



**HAL**  
open science

# A Bayesian hierarchical approach to regional frequency analysis

Benjamin Renard

► **To cite this version:**

Benjamin Renard. A Bayesian hierarchical approach to regional frequency analysis. *Water Resources Research*, 2011, 47, p. W11513 - p. 10.1029/2010WR010089 . hal-00662931

**HAL Id: hal-00662931**

**<https://hal.science/hal-00662931v1>**

Submitted on 25 Jan 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# <sup>1</sup> A Bayesian Hierarchical Approach To Regional <sup>2</sup> Frequency Analysis

B. Renard<sup>1</sup>

---

B. Renard, Cemagref, UR HHLY, Hydrology-Hydraulics, 3 bis quai Chauveau - CP 220, F-69336 Lyon, France. ([benjamin.renard@cemagref.fr](mailto:benjamin.renard@cemagref.fr))

<sup>1</sup>UR HHLY, Hydrology-Hydraulics, 3 bis  
quai Chauveau, CP 220, F-69336 Lyon,  
France.

3 **Abstract.** Regional Frequency Analysis (RFA) has a long history in Hy-  
4 drology, and numerous distinct approaches have been proposed over the year  
5 to perform the estimation of some hydrologic quantity at a regional level.  
6 However, most of these approaches still rely on strong hypotheses that may  
7 limit their application and complicate the quantification of predictive un-  
8 certainty. The objective of this paper is to propose a general Bayesian hi-  
9 erarchical framework to implement RFA schemes that avoid these difficul-  
10 ties. The proposed framework is based on a two-level hierarchical model. The  
11 first level of the hierarchy describes the joint distribution of observations. An  
12 arbitrary marginal distribution, whose parameters may vary in space, is as-  
13 sumed for at-site series. The joint distribution is then derived by means of  
14 an elliptical copula, therefore providing an explicit description of the spa-  
15 tial dependence between data. The second level of the hierarchy describes  
16 the spatial variability of parameters using a regression model that links the  
17 parameter values with covariates describing site characteristics. Regression  
18 errors are modeled with a Gaussian spatial field which may exhibit spatial  
19 dependence. This framework enables performing prediction at both gauged  
20 and ungauged sites and, importantly, rigorously quantifying the associated  
21 predictive uncertainty. A case study based on annual maxima of daily rain-  
22 fall demonstrates the applicability of this hierarchical approach. Although  
23 numerous avenues for improvement can already be identified (amongst which  
24 the inclusion of temporal covariates to model time variability), the proposed

25 model constitutes a general framework for implementing flexible RFA schemes  
26 and quantifying the associated predictive uncertainty.

## 1. Introduction

### 1.1. Standard implementations of regional frequency analysis

27 The purpose of Regional Frequency Analysis (RFA) is to estimate the distribution of  
28 some hydrologic variable (e.g. annual maximum rainfall or runoff) using data from several  
29 sites. Compared with standard at-site frequency analysis, RFA attempts to improve the  
30 precision of estimates by sharing the information stemming from similar sites. Moreover,  
31 RFA enables estimation at ungauged or poorly gauged sites by transferring the information  
32 arising from neighboring gauging stations.

33 Numerous approaches have been proposed to implement RFA schemes. Amongst them,  
34 the index flood method proposed by *Dalrymple* [1960] is still widely used in engineering  
35 practice. It is based on a scale invariance hypothesis: within an homogeneous region,  
36 distributions from all sites are assumed identical, up to a scale factor, termed the index  
37 flood. The implementation of the index flood method can be summarized in three steps  
38 [e.g., *Hosking and Wallis*, 1997]: (i) delineation of an homogeneous region; (ii) estimation  
39 of the common regional distribution, based on scaled at-site data (i.e. divided by the index  
40 flood); (iii) transfer of information to ungauged or poorly gauged site using a regression  
41 model linking the index flood values with site characteristics.

42 The index flood method is widely used due to its ease of implementation and its robustness.

43 However, its basic implementation is affected by several limitations:

- 44 • The delineation of homogeneous regions, where the scale invariance assumption holds,  
45 is far from obvious.
- 46 • The scale invariance assumption might simply be too restrictive in some cases.

47 • In most cases, the index flood is defined as the mean or the median of at-site data,  
48 but the physical reasons behind this choice are unclear.

49 • The regional distribution is estimated by pooling scaled at-site data, and treating  
50 them as if they were independent, which is rarely the case.

51 • Standard regression methods like ordinary least squares might be statistically in-  
52 efficient because index flood values are statistics (as opposed to observations), and are  
53 therefore affected by estimation errors that may be dependent in space and whose prop-  
54 erties may vary from site to site.

55 • The previous points make the quantification of the total predictive uncertainty chal-  
56 lenging.

57 A wealth of research has been carried out over the years, either to improve the imple-  
58 mentation of the index flood method, or to generalize it by abandoning some of its most  
59 restrictive assumptions (in particular, scale invariance). A non-exhaustive list of exam-  
60 ples includes the work by *Ouarda et al.* [2001] on the concept of homogeneous region,  
61 the studies by *Stedinger* [1983] and *Hosking and Wallis* [1988] on estimating the regional  
62 distribution with spatially dependent data, or the extension to peak-over-threshold series  
63 of *Madsen and Rosbjerg* [1997] and *Ribatet et al.* [2007].

64 The transfer of information from gauged to ungauged site and the quantification of the  
65 associated predictive uncertainty has been a topic of particular attention [e.g., *Stedinger*  
66 *and Tasker*, 1985, 1986; *Reis et al.*, 2005; *Kjeldsen and Jones*, 2009a; *Micevski and Kucz-*  
67 *era*, 2009]. Indeed, this transfer relies on a regression model linking the index flood values  
68 (estimated at gauged sites) and site characteristics. *Stedinger and Tasker* [1985, 1986]  
69 introduced the Generalized Least Squares (GLS) approach to account for both the het-

70 eroscedastic and spatially dependent nature of sampling errors (i.e., errors in estimating  
71 the index flood at gauged sites) and the existence of regression errors. *Robson and Reed*  
72 [1999] and *Kjeldsen and Jones* [2009a] further generalized the GLS approach by consid-  
73 ering spatially dependent regression errors.

74 Despite these advances, GLS-based transfer of information from gauged to ungauged site  
75 still relies on the following two-step procedure:

76 1. Local estimation: an index flood is first chosen (e.g. the at-site mean or median),  
77 and is estimated at each gauged site. This estimation is affected by sampling errors,  
78 which are spatially dependent and whose variance varies from site to site. Consequently,  
79 the covariance matrix  $\hat{\Sigma}$  of sampling errors is also estimated;

80 2. Regional estimation: A regression model is estimated to link the index flood value  
81 with site/catchment characteristics. Importantly, the estimation of the regression model  
82 accounts for the existence of sampling errors in a GLS framework, and is hence performed  
83 conditionally on  $\hat{\Sigma}$ . Also note that estimating a regression model is not restricted to esti-  
84 mating the regression coefficients, but also involves estimating the properties of regression  
85 errors, i.e. their covariance matrix  $\hat{\Gamma}$ . This is of primary importance since this matrix  
86 plays an important role in the predictive uncertainty at ungauged sites.

87 Such a two-step procedure might be problematic in the context of quantifying the total  
88 predictive uncertainty. Indeed, estimates at step 2 are conditional on estimates at step 1.  
89 In particular, the covariance matrix  $\hat{\Sigma}$  is an estimate, and may itself be in error (see e.g.  
90 the discussion in *Stedinger and Tasker* [1985] and *Kroll and Stedinger* [1998] on desirable  
91 properties of  $\hat{\Sigma}$ ). Such error may then propagates to step 2. Consequently, it would  
92 be desirable to avoid separating the inference process in two separate steps: this can be

93 achieved using hierarchical models (see following section 1.2).  
94 In addition to these issues related to the estimation procedure, the assumptions underlying  
95 index flood approaches remain questionable. Indeed, the scale invariance assumption is  
96 convenient because it merges all spatial variability into a single parameter (the index-flood  
97 parameter). However, this assumption forces the coefficient of variation of data to remain  
98 constant throughout an homogeneous region: this might be too restrictive in some regions,  
99 or it might force to drastically reduce the spatial extent of the region to ensure the scale  
100 invariance hypothesis is met. Several alternatives have therefore been proposed to move  
101 beyond this restrictive framework, in particular region of influence approaches [*Burn*, 1990]  
102 and recent developments [*Kjeldsen and Jones*, 2009b], normalized quantile regression [*Fill*  
103 *and Stedinger*, 1998] or empirical Bayes procedures [*Kuczera*, 1982a, b, 1983] (see also the  
104 discussion provided by *Griffis and Stedinger* [2007]).

## 1.2. Bayesian hierarchical models

105 An alternative approach, based on Bayesian hierarchical models, has been explored  
106 more recently. In particular, *Wikle et al.* [1998] proposed a general hierarchical frame-  
107 work to describe the spatial variability of the distribution of some environmental variable.  
108 The principle of a hierarchical model is to use several modeling layers. For instance, a  
109 first layer may assume that the data follow some distribution with unknown parameters,  
110 while a second layer may model the variability of those parameters in space, using some  
111 regression model. This closely corresponds to the successive steps involved in the stan-  
112 dard implementation of RFA approaches. However, the main advantage of a hierarchical  
113 model is that all unknown quantities can be inferred simultaneously, therefore accounting  
114 for possible interactions between estimation errors made at different layers and yielding



115 a more rigorous quantification of the predictive uncertainty. In other words, hierarchical  
116 models allow describing the local variability of data *together* with their regional coherence,  
117 without separating the inference process in several steps. Moreover, such models are more  
118 general than the model underlying the index flood approach, since they do not require  
119 assuming scale invariance.

120 Several applications of Bayesian hierarchical models in a hydrological context have been  
121 proposed in the literature. For instance, *Cooley et al.* [2007] described the spatial variabil-  
122 ity of extreme rainfall, while *Aryal et al.* [2009] extended this description to both spatial  
123 and temporal variability. Similarly, *Lima and Lall* used Bayesian hierarchical models to  
124 describe daily rainfall occurrences [*Lima and Lall*, 2009] or runoff extremes [*Lima and*  
125 *Lall*, 2010] in a regional context. In addition to these hydrological applications, similar  
126 Bayesian hierarchical models have been used in other fields, e.g. for extreme wind speed  
127 modeling [*Coles and Casson*, 1998; *Casson and Coles*, 1998, 2000].

128 Despite improving the standard implementation of RFA approaches, all Bayesian hierar-  
129 chical models described above rely on an assumption of conditional independence: data  
130 are assumed spatially independent given the values of their distribution's parameters.  
131 This would be a valid assumption if most spatial covariation in the data was explained  
132 by the spatial covariation in the parameters. However, it is questionable since spatial  
133 dependence between data in the one hand and parameters in the other hand arise from  
134 distinct processes, as noted by *Cooley et al.* [2007]: in a nutshell, data dependence can be  
135 interpreted as *weather* spatial dependence, while dependence between parameters (also  
136 termed process dependence) relates to *climate* spatial dependence. Weather should ex-  
137 hibit spatial dependence (at least within a short distance range), even if the climate were

138 perfectly known.

139 Examples of spatial hierarchical models explicitly accounting for intersite dependence in  
140 the observations are very few. *Perreault* [2000] proposed such a model to detect a re-  
141 gional step-change in annual runoff. Alternatively, *Micevski et al.* [2006] and *Micevski*  
142 [2007] proposed a Bayesian hierarchical regional flood model accounting for data depen-  
143 dence. However, in both cases, the explicit description of data dependence was rooted to  
144 a particular distributional assumption: Gaussian data were assumed by *Perreault* [2000],  
145 while *Micevski et al.* [2006] and *Micevski* [2007] used a mixture of log-normal distributions.  
146 Unfortunately, Gaussian-related assumptions may be too restrictive for other hydrologic  
147 variables or in other geographical contexts.

### 1.3. Objectives

148 Building on previous work described in the preceding sections, this paper aims to derive  
149 a general Bayesian hierarchical framework for RFA. In particular, this framework should  
150 enable an explicit description of spatial dependence between data, without relying on  
151 Gaussian-related distributional assumptions. This is achieved by means of the elliptical  
152 copula family [*Genest and Favre, 2007*], which constitutes a convenient tool to model  
153 dependence in a highly dimensional and non-Gaussian context.

154 The paper is organized as follows. Section 2 describes the two-level hierarchical framework,  
155 with level 1 modeling the joint distribution of observations and level 2 modeling the  
156 variation of the distribution' parameters in space. Section 3 describes the inference of the  
157 hierarchical model and its use for prediction at both gauged and ungauged sites. Section  
158 4 illustrates the application of the proposed framework for the regional estimation of

159 extreme rainfall. Avenues for further improvement are identified and discussed in section  
160 5, before summarizing the main results in section 6.

## 2. A Bayesian hierarchical modeling framework

### 2.1. Data level

161 Let  $Y(s, t)$  denote the variable of interest at site  $s$  and time  $t$ . For instance,  $Y(s, t)$  may  
162 represent the annual maximum daily rainfall at site  $s$  and year  $t$ . Let us further assume  
163 that observations are available at sites  $\tilde{s}_1, \dots, \tilde{s}_M$  and times  $t=1:T$ . The  $(T \times M)$  observation  
164 matrix is denoted by  $\tilde{\mathbf{y}} = (\tilde{\mathbf{y}}(\tilde{s}_1), \dots, \tilde{\mathbf{y}}(\tilde{s}_M))$ , with  $\tilde{\mathbf{y}}(\tilde{s}_i) = (\tilde{y}(\tilde{s}_i, t))_{t=1:T}$  denoting the time  
165 series of observations at a given site  $\tilde{s}_k$ . The shorthand notation  $\tilde{\mathbf{y}}(t) = (\tilde{y}(\tilde{s}_i, t))_{i=1:M}$  is  
166 also used to denote the (spatial) vector of observations at a given time  $t$ . Note that this  
167 notation assumes that all sites share the same observation period. Although considering  
168 non-concomitant observation periods would not affect the modeling hypotheses, it would  
169 certainly create tedious notation and complicate the model implementation (see discussion  
170 in section 5.1).

#### 2.1.1. At-site distribution

172 At an arbitrary site  $s$  within the area of study, and at any time  $t$ ,  $Y(s, t)$  is assumed to be  
173 a realization from a given distribution whose parameters vary in space (see left-hand-side  
174 of Figure 1):

$$Y(s, t) \sim p(\boldsymbol{\theta}(s)) \quad (1)$$

175 RFA involves estimating the parameter vector  $\boldsymbol{\theta}(s)$  at any (gauged or ungauged) site  $s$ .  
176 In order to avoid confusion with other parameters that will be introduced later on, the  
177 term "D-parameters" is systematically used to denote the parameters  $\boldsymbol{\theta}(s) = (\theta_k(s))_{k=1:D}$

178 of the parent distribution. Note that D-parameters are allowed to vary in space but not in  
179 time. An extension of the framework for allowing time-varying D-parameters is possible  
180 but is left for future work, since this paper mainly focuses on spatial aspects.

### 181 **2.1.2. Joint distribution**

182 In a spatial context, the derivation of the likelihood of observations  $\tilde{\mathbf{y}}$  requires knowing  
183 the multivariate distribution of the spatial observation vector  $\tilde{\mathbf{y}}(t)$  at any time  $t$ . An  
184 assumption often made (explicitly or implicitly) by some regional estimation methods is  
185 that the data are spatially independent. In this case, the multivariate pdf is simply equal  
186 to the product of marginal pdfs in (1). Implications of this assumption are discussed  
187 by e.g. *Stedinger* [1983], *Hosking and Wallis* [1988], *Madsen and Rosbjerg* [1997] and  
188 *Renard and Lang* [2007]. In particular, these authors demonstrate that the variance (i.e.  
189 the uncertainty) of estimated quantities is underestimated when data do not support the  
190 independence assumption.

191 One of the main objectives of the framework presented in this paper is to overcome this  
192 limitation, by explicitly modeling spatial dependence. The approach taken to achieve  
193 this objective has to account for two important points: (i) it has to be applicable for an  
194 arbitrary choice of marginal distribution in (1); (ii) since regional analysis may involve  
195 hundreds of sites, it has to remain practical with high-dimensional data.

196 Point (i) above makes copulas a natural candidate to model dependence in the context  
197 of this framework. Indeed, the copula theory is based on the description of the depen-  
198 dence structure independently of marginal distributions [e.g., *Favre et al.*, 2004]. Point  
199 (ii) suggests focusing on the elliptical copula family, because of its ability to describe high-  
200 dimensional datasets. A thorough description of the elliptical copula family can be found

201 in the paper by *Genest and Favre* [2007], and a summary of the main characteristics of  
202 two particular members (the Gaussian and the Student copulas) is given in Appendix A.  
203 A  $M$ -dimensional elliptical copula is parameterized by a  $(M \times M)$  symmetric dependence  
204 matrix  $\Sigma$  describing pairwise dependences. Additional parameters  $\boldsymbol{\eta}$  may be used for  
205 some members of the family (e.g. the Student copula) to describe the strength of tail  
206 dependence. The multivariate distribution is finally derived by combining this depen-  
207 dence structure with the marginal distributions given by equation (1) (see Appendix A  
208 for details). An example of application of an elliptical copula (the Gaussian copula) to  
209 regional frequency analysis is given by *Renard and Lang* [2007].

210 Formally, it is assumed that the multivariate distribution of data from any set of  
211  $M$  sites can be derived from an  $M$ -dimensional elliptical copula, with pairwise de-  
212 pendence matrix  $\Sigma$ , additional dependence parameters  $\boldsymbol{\eta}$  and marginal distributions  
213  $\{p(\boldsymbol{\theta}(s_1)), \dots, p(\boldsymbol{\theta}(s_M))\}$ :

$$(Y(s_1, t), \dots, Y(s_M, t)) \sim EC_M(\Sigma, \boldsymbol{\eta}, \{\boldsymbol{\theta}(s_1), \dots, \boldsymbol{\theta}(s_M)\}) \quad (2)$$

214 The symbol  $EC_M$  stands for "M-dimensional Elliptical Copula", and analytical formulas  
215 for the corresponding joint pdf are given in Appendix A .

216 In order to simplify the model, it is further assumed that the dependence between data  
217 from two sites solely depends on the inter-site distance. An analogy can be drawn with  
218 the common treatment of stationary and isotropic spatial random fields in geostatistics  
219 [e.g., *Chiles and Delfiner*, 1999]. It follows that the elements of the pairwise dependence  
220 matrix  $\Sigma$  can be expressed as a function of the inter-site distance, parameterized by some  
221 vector  $\boldsymbol{\psi}$  (see Figure 1):

$$\Sigma(i, j) = \Psi(\|s_i - s_j\|; \boldsymbol{\psi}) \quad (3)$$

222 Note that the inter-site distance is not necessarily Euclidean in the bi-dimensional xy-  
223 space (geographical coordinates). For instance, a distance in the tri-dimensional space  
224 xy+elevation may be useful if inter-site dependence is lower for sites having contrasted  
225 elevations. More generally, the distance may depend on covariates other than geographi-  
226 cal coordinates [e.g., *Cooley et al., 2007; Blanchet and Davison, 2011*]. However, for the  
227 sake of simplicity, this is not made explicit in the notation of equation (3). This topic is  
228 further discussed in section 5.8.

229 A valid dependence-distance function  $\Psi(\cdot; \boldsymbol{\psi})$  must ensure the positive-definiteness of the  
230 dependence matrix  $\boldsymbol{\Sigma}$ . Such a function can be chosen amongst the numerous covariogram  
231 models (e.g., exponential, spherical, Gaussian) existing in geostatistics. By analogy with  
232 the covariogram used in geostatistics, the dependence-distance model in (3) is termed  
233 "dependogram". Note that a distinct naming convention is used because the elements  
234 of the matrix  $\boldsymbol{\Sigma}$  are not equal to the covariances between data pairs, but rather to the  
235 dependence coefficients of the elliptical copula.

