed in order to meet the following specification:

$$\begin{aligned} &\Pr\left(\left| \frac{\beta_{j,2000} - \beta_{j,1960}}{\beta_{j,1960}} \right| \le 100\% \right) = 0.95 \\ &\Leftrightarrow \Pr\left(\left| \frac{\beta_j (1 + \delta(2000 - 1960)) - \beta_j (1 + \delta(1960 - 1960))}{\beta_j (1 + \delta(1960 - 1960))} \right| \le 100\% \right) = 0.95 \\ &\Leftrightarrow \Pr\left(\left| 40\delta \right| \le 100\% \right) = 0.95 \\ &\Leftrightarrow \Pr\left(\left| \delta \right| \le 1/40 \right) = 0.95 \end{aligned} \qquad (11)$$

This leads to the prior standard deviation for δ being set to 1/80, so that the change in 40 years, expressed as a percentage, lies in the interval [-100%;100%] with a probability larger than 0.95.

Empirical knowledge for the local location and scale parameters is harder to derive, because these parameters are more closely related to the specific characteristics of the watershed (area, topography, geology, etc.). Consequently, the data derived from non-shared years are used for this purpose. Maximum likelihood (ML) estimates of the local stationary GEV distributions are computed. Because of the limited sample size available for some sites, the use of the Fisher information matrix to derive the distribution of the estimates may lead to

biased results. Consequently, bootstrap samples of length 10 000 were generated for each site, and a Gamma distribution was fitted to the sample of ML estimates obtained:

$$Ga\left(x;a,b\right)=\frac{b^a}{\Gamma(a)}x^{a-1}\exp(-bx),\ x>0 \tag{12}$$

Table 2 gives details of the values obtained, and the number of data available for prior elicitation.

Finally, the prior independence of all parameters is assumed, which leads to the following prior distribution:

$$\begin{aligned}&\pi(\boldsymbol{\theta})=\pi(\alpha_1,\beta_1,...,\alpha_6,\beta_6,\xi,\delta)\\&=N(\xi;0,0.3^2)N\left(\delta;0,(1/80)^2\right)\prod_{j=1}^{6}Ga\left(\alpha_j;a_j^{(\alpha)},b_j^{(\alpha)}\right)Ga\left(\beta_j;a_j^{(\beta)},b_j^{(\beta)}\right)\end{aligned} \tag{13}$$

MCMC simulations were performed on the non-normalized posterior distribution $p(\boldsymbol{X}\mid\boldsymbol{\theta})\pi(\boldsymbol{\theta})$ using N_{Gibbs}=1000 iterations of the Gibbs/Metropolis sampler (N_{Metro}=100) followed by N_{iter}=100 000 iterations of the Metropolis one, as explained in section III. For each parameter, the starting point was set to the prior mean and the starting variance was set to the prior variance, which allows the first 1000 Gibbs/Metropolis iterations to be performed. The Metropolis sampler was then used with a starting point and a starting variance matrix derived from the last 500 Gibbs/Metropolis iterations.

IV.2. Assessment of MCMC runs convergence

As a first step, convergence of the MCMC simulations toward the posterior distribution has to be assessed. Unlike section III.2, the target distribution is here unknown, thus preventing any conclusion to be made concerning convergence by simply examining a single output of the MCMC runs. Gelman et al. (1995) suggested using parallel simulations performed with different starting points. Ten Gibbs/Metropolis simulations of length 1000 were thus used with overdispersed starting points. Figure 4 shows the random walks of parallel sequences in

the planes formed: (a) by scale and position parameters of the first station and (b) by shape and trend regional parameters. It can be observed that simulated points join the same area after a few iterations, which indicates that mean and variance characteristics computed from the last 500 Gibbs/Metropolis iterations will give rather good starting characteristics for the subsequent Metropolis iterations.

In order to quantify the minimum number of iterations needed to reach convergence, the approach suggested by Gelman et al. (1995) was applied. It is based on the computation of the $\sqrt{\hat{R}}$ statistic, where $\hat{R}$ estimates the ratio of total variance by within-sequences variance. If convergence is reached, then the between-sequences variance should be negligible, thus leading to a ratio close to one. For a simulation of length *n*, this ratio is estimated thanks to the second half of simulated values. Figure 5a shows the results for the 4 preceding parameters. It confirms that the Gibbs/Metropolis algorithm is very fast in reaching convergence, as the $\sqrt{\hat{R}}$ statistic declines to an acceptable value with only 50 iterations (values below 1.2 are usually considered as acceptable). Because only the last 500 generated values will be used for computing the mean and the variance matrix, it can be considered that these estimates will be adequate as starting characteristics of the subsequent Metropolis iterations.

Once this first step has been achieved by Gibbs/Metropolis sampling, additional Metropolis runs have to be performed. A high number of values is required here, firstly because of the tied values generated, secondly because this sample will be used to perform the inference on the posterior distribution. Once again, the approach of Gelman et al. (1995) was applied, with ten overdispersed starting points, and a starting variance matrix computed on the last 500 Gibbs/Metropolis iterations. Figure 5b shows the value of the $\sqrt{\hat{R}}$ statistic for different run lengths. After 9000 iterations, it can be considered that convergence is reached. As the total number of iterations is equal to 100 000, this will give a sufficient number of

values to perform inference on quantities of interest. More accurately, the following computations will be performed with the last 50 000 generated values.

IV.3. Regional trend estimation

Figure 6 shows the estimated posterior marginal distributions of two local parameters, namely the scale and the location parameters of the first station. Vertical lines represent locally-estimated parameters using the maximum likelihood method on the shared 34 years, and the thin lines represent the prior distributions. It can be observed that the results are consistent for the two parameters, although the Bayesian estimate of α_1 is larger than the corresponding ML-estimate, because of the influence of the prior distribution. Moreover, the posterior distribution of α_1 shows a slight departure from normality, which can easily be taken into account with the Bayesian framework, because no Gaussian asymptotic assumption is made. Figure 7 deals with the two regional parameters. Local and regional estimates are consistent for the trend parameter δ (right panel), which indicates that the assumption of a shared regional trend is plausible. This is also illustrated by Figure 8, which shows for each of the six stations the 34 years of data used for inference, together with the estimated regional trend. Moreover, the variance of the regional parameter is small compared to the range of locally-estimated values. This indicates that such a regional model is likely to reduce the estimation uncertainty. Focusing on the shape parameter estimate ξ, the assumption of an identical shape parameter over the region may appear doubtful (Figure 7, left panel). However, only the point estimates are shown, and the uncertainty is known to be very large for this parameter, especially with only 34 data points.

In order to derive a more accurate comparison between local and regional estimates, the posterior distribution in the regional model is compared to the posterior distributions of the

local models $M_1^{(i)}$. The prior distributions are specified as explained in section IV.1, which means that for a particular site *i*:

$$\begin{aligned}\pi(\boldsymbol{\theta}) &= \pi(\alpha_i, \beta_i, \xi_i, \delta_i) \\ &= N(\xi_i; 0, 0.3^2) N\left(\delta_i; 0, (1/80)^2\right) Ga\left(\alpha_i; a_i^{(\alpha)}, b_i^{(\alpha)}\right) Ga\left(\beta_i; a_i^{(\beta)}, b_i^{(\beta)}\right)\end{aligned} \quad (14)$$

For each site, MCMC simulations are performed on the non-normalized posterior distribution $p(\boldsymbol{X}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})$, as explained in section IV.1: the starting point and the starting variances of the Gibbs/Metropolis sampler are derived from equation (14) as the prior means and variances of each parameter. After 1000 iterations, the empirical mean and the empirical variance matrix are computed from the last 500 generated values, and are used to derive the starting characteristics of the Metropolis sampler, which will be run for 100 000 iterations. Figure 9 shows the box plots of the local trend and shape parameters, compared with the regional ones. Boxes extend from the first up to the third quartile, and the median is denoted by a line inside the box. The whiskers denote the quantiles with probabilities 0.05 and 0.95. The median values obtained for the shape parameters display a variability which is comparable to that obtained with the ML-estimates (vertical lines of Figure 7). However, the great uncertainty associated with the estimation of this parameter implies that all the 90% probability intervals are overlapping. A more complete investigation is therefore needed in order to evaluate the plausibility of the hypothesis of a regional shape parameter: this will be done in the next section IV.4. By contrast, the estimated median values of the trend parameters seem more consistent for the six stations. However, the uncertainty associated with each of the six at-site estimates remains strong, which explains the difficulty of detecting a trend in extreme data. For both parameters, the posterior variability of the regional estimates is strongly reduced compared to the local estimates: this emphasizes the benefit of a regional model, which increases the estimates precision by combining data from different locations.

