



# Extra-Parametrized Extreme Value Copula: Extension to a Spatial Framework

Julie Carreau, Gwladys Toulemonde

## ► To cite this version:

Julie Carreau, Gwladys Toulemonde. Extra-Parametrized Extreme Value Copula: Extension to a Spatial Framework. Spatial Statistics, 2020, 40, 10.1016/j.spasta.2020.100410 . hal-02419118

**HAL Id: hal-02419118**

**<https://inria.hal.science/hal-02419118>**

Submitted on 19 Dec 2019

**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.

# Extra-Parametrized Extreme Value Copula : Extension to a Spatial Framework

J. Carreau<sup>1\*</sup> and G. Toulemonde<sup>2</sup>

Julie.Carreau@ird.fr, (phone) +33467149027, (fax) +33467144774

<sup>1</sup> HydroSciences Montpellier, CNRS/IRD/UM, Université de Montpellier - Case 17, 163 rue Auguste Broussonet 34090 Montpellier, France

<sup>2</sup> Institut Montpellierain Alexander Grothendieck, Université de Montpellier, CNRS, Inria.

## Abstract

Hazard assessment at a regional scale may be performed thanks to a spatial model for maxima including a characterization of the spatial dependence structure. Such a spatial model can be obtained by combining the generalized extreme-value (GEV) distribution for the univariate marginal distributions with extreme-value copulas to describe their dependence structure, as justified by the theory of multivariate extreme values. However, most high-dimensional copulas are too simplistic for a spatial application. Recently, a class of flexible extreme-value copulas called extra-parametrized Gumbel or XGumbel was proposed. The XGumbel copula combines two Gumbel copulas with weight parameters, termed extra-parameters, taking values in the unit hyper-cube. In a multisite study, the copula dimension being the number of sites, the XGumbel copula quickly becomes over-parametrized. In addition, interpolation to ungauged locations is not easily achieved. We develop an extension of the XGumbel copula to the spatial framework. Our case study consists of annual maxima of daily precipitation totals at 177 gauged stations over a 57 year period in the French Mediterranean. The pattern of decrease with the distance of the strength of extremal dependence, as described by the upper tail dependence coefficient, indicates asymptotic dependence as it stabilizes at a non-zero value. In addition, non-stationarity is detected by looking at maps of the upper tail dependence coefficients with respect to a reference station. We propose a spatial model for maxima that combines a spatial regression for GEV marginals built with a vector generalized linear model and the spatialized XGumbel copula defined thanks to a spatial mapping for the extra-parameters. The mapping is designed shaped as a disk according to bivariate properties of the XGumbel copula. An Approximate Bayesian Computation (ABC) scheme that seeks to reproduce upper tail dependence coefficients for distance classes is used to infer the parameters. Evaluation of the proposed spatial model for maxima and comparison with a Brown-Resnick process, a spatial process adapted for pointwise maxima, are carried out on our case study. The spatial regression for GEV marginals is assessed in terms of return levels at six stations kept aside for validation purposes. The posterior distribution of the ABC scheme for the spatialized XGumbel yields, except in one instance, interpretable parameters. The comparisons between the spatialized XGumbel copula and the Brown-Resnick process show the following. The pattern of decrease of the upper tail dependence coefficients with the distance is well reproduced in both cases. However, the spatialized XGumbel copula has a clear non-stationary pattern in terms of the upper tail dependence coefficient inherited from the shape of the extra-parameter mapping while the Brown-Resnick process has none, by construction. Differences due to the non-stationarity of the dependence structure or the lack thereof are also notable on data-scale simulations and on 3-dimensional joint tail dependence coefficient estimates, a quantity involving trivariate distributions that could be of interest for a regional hazard analysis.