236 It is stressed that the use of an elliptical copula to model spatial dependence is mo-  
237 tivated by practical considerations. Although the elliptical family is quite flexible (in  
238 particular, it encompasses tail-dependent and tail-independent models), there can be no  
239 guarantee that an elliptical copula will be able to model the data at hand. In particu-  
240 lar, a multivariate distribution derived from an elliptical copula is not compatible with  
241 the family of multivariate extreme value distributions [*Mikosch, 2005*]. It follows that  
242 the use of the elliptical copula model in extrapolation to estimate low probabilities (e.g.,

243  $Pr(Y(s_1, t) > u \cap \dots \cap Y(s_M, t) > u)$  for some large  $u$ ) is likely to yield inaccurate re-  
244 sults. However, the aim of the framework presented in this paper is not to estimate the  
245 probability of extreme multivariate events, but rather to estimate the distribution of the  
246 variable of interest at any site, while accounting for the existence of dependences at ob-  
247 served levels. The implications of the elliptical copula assumption are more thoroughly  
248 discussed in section 5.2.

## 2.2. Process level

249 Section 2.1 described the construction of the multivariate distribution of data by spec-  
250 ifying the following three components: (i) the at-site distribution (1); (ii) the elliptical  
251 copula (2) used to model spatial dependence; (iii) the dependogram (3). The second level  
252 of the hierarchical framework aims to describe the variation of the D-parameters in space,  
253 by means of a Gaussian spatial process whose mean depends on covariates describing the  
254 site (or catchment) characteristics (see right-hand-side of Figure 1).

### 2.2.1. D-parameter regression model

256 Let  $\boldsymbol{\theta}(s) = (\theta_k(s))_{k=1:D}$  be the D-parameter vector. A regression model is used to  
257 describe the spatial variation of each D-parameter  $\theta_k(s)$  as follows:

$$g_k(\theta_k(s)) = h_k(\mathbf{x}_k(s); \boldsymbol{\beta}_k) + \epsilon_k(s) \quad (4)$$

258 Equation (4) uses the following components:

- 259 1. One-to-one function  $g_k$  is termed the "link function" by analogy with generalized  
260 linear models [e.g., *Dobson, 2001*]. The identity function may be used in most cases.  
261 However, alternative functions might be useful for some D-parameters, e.g. a logarithm

262 function to ensure positivity or a logit function for a D-parameter comprised between zero  
263 and one.

264 2. Vector  $\mathbf{x}_k(s)$  represents the set of covariates used to describe the characteristics  
265 of site (or catchment)  $s$  (e.g., elevation, distance to sea, catchment size, etc.). As in  
266 any regression framework, the choice of relevant covariates is of primary importance, and  
267 preliminary analyzes are often useful e.g. to eliminate highly correlated covariates.

268 3. Function  $h_k(\cdot; \boldsymbol{\beta}_k)$  is the regression function, parameterized by a vector of regression  
269 parameters  $\boldsymbol{\beta}_k$ . The most common choice of regression function is the linear function  
270  $h_k(\mathbf{x}_k(s); \boldsymbol{\beta}_k) = \mathbf{x}_k(s)\boldsymbol{\beta}_k$ , but alternative functions may be used.

271 4.  $\epsilon_k(s)$  is the residual of the regression. It stems from the imperfect nature of the  
272 regression model, and is therefore termed "regression error". In the hierarchical model-  
273 ing context described in this paper, regression errors are treated as latent variables, i.e.  
274 unobserved variables that need to be inferred.

### 275 **2.2.2. Regression errors**

276 Regression errors  $\boldsymbol{\epsilon}_k = (\epsilon_k(s))_{s=1:M}$  are commonly assumed to be spatially independent  
277 [e.g., *Stedinger and Tasker*, 1985, 1986; *Reis et al.*, 2005]. However, recent work by  
278 *Kjeldsen and Jones* [2009a] suggests that regression errors can be significantly dependent,  
279 especially for nearby sites and/or regression models with poor predictive ability. Ignoring  
280 spatial dependences between regression errors possibly affects regional frequency analysis  
281 in two distinct ways: (i) it may impact the accuracy and/or precision of estimates for the  
282 regression parameters; (ii) it may impact the predictive variance (i.e. the uncertainty in  
283 quantities estimated at ungauged sites) by not taking advantage of the regression errors  
284 estimated as nearby sites. Consequently, following *Kjeldsen and Jones* [2009a], regression



285 errors are assumed to be a realization from a spatial Gaussian field with mean zero and  
286 covariance matrix  $\mathbf{\Gamma}_k$ :

$$(\epsilon_k(s_1), \dots, \epsilon_k(s_M)) \sim N_M(0; \mathbf{\Gamma}_k) \quad (5)$$

287 A covariogram model is used to relate pairwise covariances to inter-site distances:

$$\Gamma_k(i, j) = \Upsilon_k(\|s_i - s_j\|; \mathbf{v}_k) \quad (6)$$

288 In a Bayesian hierarchical context, the Gaussian distribution used to describe regression  
289 errors is termed the hyper-distribution. Parameters  $\mathbf{v}_k$  defining the covariance matrix are  
290 termed the hyper-parameters (see Figure 1). Note that as in previous section 2.1.2, the  
291 inter-site distance is not necessarily Euclidean and can be interpreted in a wider sense.  
292 Moreover, note that regression errors  $\epsilon_k(s)$  and  $\epsilon_q(s)$  related to two distinct D-parameters  
293  $\theta_k(s)$  and  $\theta_q(s)$  will be described by independently using equations (5-6) twice. This  
294 implies that an assumption of independence between regression errors  $\epsilon_k(s)$  and  $\epsilon_q(s)$  is  
295 effectively made. This assumption will be further discussed subsequently (see sections 4  
296 and 5.3).

### 2.3. Remarks

297 In the hierarchical framework presented in previous sections, two distinct dependence  
298 structures are used: the dependogram (3) aims at describing dependence between obser-  
299 vations, while the covariogram (6) aims at describing dependence between the parameters  
300 of their distribution (or more accurately, between the errors of the regression linking the  
301 D-parameters with catchment/site characteristics). As explained in the introduction, the  
302 former structure relates to weather spatial dependence, while the latter relates to climate

303 spatial dependence [see also the discussion in *Cooley et al.*, 2007]. These dependence  
304 structures play distinct roles in the context of regional frequency analysis: climate spa-  
305 tial dependence allows transferring information from gauged to ungauged sites, while the  
306 main effect of weather dependence is to diminish the information content of data collected  
307 at nearby sites. Consequently, both dependence structures are likely to have an impact  
308 on predictions and should therefore be accounted for. Note that the assumption of con-  
309 ditional independence frequently made in similar Bayesian hierarchical frameworks (see  
310 introduction) corresponds to forcing the dependence matrix  $\Sigma$  in equation (3) to unity.  
311 Moreover, the regression model of equation (4) yields several interesting particular cases  
312 when the regression errors are forced to zero:

313 1. A purely local D-parameter (i.e. whose value is site-specific) can be obtained with  
314 the regression model  $\theta_k(\tilde{s}) = \beta_k^{(\tilde{s})}$ . This corresponds to introducing as many regression  
315 parameters  $\beta_k^{(\tilde{s})}$  as there are sites, and use them to model a site effect. An obvious  
316 drawback of this approach is that the estimation at ungauged site  $s$  is not directly possible  
317 since the site effect  $\beta_k^{(s)}$  is unknown.

318 2. A regional D-parameter (i.e. having an identical value for all sites within the region)  
319 corresponds to the regression model  $\theta_k(\tilde{s}) = \beta_k$ . This is a rather strong assumption,  
320 which may yield an underestimation of the predictive uncertainty. Indeed, in the absence  
321 of regression errors, the uncertainty in the estimation of  $\theta_k$  is identical for all sites (gauged  
322 or not).

323 3. The strong assumption that parameter  $\theta_k(\tilde{s})$  is identical for all sites can be relaxed  
324 by using a regression model  $g_k(\theta_k(\tilde{s})) = h_k(\mathbf{x}_k(\tilde{s}); \boldsymbol{\beta}_k)$ . This corresponds to the "covariate  
325 modeling" approach proposed by several authors [e.g., *Katz et al.*, 2002; *Maraun et al.*,

2009]. However, the possible underestimation of predictive uncertainty in the absence of  
regression errors still holds.

Lastly, the Bayesian hierarchical framework shares several similarities with the GLS  
approach proposed by *Stedinger and Tasker* [1985, 1986] and its latest developments by  
*Kjeldsen and Jones* [2009a]. Indeed, GLS also uses two spatial dependence structures: (i)  
the covariance matrix of sampling estimation errors (e.g., errors in estimating an index  
flood at gauged sites), which results from the spatially dependent nature of data; (ii) the  
covariance matrix of regression errors (termed model errors in the GLS approach). The  
framework presented in this paper therefore borrows from GLS the objective of modeling  
the spatial variability in the distinct sources of errors affecting RFA. However, it differs  
in its implementation. GLS first estimates the spatial variability of sampling errors, and  
then uses this estimation to fit the regression in a second step. By contrast, the proposed  
framework performs the inference in a single step, which facilitates the quantification of  
the total predictive uncertainty (see discussion in the introduction section). Moreover, the  
regression is applied on each parameter, while in general GLS is rather applied to a single  
hydrologic quantity (e.g. an index flood or a quantile, but see *Tasker and Stedinger* [1989]  
and *Griffis and Stedinger* [2007] for exceptions). This is a more general approach, at least  
in principle; however what level of model complexity can be identified given the limited  
information content of data remains an open question (this will be further discussed in  
section 5.5).

### 3. Estimation and prediction

### 3.1. Posterior distribution

346 The application of the hierarchical framework described in section 2 requires estimating  
 347 the following unknown quantities:

- 348 1. The dependogram parameters  $\boldsymbol{\psi}$
- 349 2. Additional parameters of the elliptical copula  $\boldsymbol{\eta}$
- 350 3. The regression parameters  $\boldsymbol{\beta}_k$ , for  $k = 1:D$
- 351 4. The regression errors  $\boldsymbol{\epsilon}_k$ , for  $k = 1:D$  (latent variables)
- 352 5. The covariogram parameters  $\boldsymbol{v}_k$ , for  $k = 1:D$  (hyper-parameters)

353 The posterior distribution of these quantities, given observations  $\tilde{\boldsymbol{y}}$  and covariates  $\boldsymbol{x}$ ,  
 354 can be derived as follows:

$$p(\boldsymbol{\psi}, \boldsymbol{\eta}, \boldsymbol{\beta}_{k=1:D}, \boldsymbol{\epsilon}_{k=1:D}, \boldsymbol{v}_{k=1:D} | \tilde{\boldsymbol{y}}, \boldsymbol{x})$$

$$\propto p(\tilde{\boldsymbol{y}} | \boldsymbol{\psi}, \boldsymbol{\eta}, \boldsymbol{\beta}_{k=1:D}, \boldsymbol{\epsilon}_{k=1:D}, \boldsymbol{v}_{k=1:D}, \boldsymbol{x}) p(\boldsymbol{\psi}, \boldsymbol{\eta}, \boldsymbol{\beta}_{k=1:D}, \boldsymbol{\epsilon}_{k=1:D}, \boldsymbol{v}_{k=1:D} | \boldsymbol{x}) \quad (7a)$$

$$= p(\tilde{\boldsymbol{y}} | \boldsymbol{\psi}, \boldsymbol{\eta}, \boldsymbol{\beta}_{k=1:D}, \boldsymbol{\epsilon}_{k=1:D}, \boldsymbol{x}) p(\boldsymbol{\psi}, \boldsymbol{\eta}, \boldsymbol{\beta}_{k=1:D}) p(\boldsymbol{\epsilon}_{k=1:D}, \boldsymbol{v}_{k=1:D}) \quad (7b)$$

$$= p(\tilde{\boldsymbol{y}} | \boldsymbol{\psi}, \boldsymbol{\eta}, \boldsymbol{\beta}_{k=1:D}, \boldsymbol{\epsilon}_{k=1:D}, \boldsymbol{x}) p(\boldsymbol{\psi}, \boldsymbol{\eta}, \boldsymbol{\beta}_{k=1:D}) p(\boldsymbol{\epsilon}_{k=1:D} | \boldsymbol{v}_{k=1:D}) p(\boldsymbol{v}_{k=1:D}) \quad (7c)$$

355 The following assumptions have been made to derive this posterior distribution:

- 356 1. Equation (7a) is a direct application of Bayes theorem.
- 357 2. Equation (7b) assumes: (i) the prior distribution does not depend on covariates  $\boldsymbol{x}$ ;
- 358 (ii) prior independence between  $\boldsymbol{\psi}, \boldsymbol{\eta}, \boldsymbol{\beta}_{k=1:D}$  in the one hand, and  $\boldsymbol{\epsilon}_{k=1:D}, \boldsymbol{v}_{k=1:D}$  in the  
 359 other hand. This assumption aims at isolating the hierarchical components of the model,  
 360 i.e.  $\boldsymbol{\epsilon}_{k=1:D}$  and  $\boldsymbol{v}_{k=1:D}$ . Moreover, it is noted that the likelihood of observations can  
 361 be derived without using the hyper-parameters  $\boldsymbol{v}_{k=1:D}$ . Indeed, regression errors  $\boldsymbol{\epsilon}_{k=1:D}$

362 suffice to apply the regression model (4) and hence to derive D-parameters (see detailed  
 363 derivation of the likelihood in following equation (9)).

364 3. Equation (7c) is an application of conditional probability rules.

365 The posterior distribution (7) is made up of the following components:

366 1. The terms  $p(\boldsymbol{\psi}, \boldsymbol{\eta}, \boldsymbol{\beta}_{k=1:D})$  and  $p(\mathbf{v}_{k=1:D})$  are priors for the inferred parameters and  
 367 hyper-parameters, respectively.

2. The term  $p(\boldsymbol{\epsilon}_{k=1:D}|\mathbf{v}_{k=1:D})$  represents the hierarchical part of the model. Assuming  
 regression errors  $\boldsymbol{\epsilon}_k$  related to different D-parameters  $\theta_k$  are mutually independent yields:

$$p(\boldsymbol{\epsilon}_{k=1:D}|\mathbf{v}_{k=1:D}) = \prod_{k=1}^D f_N(\epsilon_k(\tilde{s}_1), \dots, \epsilon_k(\tilde{s}_M)|\mathbf{0}, \boldsymbol{\Gamma}_k(\mathbf{v}_k)) \quad (8)$$

368 where  $\boldsymbol{\Gamma}_k(\mathbf{v}_k)$  is the covariance matrix derived from the covariogram model (6) and  
 369  $f_N(\mathbf{z}|\boldsymbol{\mu}, \boldsymbol{\Gamma})$  represents the pdf of a multivariate normal distribution with mean  $\boldsymbol{\mu}$  and  
 370 covariance matrix  $\boldsymbol{\Gamma}$ . Note that the assumption that there is no cross-correlation between  
 371 regression errors related to different D-parameters may be restrictive. This assumption  
 372 will be evaluated in the case study (section 4) and will be further discussed in section 5.3.

373 3. The term  $p(\tilde{\mathbf{y}}|\boldsymbol{\psi}, \boldsymbol{\eta}, \boldsymbol{\beta}_{k=1:D}, \boldsymbol{\epsilon}_{k=1:D}, \mathbf{x})$  is the likelihood of observations. Its derivation  
 374 requires further explanation. At a given time step  $t$ , the likelihood of (spatial) observations  
 375  $\tilde{\mathbf{y}}(t)$  can be computed as follows:

$$p(\tilde{\mathbf{y}}(t)|\boldsymbol{\psi}, \boldsymbol{\eta}, \boldsymbol{\beta}_{k=1:D}, \boldsymbol{\epsilon}_{k=1:D}, \mathbf{x}) = f_{EC}(\tilde{\mathbf{y}}(t)|\boldsymbol{\Sigma}(\boldsymbol{\psi}), \boldsymbol{\eta}, \{[\theta_k(\mathbf{x}_k(\tilde{s}_1), \boldsymbol{\beta}_k, \epsilon_k(\tilde{s}_1))]_{k=1:D}, \dots, [\theta_k(\mathbf{x}_k(\tilde{s}_M), \boldsymbol{\beta}_k, \epsilon_k(\tilde{s}_M))]_{k=1:D}\}) \quad (9)$$

376 In equation (9),  $\Sigma(\boldsymbol{\psi})$  is the dependence matrix of the elliptical copula, derived from  
 377 the dependogram model (3). For a given observation site  $\tilde{s}_j$ , the  $k$ th D-parameter  
 378  $\theta_k(\mathbf{x}_k(\tilde{s}_j), \boldsymbol{\beta}_k, \epsilon_k(\tilde{s}_j))$  is derived by applying the regression equation (4), i.e.:

$$\theta_k(\mathbf{x}_k(\tilde{s}_j), \boldsymbol{\beta}_k, \epsilon_k(\tilde{s}_j)) = g_k^{-1}(h_k(\mathbf{x}_k(\tilde{s}_j), \boldsymbol{\beta}_k) + \epsilon_k(\tilde{s}_j)) \quad (10)$$

379 Lastly,  $f_{EC}(\mathbf{z}|\Sigma, \boldsymbol{\eta}, \{\boldsymbol{\theta}(s_1), \dots, \boldsymbol{\theta}(s_M)\})$  represents the multivariate pdf of a  $M$ -dimensional  
 380 vector  $\mathbf{z}$  derived from the elliptical copula with pairwise dependence matrix  $\Sigma$ , additional  
 381 dependence parameters  $\boldsymbol{\eta}$  and marginal distributions  $\{p(\boldsymbol{\theta}(s_1)), \dots, p(\boldsymbol{\theta}(s_M))\}$ , as detailed  
 382 in Appendix A.

383 Assuming temporal independence between observations  $\tilde{\mathbf{y}}(t)$ , the likelihood of the whole  
 384 observation matrix  $\tilde{\mathbf{y}}$  is simply obtained as follows:

$$p(\tilde{\mathbf{y}}|\boldsymbol{\psi}, \boldsymbol{\eta}, \boldsymbol{\beta}_{k=1:D}, \boldsymbol{\epsilon}_{k=1:D}, \mathbf{x}) = \prod_{t=1}^T p(\tilde{\mathbf{y}}(t)|\boldsymbol{\psi}, \boldsymbol{\eta}, \boldsymbol{\beta}_{k=1:D}, \boldsymbol{\epsilon}_{k=1:D}, \mathbf{x}) \quad (11)$$

### 3.2. Inference

385 The posterior distribution (7) poses a computational challenge because its dimension  
 386 grows with the number of sites. This is due to the explicit modeling of regression errors  
 387 through latent variables. Consequently, the number of quantities to be inferred from the  
 388 posterior can amount to hundreds. This is typical of hierarchical models with latent vari-  
 389 ables used to describe unobserved processes [e.g., Clark, 2005]. A Markov Chain Monte  
 390 Carlo (MCMC) sampler is used to address this difficulty. It is stressed that MCMC sam-  
 391 pling from high-dimensional posteriors is challenging but by no means insurmountable.  
 392 Successful examples are provided by e.g. Crainiceanu et al. [2003]; Vrugt et al. [2008];

393 *Thyer et al.* [2009]; *Reichert and Mieleitner* [2009] in Hydrology, not to mention numer-  
394 ous applications in other fields [e.g., *Clark*, 2003, in ecology][*Storz and Beaumont*, 2002,  
395 in genetics,etc.].