This benefit may need to be mitigated by the assumption of independence between stations, which was made for convenience. This issue will be addressed in the discussion section.

IV.4. Model checking

In order to assess the plausibility of the hypotheses used to construct the M_1 model (GEV distribution, regional trend and shape parameters, independence between stations), the present section compares the posterior predictive distribution to the observed distribution of the data. If the model is appropriate, a new sample of data simulated from this model should be similar to the observed data. More accurately, a sample can be replicated from the posterior predictive density, which is defined as follows:

$$p(\boldsymbol{y}^* \mid \boldsymbol{X}) = \int p(\boldsymbol{y}^* \mid \boldsymbol{\theta})\, p(\boldsymbol{\theta} \mid \boldsymbol{X}) d\boldsymbol{\theta}\,, \qquad (15)$$

where $\boldsymbol{y}^*$ is an unknown but observable future value. By simulating a large number of samples arising from the model, the observed data can thus be located within the overall distribution of data that could have been observed (Gelman et al., 1995). In order to quantify how different the observed and the replicated data are, both having an identical sample size, the characteristics of their distributions are summarized in a test statistic $T(\boldsymbol{y})$. It is then possible to compute the *p*-value corresponding to the observed statistic $T(\boldsymbol{X})$, once replaced in the sample of $\left(T(\boldsymbol{y}^{(r)})\right)_{r=1,\ldots,N_{rep}}$ replicated from the posterior predictive distribution. More accurately, this *p*-value is estimated as the proportion of $\left(T(\boldsymbol{y}^{(r)})\right)_{r=1,\ldots,N_{rep}}$ which exceeds the observed statistic $T(\boldsymbol{X})$. An extension of this method involves discrepancy measures, *i.e.* statistics depending both on the data and parameters, rather than classical test statistics.

In the present case study, the problem is multivariate. The computation of the *p*-values will thus be made for each station. The choice of the test statistic has to reflect the departures from the model that can be expected. In the present case study, the main hypotheses are as follows:

- ✓ The local data follow a GEV distribution: this is a standard hypothesis when dealing with block maxima, because this distribution arises from extreme value theory. Test quantities related to the first three empirical moments are computed, namely the empirical mean, standard deviation and skewness.
- ✓ All the shape parameters are identical: this implies that all sites should have a comparable tail behavior. The skewness, which only depends on the shape parameter for GEV data, will be a first indicator. The maximum value of the sample is also computed, because it is very sensitive to the value of the shape parameter.
- ✓ All the stations are affected by a common trend: Kendall's tau coefficient between time and annual maxima is computed.
- ✓ The stations are independent: Kendall's tau coefficient between the annual maxima of two sites is derived.

The results are presented in Table 3. Values smaller than 0.05 or larger than 0.95 are denoted by an asterisk. The first three moments of the generated data are consistent with the observed ones. Moreover, the skewness and the maximum have acceptable *p*-values, which indicates that the regional shape parameter hypothesis is compatible with the observations. The regional trend hypothesis seems acceptable for five of the six stations. At the first station, the observed values lead to a larger trend than that simulated by the model. In contrast, the independence between stations is clearly doubtful: as an illustration, three of the five Kendall's tau coefficients between the first station and the others have a *p*-value smaller than 0.05. This is not surprising, as choosing a homogeneous region generally implies selecting dependent sites. The potential effect of this hypothesis, which is made for convenience, will be addressed in the discussion section.

IV.5. Frequency analysis

Although the model seems acceptable for describing the data, this does not imply that other models would not be able to do as well. The regional trend model M_1 is therefore compared with the corresponding stationary model M_0, which means that at-site annual maxima are assumed to arise from stationary GEV distributions, $X_t^{(i)} \sim GEV(\alpha_i, \beta_i, \xi)$.

Marginal likelihoods were first computed for models M_0 and M_1, thanks to the Chib approach presented in Appendix 1. The logarithm of the marginal likelihood was approximately equal to –863.2 for the stationary model and –859.8 for the regional trend model. Secondly, when equal prior probabilities for each model were assumed, the posterior probabilities were equal to (Equation (4)):

$$\hat{p}(M_0 \mid \boldsymbol{X}) = \frac{0.5\, p(\boldsymbol{X} \mid M_0)}{0.5\, p(\boldsymbol{X} \mid M_0) + 0.5\, p(\boldsymbol{X} \mid M_1)} = \frac{1}{1 + \exp(\log p(\boldsymbol{X} \mid M_1) - \log p(\boldsymbol{X} \mid M_0))} \approx 0.03 \quad (16)$$

and $\hat{p}(M_1 \mid \boldsymbol{X}) \approx 0.97$

The model assuming a consistent regional trend has a high probability, which means that a stationary regional model is far less suitable for describing these data. The Bayes factor between the two models is thus approximately equal to 30, which suggests, according to the scale of Kass and Raftery (1995), a strong evidence for the regional trend model.

The posterior distribution of the 0.9-quantile was then derived for each of the six sites, by using the model averaging procedure described in section II.2. The temporal changes in the posterior means are shown in Figure 10, together with 90% probability intervals. The quantile has decreased by approximately 10% for these six stations between years 1960 and 2000. Extrapolation to the future has to be considered with caution, because it implicitly assumes that the estimated trend will persist for ever. If the cause of the change is suspected, the use of a more suitable covariate than time may be able to deliver more realistic extrapolations.

The frequency analysis obtained with a regional model was finally compared with local analyses performed on each site. More accurately, for a given site *i*, the posterior distributions of the parameters in models $M_0^{(i)}$ and $M_1^{(i)}$ were estimated, leading to the posterior probabilities of each model and to the related Bayes factor. The results are presented in Table 4. It can be observed that the posterior probability of the stationary model remains far from zero or one, which indicates that at the at-site scale, there is no clear evidence for one model or the other. Consequently, Bayes factors are close to one, except for site 6, where the evidence for stationarity can be considered as "positive", according to the scale of Kass and Raftery (1995). This result emphasizes the advantage of regional analysis: by increasing the accuracy of regional estimates, several weak but consistent changes within a homogeneous region can result in a regional trend with high significance. The temporal changes of the 0.9-quantile were also studied at the local scale, and are shown in Figure 11. Bayesian model averaging is especially useful in this case: because of the difficulty of choosing between the two concurrent models, it leads to a more realistic quantification of the uncertainties, especially when extrapolating to the future. Compared to the quantiles obtained with the regional model, the probability intervals are thus wider, which reflects the stronger uncertainty related to local estimates. It can also be noticed that the posterior mean values of the quantiles obtained with the regional model remain inside the probability intervals obtained with the local models: regional and local estimates are thus consistent, which illustrates the adequacy of the regional model to describe these data.

IV.6. Discussion

The preceding case study provides an illustration of the usefulness of the Bayesian framework and the MCMC methods for the estimation of complex models. Estimation and detection of a trend in extreme data is generally acknowledged to be a difficult problem, because of the strong variability of such variables. The approach proposed here uses the

paradigm of regional analysis, which maintains that the estimation can be improved by gathering data from several sites in a homogeneous region. The more basic regional estimation scheme is based on the well-known index flood method (Dalrymple, 1960): within a homogeneous hydrological region, all scaled annual maxima series are assumed to have the same distribution. In practice, scaling is performed by dividing data by the sample means, as an example. In a second step, a unique distribution is fitted to this sample of normalized data. Unfortunately, this method implies a strong underestimation of the uncertainties. The alternative method is to translate the index flood assumption onto the local distributions parameters, as suggested by Katz et al. (2002). An example of such a study can be found in Buishand (1991), who used a two-steps estimation procedure. A more efficient estimation scheme would be to simultaneously estimate all parameters, which can be done by means of MCMC methods.

Moreover, a common drawback of almost, if not all the regional methods for extremes in hydrology, is that the spatial dependence of the data is ignored. This issue has been addressed by several authors, namely by Stedinger (1983), Hosking and Wallis (1988; 1997), or Madsen and Rosbjerg (1997). All these studies aim at quantifying the effect of inter-site dependence on the estimates' accuracy, rather than including it in a formal statistical model. A shared conclusion is that the existence of spatial dependence increases the variance of the estimates, but leaves the bias unchanged. Moreover, the estimates accuracy, in terms of RMSE, remains larger in the case of a regional analysis with dependent sites than in the case of at-site analysis. This conclusion is even valid in moderately heterogeneous regions, *i.e.* regions where the regional distribution is mis-specified. However, these conclusions have to be mitigated in the present case study. First of all, these authors used a different estimation method, namely the probability-weighted moments or the *L*-Moments. Secondly, these results have been obtained using Monte Carlo simulations, and their validity needs to be verified in

other simulation contexts. More accurately, dependent data have been simulated with a Gaussian copula: a multivariate Gaussian data set is first generated, and marginal data are transformed in order to fit an extreme value distribution. Although this dependence structure seems well-supported by some real-world data, it can not be considered as universal: for instance, Hosking and Wallis (1988) provide an example where very high flood events are more dependent than moderate ones. They explain this structure by the tendency of extreme floods to be generated by meteorological conditions that affect large areas. Based on a simulation study which mimics this structure, they found that the loss of accuracy is larger than in the case of the constant correlation produced by the Gaussian copula. Notice that in some regions, the opposite structure might be observed (smaller dependence for very high flood events), because of the tendency of extreme floods to be generated by convective storms, which only affect a limited area. Finally, a major difference between these studies and the case study presented in this paper is the existence of a trend. This additional parameter has a strong variability, which may affect the influence of the spatial dependence. Consequently, additional studies involving contrasting dependence structures in a non-stationary context would be of great interest.