**keywords :** Gumbel copula, spatial extremes, heavy precipitation, ABC, non-stationarity

# 1 Introduction

The French Mediterranean is exposed to intense rainfall events called Cevenol events. These regularly cause flooding leading to important material damages and fatalities (Delrieu *et al.*, 2005; Braud *et al.*, 2014). Hazard assessment is conventionally performed by determining at-site  $T$  year return levels - the rainfall intensity level that is expected to be exceeded on average once per  $T$  years at a given site, see for instance Carreau *et al.* (2017). However, planning for flood risk mitigation is generally made at a regional scale. Therefore, a quantity of interest might rather be the probability that, conditionally on the fact that rainfall intensity at a given site has reached a high level, high intensity levels are likely to be reached at nearby sites. To estimate such a probability, characterization of the dependence of intense rainfall events in space, that is knowledge on spatial patterns of extreme events, is required. To this end, a spatial model for maxima over blocks of observations may be used.

Extreme value theory developed a sound theoretical framework to model the distribution of maxima over sufficiently large blocks of observations (Coles, 2001). Their univariate marginal distributions can be approximated by the Generalized Extreme Value (GEV) distribution (Fisher & Tippett, 1928; Gnedenko, 1943; Gumbel, 1958). In the multivariate case, theoretically justified distributions for componentwise maxima are the so-called Multivariate Extreme Value (MEV) distributions. The extension to the spatial setting leads to max-stable processes whose finite dimensional margins are MEV (de Haan, 1984). MEV distributions are either asymptotically dependent which entails that the dependence level remains constant at extreme levels or strictly independent (no dependence whatever the level).

MEV distributions and max-stable processes, unlike the GEV, do not have a unique finite dimensional parametrization (Beirlant *et al.*, 2004). MEV distributions can be constructed by associating GEV margins with MEV copulas. Some MEV copulas such as the Gumbel copula exist in high dimension but are limited in their ability to reproduce complex dependencies. Moreover, interpolation to ungauged locations is not straightforward. Several parametric models for max-stable processes have been proposed, see Davison *et al.* (2012) for a recent review. For small study regions, a single parametric model may be used, for instance see Thibaud *et al.* (2013). However, in order to account for differences in dependence structures resulting from non-stationarities, larger study regions may be split into smaller sub-regions (Blanchet & Davison, 2011; Blanchet & Creutin, 2017).

As the complete log-likelihood is often intractable in high dimension, let alone in the spatial framework, pairwise log-likelihood inference is a common practice, in particular for max-stable processes (Davison *et al.*, 2012). Another possibility is Approximate Bayesian Computation (ABC) likelihood free inference that selects parameters such that the model reproduces statistics of interest (see Beaumont (2010) for instance). By simulating from the model for candidate parameters drawn from a prior distribution, ABC schemes constitute the so-called *reference table* that contains the statistics of interest. The posterior distribution consists of the candidate



parameters that yielded statistics sufficiently similar to the observations'. ABC schemes for max-stable processes rely on summary statistics containing information on the extremal dependence structure (Erhardt & Smith, 2012; Erhardt & Sisson, 2016; Lee *et al.*, 2018).

In this work, we propose a spatial model for maxima that rely on the extension to the spatial framework of the class of extra-parametrized MEV copulas (Durante & Salvadori, 2010; Salvadori & De Michele, 2010). The extra-parameters characterize each dimension thereby introducing additional flexibility in the dependence structure. We focus on extra-parametrized Gumbel (*XGumbel*) copulas, see Section 2. In Section 3, we present our case study, annual maxima of daily precipitation at 177 gauged stations over a 57 year period in the French Mediterranean. Our proposed spatial model, described in Section 4, combines the extension to the spatial framework of the XGumbel copula with a spatial regression model for the GEV marginals. This way, MEV distributions are defined for any set of sites, whether gauged or ungauged. The spatial extension is achieved by defining the extra-parameters of the XGumbel copula as a mapping of geographical covariates. An ABC scheme is designed to perform the inference. Evaluation on our precipitation case study is carried out in Section 5. The spatialized XGumbel copula is compared with the Brown-Resnick process, a max-stable process commonly used to model environmental extremes (Brown & Resnick, 1977; Davison *et al.*, 2012).