396 The MCMC sampler used in this paper is made up of two stages. Stage one makes use of  
397 an adaptive block Metropolis algorithm with univariate Gaussian jump distributions (i.e.  
398 components of the parameter vector are updated one at a time). The adaption strategy  
399 adjusts the jump variances to produce an adequate jump rate [see *Renard et al.*, 2006, for  
400 a detailed description]. This first sampler is used to perform a preliminary exploration  
401 of the posterior distribution properties (notably in terms of posterior covariance). In  
402 the second stage, a standard Metropolis sampler [*Metropolis and Ulam*, 1949; *Metropolis*  
403 *et al.*, 1953] is used, with a Gaussian jump distribution whose covariance matrix is speci-  
404 fied using the preliminary exploration performed at stage one. Convergence is assessed by  
405 evolving four parallel chains and verifying that the Gelman-Rubin criteria [*Gelman et al.*,  
406 1995] are close to one for all inferred quantities.

407 Note that additional computing efficiency might be achieved in some cases by using con-  
408 jugate priors. However, this possibility is not investigated in this paper.

### 3.3. Prediction at gauged site

409 Once inference has been performed using the MCMC strategy outlined above, the next  
410 step is to use the model to predict some quantity of interest. In this section, focus is  
411 on prediction at a gauged site  $\tilde{s}$ . A typical quantity of interest is the  $p$ -quantile of the  
412 distribution of observations at site  $\tilde{s}$ , which can be directly derived from D-parameters  
413  $\boldsymbol{\theta}(\tilde{s})$ .

414 In a Bayesian context, another interesting byproduct of the posterior distribution is the

415 predictive distribution of a (future) observation. In general, if  $\Theta$  denotes the vector of all  
 416 parameters subject to inference, the pdf of this predictive distribution, evaluated at some  
 417 value  $w$ , is mathematically defined as follows [Gelman *et al.*, 1995]:

$$p(w|\tilde{\mathbf{y}}) = \int p(w|\Theta)p(\Theta|\tilde{\mathbf{y}})d\Theta \quad (12)$$

418 In practice, this integration is not performed analytically but is approximated using the  
 419 MCMC replicates from the posterior (7). The approximation algorithm simply consists  
 420 in generating a value from the at-site distribution (1) for each MCMC replicate. The  
 421 resulting sample is a realization from the predictive distribution and can therefore be  
 422 used to estimate its characteristics (mean, variance, probability interval, etc.).

423 Let us assume that MCMC sampling generated a set of  $n_{sim}$  replicates from the posterior  
 424 distribution,  $(\boldsymbol{\psi}^{(j)}, \boldsymbol{\eta}^{(j)}, \boldsymbol{\beta}_{k=1:D}^{(j)}, \boldsymbol{\epsilon}_{k=1:D}^{(j)}, \mathbf{v}_{k=1:D}^{(j)})_{j=1:n_{sim}}$ . These replicates can be used to  
 425 make a prediction at gauged site  $\tilde{s}$  using the following algorithm:

426 Do  $j = 1 : n_{sim}$

427 1. compute D-parameters from the regression model (4),  $\theta_k^{(j)}(\tilde{s}) = g_k^{-1} \left( h_k(\mathbf{x}_k(\tilde{s}), \boldsymbol{\beta}_k^{(j)}) + \epsilon_k^{(j)}(\tilde{s}) \right)$   
 428 for  $k = 1 : D$

429 2. compute derived quantity  $z^{(j)} = z(\boldsymbol{\theta}^{(j)}(\tilde{s}))$  (e.g., the T-year quantile) or generate a  
 430 value from the at-site distribution  $w^{(j)} \sim p(\boldsymbol{\theta}^{(j)}(\tilde{s}))$

431 The samples  $(z^{(j)})_{j=1:n_{sim}}$  and  $(w^{(j)})_{j=1:n_{sim}}$  can be considered as realizations from the  
 432 posterior distribution of the quantity of interest and from the predictive distribution,  
 433 respectively.

### 3.4. Prediction at ungauged site



434 The fundamental difference with the prediction approach presented in previous sec-  
 435 tion 3.3 is that the regression error  $\epsilon_k(s)$  has not been inferred for an ungauged site  $s$ .  
 436 However, the properties of the hyper-distribution (5) have been inferred. Since the hyper-  
 437 distribution is the joint distribution of a vector of regression errors, it can be used to  
 438 indirectly infer the properties of the regression error  $\epsilon_k(s)$  for the target ungauged site  $s$ .  
 439 More precisely, the predictive distribution of the regression error  $\epsilon_k(s)$ , given observed  
 440 data  $\tilde{\mathbf{y}}$  and covariates  $\mathbf{x}$ , can be computed as follows:

$$p(\epsilon_k(s)|\tilde{\mathbf{y}}, \mathbf{x}) = \int p(\epsilon_k(s), \epsilon_k(\tilde{s}_1), \dots, \epsilon_k(\tilde{s}_M), \mathbf{v}_k | \tilde{\mathbf{y}}, \mathbf{x}) d\epsilon_k(\tilde{s}_1) \dots d\epsilon_k(\tilde{s}_M) d\mathbf{v}_k \quad (13a)$$

$$= \int p(\epsilon_k(s) | \epsilon_k(\tilde{s}_1), \dots, \epsilon_k(\tilde{s}_M), \mathbf{v}_k, \tilde{\mathbf{y}}, \mathbf{x}) p(\epsilon_k(\tilde{s}_1), \dots, \epsilon_k(\tilde{s}_M), \mathbf{v}_k | \tilde{\mathbf{y}}, \mathbf{x}) d\epsilon_k(\tilde{s}_1) \dots d\epsilon_k(\tilde{s}_M) d\mathbf{v}_k \quad (13b)$$

441 The second term in equation (13b) is the posterior distribution of  $(\epsilon_k(\tilde{s}_1), \dots, \epsilon_k(\tilde{s}_M), \mathbf{v}_k)$ .  
 442 It can therefore be directly approximated using the MCMC replicates. The first term  
 443 represents the distribution of the regression error  $\epsilon_k(s)$  for the target ungauged site  $s$ ,  
 444 conditional on regression errors at gauged sites  $\epsilon_k(\tilde{s}_1) \dots \epsilon_k(\tilde{s}_M)$  and hyper-parameters  $\mathbf{v}_k$ .  
 445 Following equation (5), the joint distribution of  $\epsilon_k(s)$  and  $\epsilon_k(\tilde{s}_1) \dots \epsilon_k(\tilde{s}_M)$  (conditional  
 446 on  $\mathbf{v}_k$ ) is a multivariate Gaussian distribution with mean zero and covariance matrix  
 447  $\mathbf{\Gamma}_k = \mathbf{\Gamma}_k(\mathbf{v}_k)$  derived from the covariogram (6). Let us partition this covariance matrix as  
 448 follows:

$$\mathbf{\Gamma}_k = \begin{pmatrix} \sigma_k^2 & \mathbf{\Lambda}_k \\ \mathbf{\Lambda}_k^t & \mathbf{\Omega}_k \end{pmatrix} \quad (14)$$

449 where  $\sigma_k^2$  is the marginal variance of regression errors,  $\mathbf{\Omega}_k = \mathbf{\Omega}_k(\mathbf{v}_k)$  is the  $M \times M$   
 450 covariance matrix of regression errors at gauged sites  $\epsilon_k(\tilde{s}_1) \dots \epsilon_k(\tilde{s}_M)$ , and  $\mathbf{\Lambda}_k = \mathbf{\Lambda}_k(\mathbf{v}_k)$

451 is the  $1 \times M$  vector of covariances between  $\epsilon_k(s)$  and  $\epsilon_k(\tilde{s}_1) \dots \epsilon_k(\tilde{s}_M)$ . Using a well-known  
 452 formula for conditional Gaussian distributions, it follows that the conditional distribution  
 453 in equation (13b) is a univariate Gaussian distribution with mean  $\mu_{k,cond}$  and variance  
 454  $\sigma_{k,cond}^2$ :

$$\begin{aligned}
 p(\epsilon_k(s) | \epsilon_k(\tilde{s}_1), \dots, \epsilon_k(\tilde{s}_M), \mathbf{v}_k, \tilde{\mathbf{y}}, \mathbf{x}) &= N(\mu_{k,cond}; \sigma_{k,cond}^2) \\
 \mu_{k,cond} &= \mathbf{\Lambda}_k \mathbf{\Omega}_k^{-1} (\epsilon_k(\tilde{s}_1), \dots, \epsilon_k(\tilde{s}_M))^t \\
 \sigma_{k,cond}^2 &= \sigma_k^2 - \mathbf{\Lambda}_k \mathbf{\Omega}_k^{-1} \mathbf{\Lambda}_k^t
 \end{aligned} \tag{15}$$

455 The conditional distribution in equation (15) complements the regression to transfer  
 456 information from gauged to ungauged sites. Indeed, the vector  $\mathbf{\Lambda}_k$  acts as a weight vector  
 457 favoring gauged sites nearby the target ungauged site  $s$ . If  $\epsilon_k(s)$  is independent from all  
 458  $\epsilon_k(\tilde{s}_j)$ 's, the conditional distribution is equal to the marginal distribution of regression  
 459 errors, i.e. a Gaussian distribution with mean  $\mu_{k,cond} = 0$  and variance  $\sigma_{k,cond}^2 = \sigma_k^2$ . In  
 460 this case, the transfer of information to ungauged site  $s$  does not favor any particular  
 461 gauged site  $\tilde{s}$ .

462 The explanations given above lead to the following algorithm for prediction at an ungauged  
 463 site  $s$ :

464 Do  $j = 1 : n_{sim}$

- 465 1. compute  $\mathbf{\Omega}_k^{(j)} = \mathbf{\Omega}_k(\mathbf{v}_k^{(j)})$ ;  $\mathbf{\Lambda}_k^{(j)} = \mathbf{\Lambda}_k(\mathbf{v}_k^{(j)})$ ;  $\sigma_k^{2(j)} = \sigma_k^2(\mathbf{v}_k^{(j)})$ .
- 466 2. compute  $\mu_{k,cond}^{(j)}$  and  $\sigma_{k,cond}^{(j)}$  according to equation (15).
- 467 3. generate regression error  $\epsilon_k^{(j)}(s)$  from the conditional distribution (15) for  $k = 1 : D$   
 468 (Gaussian distribution with mean  $\mu_{k,cond}^{(j)}$  and standard deviation  $\sigma_{k,cond}^{(j)}$ ).

- 469 4. compute D-parameters from the regression model (4),  $\theta_k^{(j)}(s) = g_k^{-1} \left( h_k(\mathbf{x}_k(s), \boldsymbol{\beta}_k^{(j)}) + \epsilon_k^{(j)}(s) \right)$   
470 for  $k = 1 : D$
- 471 5. compute derived quantity  $z^{(j)} = z(\boldsymbol{\theta}^{(j)}(s))$  or generate a value from the at-site  
472 distribution  $w^{(j)} \sim p(\boldsymbol{\theta}^{(j)}(s))$

#### 4. Case study: Mediterranean extreme rainfall

473 The application of the Bayesian hierarchical approach is illustrated with a case study  
474 involving extreme rainfall data. The main objectives of this application are to demon-  
475 strate the feasibility of a Bayesian hierarchical approach and to assess the impact of some  
476 modeling hypotheses (in particular the choices of the data dependence model and the  
477 regression model). Note that a synthetic case study was also performed to verify the  
478 internal consistency of the modeling framework (not shown).

##### 4.1. Data

479 Annual maxima from 87 series of daily rainfall are used in this study (Figure 2). Rain-  
480 gauges, whose elevations range from 1 to 1102 m., are located in the French Mediterranean  
481 area. This region is characterized by intense rainfall in autumn and is delimited by three  
482 mountainous areas: the Pyreneans (South-West), the Alps (East) and the Cevennes (Cen-  
483 ter). The thin lines in Figure 2 represent six homogeneous regions defined by *Pujol et al.*  
484 [2007].

485 Sixty raingauges are used for estimation, the remaining 27 series being used for model  
486 validation. Data are available over the period 1955-2004. Years with more than 15 days  
487 of missing values are treated as missing data. Since the treatment of missing data during  
488 estimation is not obvious, the 60 estimation series were chosen to ensure the completeness

489 of annual maxima series over the period 1955-2004 (this point is discussed in more details  
 490 in section 5.1).

## 4.2. Model specification

### 4.2.1. Data level

491 Let  $(y(\tilde{s}_i, t))_{i=1:60, t=1:50}$  denote the observed annual maxima. At any site  $s$ , the random  
 492 variable  $(Y(s, t))$  is assumed to follow a Generalized Extreme Value (GEV) distribution  
 493 with parameters  $(\mu(s), \lambda(s), \xi(s))$ . The pdf of a GEV distribution with location parameter  
 494  $\mu$ , scale parameter  $\lambda$  and shape parameter  $\xi$  is:

$$p(y|\mu, \lambda, \xi) = \frac{1}{\lambda} \left[ 1 - \xi \left( \frac{y - \mu}{\lambda} \right) \right]^{\frac{1}{\xi} - 1} \exp \left\{ - \left[ 1 - \xi \left( \frac{y - \mu}{\lambda} \right) \right]^{1/\xi} \right\}$$

$$\lambda > 0, \xi \neq 0, 1 - \xi \left( \frac{y - \mu}{\lambda} \right) > 0 \quad (16)$$

496 The joint distribution of observations  $\tilde{\mathbf{y}}(t)$  is derived using three distinct assumptions on  
 497 the dependence structure. The first assumption corresponds to using a Gaussian copula  
 498 [e.g., *Renard and Lang, 2007*] with dependence matrix  $\Sigma$ :

$$(Y(\tilde{s}_1, t), \dots, Y(\tilde{s}_M, t)) \sim GCop_M(\Sigma, \{\mu(\tilde{s}_i), \lambda(\tilde{s}_i), \xi(\tilde{s}_i)\}_{i=1:M}) \quad (17)$$

499 The second assumption corresponds to using a Student copula with dependence matrix  
 500  $\Sigma$  and tail dependence coefficient  $\nu$ :

$$(Y(\tilde{s}_1, t), \dots, Y(\tilde{s}_M, t)) \sim SCop_M(\Sigma, \nu, \{\mu(\tilde{s}_i), \lambda(\tilde{s}_i), \xi(\tilde{s}_i)\}_{i=1:M}) \quad (18)$$

501 Lastly, the third assumption corresponds to assuming spatial independence between  
 502 data. It can be viewed as a special case of Equation (17) with the dependence matrix  $\Sigma$

503 forced to unity.

504 In both the Student and Gaussian copula cases, the dependogram is formulated as a  
505 weighted sum of two exponential functions (with  $\psi_1 < \psi_2$ ) [e.g., *Kjeldsen and Jones,*  
506 2009a]:

$$\Sigma(i, j) = \psi_0 \exp(-\psi_1 \|\tilde{s}_i - \tilde{s}_j\|) + (1 - \psi_0) \exp(-\psi_2 \|\tilde{s}_i - \tilde{s}_j\|) \quad (19)$$

#### 507 4.2.2. Process level

508 The variation of D-parameters ( $\mu(s), \lambda(s), \xi(s)$ ) in space is described using several mod-  
509 eling assumptions. First, an index-flood-like approach, assuming scale invariance, is used.  
510 This assumption induces strong constraints on the parameters of at-site distributions.  
511 More precisely, a GEV distribution complying with the scale invariance assumption can  
512 be reparameterized as follows [e.g., *Ribatet et al., 2007*]:

$$Y(s, t) \sim GEV(\delta(s)\mu, \delta(s)\lambda, \xi) \quad (20)$$

513 Equation (20) states that (i) the shape parameter is constant throughout the region; (ii)  
514 the ratio between the location and the scale parameters is constant throughout the region,  
515 which implies that annual maxima from all sites have the same coefficient of variation. In  
516 other terms, the whole spatial variability of annual maxima is accounted for by the index  
517 flood parameter  $\delta(s)$ , with other parameters  $\mu, \lambda, \xi$  being assumed regional.

518 Two distinct regression models are used in this index flood approach. The first model,  
519  $\mathcal{M}_1$ , does not use any covariate and simply describes the spatial variation of the index  
520 flood D-parameter  $\delta(s)$  using a Gaussian spatial field:

$$\log(\delta(\tilde{s}_i)) = \epsilon(\tilde{s}_i); (\epsilon(\tilde{s}_1), \dots, \epsilon(\tilde{s}_M)) \sim N(0, \mathbf{\Gamma}) \quad (21)$$

521 The second model,  $\mathcal{M}_2$ , uses the raingauge elevation as covariate. Preliminary inves-  
 522 tigations indicate that the effect of elevation depends on the region. Consequently, the  
 523 model uses region-specific relationships between the index flood D-parameter  $\delta(\tilde{s}_i)$  and  
 524 elevation:

$$\begin{aligned} \log(\delta(\tilde{s}_i)) &= \beta_0^{(j)} + \beta_1^{(j)} * \text{elevation}(\tilde{s}_i) + \epsilon(\tilde{s}_i) \\ (\epsilon(\tilde{s}_1), \dots, \epsilon(\tilde{s}_M)) &\sim N(0, \mathbf{\Gamma}) \end{aligned} \quad (22)$$

525 where the superscript  $(j)$  is used to denote the region of site  $\tilde{s}_i$ . In order to ensure the  
 526 identifiability of the model, the following additional constraint is applied:

$$\sum_{j=1}^6 \beta_0^{(j)} = 0 \quad (23)$$

527 In a second step, the index-flood approach described above is made less restrictive  
 528 by abandoning the scale invariance assumption. More precisely, the following at-site  
 529 distribution is assumed:

$$Y(s, t) \sim GEV(\mu(s), \lambda(s), \xi) \quad (24)$$

530 Compared to the index-flood equation (20), equation (24) does not assume a constant  
 531 ratio between the location and the scale parameters throughout the region. However, it  
 532 still assumes a constant shape parameter. The spatial variation of the location  $\mu(s)$  and  
 533 scale  $\lambda(s)$  D-parameters is described using region-specific relationships with elevation:

$$\log(\mu(\tilde{s}_i)) = \beta_0^{(j)} + \beta_1^{(j)} * \text{elevation}(\tilde{s}_i) + \epsilon_\mu(\tilde{s}_i)$$

$$(\epsilon_\mu(\tilde{s}_1), \dots, \epsilon_\mu(\tilde{s}_M)) \sim N(0, \mathbf{\Gamma}_\mu)$$

$$\log(\lambda(\tilde{s}_i)) = \alpha_0^{(j)} + \alpha_1^{(j)} * \text{elevation}(\tilde{s}_i) + \epsilon_\lambda(\tilde{s}_i)$$

$$(\epsilon_\lambda(\tilde{s}_1), \dots, \epsilon_\lambda(\tilde{s}_M)) \sim N(0, \mathbf{\Gamma}_\lambda) \tag{25}$$

534 The model defined by equations (24)-(25) is noted  $\mathcal{M}_3$ .

535 In all models  $\mathcal{M}_1$ - $\mathcal{M}_3$ , the covariance matrixes  $\mathbf{\Gamma}$  in equations (21), (22) and (25) are  
 536 parameterized as follows:

$$\Gamma(i, j) = \sigma^2 [v_0 \exp(-v_1 \|\tilde{s}_i - \tilde{s}_j\|) + (1 - v_0) \exp(-v_2 \|\tilde{s}_i - \tilde{s}_j\|)] \tag{26}$$

537 Finally, vague priors are specified for all parameters by using uniform distributions with  
 538 large support. The only exception is the shape parameter, for which a Gaussian prior with  
 539 mean zero and standard deviation 0.3 is specified: this is similar to the "Geophysical prior"  
 540 used by *Martins and Stedinger* [2000]. Prior specification is further discussed in section  
 541 5.6.