V. Conclusion and perspectives

The aim of this article was to explore a non-stationary regional model for describing annual maximum discharges of several gauging stations within a homogeneous region. The estimation of the parameters was achieved by means of the Bayesian framework and MCMC methods, which are convenient tools for overcoming the issues related to such complex models.

The regional approach adopted in this study consists in considering that some parameters should be identical for all sites within the region. The existence of such regional parameters

can be viewed as a consequence of standard regional methods (e.g. the index flood method). A regional model has the advantage of avoiding a normalization step, which leads to an underestimation of the uncertainty. Moreover, in a non-stationary context, prior beliefs about the structure of a change are often available. For instance, if a change is related to a large-scale climatic phenomenon, it is natural to search for consistency within a homogeneous region. Finally, from a statistical point of view, regional analysis is a way of decreasing the estimation uncertainties, by combining data from different locations. This is particularly useful in the case of extreme value models, whose parameters are known to be particularly difficult to estimate. The case study described in this paper illustrates the benefit of a regional inference compared to at-site analyses, especially in terms of trend detection.

The Bayesian framework appears as a suitable method for the inference of this kind of model, because it presents several advantages compared to classical likelihood-based approaches. Firstly, the prior distribution is a convenient way to incorporate all the available data and knowledge in the analysis. Secondly, the posterior distribution allows a complete and comprehensive inference to be derived, without using any asymptotical assumption, which may be irrelevant for small samples. Moreover, in the case of extreme value distributions, the end-point of the density is a function of the parameters, because of the condition $1-\xi\left(\frac{x-\beta}{\alpha}\right)>0$. Consequently, some parameter values are incompatible with the data. A blind application of the asymptotic normality argument may thus lead in some cases to the assignment of a significant non-zero density to impossible parameters values. This drawback does not exist in the Bayesian framework, because the posterior density would be strictly equal to zero for such parameter values. Finally, an interesting characteristic of the Bayesian analysis is the possibility of taking into account modeling uncertainties thanks to the Bayesian model averaging approach.

In spite of these advantages, the development of Bayesian models used to be limited by numerical difficulties. Much progress on MCMC methods has been achieved since then, and has provided a number of tools for overcoming the numerical issues. Consequently, MCMC methods can find applications in a very wide range of situations, and are especially useful in problems with a large number of dimensions, where standard optimization methods can fail. However, some care is needed in order to assess the algorithms' convergence or to prevent numerical errors. Accordingly, the use of MCMC methods requires familiarity with the algorithms' functioning, as using them as a "black box" tool can lead to misinterpreted results. In this paper, both an adaptive sampler and a Metropolis algorithm were used. The adaptive Gibbs/Metropolis algorithm might suffer from theoretical limitations, and should therefore be used with caution.

The usefulness of the regional model studied in this paper has been illustrated by the case study. However, several improvements may be achieved, in order to consider less restrictive hypotheses. Firstly, the equality of some parameters over the whole region is a strong assumption. A more plausible hypothesis consists in assuming that such parameters are the realization of a spatial process. In this way, these parameters do not need to be equal anymore, but should respect structural relationships which ensure smooth variations in the region. Once again, the Bayesian approach is a suitable tool for deriving such models, by means of a Bayesian hierarchical model. Perreault (2000) developed this approach in the Gaussian case in order to model consistent step-changes in several sites. The model proposed by Cooley (2005) is a hierarchical regional model for extreme precipitations, where the variation of the parameters in the region is viewed as the result of a spatial process, which includes covariates such as altitude. The application of this kind of model to a non-stationary context would be a very interesting development.

As discussed in the case study, the spatial independence hypothesis is also a restrictive assumption, whose impact is not obvious to anticipate, especially in a non-stationary context. A major improvement would thus be to model dependence explicitly. Unfortunately, a number of problems still remain from a theoretical point of view in the case of extreme data. Multivariate models for extremes are still challenging, as most of the existing approaches are based on a too poor description of dependences, or are too difficult to handle in real-life problems. Specific methods have thus been studied for modeling spatial extremes (Schalther and Tawn, 2003; Naveau et al., 2005; De Haan and Pereira, 2006). It has to be explored how these approaches can be included in regional models.

More generally, recent developments in statistics of extremes provide a number of tools for hydrological applications. Deriving adequate and explicit statistical models together with convenient inference methods is likely to comply with more statistical rigor and to greatly improve the description of hydrological phenomena.

VI. Bibliography

Bernier, J. 1994. Statistical detection of changes in geophysical series. Engineering Risk in Natural Resources Management **275**:159 - 176.

Bortot, P., S. Coles, and J. A. Tawn. 2000. The multivariate Gaussian tail model: an application to oceanographic data. Applied Statistics **49**:31-49.

Buishand, T. A. 1991. Extreme Rainfall Estimation by Combining Data from Several Sites. Hydrological Sciences Journal-Journal Des Sciences Hydrologiques **36**:345-365.

Chib, S. 1995. Marginal Likelihood From the Gibbs Output. Journal of American Statistical Association **90**:1313-1321.

Coles, S. 2001. An Introduction to Statistical Modeling of Extreme Values. Springer-Verlag. 210 p. Springer-Verlag, London.

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

Coles, S., L. R. Pericchi, and S. Sisson. 2003. A fully probabilistic approach to extreme rainfall modelling. Journal of Hydrology **273**:35-50.

Coles, S., and J. A. Tawn. 1990. Statistics of coastal flood prevention. Philosophical Transactions of the Royal Society of London Series a - Mathematical Physical and Engineering Sciences **332**:457-476.

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

Cooley, D. 2005. Statistical Analysis of Extremes Motivated by Weather and Climate Studies: Applied and Theoretical Advances. University of Colorado. 122 p.

Dalrymple, T. 1960. Flood frequency analyses. *in* Water-supply paper 1553-A. US Geological Survey.

De Haan, L., and T. T. Pereira. 2006. Spatial extremes: Models for the stationary case. Annals of Statistics **34**:146-168.

El Adlouni, S., A. C. Favre, and B. Bobee. 2006. Comparison of methodologies to assess the convergence of Markov chain Monte Carlo methods. Computational Statistics and Data Analysis. In press.

Favre, A. C., S. El Adlouni, L. Perreault, N. Thiemonge, and B. Bobee. 2004. Multivariate hydrological frequency analysis using copulas. Water Resources Research **40**.

Fisher, R. A., and L. H. Tippett. 1928. Limiting forms of the frequency distribution of the largest or smallest member of a sample. Cambridge Phil. Soc. **24**.

Gelfand, A. E., and A. F. M. Smith. 1990. Sampling-based approaches to calculating marginal densities. Journal of American Statistical Association **85**:398-409.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 1995. Bayesian data analysis. C. Hall. 526 p. Chapman & Hall.

Gelman, A., G. O. Roberts, and W. R. Gilks. 1996. Efficient Metropolis Jumping Rules. Pages 599-607 *in* Bayesian Statistics 5. Oxford University Press.

Geman, S., and D. Geman. 1984. Stochastic relaxation, Gibbs distributions, and the bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence **6**:721-741.

Gilks, W. R., N. G. Best, and K. K. C. Tan. 1995. Adaptive Rejection Metropolis Sampling within Gibbs Sampling. Applied Statistics-Journal of the Royal Statistical Society Series C **44**:455-472.

Gilks, W. R., and P. Wild. 1992. Adaptive Rejection Sampling for Gibbs Sampling. Applied Statistics-Journal of the Royal Statistical Society Series C **41**:337-348.

Haario, H., E. Saksman, and J. Tamminen. 2001. An adaptive Metropolis algorithm. Bernoulli **7**:223-242.

Haario, H., E. Saksman, and J. Tamminen. 2005. Componentwise adaptation for high dimensional MCMC. Computational Statistics **20**:265-273.