## 2 Extra-parametrized Gumbel copula

### 2.1 Multivariate definition

The multivariate XGumbel copula  $C_\psi(\cdot)$ , defined as

$$C_\psi(\mathbf{u}) = C_{\beta_A}(\mathbf{u}^{\mathbf{a}})C_{\beta_B}(\mathbf{u}^{\mathbf{1}-\mathbf{a}}), \quad \beta_A, \beta_B \geq 1, \quad \mathbf{a} = (a_1, \dots, a_d) \in [0, 1]^d, \quad (1)$$

is a distribution function on the unit hypercube  $[0, 1]^d$  with parameter vector  $\psi = (\beta_A, \beta_B, \mathbf{a})$ . The parameters  $\beta_A, \beta_B \geq 1$  are inherited from the two Gumbel copulas,  $C_{\beta_A}(\cdot)$  and  $C_{\beta_B}(\cdot)$ , whose general form is

$$C_\beta(\mathbf{u}) = \exp \left\{ - \left[ \sum_{i=1}^d (-\ln u_i)^\beta \right]^{1/\beta} \right\}, \quad \beta \geq 1. \quad (2)$$

Note that the case  $\beta = 1$  corresponds to the independent copula. As they affect all  $d$  dimensions in the same fashion, the two parameters  $\beta_A$  and  $\beta_B$  can be thought of as global parameters. The extra-parameter vector  $\mathbf{a} = (a_1, \dots, a_d) \in [0, 1]^d$  appears as componentwise exponents in Eq. (1). As each dimension is characterized separately, extra-parameters may be thought of, in contrast to  $\beta_A$  and  $\beta_B$ , as local parameters. As can be seen from Eq. (1), if the values of  $\beta_A$  and  $\beta_B$  are swapped, the same copula  $C_\psi$  is obtained by replacing  $\mathbf{a}$  with  $\mathbf{1} - \mathbf{a}$ . To remove this identifiability issue, we fix  $\beta_A \leq \beta_B$ .

As it fulfills the max-stability property, i.e.  $C(u_1^t, \dots, u_d^t) = C^t(u_1, \dots, u_d) \forall t > 0$ , the Gumbel copula is a multivariate extreme value (MEV) copula. By the definition in Eq. (1), it follows that  $C_\psi$  is a MEV copula as well (Salvadori & De Michele, 2010). The multivariate XGumbel copula may be obtained by a constructive approach as follows (see e.g. Liebscher (2008)). Let  $\mathbf{U} \sim C_{\beta_A}$  and  $\mathbf{V} \sim C_{\beta_B}$ , then  $\max(\mathbf{U}^{1/a}, \mathbf{V}^{1/(1-a)})$  is distributed according to Eq. (1).

## 2.2 Bivariate properties

A MEV copula can be defined with the Pickands function conventionally denoted by  $A$  (Pickands, 1981; Marcon *et al.*, 2017). In the bivariate case, a copula  $C$  is a MEV copula if and only if there exists a convex function  $A : [0, 1] \mapsto [1/2, 1]$  such that

$$\mathbb{P}(U_1 \leq u_1, U_2 \leq u_2) = C(u_1, u_2) = \exp \left[ \ln(u_1 u_2) A \left( \frac{\ln(u_2)}{\ln(u_1 u_2)} \right) \right], \quad (3)$$

with  $U_1$  and  $U_2$  two uniform random variables on the interval  $[0, 1]$  and  $0 \leq u_1, u_2 \leq 1$ . The following properties must be fulfilled :  $\min((1-t), t) \leq A(t) \leq 1$ , for all  $t \in [0, 1]$ ,  $A(0) = A(1) = 1$ ,  $-1 \leq A'(0) \leq 0$ ,  $0 \leq A'(1) \leq 1$  and  $A'' \geq 0$ .