### 4.3. Estimation and prediction

542 In this section, inference is performed using the regression model  $\mathcal{M}_3$  coupled with a  
 543 Gaussian copula assumption for describing data dependence (equation (17)). This model  
 544 is noted  $\mathcal{M}_3$ -GCop.

#### 4.3.1. Parameter estimates

546 Figure 3a shows the posterior pdf of the regional shape parameter  $\xi$ . The median  
547 value of about -0.12 corresponds to an heavier tail than the Gumbel distribution: this  
548 is consistent with previous studies in the same area [Neppel *et al.*, 2007]. Moreover, this  
549 estimation is quite precise: this is a consequence of the hypothesis that this parameter is  
550 constant throughout the region.

551 Figure 3b-c shows the posterior pdfs of a few regression errors, for both location (b)  
552 and scale (c) regressions (*latent variables* in the vocabulary of hierarchical modeling).  
553 Recall that those errors are assumed mutually independent in the inference framework (see  
554 section 3). This assumption can be evaluated here by computing the correlation between  
555 series of estimated regression errors  $\hat{\epsilon}_\mu(\tilde{s}_1), \dots, \hat{\epsilon}_\mu(\tilde{s}_M)$  and  $\hat{\epsilon}_\lambda(\tilde{s}_1), \dots, \hat{\epsilon}_\lambda(\tilde{s}_M)$  (maximum-  
556 posterior estimates are used). A correlation of about 0.52 is found, which suggests that  
557 the assumption that both error processes are not cross-correlated might not be realistic.  
558 This is further discussed in section 5.3.

559 The standard deviations of regression errors are represented in Figure 3d. Those terms  
560 (*hyper-parameters* in the vocabulary of hierarchical modeling) are of primary importance  
561 since they control the predictive uncertainty at ungauged sites. In this case study, the  
562 estimated hyper-standard deviations are similar for both location and scale regressions.  
563 They roughly correspond to a standard error of 15% in the regressions.

#### 564 4.3.2. Data and parameter dependences

565 Figure 4 shows the estimated dependence-distance relationships, for both the data de-  
566 pendence model (19) and the regression errors dependence model (26). The dependence  
567 between data (Figure 4a) is precisely estimated, and confirm that data are not spatially  
568 independent. Conversely, the dependence between regression errors (Figure 4b-c) appears



569 more difficult to estimate, with large posterior intervals denoting a lower precision. How-  
570 ever, the strength of dependence is rather limited here, with the correlograms dropping  
571 to near-zero values at relatively short distances of about 10-20 km.

### 572 **4.3.3. Prediction**

573 The 100-year daily rainfall is predicted on a 40\*50 grid (yielding 10\*10 km cells) using  
574 the procedure detailed in section 3.4. Since the highest raingauge has an elevation of only  
575 1102 m, high-elevation areas from the Alps and the Pyreneans (whose maximal elevations  
576 reach 4810 m and 3404 m, respectively) were excluded from the prediction (white areas  
577 in Figure 5). The left panel of Figure 5 shows the 100-year daily rainfall estimated on  
578 the grid using the posterior median. An area with higher  $R_{0.99}$  values (reaching 400 mm  
579 at some grid points) is located on the Cevennes mountain range: this area is well-known  
580 to be affected by the highest rainfall intensities in France. Moreover, the rainfall quantile  
581 abruptly drops to smaller values ( $\approx 100$ -150 mm) downwind of the Cevennes.

582 The Bayesian framework used in this paper enables a direct assessment of the uncertainties  
583 affecting predictions. The right panel of Figure 5 shows the uncertainty in predicted 100-  
584 year daily rainfall, measured by the posterior coefficient of variation. This uncertainty  
585 appears relatively uniform over the region, and mostly corresponds to 15-20% coefficients  
586 of variations. Note however the high-uncertainty area appearing on the foothill of the  
587 Alps (East): this is due to the low number of calibration sites in this region (only three  
588 sites), which does not enable a precise estimation of the elevation effect.

## **4.4. Impact of the data dependence model**

589 In this section, the impact of the model used to describe spatial dependence between  
590 data is evaluated. To this aim, inference is performed with the three dependence models

591 described in section 4.2.1, coupled with the regression model  $\mathcal{M}_3$ . The three resulting  
592 models are noted  $\mathcal{M}_3$ -Inde (independence assumption),  $\mathcal{M}_3$ -GCop (Gaussian copula as-  
593 sumption) and  $\mathcal{M}_3$ -SCop (Student copula assumption).

#### 594 4.4.1. Parameter estimates

595 The regional shape parameter is only moderately impacted by the data dependence  
596 model (Figure 6a). The posterior pdf of model  $\mathcal{M}_3$ -SCop is slightly shifted and is more  
597 variable compared with models  $\mathcal{M}_3$ -Inde and  $\mathcal{M}_3$ -GCop. A similar observation can be  
598 made for the hyper-standard deviation of location regression errors (Figure 6b). Posterior  
599 pdfs for the hyper-standard deviation of scale regression errors (Figure 6c) have a similar  
600 variance but a slightly different mode.

#### 601 4.4.2. Data and parameter dependences

602 Figure 7a shows the estimated dependence between data. The dependogram drops  
603 immediately to zero for  $\mathcal{M}_3$ -Inde due to the spatial independence assumption. Dependo-  
604 grams for  $\mathcal{M}_3$ -GCop and  $\mathcal{M}_3$ -SCop are virtually indistinguishable. However, the Student  
605 copula also depends of an additional parameter controlling the strength of asymptotic  
606 dependence. The posterior pdf of this parameter is shown in Figure 7a, and corresponds  
607 to a rather limited tail dependence.

608 Stronger differences are observed for the dependence between regression errors (Figure  
609 7b-c). In particular, the dependogram for  $\mathcal{M}_3$ -GCop drops faster to zero for location  
610 regression errors (Figure 7b). Moreover, the choice of the data dependence model appears  
611 to impact the uncertainty in the estimation of the regression errors correlograms (for both  
612 location and scale regressions, Figure 7b-c):  $\mathcal{M}_3$ -Inde yields wider posterior intervals than  
613  $\mathcal{M}_3$ -SCop and  $\mathcal{M}_3$ -GCop.

### 614 4.4.3. Prediction

615 Results described in previous sections 4.4.1 and 4.4.2 show that the impact of the model  
616 for data dependence is quite complex. A more integrated assessment can be made by  
617 comparing the prediction of 100-year quantiles obtained with the three models. Figure 8  
618 maps the relative difference (in percent) between (i)  $\mathcal{M}_3$ -Inde (top panels) and  $\mathcal{M}_3$ -SCop  
619 (bottom panels) estimates in the one hand; and (ii)  $\mathcal{M}_3$ -GCop estimates in the other  
620 hand. The Gaussian Copula is therefore considered as a benchmark model, and the two  
621 other dependence assumptions are compared to this benchmark.

622 Overall, assuming spatial independence between data yields only minor change in  $R_{0.99}$   
623 estimates (top left panel). Exceptions are located in the Cevennes ( $R_{0.99}$  values are  $\approx$   
624 10-15% higher), and on the foothill of the Alps ( $R_{0.99}$  values are  $\approx$  10-35% higher in the  
625 North, and  $\approx$  10-35% lower in the South). Similarly, changes in the uncertainty of  $R_{0.99}$   
626 estimates are minor (top right panel). However, uncertainty reductions dominate, with  
627 decreases in the range 0-25%. This is consistent with the expected behavior when spatial  
628 dependence is ignored (see discussion in section 2.1.2): the variance of estimates may be  
629 underestimated, i.e. this decrease in uncertainty may be unduly optimistic. Note that  
630 once again, the foothill of the Alps is an exception to this overall decrease.

631 Overall, replacing the Gaussian copula model by a Student copula also yields minor change  
632 in  $R_{0.99}$  estimates (bottom left panel), except in the Alps region where strong increases are  
633 observed ( $\approx$  40%, culminating at  $\approx$  100% for a couple of pixels). In terms of uncertainty  
634 (bottom right panel), increases now dominate, but are still moderate (in the range 0-20%),  
635 with the Alps showing an opposite behavior (uncertainty reductions in the range 20-40%).

### 4.5. Impact of the regression model

636 In this section, the impact of the regression model is evaluated. Inference is performed  
637 with the three regression models described in section 4.2.2, coupled with the Gaussian  
638 copula model for data dependence. The three resulting models are noted  $\mathcal{M}_1$ -GCop (scale  
639 invariance assumption, no elevation effect),  $\mathcal{M}_2$ -GCop (scale invariance assumption with  
640 elevation effect) and  $\mathcal{M}_3$ -GCop (elevation effect, scale invariance is not assumed).

#### 641 4.5.1. Parameter estimates

642 The assumption of scale invariance notably impacts the shape parameter estimates  
643 (Figure 9a), with a marked shift between scale-invariant models  $\mathcal{M}_1$ -GCop and  $\mathcal{M}_2$ -GCop  
644 in the one hand, and model  $\mathcal{M}_3$ -GCop in the other hand. Conversely, the completeness  
645 of the regression model seems to be the primary factor of influence for the hyper-standard  
646 deviations (Figure 9b-c): model  $\mathcal{M}_1$ -GCop, which ignores the elevation effect, yields a  
647 markedly higher hyper-standard deviation than models  $\mathcal{M}_2$ -GCop and  $\mathcal{M}_3$ -GCop, which  
648 use elevation as covariate. This is an expected result, since improving the regression model  
649 results in decreasing the standard deviations of regression errors. Note that in Figure 9b-  
650 c, the posterior pdfs for  $\mathcal{M}_1$ -GCop and  $\mathcal{M}_2$ -GCop are related to the regression for the  
651 index flood parameter  $\delta(s)$ , as defined in equation (21). Since this regression acts on both  
652 the location and scale D-parameters (see Equation (20)), those pdfs are repeated in both  
653 panels b and c.

#### 654 4.5.2. Data and parameter dependences

655 Figure 10a shows that data dependence is virtually identical with all regression models.  
656 This is consequence of using a copula formalism, with dependence being modeled inde-  
657 pendently of marginal distributions. Since the regression model only applies on marginal  
658 distributions, it does not impact the estimation of dependence between observed data.

659 Conversely, the dependence between regression errors is strongly impacted by the choice  
660 of a regression model (Figure 10b-c). The amount of dependence is markedly higher with  
661 simple model  $\mathcal{M}_1$ -GCop because the whole spatial variability in the distribution of an-  
662 nual maxima is accounted for by the regression errors. Models  $\mathcal{M}_2$ -GCop and  $\mathcal{M}_3$ -GCop  
663 yield lower dependences, suggesting that the elevation effect explains a major part of  
664 the spatial variability. This result is consistent with the findings of *Kjeldsen and Jones*  
665 [2009a] who observed a decrease in the dependence between regression errors when the  
666 regression model is improved. Lastly, model  $\mathcal{M}_3$ -GCop yield higher uncertainties than  
667 model  $\mathcal{M}_2$ -GCop. This might be due to the estimation of two distinct regression models  
668 for location and scale D-parameters in model  $\mathcal{M}_3$ -GCop, while model  $\mathcal{M}_2$ -GCop uses a  
669 single regression acting on the index flood parameter  $\delta(s)$ .

### 670 4.5.3. Prediction

671 Maps of 100-year quantiles are compared in a similar way to section 4.4.3. Figure 11  
672 maps the relative difference (in percent) between (i)  $\mathcal{M}_1$ -Gcop (top panels) and  $\mathcal{M}_2$ -  
673 GCop (bottom panels) estimates in the one hand; and (ii)  $\mathcal{M}_3$ -GCop estimates in the  
674 other hand. The regression model  $\mathcal{M}_3$  is therefore considered as a benchmark model, and  
675 the two other regression models are compared to this benchmark.

676 The simple index-flood model  $\mathcal{M}_1$  (which ignores the elevation effect) yields marked dif-  
677 ferences with the benchmark model  $\mathcal{M}_3$  in terms of  $R_{0.99}$  values (upper left panel): higher  
678 values are observed in the Cevennes (up to  $\approx +100\%$ ) and in the Alps (up to  $\approx +30\%$ ),  
679 while lower values are observed in the south-eastern corner of the domain (up to  $\approx -$   
680  $50\%$ ). Marked differences also appear for the uncertainty in  $R_{0.99}$  estimates (upper right  
681 panel), with an overall increase exceeding  $\approx 50\%$ , except in the Alps where the uncer-

682 tainty decreases (up to  $\approx 50\%$ ). Note however that the difference between  $\mathcal{M}_1$  and  $\mathcal{M}_3$   
683 uncertainties is not uniform and follows an interesting spatial pattern: uncertainties are  
684 similar for both models nearby calibration sites, while  $\mathcal{M}_1$  estimates become far more  
685 uncertain than  $\mathcal{M}_3$  estimates in poorly-gauged areas. This is a consequence of regression  
686 errors showing a significant amount of spatial dependence with model  $\mathcal{M}_1$  (see section  
687 4.5.2 and Figure 10b-c). This spatial dependence improves the efficiency of the transfer of  
688 information from gauged sites to nearby ungauged sites. However, when moving further  
689 away from the calibration sites, this dependence vanishes and the uncertainty becomes  
690 primarily controlled by the hyper-standard deviation of regression errors. This hyper-  
691 standard deviation is markedly higher for model  $\mathcal{M}_1$  (see section 4.5.1 and Figure 9b-c),  
692 which explains its higher uncertainty in poorly-gauged areas.

693 Qualitatively similar results are found for the index-flood model  $\mathcal{M}_2$  (which includes an  
694 elevation effect). However, differences with the model  $\mathcal{M}_3$  in terms of  $R_{0.99}$  values are  
695 smaller (in the range  $\approx \pm 20\%$ , lower left panel). The uncertainty in  $R_{0.99}$  estimates is  
696 also larger with model  $\mathcal{M}_2$  than with model  $\mathcal{M}_3$  (lower right panel), with increases in the  
697 range  $\approx 30\text{-}40\%$  being quite evenly distributed in space.

#### 698 **4.5.4. Assessment of the scale invariance assumption**

699 The scale invariance assumption underlying models  $\mathcal{M}_1$  and  $\mathcal{M}_2$  can be appraised by  
700 computing the coefficients of variation (CV) of data from each site. As explained in section  
701 4.2.2, scale invariance implies that the CV remains constant throughout the studied area.  
702 Figure 12 shows the empirical CV computed on each site (sites are reordered by increasing  
703 CV). The dashed horizontal lines show a 90% posterior interval of the CV resulting from  
704 model  $\mathcal{M}_2$ -GCop. Numerous empirical CVs are well outside this interval, which casts

705 doubts on the validity of the scale invariance assumption. Conversely, model  $\mathcal{M}_3$ -GCop  
706 (vertical bars) is able to track more closely the varying CVs at different sites.

## 4.6. Validation

707 The validation of predictions arising from the Bayesian hierarchical framework is per-  
708 formed by using tools and concepts borrowed from the field of probabilistic forecasting.  
709 Such tools are of particular interest when the quantification of uncertainty is a primary  
710 concern, which is the case in this paper. In particular, the reliability and the precision of  
711 predictions are evaluated: reliability refers to the ability to derive predictive distributions  
712 that are consistent with observations from validation sites, while precision refers to the  
713 amount of uncertainty in predictions. Both concepts are complementary: predictions can  
714 be reliable but not precise (uncertainties are large, but predictions remain consistent with  
715 observations) or alternatively precise but not reliable (e.g. due to an underestimation of  
716 uncertainties). An ideal predictive framework would yield predictions that are reliable  
717 and as precise as possible.

### 4.6.1. Reliability of predictions

718 The reliability of predictions can be assessed by comparing annual maxima from the  
719 27 validation sites with their predictive distribution (derived using the "ungauged site"  
720 procedure described in section 3.4). More precisely, this comparison is based on the  
721 predictive QQ-plot used by e.g. Dawid [1984], Gneiting et al. [2007], Laio and Tamea  
722 [2007] or Thyer et al. [2009]. Let  $F_s$  be the cdf of the predictive distribution at site  
723  $s$ , and  $\{Y(s, t)\}_{t=1:T_s}$  validation data at site  $s$ . Under the assumption that validation  
724 data are realizations from the predictive distribution, the  $p$ -values  $\{F_s(Y(s, t))\}_{t=1:T_s}$  are  
725 realizations from a uniform distribution on  $[0;1]$ . This can be evaluated with a QQ-plot  
726

727 comparing the empirical cdf of  $p$ -values with the cdf of a uniform distribution. This  
728 comparison can be performed individually for each validation site, or by pooling  $p$ -values  
729 from all validation sites.

730 Figure 13 shows the site-specific predictive QQ-plots for each model (first five panels), then  
731 compares the pooled-sites-predictive QQ-plots of the five models in a single plot (lower  
732 right panel). Overall, the reliability of the predictive distributions appears acceptable  
733 with all regression models, although some specific sites show signs of departure from the  
734 1:1 line. Moreover, when all sites are pooled together, all models yield nearly-diagonal pp-  
735 plots (lower right panel), suggesting that all methods provide equally reliable predictions.  
736 This result might appear surprising at first sight, given the differences between models  
737 highlighted in previous sections 4.4 and 4.5. However, there is no contradiction in this  
738 statement: in the framework of probabilistic prediction, several predictive distributions  
739 can be equally reliable but distinct from each other. In particular, the fact that the overall  
740 reliability is similar does not mean that, for a given site, the predictive distributions will  
741 themselves be similar. This is illustrated in Figure 14, where the model  $\mathcal{M}_2$ -GCop yields  
742 a predictive distribution markedly different from the other four models. Moreover, the  
743 probabilistic predictions may also differ in their precision (also termed "sharpness" in the  
744 field of probabilistic prediction [see *Gneiting et al.*, 2007, for further discussion]).

#### 745 **4.6.2. Precision of predictions**

746 Given that all models were found equally reliable, one would favor the one yielding the  
747 most precise predictions. In order to evaluate the predictive precision, Figure 15 shows  
748 the posterior coefficient of variation of the estimated 100-year rainfall  $R_{0.99}$ , as a function  
749 of the distance between the validation site and the nearest calibration site. In general,



750 index-flood models  $\mathcal{M}_1$ -GCop and  $\mathcal{M}_2$ -GCop appear less precise than models based on  
751 the regression  $\mathcal{M}_3$ . Also note that model  $\mathcal{M}_2$ -GCop yields very imprecise predictions  
752 (CV>0.3) for three validation sites, which are all located in the same region (South-  
753 Eastern Mediterranean coast).

754 Moreover, an apparent trend suggests that the  $\mathcal{M}_1$ -predictions are more precise for val-  
755 idation sites located close to a calibration site than for isolated sites. For this model,  
756 Figure 15 suggests that if a calibration site is available at about 10 km, the predictive  
757 uncertainty is divided by almost two compared with isolated sites. This confirms an  
758 observation previously made in section 4.5.3, and illustrates a positive outcome of mod-  
759 eling spatial dependence between regression errors. On the other hand, Figure 15 also  
760 shows that when the regression model is improved ( $\mathcal{M}_3$ ), the predictive precision is as  
761 good as model  $\mathcal{M}_1$  nearby calibration sites, despite the fact that spatial dependence be-  
762 tween regression errors is negligible (see section 4.5.2). Moreover,  $\mathcal{M}_3$ -predictions are  
763 more precise for isolated sites. This important observation suggests that improving the  
764 regression to make errors as spatially independent as possible is the most sensible strategy.