Hastings, W. K. 1970. Monte Carlo sampling methods using Markov chains and their applications. Biometrika **57**:97-109.

Hoeting, J. A., D. Madigan, A. E. Raftery, and C. T. Volinsky. 1999. Bayesian model averaging: A tutorial. Statistical Science **14**:382-401.

Hosking, J. R. M., and J. R. Wallis. 1988. The effect of intersite dependence on regional flood frequency analysis. Water Resources Research **24**:588-600.

Hosking, J. R. M., and J. R. Wallis. 1997. Regional Frequency Analysis: an approach based on L-Moments. 226 p. Cambridge University Press, Cambridge, UK.

Jenkinson, A. F. 1955. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. Quarterly Journal of the Royal Meteorological Society **81**:158-171.

Kass, R. E., and A. E. Raftery. 1995. Bayes Factors. Journal of the American Statistical Association **90**:773-795.

Katz, R. W., M. B. Parlange, and P. Naveau. 2002. Statistics of extremes in hydrology. Advances in Water Resources **25**:1287-1304.

Khodja, H., H. Lubès-Niel, J. M. Sabatier, E. Servat, and J. E. Paturel. 1998. Analyse spatio-temporelle de données pluviométriques en Afrique de l'Ouest. Recherche d'une rupture en moyenne. Une alternative intéressante: les tests de permutations. Revue de Statistique Appliquée **46**:95 -110.

Lee, A. F. S., and S. M. Heghinian. 1977. A shift of the mean level in a sequence of independant normal random variables - a Bayesian approach. Technometrics **19**:503-506.

Lu, Z. Q., and L. M. Berliner. 1999. Markov switching time series models with application to a daily runoff series. Water Resources Research **35**:523-534.

Madsen, H., and D. Rosbjerg. 1997. The partial duration series method in regional index-flood modeling. Water Resources Research **33**:737-746.

Marshall, L., D. Nott, and A. Sharma. 2004. A comparative study of Markov chain Monte Carlo methods for conceptual rainfall-runoff modeling. Water Resources Research **40**.

Martins, E. S., and J. R. Stedinger. 2000. Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. Water Resources Research **36**:737-744.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. 1953. Equation of state calculations by fast computing machines. Journal of chemical physics **21**:1087-1092.

Metropolis, N., and S. Ulam. 1949. The Monte Carlo method. Journal of the American Statistical Association **44**:335-341.

Moyeed, R. A., and R. T. Clarke. 2005. The use of Bayesian methods for fitting rating curves, with case studies. Advances in Water Resources **28**:807-818.

Naveau, P., D. Cooley, and P. Poncet. 2005. Spatial extremes analysis in climate studies. *in*. Extreme Value Analysis, Gothenburg, Sweden.

Ouarda, T. B. M. J., M. Lang, B. Bobee, J. Bernier, and P. Bois. 1999. Analysis of regional flood models utilized in France and Quebec. Revue des Sciences de l'Eau **12**:155-182.

Parent, E., and J. Bernier. 2003. Encoding prior experts judgments to improve risk analysis of extreme hydrological events via POT modeling. Journal of Hydrology **283**:1-18.

Perreault, L. 2000. Analyse bayésienne rétrospective d'une rupture dans les séquences de variables aléatoires hydrologiques. Thesis. ENGREF, INRS-Eau. 200 p.

Perreault, L., J. Bernier, B. Bobee, and E. Parent. 2000a. Bayesian change-point analysis in hydrometeorological time series. Part 1. The normal model revisited. Journal of Hydrology **235**:221-241.

Perreault, L., J. Bernier, B. Bobee, and E. Parent. 2000b. Bayesian change-point analysis in hydrometeorological time series. Part 2. Comparison of change-point models and forecasting. Journal of Hydrology **235**:242-263.

Perreault, L., and V. Fortin. 2003. Mixture and Hidden Markov models for peak flow analysis. *in*. Seizièmes entretiens du centre Jacques Cartier, Lyon, France.

Perreault, L., M. Hache, M. Slivitzky, and B. Bobee. 1999. Detection of changes in precipitation and runoff over eastern Canada and US using a Bayesian approach. Stochastic Environmental Research and Risk Assessment **13**:201-216.

Perreault, L., E. Parent, J. Bernier, B. Bobee, and M. Slivitzky. 2000c. Retrospective multivariate Bayesian change-point analysis: A simultaneous single change in the mean of several hydrological sequences. Stochastic Environmental Research and Risk Assessment **14**:243-261.

Reis, D. S., and J. R. Stedinger. 2005. Bayesian MCMC flood frequency analysis with historical information. Journal of Hydrology **313**:97-116.

Renard, B., M. Lang, and P. Bois. 2006. Statistical analysis of extreme events in a non-stationary context via a Bayesian framework. Stochastic Environmental Research and Risk Assessment. Under press.

Ribatet, M., E. Sauquet, J. M. Gresillon, and T. B. M. J. Ouarda. 2006. A regional Bayesian POT model for flood frequency analysis. Stochastic Environmental Research and Risk Assessment. accepted.

Ritter, C., and M. A. Tanner. 1992. Facilitating the Gibbs Sampler - the Gibbs Stopper and the Griddy-Gibbs Sampler. Journal of the American Statistical Association **87**:861-868.
Roberts, G. O., and W. R. Gilks. 1994. Convergence of Adaptive Direction Sampling. Journal of Multivariate Analysis **49**:287-298.
Schalther, M., and J. A. Tawn. 2003. A dependence measure for multivariate and spatial extreme values: Properties and inference. Biometrika **90**:139-156.
Stedinger, J. R. 1983. Estimating a regional flood frequency distribution. Water Resources Research **19**:503-510.
Tanner, M. A. 1996. Tools for Statistical Inference. Springer-Verlag. 208 p. New York.
Tapsoba, D., M. Hache, L. Perreault, and B. Bobee. 2004. Bayesian Rainfall Variability Analysis in West Africa along Cross Sections in Space-Time grid Boxes. Journal of Climate **17**:1069-1082.
Thyer, M., and G. Kuczera. 2000. Modeling long-term persistence in hydroclimatic time series using a hidden state Markov model. Water Resources Research **36**:3301-3310.
Thyer, M., and G. Kuczera. 2003a. A hidden Markov model for modelling long-term persistence in multi-site rainfall time series 1. Model calibration using a Bayesian approach. Journal of Hydrology **275**:12-26.
Thyer, M., and G. Kuczera. 2003b. A hidden Markov model for modelling long-term persistence in multi-site rainfall time series. 2. Real data analysis. Journal of Hydrology **275**:27-48.
Thyer, M., G. Kuczera, and Q. J. Wang. 2002. Quantifying parameter uncertainty in stochastic models using the Box-Cox transformation. Journal of Hydrology **265**:246-257.
Tierney, L. 1994a. Markov-Chains for Exploring Posterior Distributions. Annals of Statistics **22**:1701-1728.
Tierney, L. 1994b. Markov-Chains for Exploring Posterior Distributions - Rejoinder. Annals of Statistics **22**:1758-1762.
Tierney, L., and A. Mira. 1999. Some adaptive Monte Carlo methods for Bayesian inference. Statistics in Medicine **18**:2507-2515.
Wang, Q. J. 2001. A Bayesian joint probability approach for flood record augmentation. Water Resources Research **37**:1707-1712.

Appendix 1: the Chib method

Marginal likelihood estimation will be based on the following observation, which holds for every $\boldsymbol{\theta} = (\theta_1, ..., \theta_d)$ where π and p are defined:

$$p(\boldsymbol{X} \mid M) = \frac{\pi(\boldsymbol{\theta})\, p(\boldsymbol{X} \mid \boldsymbol{\theta})}{p(\boldsymbol{\theta} \mid \boldsymbol{X})} = \frac{f(\boldsymbol{\theta})}{p(\boldsymbol{\theta} \mid \boldsymbol{X})} \tag{17}$$

In this equation, f is the non-normalized target distribution, which can always be computed at an arbitrary $\boldsymbol{\theta}$. The problem thus consists in estimating the posterior density of $\boldsymbol{\theta}$, by using samples generated by MCMC methods.