For the bivariate XGumbel copula, the Pickands function, illustrated in Fig. 1a, is

$$A(t) = \underbrace{\left[ a_1^{\beta_A} (1-t)^{\beta_A} + a_2^{\beta_A} t^{\beta_A} \right]^{1/\beta_A}}_{A_{\beta_A}(t)} + \underbrace{\left[ (1-a_1)^{\beta_B} (1-t)^{\beta_B} + (1-a_2)^{\beta_B} t^{\beta_B} \right]^{1/\beta_B}}_{A_{\beta_B}(t)}. \quad (4)$$

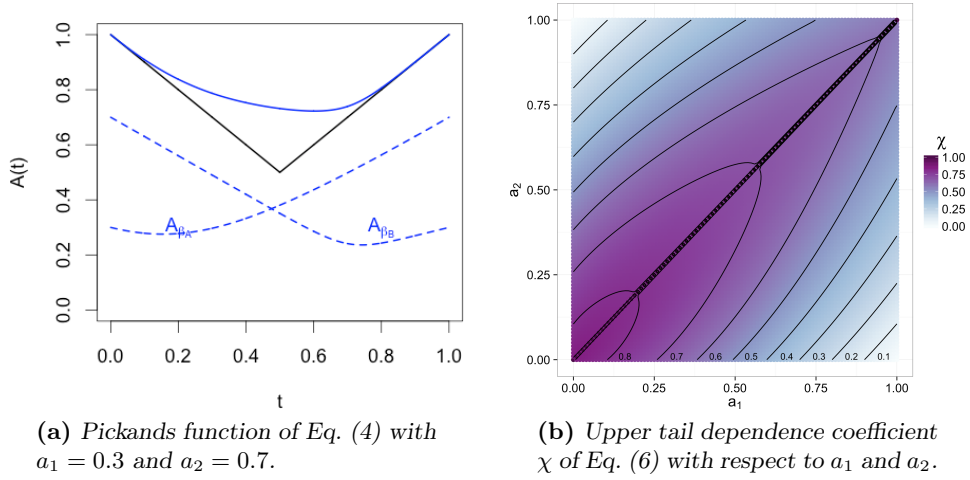
The Pickands function completely characterizes bivariate extremal dependence. It is equal to 1 in case of independence and equal to  $\min((1-t), t)$  in case of perfect dependence. In between, the strength and the shape of the dependence, in particular the asymmetry, may vary. Note that the XGumbel copula is symmetrical when  $a_1 = a_2$  or when  $\beta_A = \beta_B$  and  $a_1 = 1 - a_2$ .

For 2-dimensional MEV distributions, the strength of extremal dependence may be summarized by the upper tail dependence coefficient  $\chi$  defined as

$$\chi = \chi(u) = \mathbb{P}(U_2 > u \mid U_1 > u) = 2(1 - A(1/2)), \quad \forall 0 < u < 1. \quad (5)$$

In case of asymptotic independence, which necessarily corresponds to strict independence in a max-stable context,  $\chi = 0$ . Otherwise,  $0 < \chi \leq 1$  indicates the strength of the asymptotic dependence (Sibuya, 1960; Coles *et al.*, 1999).

The upper tail dependence coefficient of the bivariate XGumbel copula is defined as



**Figure 1:** Bivariate properties of the XGumbel copula of Eq. (1) with  $\beta_A = 2$  and  $\beta_B = 5$ .

$$\chi = 2 - [(a_1^{\beta_A} + a_2^{\beta_A})^{1/\beta_A} + ((1 - a_1)^{\beta_B} + (1 - a_2)^{\beta_B})^{1/\beta_B}]. \quad (6)$$