765

## 5. Discussion

766 The case study described in section 4 demonstrates the feasibility of a Bayesian hier-  
767 archical approach to regional frequency analysis. Despite this encouraging preliminary  
768 investigation, numerous avenues for improvement can already be identified. This section  
769 discusses the main issues to be addressed.

### 5.1. Treatment of missing data

770 In the case study of section 4, calibration sites were selected in order to avoid missing  
771 values in the calibration dataset. This was made for the sake of simplicity, but it cannot  
772 be considered as an acceptable practical solution. Indeed, most regional analyses make  
773 use of at-site data with different lengths. Restricting the calibration dataset to years  
774 shared by all sites results in a loss of information. Moreover, data from poorly-gauged  
775 sites still yield a substantial information on the at-site distribution. Combining this in-  
776 formation with the regional information transferred from nearby sites is likely to improve  
777 the inference.

778 The treatment of missing data is challenging due to the explicit modeling of spatial de-  
779 pendence between data. Indeed, the likelihood equation (9) requires writing the joint dis-  
780 tribution of a (spatial) vector  $\tilde{\mathbf{y}}(t)$  at any time  $t$ . A first possibility would be to consider  
781 that the size of vector  $\tilde{\mathbf{y}}(t)$  varies with  $t$ , depending on the available data. Unfortunately,  
782 this complicates the implementation of the model and increases its computational cost,  
783 due to the manipulation and inversion of the dependence matrix  $\Sigma$ , whose dimension  
784 would also vary with  $t$  in this case. An alternative and possibly more efficient approach  
785 would be to consider models for missing data and data augmentation algorithms [see e.g.  
786 *Gelman et al.*, 1995].

## 5.2. Sensitivity to the model for data dependence

787 As stressed in section 2.1.2, the use of an elliptical copula to model data dependence  
788 corresponds to a parametric assumption on the structure of dependence. As such, it might  
789 be an inappropriate model for some data, especially in the case of extreme data [*Mikosch*,  
790 2005]. This is a crucial issue if the dependence model is to be used for estimating rare  
791 multivariate events (e.g.,  $\Pr(\text{annual maxima from all raingauges larger than 100 mm})$ ).

792 The literature suggests that such estimations are highly sensitive to the choice of the  
793 dependence model [e.g. *Coles et al.*, 1999; *Renard and Lang*, 2007]. However, the impact  
794 of the dependence model on the estimation at ungauged site is not as clear. Indeed,  
795 prediction at ungauged site does not directly use the data dependence model (see section  
796 3.4). Yet, the latter may indirectly influence predictions when inappropriate, by inducing  
797 bias in parameter estimates and/or inadequate quantification of uncertainties (posterior  
798 variance).

799 The preliminary investigation carried out in section 4.4 suggests that the choice of a  
800 data dependence model is of second order importance compared with the choice of an  
801 appropriate regression model to describe spatial variability. However, this result is based  
802 on a single case study and can therefore not be considered as a generality. Consequently,  
803 the sensitivity of predictions to the data dependence model needs to be further evaluated.  
804 In particular, the impact of the spatial dependence model might be more pronounced for  
805 datasets showing a higher level of asymptotic dependence. Including multivariate extreme  
806 models recently proposed in the literature [e.g. *Keef et al.*, 2009; *Padoan et al.*, 2010] would  
807 also constitute an improvement.

### 5.3. The role of spatial- and cross-dependence between regression errors

808 The framework developed in this paper follows the recent work by *Kjeldsen and Jones*  
809 [2009a] in assuming that regression errors are spatially dependent. The case study of  
810 section 4 illustrates the benefit of this assumption: when the regression model is poor  
811 and fails to capture key relationships with spatial covariates (model  $\mathcal{M}_1$ ), spatial depen-  
812 dence between regression errors is significant and improves the predictive precision nearby  
813 calibration sites. However, this spatial dependence quickly vanishes when the regression

814 model is improved (models  $\mathcal{M}_2$  and  $\mathcal{M}_3$ ). This illustrates the relationship between the  
815 quality of the regression and the existence of spatial dependence in regression errors: in a  
816 nutshell, spatial dependence acts as a surrogate for the spatial variability "missed" by the  
817 regression model. This suggests that improving the regression to make regression errors  
818 as spatially independent as possible is a better strategy than attempting to refine the  
819 description of spatial dependence. However, in cases where no satisfying regression can  
820 be uncovered, the explicit modeling of spatial dependence appears beneficial in terms of  
821 predictive precision.

822 Moreover, cross-correlation between regression errors related to distinct D-parameters is  
823 neglected in this framework. Unfortunately, the case study suggests that such cross-  
824 correlation exists (see section 4.3.1). Its impact on predictions is unclear at this stage and  
825 requires further evaluation.

826 An explicit modeling of cross-correlation was adopted by *Tasker and Stedinger* [1989],  
827 in the case of spatially independent regression errors. This paper somehow considers  
828 the opposite option: spatial dependence between regression errors is modeled, but their  
829 cross-correlation is ignored. Which of these two options should be favored is likely case-  
830 specific, and depends on the relative strength of spatial- and cross-correlation. Moreover,  
831 it is worth noting that both types of dependence can be constrained (to some extent) by  
832 improving the model. In the one hand, improving the regression model decreases spatial  
833 dependence as discussed above. In the other hand, cross-correlation can be decreased by  
834 reparameterizing the marginal distributions: for instance, the GEV distribution could be  
835 parameterized in terms of location, CV and shape parameters (instead of location, scale  
836 and shape, see the discussion in *Stedinger and Griffis* [2011]), which may limit cross-

837 correlation between location and scale regression errors. Which of these two strategies  
838 offers the most flexibility remains an open question, that will be addressed in future work.  
839 Lastly, the third strategy would be to explicitly model both spatial- and cross-correlation  
840 in regression errors. However, this would require a description of the dependence between  
841 spatial random fields, which is a challenging task. Solutions might exist and be borrowed  
842 from the field of geostatistics [e.g. using cokriging tools, see *Goovaerts, 1998*]. However,  
843 improving the model to make either spatial- or cross-correlation negligible seems easier to  
844 implement and might be an acceptable solution in many cases.

#### 5.4. Validation procedures

845 As any predictive statistical model, a model constructed within the hierarchical frame-  
846 work presented in this paper needs to be thoroughly validated based on data that were  
847 not used for estimation. In the case study of section 4, this was achieved by comparing  
848 the predictive distribution (derived in an ungauged site context) with validation data.  
849 Such a comparison assesses the overall reliability of the predictive distribution, and may  
850 detect systematic predictive biases (e.g. the predictive mean is significantly smaller than  
851 the mean of validation data). However, it might not be sufficient to assess the reliability  
852 of predicted extreme quantiles, which are of primary interest in frequency analysis. Con-  
853 sequently, more specialized validation procedures are needed to open predicted extreme  
854 quantiles to direct validation and scrutiny. The validation tools recently proposed by  
855 *Garavaglia et al. [2011]* might be of particular interest in this respect.

#### 5.5. Model complexity and identifiability

856 The case study of section 4 compares several models differing in their flexibility (and,  
857 as a consequence, in their complexity). In particular, models  $\mathcal{M}_1$  and  $\mathcal{M}_2$  are based on  
858 the scale invariance assumption. This limits the complexity of the model by merging all  
859 spatial variability into a single D-parameter (the index-flood parameter  $\delta(s)$ ), the other  
860 D-parameters being assumed constant over the region. However, this assumption seems  
861 unrealistic in the studied region. The alternative model  $\mathcal{M}_3$  describes the spatial varia-  
862 tions in two D-parameters, namely the location and the scale parameters, while the third  
863 D-parameter (shape) remains constant. This is similar to the approach evaluated by *Ste-*  
864 *dinger and Lu* [1995], and it was found to yield reliable and more precise predictions in  
865 this case study. Consequently, it would be tempting to go even further in terms of model  
866 complexity, by using three distinct regression models for the location, scale and shape  
867 parameters.

868 However, as in any statistical model, a trade-off between descriptive and predictive power  
869 has to be found. Model complexity may come at the cost of reduced predictive ability.  
870 Moreover, non-identifiability issues may arise if the information content of the data does  
871 not support the inference of several spatial processes. Establishing an acceptable trade-off  
872 between the flexibility of the model and the level of complexity that can be identified from  
873 the data is a very challenging task. In particular, it is linked with the preceding discus-  
874 sion on validation procedures: introducing additional complexity is only beneficial if it  
875 demonstrably improves predictive reliability and/or precision, yet the power of validation  
876 procedures to detect model failures may be too limited, especially for extreme quantiles.  
877 Consequently, in the case study presented in this paper, developing more efficient valida-  
878 tion tools seems to be a prerequisite before attempting to model spatial variations in the

879 shape parameter.

880 Lastly, it is stressed that the high dimensionality of models resulting from the hierarchi-  
881 cal framework presented herein does not necessarily lead to over-parameterized or non-  
882 identifiable models. In the context of Bayesian hierarchical modeling, the issue of model  
883 complexity is far from obvious [see *Spiegelhalter et al.*, 2002, for a thorough discussion on  
884 this topic] and requires specialized criteria for model selection [e.g. *Cooley et al.*, 2007].

## 5.6. Prior specification

885 The case study of section 4 uses mostly non-informative priors. This is because the aim  
886 of this case study is to understand some properties of the proposed inference approach,  
887 rather than to seek an optimal estimation of rainfall extremes in this particular region.  
888 In this context, using precise priors might exert a strong, case-specific leverage on the  
889 conclusions. However, the specification of precise and accurate priors might indeed be  
890 beneficial to the inference. In particular, the following investigations would be of interest:

891 • Prior specification depends on the particular choices of marginal distributions and  
892 regression functions. It is therefore difficult to derive general specification guidelines.  
893 In particular, the fact that D-parameters are not inferred directly, but are only defined  
894 through the regression function (4) complicates prior specification. Methods for expressing  
895 prior knowledge (either based on expertise or on external data not used for inference) in  
896 the form of a proper prior distribution need to be developed.

897 • Using conjugate priors has the potential to ease computations by replacing (at least  
898 partly) MCMC sampling by explicit formulas. However, the choice of a conjugate fam-  
899 ily depends on the choice of marginal distributions and regression functions, making it  
900 difficult to provide general guidelines. The benefit of using conjugate priors could be

901 first evaluated for the hyper-parameters of the Gaussian hyper-distributions, for which  
902 conjugate families are known [e.g. *Gelman et al.*, 1995].

903 • The sensitivity of the inference to priors should be further evaluated, especially for  
904 complex models that are more prone to identifiability issues (see discussion in section 5.5).

### 5.7. Modeling variability in time

905 Since this paper mainly focuses on spatial variability, at-site distributions were assumed  
906 to be constant in time (see equation (1)). A natural extension of the framework would  
907 be to allow temporal variations in at-site distributions, by including covariates varying  
908 with time. The simplest application would be to include a linear trend in one (or possibly  
909 several) D-parameter(s). The covariate could be in this case the time itself, or a large-scale  
910 climatic index (e.g. NAO or PDO indexes). This would be an important development  
911 since the strong natural variability of at-site data (in particular, extreme data) limits the  
912 detectability of climatic effects. Studying the impact of climate variability/change at a  
913 regional scale is likely to improve this assessment [*Renard et al.*, 2008; *Aryal et al.*, 2009].

### 5.8. Application to runoff or other hydrological variables

914 Although the hierarchical approach presented in this paper can in principle be applied to  
915 runoff data, it is likely to yield sub-optimal predictions if standard dependence models are  
916 used. This is because runoff data are structured by the hydrologic network. Consequently,  
917 euclidean distances (e.g. between gauging stations or between catchment centroids) are  
918 in general a sub-optimal predictor of intersite dependences. Moreover, predictions should  
919 be constrained to ensure the consistency between estimates at upstream and downstream  
920 nested catchments. Several authors proposed specialized approaches to model network



921 dependences [e.g. *Gottschalk*, 1993a, b; *Sauquet et al.*, 2000; *Sauquet*, 2006; *Skoien et al.*,  
922 2006]. The applicability of such approaches within the Bayesian hierarchical framework  
923 presented in this paper will be evaluated in future work.

924 More generally, the choice of a relevant distance to explain spatial dependences is a chal-  
925 lenging task, but is also a promising avenue to improve predictions. This choice heavily  
926 depends on the spatial properties of the hydrological variable. For rainfall, a xy-Euclidean  
927 distance (as used in the case study of section 4) may seem reasonable at first sight, al-  
928 though accounting for elevation may improve predictions. Alternatively, consider apply-  
929 ing the hierarchical approach presented in this paper to snow depth values: a meaningful  
930 dependence-distance relationship is very unlikely to be derived if elevation is neglected in  
931 the definition of the inter-site distance. Recent work by *Blanchet et al.* [*Blanchet et al.*,  
932 2009; *Blanchet and Lehning*, 2010; *Blanchet and Davison*, 2011] addresses this issue, and  
933 proposes practical solutions to derive meaningful distances. Those solutions might be  
934 extended to hydrological variables other than snow depth.

## 6. Conclusion

935 Regional frequency analysis has a long history in Hydrology, and is widely applied  
936 in practice to estimate the distribution of a hydrologic variable at ungauged or poorly  
937 gauged sites. However, despite numerous methodological developments over the years  
938 [e.g. *Stedinger*, 1983; *Stedinger and Tasker*, 1985, 1986; *Hosking and Wallis*, 1988; *Mad-  
939 sen and Rosbjerg*, 1997; *Reis et al.*, 2005; *Ribatet et al.*, 2007; *Kjeldsen and Jones*, 2009a],  
940 most implementations still rely on several hypotheses that complicate the quantification  
941 of predictive uncertainty; moreover, unrealistic hypotheses may yield unreliable predic-  
942 tions. The objective of this paper was therefore to propose a general Bayesian hierarchical

943 framework to overcome some of these limitations, and to evaluate its applicability based  
944 on a case study. This framework builds on previous work by several authors [in particular,  
945 *Wikle et al.*, 1998; *Perreault*, 2000; *Micevski et al.*, 2006; *Micevski*, 2007; *Cooley et al.*,  
946 2007; *Lima and Lall*, 2009; *Aryal et al.*, 2009; *Lima and Lall*, 2010], who explored the  
947 usefulness of spatial and temporal Bayesian hierarchical models. The main features of the  
948 proposed framework are the following:

- 949 1. At-site data can be modeled with any distribution, with the sole requirement that  
950 the same distribution (but with different parameters) is used for all sites.
- 951 2. Intersite dependence is explicitly modeled by means of an elliptical copula.
- 952 3. The variation of parameters in space is described with a regression model linking  
953 parameter values and covariates.
- 954 4. Regression errors are modeled by means of a Gaussian spatial field, which allows  
955 transferring estimates from gauged to ungauged sites, while quantifying the associated  
956 predictive uncertainty.

957 A case study based on extreme rainfall data in Mediterranean France demonstrated the  
958 applicability of the framework and its reliable estimation of predictive uncertainty. Al-  
959 though numerous improvements remain to be implemented (see discussion section), these  
960 encouraging results warrant further research to develop this framework. In particular,  
961 the inclusion of temporal covariates would allow describing the evolution of hydrologic  
962 variables in both space and time. Importantly, this improved understanding of hydrologic  
963 variability would be achieved with a rigorous quantification of the associated uncertainties.

## Appendix A: The Gaussian and Student Elliptical copulas

964 The following notation is used:

965 1.  $F_1(y), \dots, F_H(y)$  are the marginal cdfs of a  $H$ -dimensional random vector  $(Y_1, \dots, Y_H)$ .

966 2.  $f_1(y), \dots, f_H(y)$  are the corresponding marginal pdfs.

967 3.  $\phi(y)$  is the cdf of the standard normal distribution.

968 4.  $\psi(y)$  is the corresponding pdf.

969 5.  $\tau_\nu(y)$  is the cdf of the Student's  $t$  distribution with  $\nu$  degrees of freedom

970 6.  $t_\nu(y)$  is the corresponding pdf

971 7.  $\Phi_\Sigma(y_1, \dots, y_H)$  is the joint cdf of a  $H$ -dimensional Gaussian distribution, with mean

972 0 and covariance matrix  $\Sigma$

973 8.  $\Psi_\Sigma(y_1, \dots, y_H)$  is the corresponding joint pdf

974 9.  $\Theta_{\Sigma, \nu}(y_1, \dots, y_H)$  is the joint cdf of a  $H$ -dimensional Student distribution, with mean

975 0, covariance matrix  $\Sigma$  and  $\nu$  degrees of freedom

976 10.  $T_{\Sigma, \nu}(y_1, \dots, y_H)$  is the corresponding joint pdf

977 The Gaussian and the Student copulas build the joint cdf of the random vector

978  $(Y_1, \dots, Y_H)$  as follows:

$$F_{Gaussian}(y_1, \dots, y_H) = \Phi_\Sigma(u_1, \dots, u_H) \tag{A1}$$

$$F_{Student}(y_1, \dots, y_H) = \Theta_{\Sigma, \nu}(v_1, \dots, v_H) \tag{A2}$$

979 where  $u_i = \phi^{-1}(F_i(y_i))$  and  $v_i = \tau_\nu^{-1}(F_i(y_i))$ .

980 The corresponding joint pdfs can be obtained by differentiating the above cdfs:

$$f_{Gaussian}(y_1, \dots, y_H) = \frac{(\prod_{i=1}^H f_i(y_i))}{(\prod_{i=1}^H \psi(u_i))} \times \Psi_\Sigma(u_1, \dots, u_H) \tag{A3}$$

$$f_{Student}(y_1, \dots, y_H) = \frac{(\prod_{i=1}^H f_i(y_i))}{(\prod_{i=1}^H t_\nu(v_i))} \times T_{\Sigma, \nu}(v_1, \dots, v_H) \quad (\text{A4})$$

981 **Acknowledgments.** The suggestions by Jerry Stedinger, two anonymous reviewers and  
982 the associate editor significantly improved the content of this paper and are gratefully  
983 acknowledged. This work is funded by the ANR research project ExtraFlo. Meteo France  
984 is gratefully acknowledged for providing the data. Many thanks to Luc Neppel for his  
985 help.

## References

- 986 Aryal, S. K., B. C. Bates, E. P. Campbell, Y. Li, M. J. Palmer, and N. R. Viney (2009),  
987 Characterizing and modeling temporal and spatial trends in rainfall extremes, *Journal*  
988 *of Hydrometeorology*, 10(1), 241–253, doi:10.1175/2008JHM1007.1.
- 989 Blanchet, J., and A. C. Davison (2011), Spatial modelling of extreme snow depth, *Annals*  
990 *of Applied Statistics. In Press.*
- 991 Blanchet, J., and M. Lehning (2010), Mapping snow depth return levels: smooth spatial  
992 modeling versus station interpolation, *Hydrology and Earth System Sciences*, 14(12),  
993 2527–2544, doi:10.5194/hess-14-2527-2010.
- 994 Blanchet, J., C. Marty, and M. Lehning (2009), Extreme value statistics of snowfall in the  
995 swiss alpine region, *Water Resources Research*, 45.
- 996 Burn, D. (1990), Evaluation of regional flood frequency analysis with a region of influence  
997 approach, *Water Resources Research*, 26(10), 2257–2265.
- 998 Casson, E., and S. Coles (1998), Extreme hurricane wind speeds: Estimation, extrapola-  
999 tion and spatial smoothing, *Journal of Wind Engineering and Industrial Aerodynamics*,  
1000 74-76, 131–140.