Chib (1995) proposed using the following equation:

$$p(\theta_1^*,...,\theta_d^* \mid \boldsymbol{X}) = p(\theta_1^* \mid \boldsymbol{X})\, p(\theta_2^* \mid \theta_1^*, \boldsymbol{X}) ... p(\theta_q^* \mid \theta_1^*, \theta_2^*, ..., \theta_{q-1}^*, \boldsymbol{X}) \\ ... \times p(\theta_d^* \mid \theta_1^*, ..., \theta_{d-1}^*, \boldsymbol{X}) \quad , \qquad (18)$$

where $\boldsymbol{\theta}^* = (\theta_1^*,...,\theta_d^*)$ is a given parameter vector with dimension *d*, preferably lying in a high density area. Each of the terms in equation (18) is a marginal density: for instance, $p(\theta_q^* \mid \theta_1^*, \theta_2^*, ..., \theta_{q-1}^*, \boldsymbol{X})$ is the first marginal of the multivariate distribution $p(\theta_q, \theta_{q+1}, ..., \theta_d \mid \theta_1^*, \theta_2^*, ..., \theta_{q-1}^*, \boldsymbol{X})$, evaluated at θ_q^*. MCMC methods can therefore be applied to the non-normalized posterior density with the first *q*-1 components being fixed, that is $f(\theta_1^*, \theta_2^*, ..., \theta_{q-1}^*, \theta_q, \theta_{q+1}, ..., \theta_d)$, in order to generate a sample from this distribution. The problem then reduces to univariate density estimation using a sample of values. In this paper, a Gaussian kernel is used for this purpose.

Appendix 2: the Metropolis sampler

The Metropolis sampler was implemented as follows:

- ✓ Choose a starting point $\boldsymbol{\theta}^{(0)}$ and a variance matrix $\boldsymbol{\Sigma}$.
- ✓ Do $i = 1, ..., N_{iter}$
  - ✓ Generate a candidate vector $\boldsymbol{\theta}^* \sim N(. \mid \boldsymbol{\theta}^{(i-1)}, \boldsymbol{\Sigma})$
  - ✓ If $p(\boldsymbol{\theta}^* \mid \boldsymbol{X}) \geq p(\boldsymbol{\theta}^{(i-1)} \mid \boldsymbol{X})$, set $\boldsymbol{\theta}^{(i)} = \boldsymbol{\theta}^*$, else accept the candidate vector ($\boldsymbol{\theta}^{(i)} = \boldsymbol{\theta}^*$) with probability $r = \dfrac{p(\boldsymbol{\theta}^* \mid \boldsymbol{X})}{p(\boldsymbol{\theta}^{(i-1)} \mid \boldsymbol{X})}$ or reject it ($\boldsymbol{\theta}^{(i)} = \boldsymbol{\theta}^{(i-1)}$) with probability 1-*r*.

In order to avoid numerical overflows, it is useful to consider the logarithm of the posterior distribution, and to compute the posterior ratio as

$r = \exp\left(\log\left(p(\boldsymbol{\theta}^* \mid \boldsymbol{X})\right) - \log\left(p(\boldsymbol{\theta}^{(i-1)} \mid \boldsymbol{X})\right)\right)$. Moreover, this ratio is invariant by multiplying the posterior distribution by a constant, which implies that the Metropolis can be applied to a non-normalized target distribution.

Appendix 3. The Adaptive Gibbs/Metropolis sampler.

This algorithm is implemented as follows:

- ✓ Choose a starting point $\boldsymbol{\theta}^{(0)} = \left(\theta_1^{(0)}, ..., \theta_d^{(0)}\right)$, and starting variances for each component $\boldsymbol{V}^{(0)} = \left(V_1^{(0)}, ..., V_d^{(0)}\right)$
- ✓ Do $j = 1, ..., N_{Gibbs}$:
  - ✓ Do q=1,…,d
    - ✓ Let $\gamma^{(0)} = \theta_q^{(j-1)}$ and $f_q^{(j)}(x) = p\left(x \mid \theta_1^{(j)}, ..., \theta_{q-1}^{(j)}, \theta_{q+1}^{(j-1)}, ..., \theta_d^{(j-1)}, \boldsymbol{X}\right)$
    - ✓ Do i=1,…,N_{Metro}
      - ✓ Generate a candidate value $\gamma^* \sim N(. \mid \gamma^{(i-1)}, V_q^{(j-1)})$
      - ✓ If $f_q^{(j)}(\gamma^*) \geq f_q^{(j)}(\gamma^{(i-1)})$, set $\gamma^{(i)} = \gamma^*$, else accept the candidate vector ($\gamma^{(i)} = \gamma^*$) with probability $r = \dfrac{f_q^{(j)}(\gamma^*)}{f_q^{(j)}(\gamma^{(i-1)})}$ or reject it ($\gamma^{(i)} = \gamma^{(i-1)}$) with probability 1-$r$.
    - ✓ Set $\theta_q^{(j)} = \gamma^{(N_{Metro})}$
    - ✓ Compute the acceptance rate τ on the previous N_{Metro} iterations
    - ✓ If τ<0.23, set $V_q^{(j)} < V_q^{(j-1)}$
    - ✓ If τ>0.44, set $V_q^{(j)} > V_q^{(j-1)}$

List of captions

Tables

Table 1: Hydrological properties of the six stations used in the case study.

Table 2: Prior specifications for the location and scale parameters. *a* and *b* are the parameters of the Gamma distributions used as prior distributions.

Table 3: Bayesian *p*-values obtained for various characteristics of the regional distribution.

Table 4: Posterior probabilities of local stationary models (compared to local trend models), and related Bayes factors.

Figures

Figure 1: Simulated values of the first component from three target distributions. Real marginal densities are denoted by a solid line. (a) Gaussian distribution (b) Independent GEV marginal distributions. (c) Independent mixed-Gaussian marginal distributions.

Figure 2: Simulated values of the last component from three target distributions. Real marginal densities are denoted by a solid line. (a) Gaussian distribution (b) Independent GEV marginal distributions. (c) Independent mixed-Gaussian marginal distributions.

Figure 3: Location of the six sites used in the case study.

Figure 4: Gibbs/Metropolis simulated random walks with ten different starting points (black squares): (a) scale and location parameters for the first site, (b) shape and trend regional parameters.

Figure 5: $\sqrt{\hat{R}}$ statistic for different run lengths. (a) Gibbs/Metropolis, (b) Metropolis.

Figure 6: Histograms of marginal simulated samples for two local parameters of site 1. Locally-estimated values are denoted by vertical lines and prior distributions by thin lines.

Figure 7: Histograms of marginal simulated samples for the two regional parameters. Locally-estimated values are denoted by vertical lines and prior distributions by thin lines.

Figure 8: Annual maxima series for the six stations studied. The line denotes the regional trend on the mean.

Figure 9: Comparison of at-site estimates (site number 1,…,6) and regional estimates for the shape (left) and the trend (right) parameters.

Figure 10: Posterior mean and 90% probability interval for the 0.9-quantile of the six stations studied. The posterior distributions arise from the regional models M_0 and M_1.

Figure 11: Posterior mean and 90% probability interval for the 0.9-quantile of the six stations studied. For each site *i*, the posterior distributions arise from the local models $M_0^{(i)}$ and $M_1^{(i)}$. For comparison purposes, the axes are identical to those of Figure 10.

Site Number	River / Station	Drainage area (km²)	Mean annual flow (m3.s-1)	10-year daily discharge (m3.s-1)
1	Dordogne / Saint Sauves d'Auvergne	87	3.6	43
2	Santoire / Condat	172	4.71	74
3	Vézère / Bugeat	143	4.43	47
4	Vézère / Uzerche	601	15.1	136
5	Corrèze / Corrèze	168	5.6	61
6	Corrèze / Brive-la-Gaillarde	947	21.2	268

Table 1

Site number	Sample size used for prior specification	Location parameter		Scale parameter	
		a	*b*	*a*	*b*
1	39	20.03	2.46	41.88	2.68
2	20	21.09	1.72	84.11	2.42
3	18	28.39	2.20	65.12	2.26
4	51	66.96	2.10	244.19	3.19
5	26	21.51	1.32	51.40	1.47
6	48	72.73	1.39	346.00	2.22

Table 2

	Bayesian *p*-value					
	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6
Mean	0.547	0.202	0.741	0.911	0.749	0.922
Standard deviation	0.400	0.154	0.711	0.936	0.919	0.939
Skewness	0.156	0.185	0.346	0.801	0.851	0.542
Maximum value	0.288	0.127	0.724	0.914	0.920	0.807
Kendall's tau with time	0.976 *	0.786	0.398	0.622	0.372	0.670
Kendall's tau with site 1	/	0.003 *	0.068	0.042 *	0.036 *	0.119

Table 3

	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6
Posterior probability $p(M_0^{(i)}\|\boldsymbol{X})$	0.66	0.30	0.38	0.53	0.28	0.79
Bayes factor $B_{0,1}$	1.95	0.43	0.61	1.11	0.39	3.86