- 1001 Casson, E., and S. Coles (2000), Simulation and extremal analysis of hurricane events,  
1002 *Applied Statistics*, 49, 227–245.
- 1003 Chiles, J.-P., and P. Delfiner (1999), *Geostatistics: Modeling Spatial Uncertainty*, Wiley  
1004 Series in Probability and Statistics.
- 1005 Clark, J. (2005), Why environmental scientists are becoming bayesians, *Ecology letters*,  
1006 8, 2–14.
- 1007 Clark, J. S. (2003), Uncertainty and variability in demography and population growth: A  
1008 hierarchical approach, *Ecology*, 84(6), 1370–1381.
- 1009 Coles, S., and E. Casson (1998), Extreme value modelling of hurricane wind speeds,  
1010 *Structural Safety*, 20, 283–296.
- 1011 Coles, S., J. Heffernan, and J. A. Tawn (1999), Dependence measures for extreme value  
1012 analyses, *Extremes*, 2, 339–365.
- 1013 Cooley, D., D. Nychka, and P. Naveau (2007), Bayesian spatial modeling of extreme  
1014 precipitation return levels, *Journal of the American Statistical Association*, 102(479),  
1015 824–840.
- 1016 Crainiceanu, C. M., J. R. Stedinger, D. Ruppert, and C. T. Behr (2003), Modeling the us  
1017 national distribution of waterborne pathogen concentrations with application to cryp-  
1018 tosporidium parvum, *Water Resources Research*, 39(9).
- 1019 Dalrymple, T. (1960), Flood frequency analysis, *U.S. Geol. Surv. Water Supply*, 1543A.
- 1020 Dawid, A. P. (1984), Statistical-theory - the prequential approach, *Journal of the Royal*  
1021 *Statistical Society Series a-Statistics in Society*, 147, 278–292.
- 1022 Dobson, A. (2001), *An Introduction to Generalised Linear Models*, Texts in statistical  
1023 science series.

- 1024 Favre, A. C., S. El Adlouni, L. Perreault, N. Thiémonge, and B. Bobee (2004), Multivari-  
1025 ate hydrological frequency analysis using copulas, *Water Resources Research*, *40*(1).
- 1026 Fill, H. D., and J. R. Stedinger (1998), Using regional regression within index flood  
1027 procedures and an empirical bayesian estimator, *Journal of Hydrology*, *210*(1-4), 128 –  
1028 145, doi:10.1016/S0022-1694(98)00177-2.
- 1029 Garavaglia, F., M. Lang, E. Paquet, J. Gailhard, R. Garçon, and B. Renard (2011), Relia-  
1030 bility and robustness of rainfall compound distribution model based on weather pattern  
1031 sub-sampling, *Hydrology and Earth System Sciences*, *15*(2), 519–532, doi:10.5194/hess-  
1032 15-519-2011.
- 1033 Gelman, A., J. Carlin, H. Stern, and D. Rubin (1995), *Bayesian data analysis*, Texts in  
1034 Statistical Science, Chapman and Hall.
- 1035 Genest, C., and A. C. Favre (2007), Metaelliptical copulas and their use in frequency  
1036 analysis of multivariate hydrological data, *Water Resources Research*, *43*.
- 1037 Gneiting, T., F. Balabdaoui, and A. E. Raftery (2007), Probabilistic forecasts, calibration  
1038 and sharpness, *Journal of the Royal Statistical Society Series B-Statistical Methodology*,  
1039 *69*, 243–268.
- 1040 Goovaerts, P. (1998), Ordinary cokriging revisited, *Mathematical geology*, *30*(1), 21–42.
- 1041 Gottschalk, L. (1993a), Correlation and covariance of runoff, *Stochastic Hydrology and*  
1042 *Hydraulics*, *7*, 85–101, doi:10.1007/BF01581418.
- 1043 Gottschalk, L. (1993b), Interpolation of runoff applying objective methods, *Stochastic*  
1044 *Hydrology and Hydraulics*, *7*, 269–281, doi:10.1007/BF01581615.
- 1045 Griffis, V., and J. Stedinger (2007), Evolution of flood frequency analysis with bulletin  
1046 *17*, *JOURNAL OF HYDROLOGIC ENGINEERING*, *12*, 283–297.

- 1047 Hosking, J., and J. R. Wallis (1988), The effect of intersite dependence on regional flood  
1048 frequency analysis, *Water Resources Research*, *24*, 588–600.
- 1049 Hosking, J., and J. R. Wallis (1997), *Regional Frequency Analysis: an approach based on*  
1050 *L-Moments*, Cambridge University Press, Cambridge, UK.
- 1051 Katz, R. W., M. B. Parlange, and P. Naveau (2002), Statistics of extremes in hydrology,  
1052 *Advances in Water Resources*, *25*(8-12), 1287 – 1304, doi:10.1016/S0309-1708(02)00056-  
1053 8.
- 1054 Keef, C., C. Svensson, and J. A. Tawn (2009), Spatial dependence in extreme river  
1055 flows and precipitation for great britain, *Journal of Hydrology. To appear.*, doi:  
1056 10.1016/j.jhydrol.2009.09.026.
- 1057 Kjeldsen, T. R., and D. A. Jones (2009a), An exploratory analysis of error components  
1058 in hydrological regression modeling, *Water Resources Research*, *45*.
- 1059 Kjeldsen, T. R., and D. A. Jones (2009b), A formal statistical model for pooled analysis  
1060 of extreme floods, *Hydrology research*, *40*(5), 465–480.
- 1061 Kroll, C. N., and J. R. Stedinger (1998), Regional hydrologic analysis: Ordinary and  
1062 generalized least squares revisited, *Water Resources Research*, *34*(1), 121–128.
- 1063 Kuczera, G. (1982a), Combining site-specific and regional information: An empirical bayes  
1064 approach, *Water Resources Research*, *18*(2), 306–314.
- 1065 Kuczera, G. (1982b), Robust flood frequency models, *Water Resources Research*, *18*(2),  
1066 315–324.
- 1067 Kuczera, G. (1983), Effect of sampling uncertainty and spatial correlation on an empirical  
1068 bayes procedure for combining site and regional information, *Journal of Hydrology*, *65*,  
1069 373–398.

- 1070 Laio, F., and S. Tamea (2007), Verification tools for probabilistic forecasts of continuous  
1071 hydrological variables, *Hydrology and Earth System Sciences*, 11(4), 1267–1277.
- 1072 Lima, C. H. R., and U. Lall (2009), Hierarchical bayesian modeling of multisite daily  
1073 rainfall occurrence: Rainy season onset, peak, and end, *Water Resources Research.*, 45.
- 1074 Lima, C. H. R., and U. Lall (2010), Spatial scaling in a changing climate: A hierarchical  
1075 bayesian model for non-stationary multi-site annual maximum and monthly streamflow,  
1076 *Journal of Hydrology*, 383(3-4), 307–318.
- 1077 Madsen, H., and D. Rosbjerg (1997), The partial duration series method in regional  
1078 index-flood modeling, *Water Resources Research*, 33(4), 737–746.
- 1079 Maraun, D., T. Osborn, and H. Rust (2009), The influence of synoptic airflow on uk  
1080 daily precipitation extremes. part i: Observed spatio-temporal relationships, *Climate*  
1081 *Dynamics*, pp. 1–15, doi:10.1007/s00382-009-0710-9.
- 1082 Martins, E., and J. Stedinger (2000), Generalized maximum likelihood gev quantile esti-  
1083 mators for hydrologic data, *Water Resources Research*, 36(3), 737–744.
- 1084 Metropolis, N., and S. Ulam (1949), The monte carlo method, *Journal of the American*  
1085 *Statistical Association*, 44, 335–341.
- 1086 Metropolis, N., A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller (1953), Equation  
1087 of state calculations by fast computing machines, *Journal of chemical physics*, 21, 1087–  
1088 1092.
- 1089 Micevski, T. (2007), Nonhomogeneity in eastern australian flood frequency data: Identi-  
1090 fication and regionalisation, Ph.D. thesis, University of Newcastle, Australia.
- 1091 Micevski, T., and G. Kuczera (2009), Combining site and regional flood information using  
1092 a bayesian monte carlo approach, *Water Resources Research*, 45.



- 1093 Micevski, T., G. Kuczera, and S. Franks (2006), A bayesian hierarchical regional flood  
1094 model, in *Proc. 30th Hydrology and Water Resources Symposium*, edited by E. Australia,  
1095 Launceston, Tasmania, Australia.
- 1096 Mikosch, T. (2005), How to model multivariate extremes if one must?, *Statistica Neer-*  
1097 *landica*, 59(3), 324–338.
- 1098 Neppel, L., P. Arnaud, and J. Lavabre (2007), Extreme rainfall mapping: Comparison be-  
1099 tween two approaches in the mediterranean area, *Comptes Rendus Geoscience*, 339(13),  
1100 820–830, doi:10.1016/j.crte.2007.09.013.
- 1101 Ouarda, T., C. Girard, G. Cavadias, and B. Bobee (2001), Regional flood frequency  
1102 estimation with canonical correlation analysis, *Journal of Hydrology*, 254(1-4), 157–  
1103 173.
- 1104 Padoan, S., M. Ribatet, and S. Sisson (2010), Likelihood-based inference for max-stable  
1105 processes, *Journal of the American Statistical Association*, 105, 263–277.
- 1106 Perreault, L. (2000), Analyse bayesienne retrospective d’une rupture dans les sequences  
1107 de variables aleatoires hydrologiques, Ph.D. thesis, ENGREF, INRS-Eau.
- 1108 Pujol, N., L. Neppel, and R. Sabatier (2007), Regional tests for trend detection in max-  
1109 imum precipitation series in the french mediterranean region, *Hydrological Sciences*  
1110 *Journal-Journal Des Sciences Hydrologiques*, 52(5), 956–973.
- 1111 Reichert, P., and J. Mieleitner (2009), Analyzing input and structural uncertainty of non-  
1112 linear dynamic models with stochastic, time-dependent parameters, *Water Resources*  
1113 *Research*, 45.
- 1114 Reis, D. S., J. R. Stedinger, and E. S. Martins (2005), Bayesian generalized least squares  
1115 regression with application to log pearson type 3 regional skew estimation, *Water Re-*

- 1116 *sources Research*, 41(10).
- 1117 Renard, B., and M. Lang (2007), Use of a gaussian copula for multivariate extreme value  
1118 analysis: some case studies in hydrology, *Advances in Water Resources*, 30(4), 897–912.
- 1119 Renard, B., V. Garreta, and M. Lang (2006), An application of bayesian analysis and  
1120 mcmc methods to the estimation of a regional trend in annual maxima, *Water Resources*  
1121 *Research*, 42(12).
- 1122 Renard, B., et al. (2008), Regional methods for trend detection: assessing field significance  
1123 and trend consistency, *Water Resources Research.*, 44.
- 1124 Ribatet, M., E. Sauquet, J.-M. Gresillon, and T. Ouarda (2007), A regional bayesian pot  
1125 model for flood frequency analysis, *Stochastic environmental research and risk assess-*  
1126 *ment*, 21(4), 327–339.
- 1127 Robson, A., and D. Reed (1999), *Flood Estimation Handbook. Volume 3: Statistical pro-*  
1128 *cedures for flood frequency estimation*, vol. 3, Wallingford.
- 1129 Sauquet, E. (2006), Mapping mean annual river discharges: Geostatistical developments  
1130 for incorporating river network dependencies, *Journal of Hydrology*, 331(1-2), 300 –  
1131 314, doi:10.1016/j.jhydrol.2006.05.018.
- 1132 Sauquet, E., L. Gottschalk, and E. Leblois (2000), Mapping average annual runoff: a  
1133 hierarchical approach applying a stochastic interpolation scheme, *Hydrological sciences*  
1134 *journal*, 45(6).
- 1135 Skoien, J. O., R. Merz, and G. Blöchl (2006), Top-kriging - geostatistics on stream net-  
1136 works, *Hydrology and Earth System Sciences*, 10(2), 277–287, doi:10.5194/hess-10-277-  
1137 2006.

- 1138 Spiegelhalter, D. J., N. G. Best, B. R. Carlin, and A. van der Linde (2002), Bayesian  
1139 measures of model complexity and fit, *Journal of the Royal Statistical Society Series*  
1140 *B-Statistical Methodology*, *64*, 583–616.
- 1141 Stedinger, J., and L. Lu (1995), Appraisal of regional and index flood quantile estimators,  
1142 *Stochastic Hydrology and Hydraulics*, *9*(1), 49–75.
- 1143 Stedinger, J. R. (1983), Estimating a regional flood frequency distribution, *Water Re-*  
1144 *sources Research*, *19*, 503–510.
- 1145 Stedinger, J. R., and V. W. Griffis (2011), Getting from here to where? flood frequency  
1146 analysis and climate, *Journal of the American Water Resources Association (JAWRA)*,  
1147 *47*(3), 506–513, doi:10.1111/j.1752-1688.2011.00545.
- 1148 Stedinger, J. R., and G. D. Tasker (1985), Regional hydrologic analysis: 1. ordinary,  
1149 weighted and generalized least squares compared, *Water Resources Research*, *21*(9),  
1150 14211432 [Correction, *Water Resour. Res.*,22(5), 844, 1986.].
- 1151 Stedinger, J. R., and G. D. Tasker (1986), Regional hydrologic analysis: 2. model-error  
1152 estimators, estimation of sigma and log-pearson type 3 distributions, *Water Resources*  
1153 *Research*, *22*(10), 1487 1499.
- 1154 Storz, J. F., and M. A. Beaumont (2002), Testing for genetic evidence of population  
1155 expansion and contraction: An empirical analysis of microsatellite dna variation using  
1156 a hierarchical bayesian model, *Evolution*, *56*(1), 154–166.
- 1157 Tasker, G. D., and J. R. Stedinger (1989), An operational gls model for hydrologic regres-  
1158 sion, *Journal of Hydrology*, *111*, 361:375.
- 1159 Thyer, M., B. Renard, D. Kavetski, G. Kuczera, S. Franks, and S. Srikanthan (2009),  
1160 Critical evaluation of parameter consistency and predictive uncertainty in hydrological

1161 modelling: a case study using bayesian total error analysis, *Water Resources Research*,  
1162 45, doi:10.1029/2008WR006825.

1163 Vrugt, J. A., C. J. F. ter Braak, M. P. Clark, J. M. Hyman, and B. A. Robinson (2008),  
1164 Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with  
1165 markov chain monte carlo simulation, *Water Resour. Res.*, 44.

1166 Wikle, C. K., L. M. Berliner, and N. Cressie (1998), Hierarchical bayesian space-time mod-  
1167 els, *Environmental and Ecological Statistics*, 5, 117–154, doi:10.1023/A:1009662704779.

**Figure 1.** Schematic of the hierarchical modeling framework

**Figure 2.** Raingauges location. The thin lines represent six homogeneous regions defined by *Pujol et al.* [2007].

**Figure 3.** Posterior pdfs of some inferred quantities: (a) regional shape parameter  $\xi$ ; (b) errors  $\epsilon_\mu$  in the regression for the location D-parameter (only 7 distributions are shown for readability); (c) errors  $\epsilon_\lambda$  in the regression for the scale D-parameter (d) hyper-standard deviations  $\sigma_\mu$  (solid black line) and  $\sigma_\lambda$  (dashed red line), corresponding to the standard deviations of location/scale regression errors.

**Figure 4.** Estimated dependence structures. (a) Data dependence. The dots represent estimated pairwise dependences for all available pairs of sites; (b) dependence in regression errors  $\epsilon_\mu$  for the location D-parameter; (c) dependence in regression errors  $\epsilon_\lambda$  for the scale D-parameter.

**Figure 5.** Estimation of the 100-year daily rainfall  $R_{0.99}$ . Left:  $R_{0.99}$  point-estimates (posterior median); Right: uncertainty in estimating  $R_{0.99}$  (measured by the posterior coefficient of variation).

**Figure 6.** Posterior pdfs of some inferred quantities with three distinct models for data dependence. (a) Regional shape parameter  $\xi$ ; (b) hyper-standard deviations  $\sigma_\mu$  (standard deviations of location regression errors); (c) hyper-standard deviations  $\sigma_\lambda$  (standard deviations of scale regression errors).

**Figure 7.** Estimated dependence structures with three distinct models for data dependence. Solid black line =  $\mathcal{M}_3$ -Inde; dashed red line =  $\mathcal{M}_3$ -GCop; dotted blue line =  $\mathcal{M}_3$ -SCop. Shaded areas represent 90% posterior intervals. (a) Data dependence; (b) dependence in regression errors  $\epsilon_\mu$  for the location D-parameter; (c) dependence in regression errors  $\epsilon_\lambda$  for the scale D-parameter.

**Figure 8.** Impact of the data dependence model on predictions of  $R_{0.99}$ . Maps show relative differences with the benchmark model  $\mathcal{M}_3$ -GCop. Top panels =  $\mathcal{M}_3$ -Inde; bottom panels =  $\mathcal{M}_3$ -SCop. Left =  $R_{0.99}$  point-estimates (posterior median); Right = uncertainty in estimating  $R_{0.99}$  (measured by the posterior coefficient of variation).

**Figure 9.** Posterior pdfs of some inferred quantities with three distinct regression models. (a) Regional shape parameter  $\xi$ ; (b) hyper-standard deviations  $\sigma_\mu$  (standard deviations of location regression errors); (c) hyper-standard deviations  $\sigma_\lambda$  (standard deviations of scale regression errors).

**Figure 10.** Estimated dependence structures with three distinct regression models. Dotted blue line =  $\mathcal{M}_1$ -GCop; solid black line =  $\mathcal{M}_2$ -GCop; dashed red line =  $\mathcal{M}_3$ -GCop. Shaded areas represent 90% posterior intervals. (a) Data dependence; (b) dependence in regression errors  $\epsilon_\mu$  for the location D-parameter; (c) dependence in regression errors  $\epsilon_\lambda$  for the scale D-parameter.

**Figure 11.** Impact of the regression model on predictions of  $R_{0.99}$ . Maps show relative differences with the benchmark model  $\mathcal{M}_3$ -GCop. Top panels =  $\mathcal{M}_1$ -GCop; bottom panels =  $\mathcal{M}_2$ -GCop. Left =  $R_{0.99}$  point-estimates (posterior median); Right = uncertainty in estimating  $R_{0.99}$  (measured by the posterior coefficient of variation).

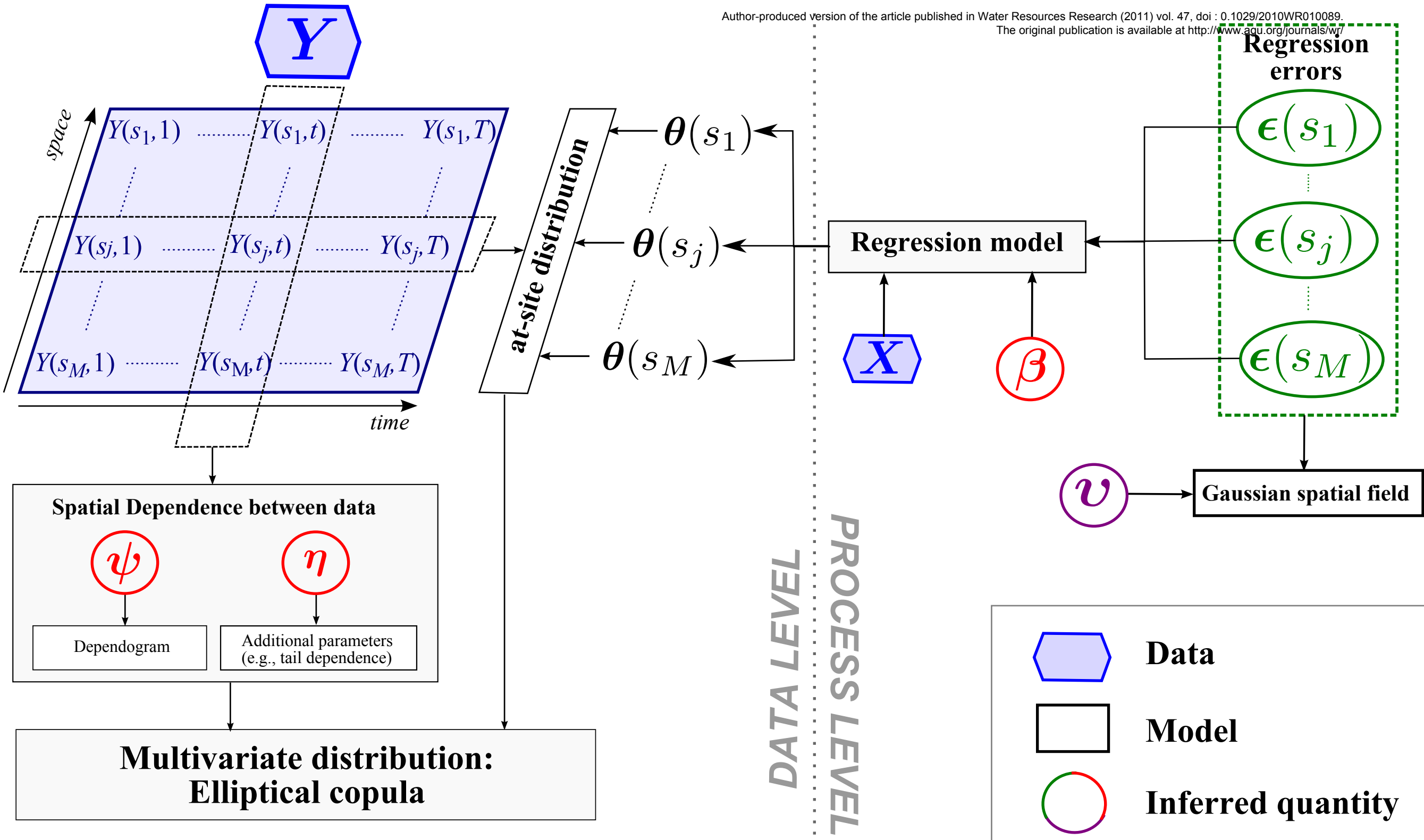
**Figure 12.** Evaluation of the scale invariance hypothesis. Points = empirical coefficients of variation of data; dashed horizontal lines = coefficient of variation estimated from  $\mathcal{M}_2$ -GCop (90% posterior interval); vertical bars = coefficient of variation estimated from  $\mathcal{M}_3$ -GCop (90% posterior intervals).

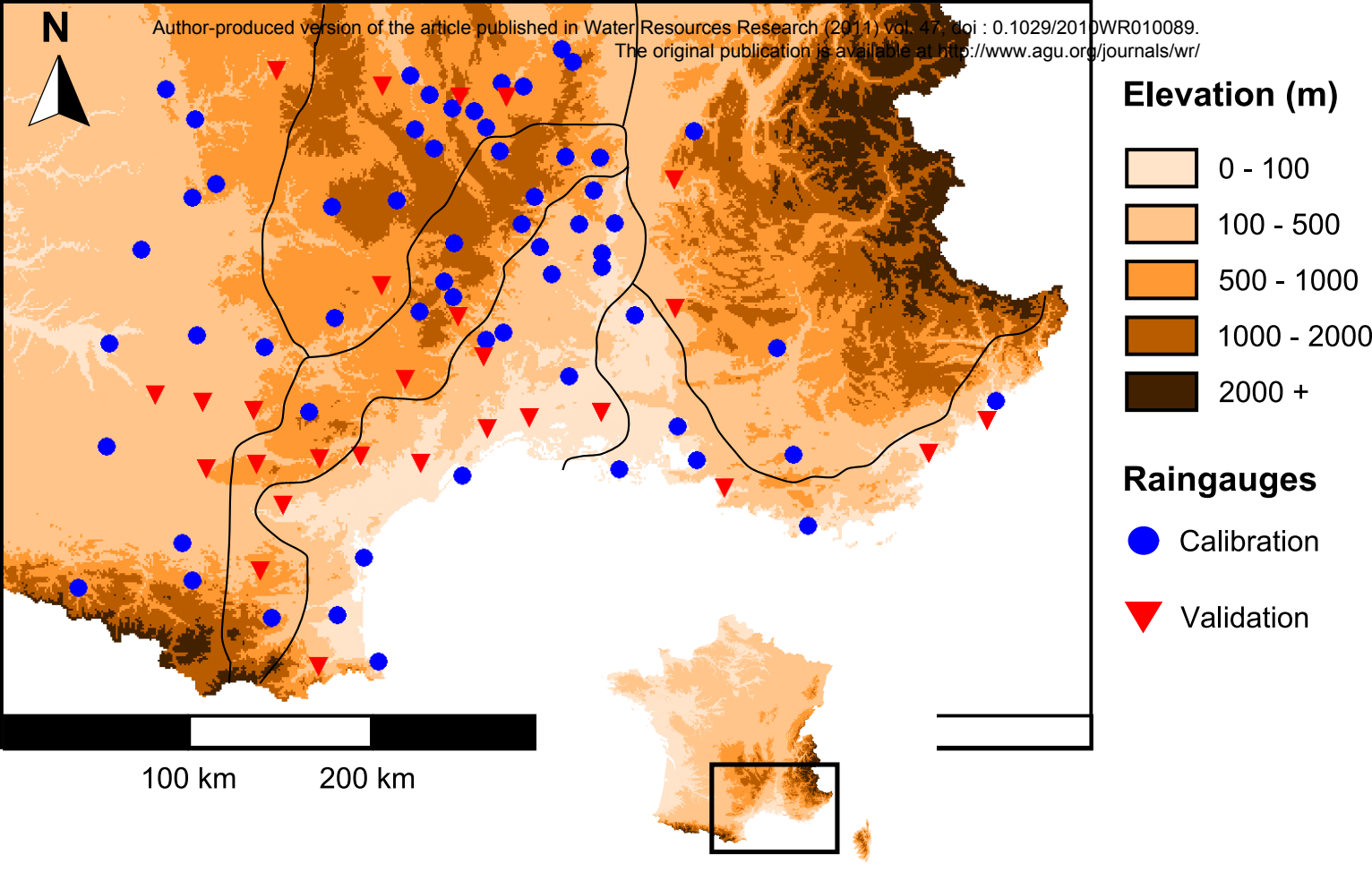
**Figure 13.** Predictive QQ-plots at validation sites for each of the five studied models. The lower right panel corresponds to merging data from all validation sites together.

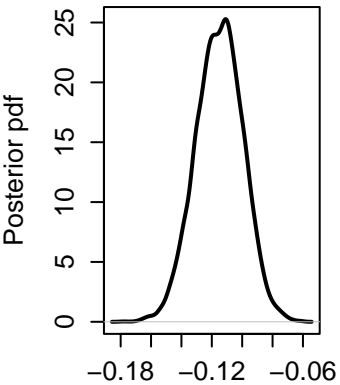
**Figure 14.** Comparison of predictive distributions obtained at one particular validation site.

**Figure 15.** Posterior coefficients of variation of  $R_{0.99}$  estimated at validation sites, as a function of the distance between the validation site and the nearest calibration site.

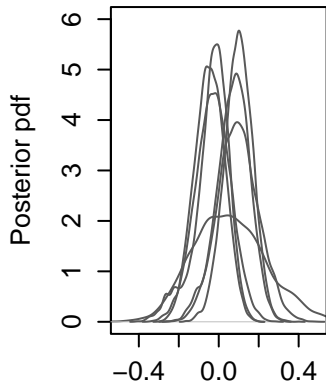




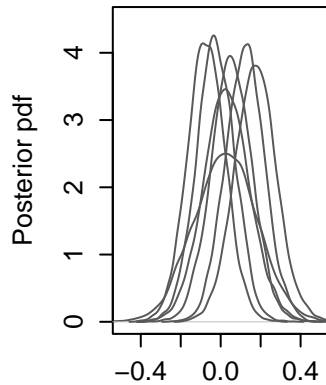




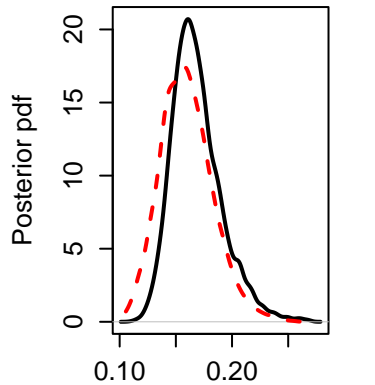
Shape  
(a)



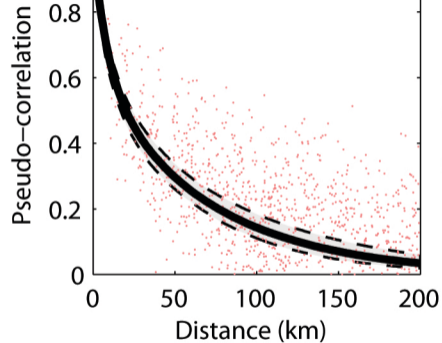
Regression errors (location)  
(b)



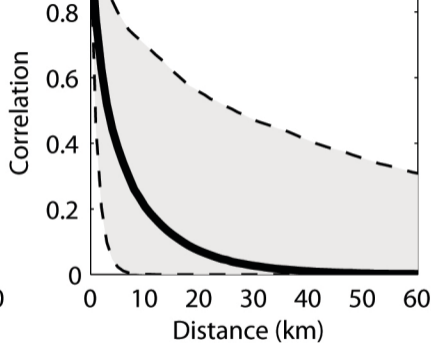
Regression errors (scale)  
(c)



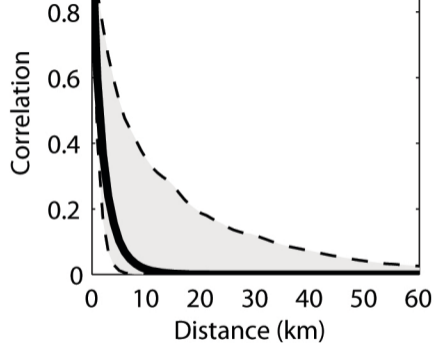
Hyper-standard deviation  
(d)



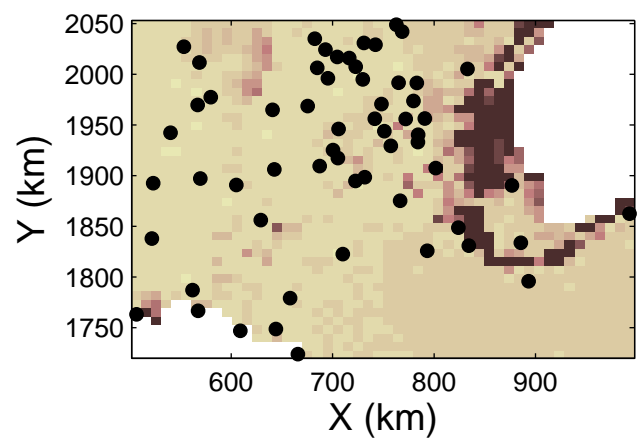
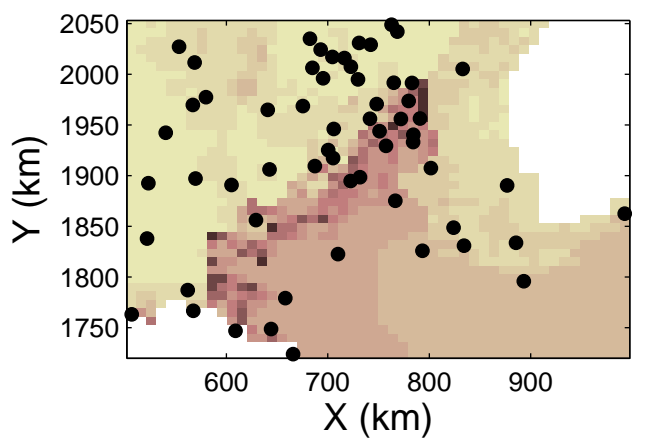
(a)



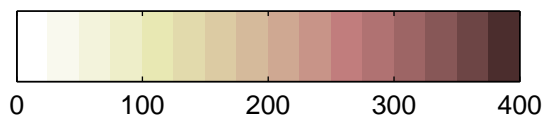
(b)



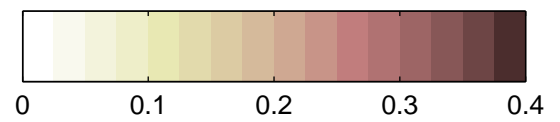
(c)

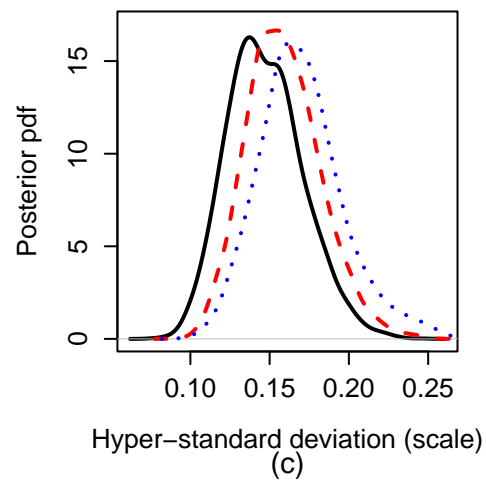
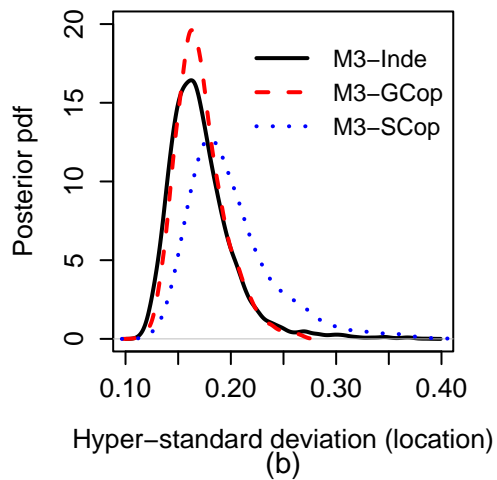
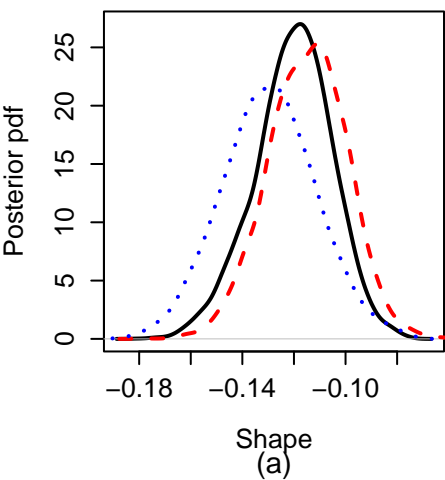


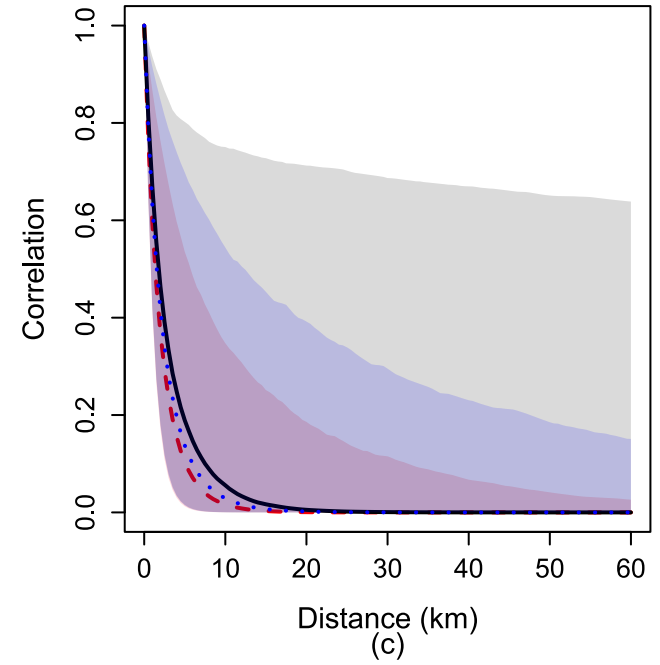
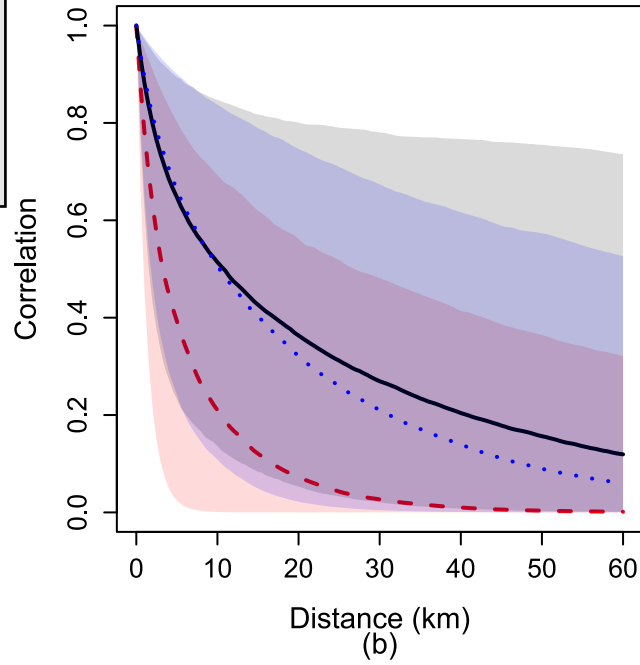
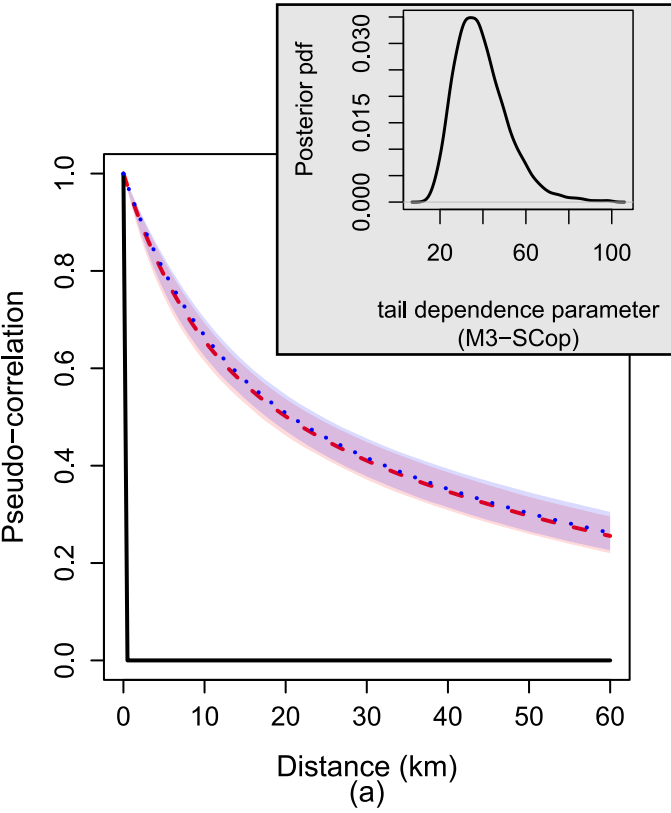
$R_{0.99}$  (mm)