Table 4

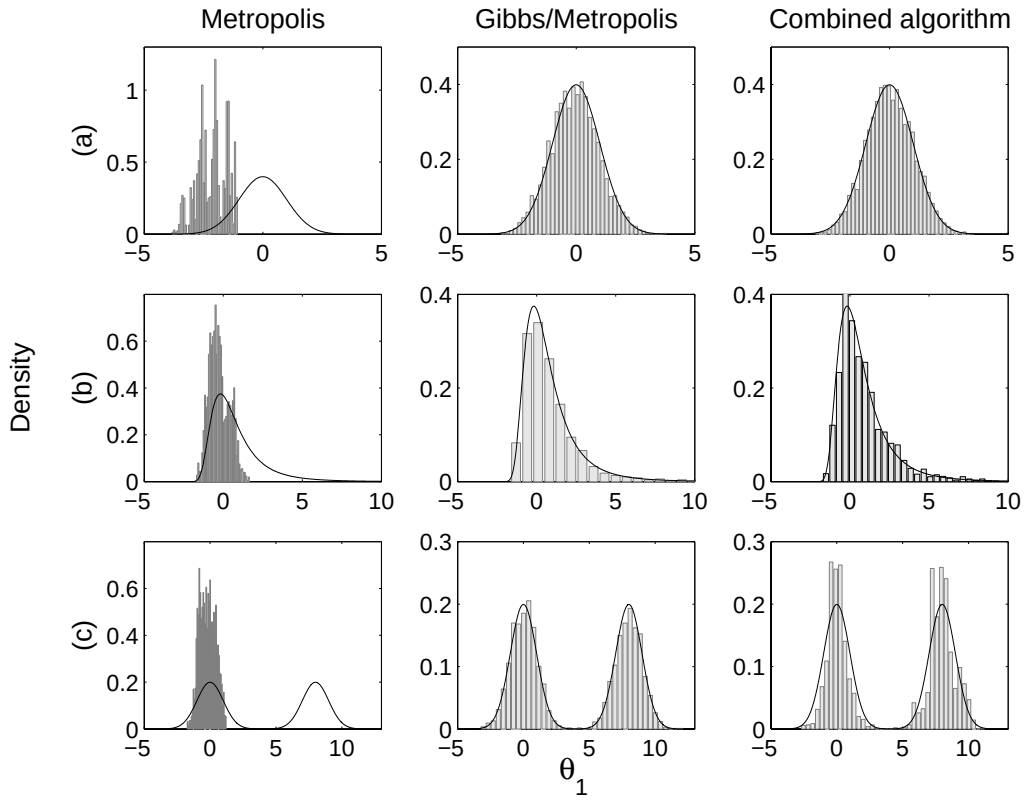

Metropolis
Gibbs/Metropolis
Combined algorithm
(a)
(b)
(c)
Density
θ1

Figure 1

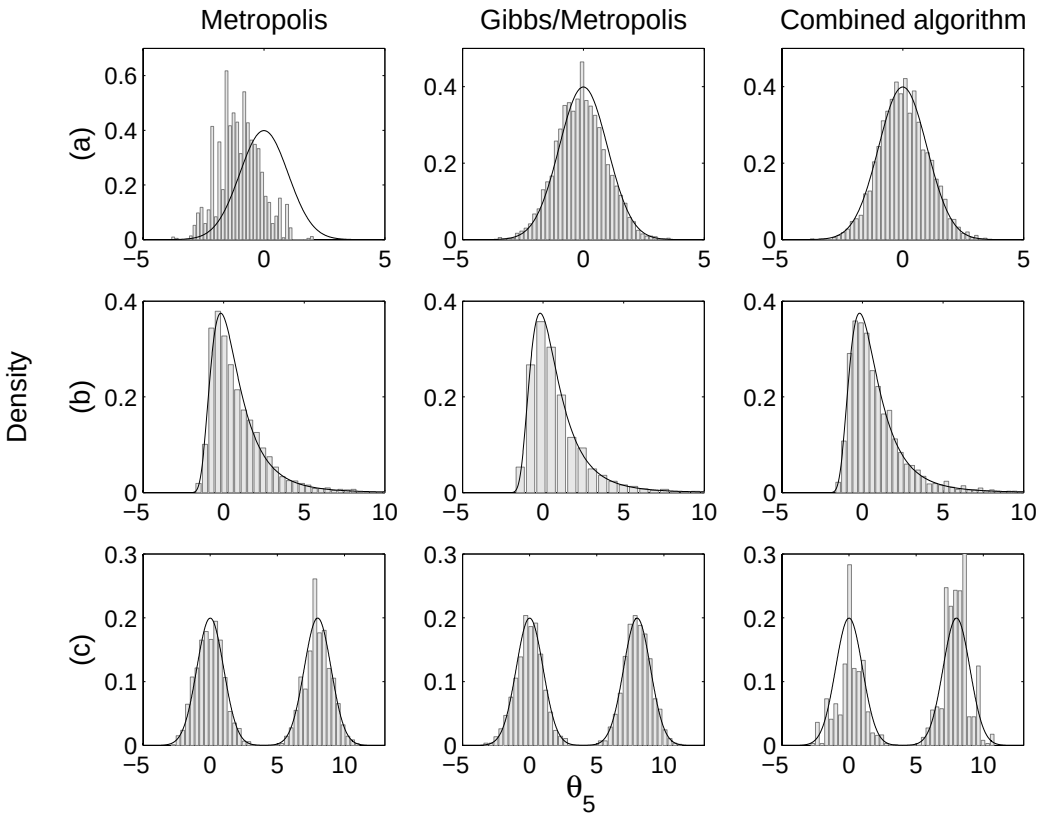

Metropolis
Gibbs/Metropolis
Combined algorithm
(a)
(b)
(c)
Density
θ5

Figure 2

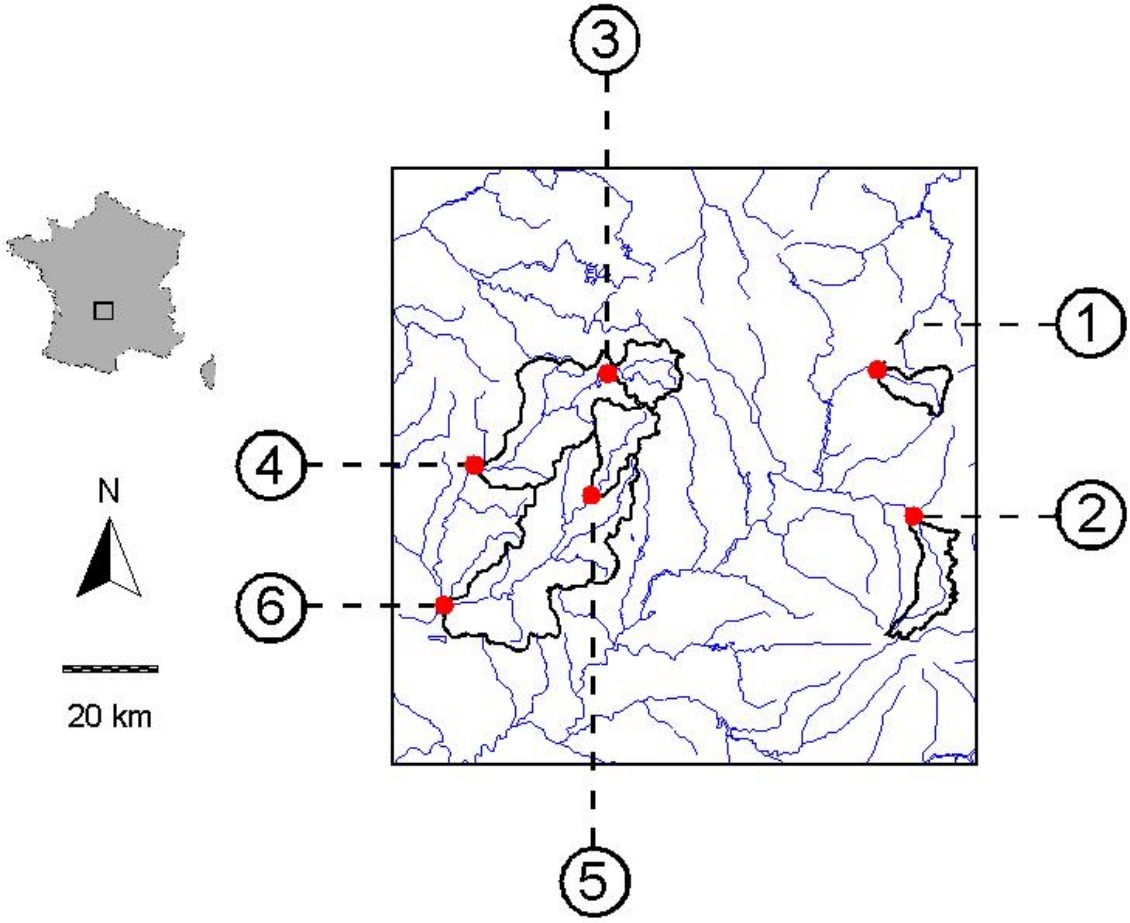


Figure 3

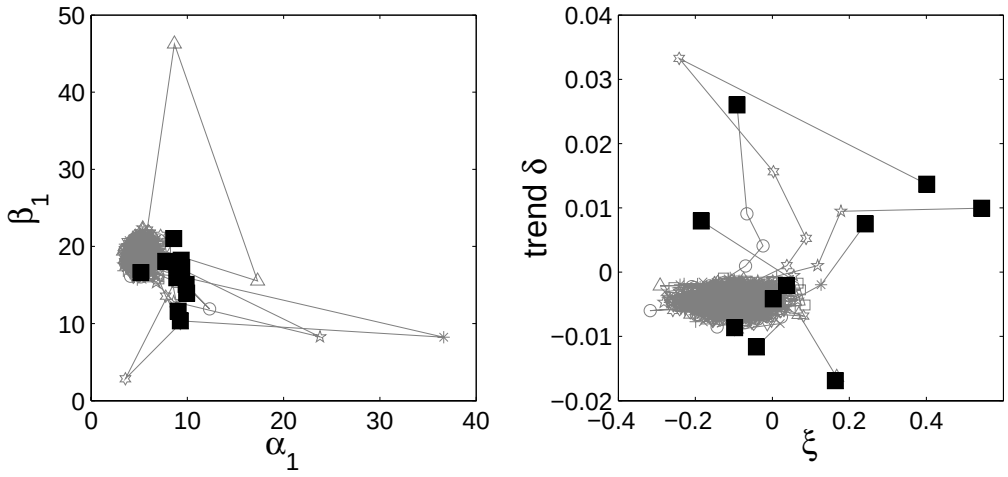
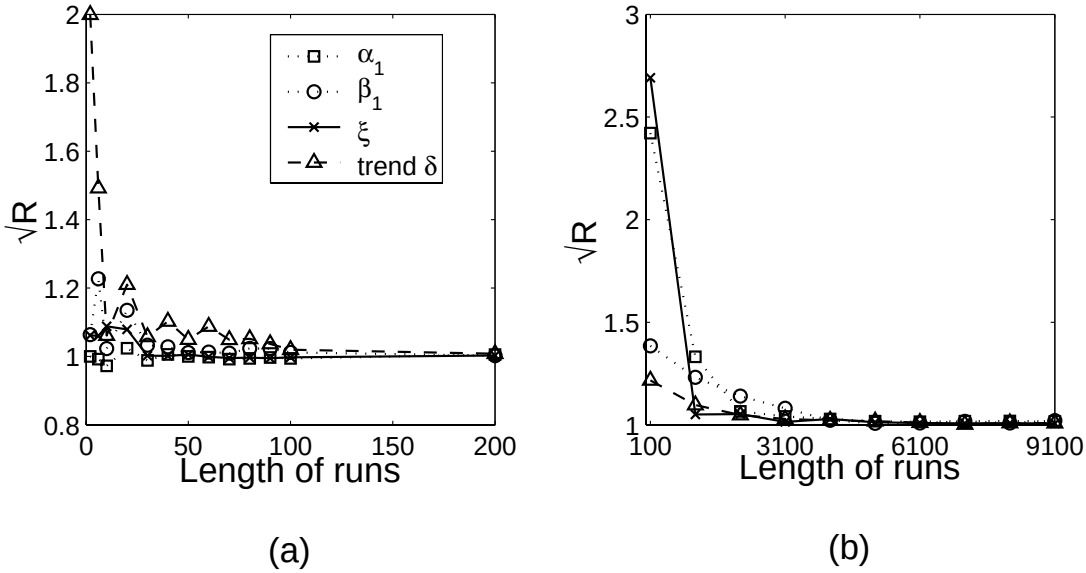

50
40
30
20
10
0
β1
0
10
20
30
40
α1
0.04
0.03
0.02
0.01
0
−0.01
−0.02
trend δ
−0.4
−0.2
0
0.2
0.4
ξ

Figure 4


2
1.8
1.6
1.4
1.2
1
0.8
√R
0
50
100
150
200
Length of runs
α1
β1
ξ
trend δ
3
2.5
2
1.5
1
√R
100
3100
6100
9100
Length of runs

(a) (b)