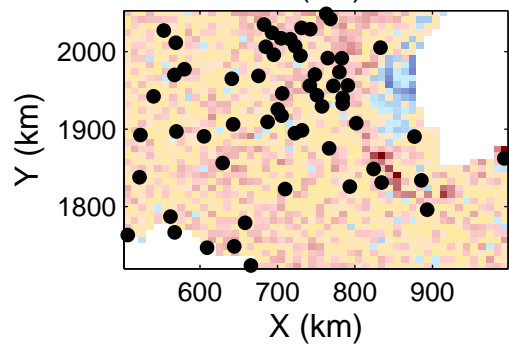
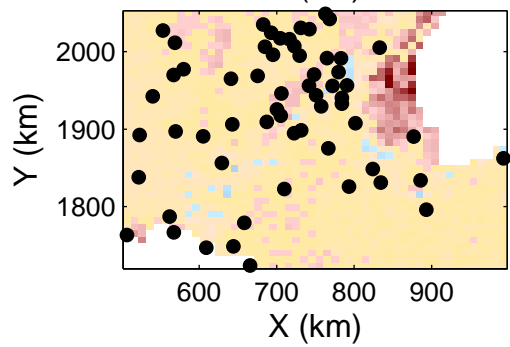
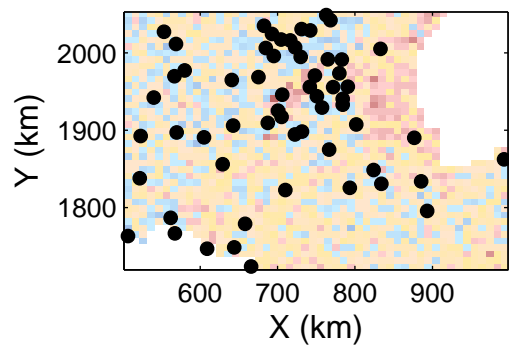
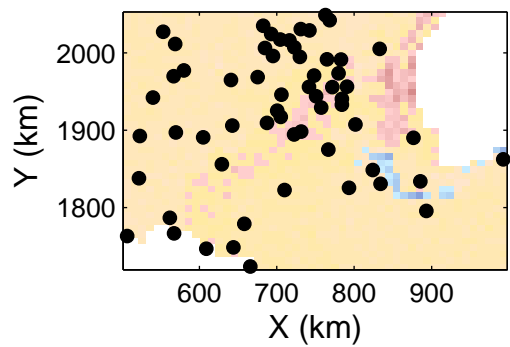


Posterior CV of  $R_{0.99}$

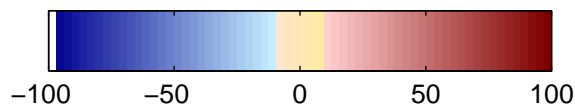






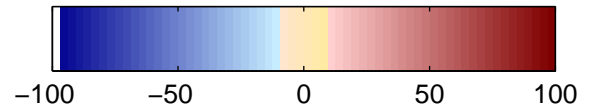


**Relative change in  $R_{0.99}$  (%)**

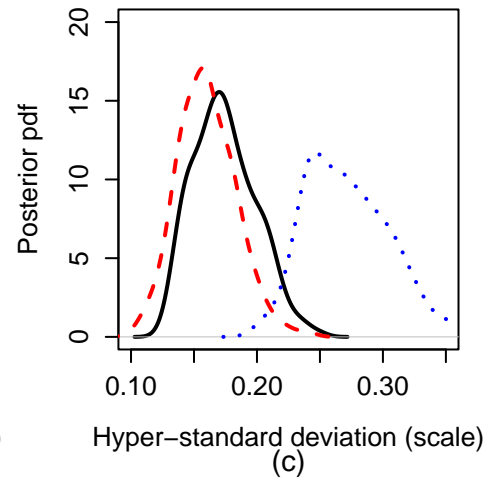
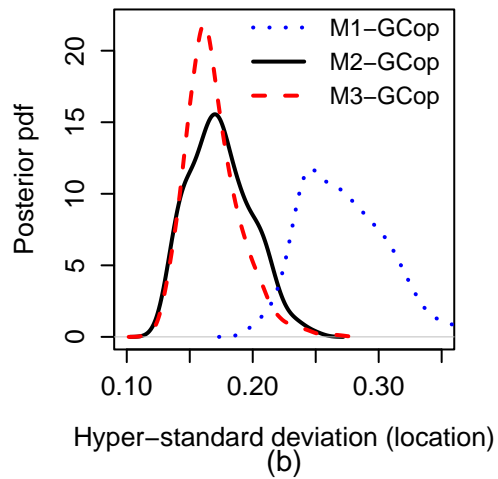
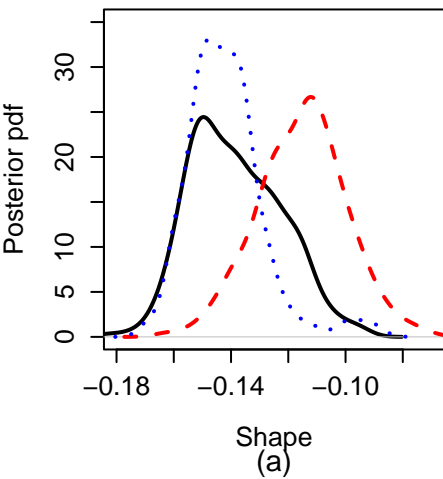


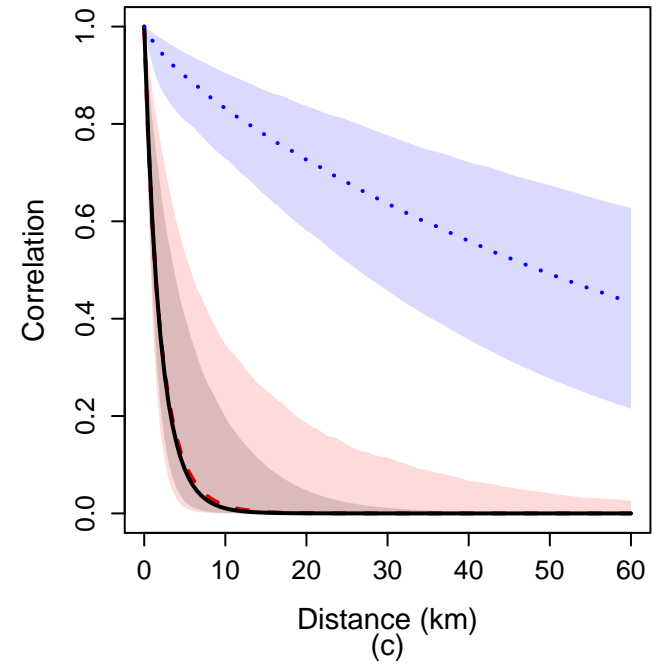
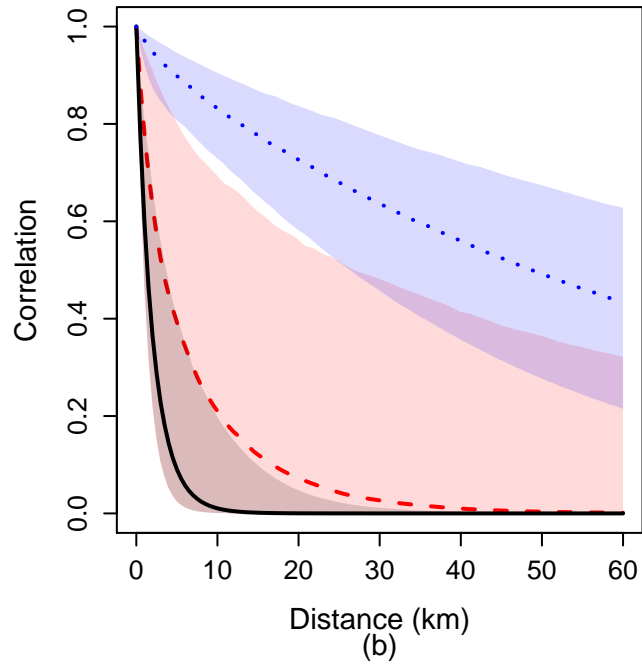
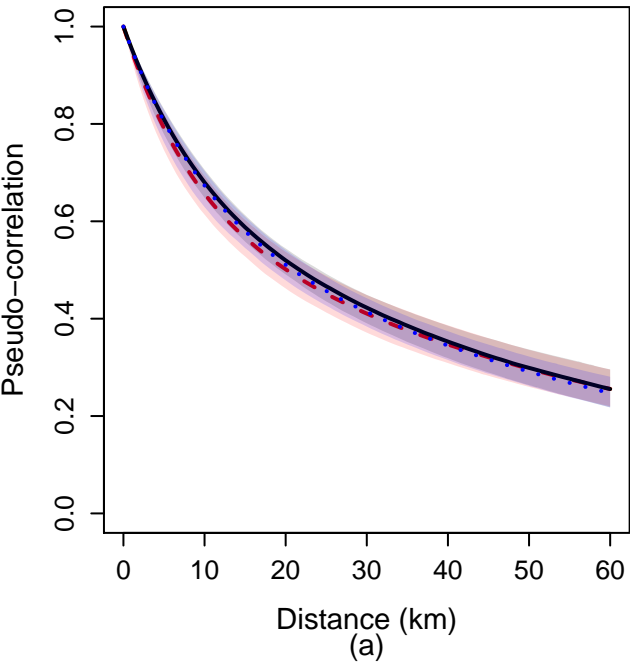
**Relative change in**

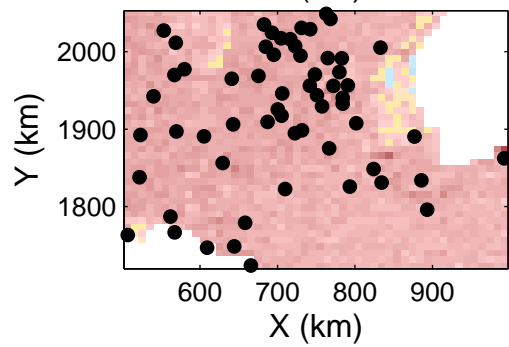
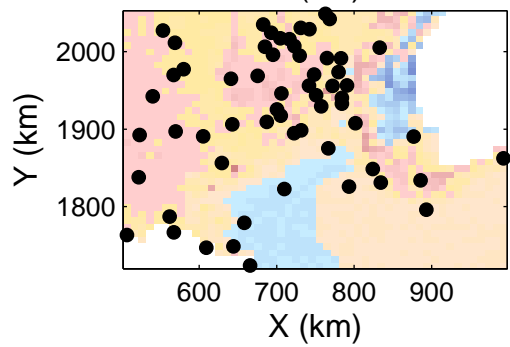
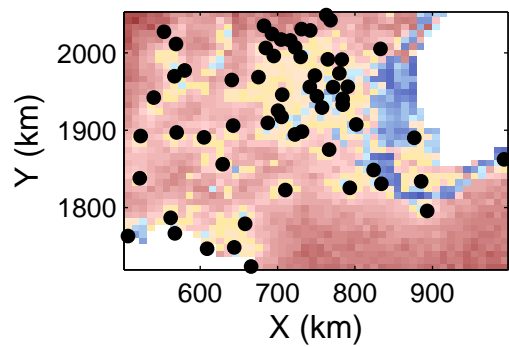
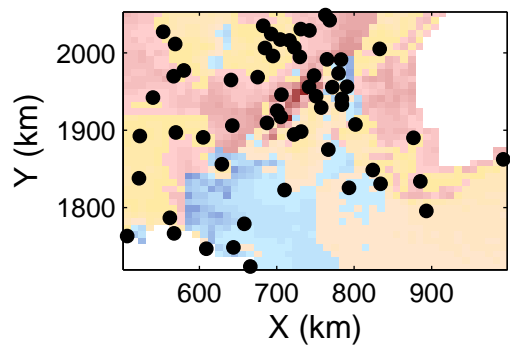
**$R_{0.99}$  uncertainty (%)**



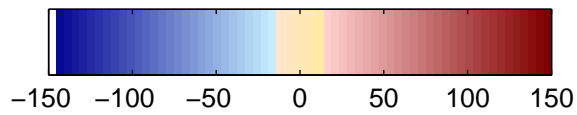






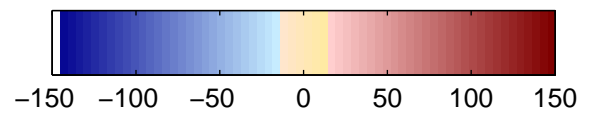


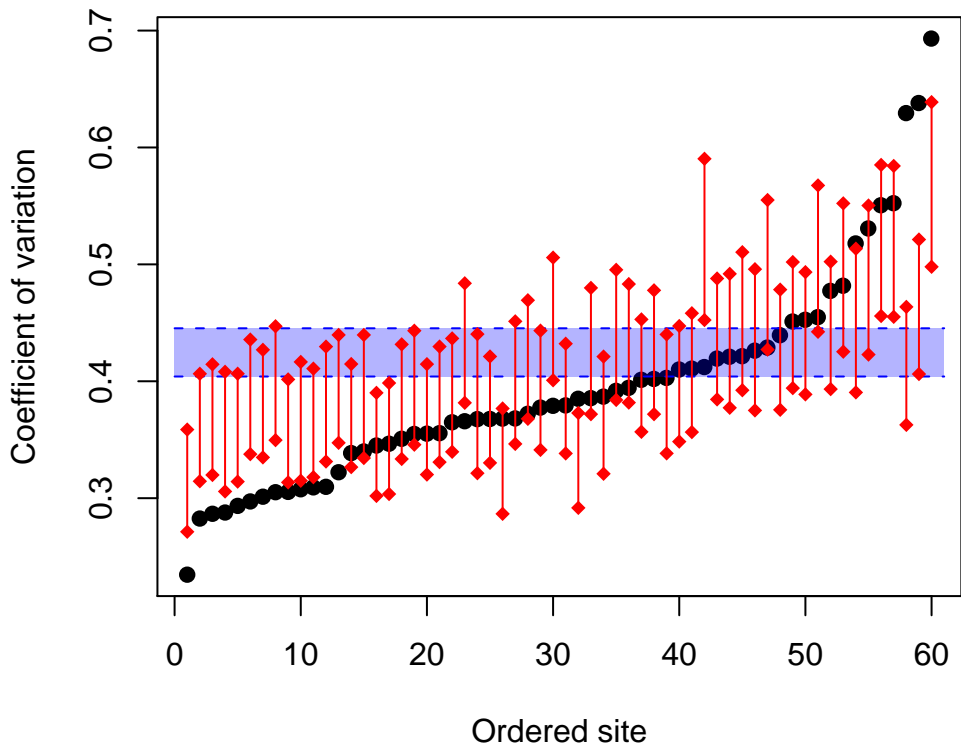
**Relative change in  $R_{0.99}$  (%)**

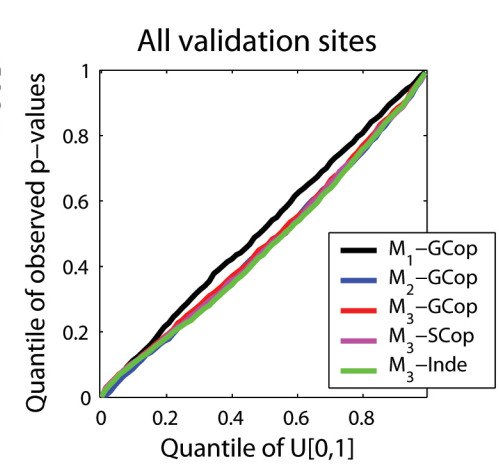
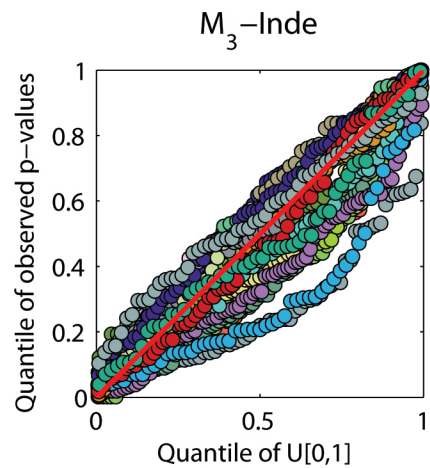
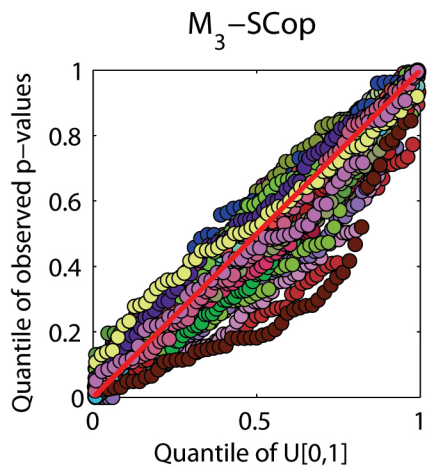
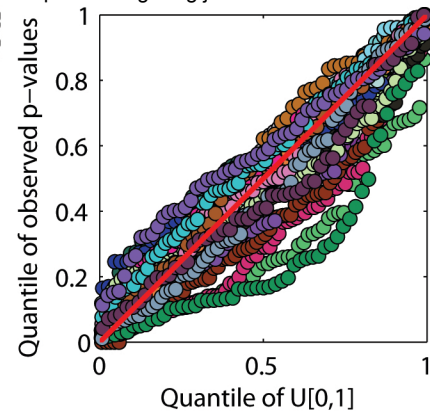
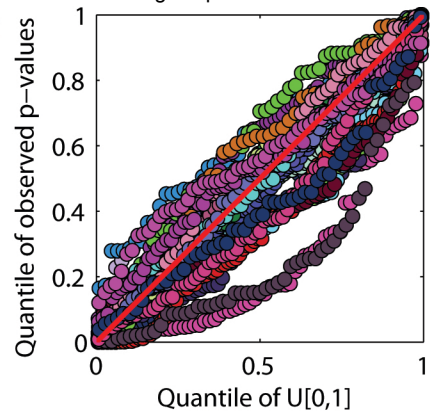
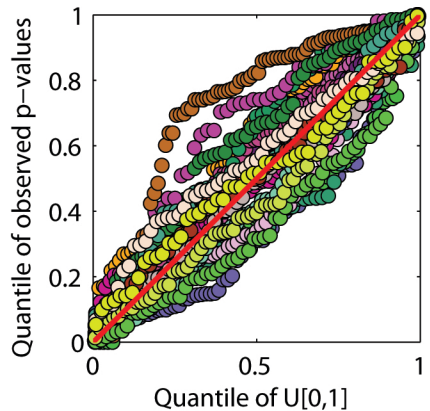


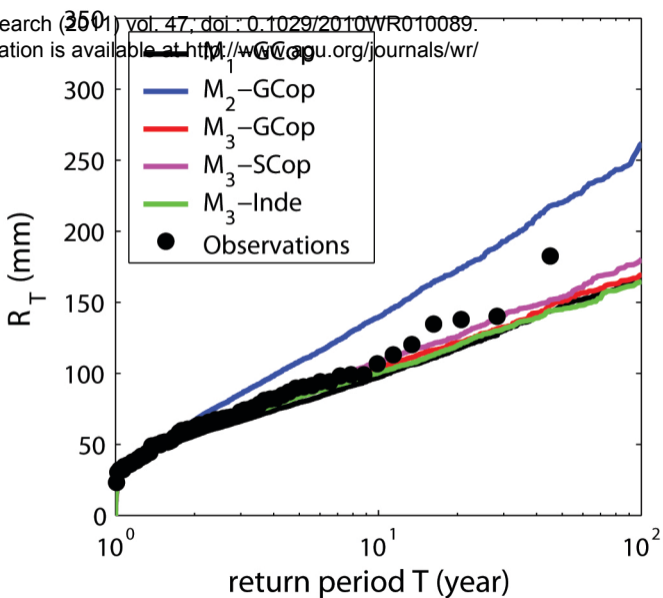
**Relative change in**

**$R_{0.99}$  uncertainty (%)**









The original publication is available at <http://www.agu.org/journals/>