Figure 5

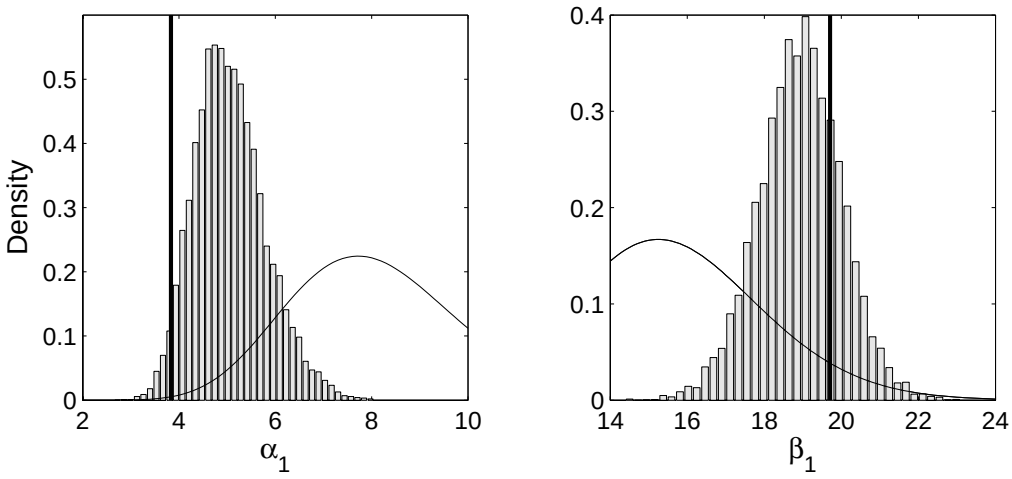


Figure 6

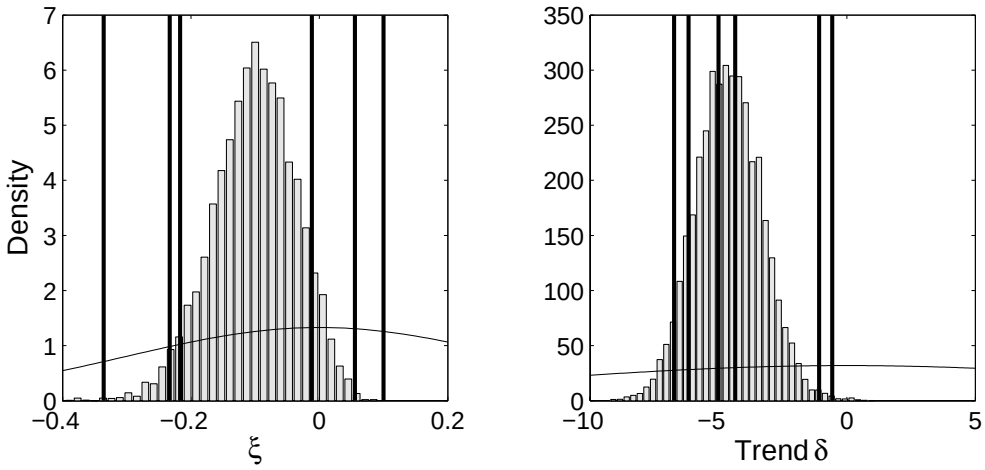


Figure 7

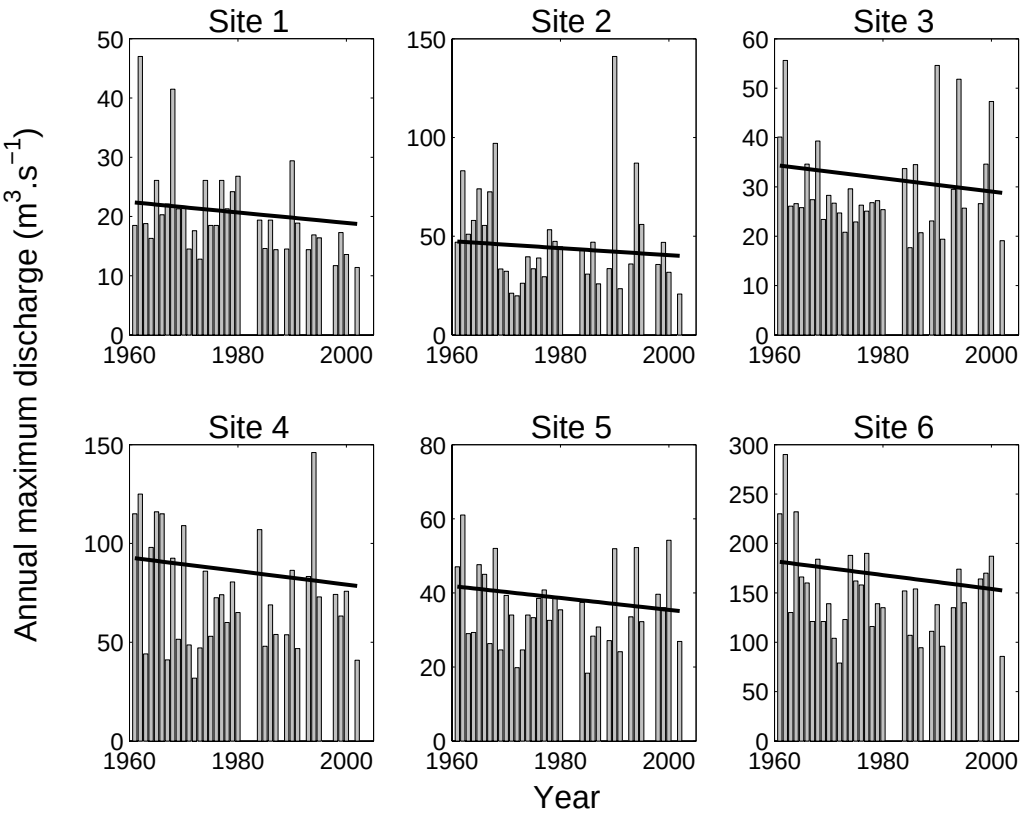

Site 1
Site 2
Site 3
Site 4
Site 5
Site 6
Annual maximum discharge (m3.s-1)
Year
1960
1980
2000

Figure 8

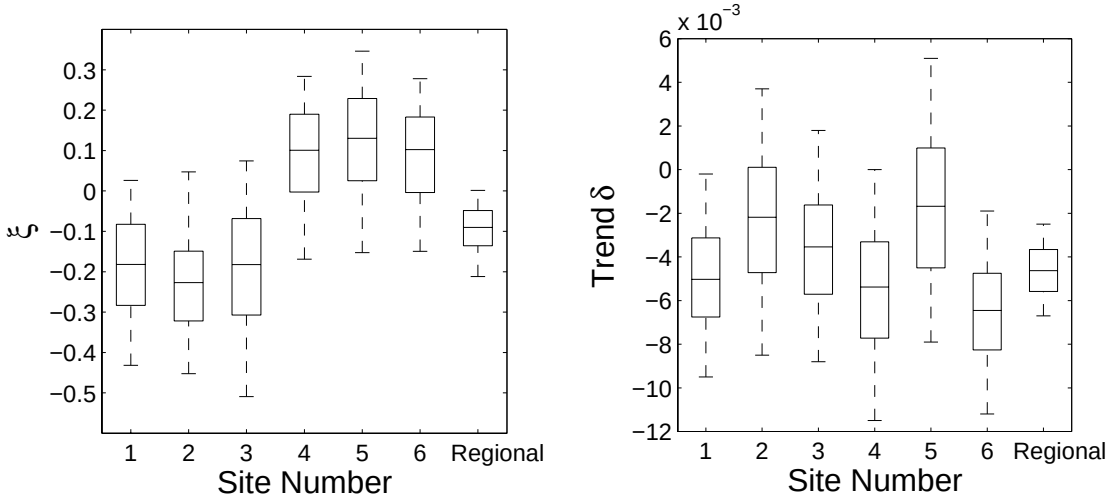

ξ
0.3
0.2
0.1
0
−0.1
−0.2
−0.3
−0.4
−0.5
1
2
3
4
5
6
Regional
Site Number
x 10^−3
Trend δ
6
4
2
0
−2
−4
−6
−8
−10
−12
1
2
3
4
5
6
Regional
Site Number

Figure 9

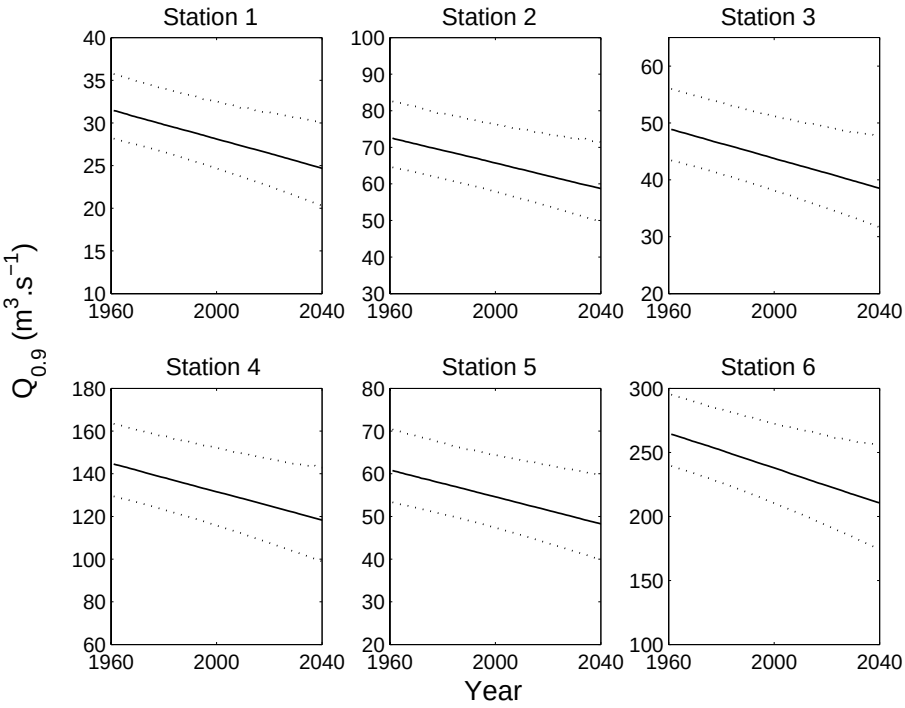

Station 1
Station 2
Station 3
Station 4
Station 5
Station 6
$Q_{0.9}$ ($m^3.s^{-1}$)
Year
1960
2000
2040

Figure 10

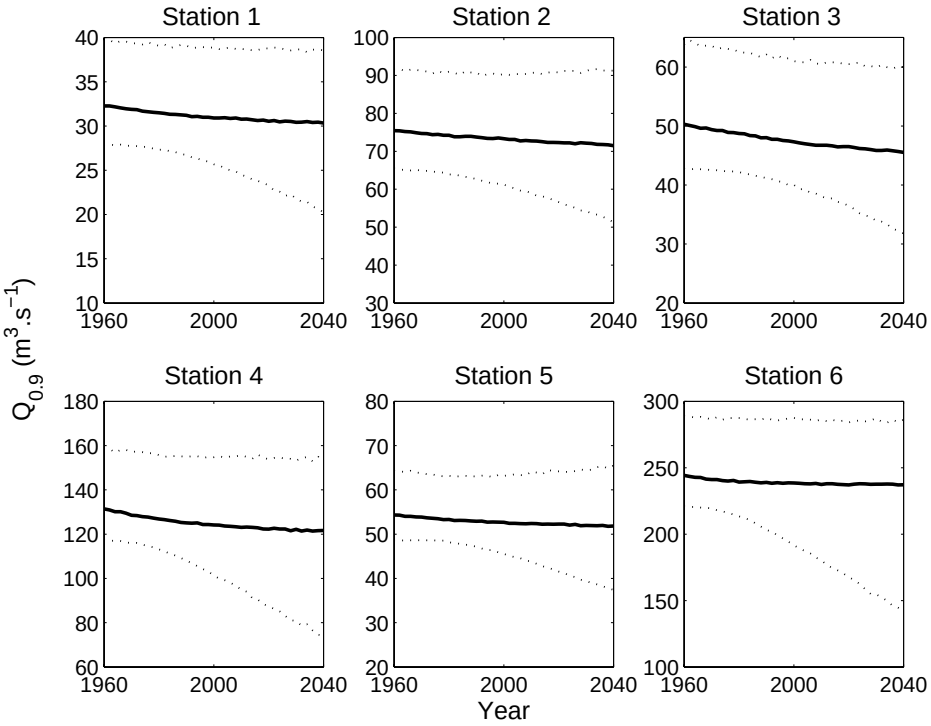

Station 1
Station 2
Station 3
Station 4
Station 5
Station 6
$Q_{0.9}$ ($m^3.s^{-1}$)
Year
1960
2000
2040

Figure 11