1 It may be deduced by combining Eq. (4) and Eq. (5). The variation of the  $\chi$  of the XGumbel  
2 copula with respect to the values of the extra-parameters  $a_1$  and  $a_2$  is illustrated in Fig. 1b for  
3  $\beta_A = 2$  and  $\beta_B = 5$ . We note that  $\chi$  is maximum when  $a_1 = a_2$  (along the first diagonal) and  
4 increases for decreasing values of the extra-parameter (in the lower left corner). In the limiting  
5 case with  $a_1 = a_2 = 0$  ( $a_1 = a_2 = 1$ ), the XGumbel copula boils down to the Gumbel copula  
6 with parameter  $\beta_B$  ( $\beta_A$ ) and  $\chi = 2 - 2^{1/\beta_B}$  ( $\chi = 2 - 2^{1/\beta_A}$ ). In addition, independence ( $\chi = 0$ )  
7 is achieved when  $a_1 = 0$  and  $a_2 = 1$  or the reverse,  $a_1 = 1$  and  $a_2 = 0$ .

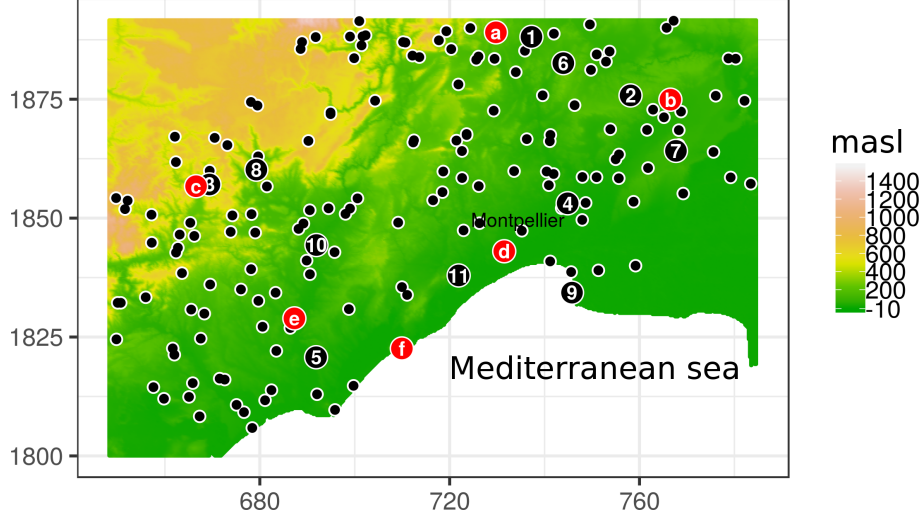
### 8 3 Precipitation data

#### 9 3.1 Study area

10 Our study area is illustrated in Fig. 2. It covers about 16 000 km<sup>2</sup> around the city of Montpellier  
11 near the Mediterranean area in the south of France. It is well-known for intense rainfall events  
12 that occur mainly in autumn (Brunet *et al.*, 2018). Owing to the Cévennes mountain range  
13 sitting in the north-west of the area, the Rhône river valley running in the eastern end that  
14 encompasses the city of Montpellier and the Mediterranean sea in the south, there is a strong  
15 variability in the distribution of heavy precipitation both in terms of intensities and of dependence  
16 structure (Blanchet & Creutin, 2017; Carreau *et al.*, 2017).

17 We selected 177 gauged stations from the Météo-France network, the French weather service,  
18 that are located within our study area. For each station, we extracted annual maxima of daily  
19 precipitation totals over a 57 year period (1958-2014). The calibration set consists of the stations  
20 depicted as black filled circles. Among these, 11 numbered stations are used for a regional hazard

analysis in § 5.3. In addition, six stations with no missing values scattered in the study region, shown as red filled circles wearing letters in Fig. 2, are kept aside for validation purposes in Section 5.



**Figure 2:** Gauged stations in the study area located in the French Mediterranean : 171 stations (in black) are used for calibration, 11 of these are numbered and serve in a regional hazard analysis, 6 stations (in red) are kept aside for validation - coordinates are in extended Lambert II projection.

### 3.2 Exploratory analyses of the dependence structure

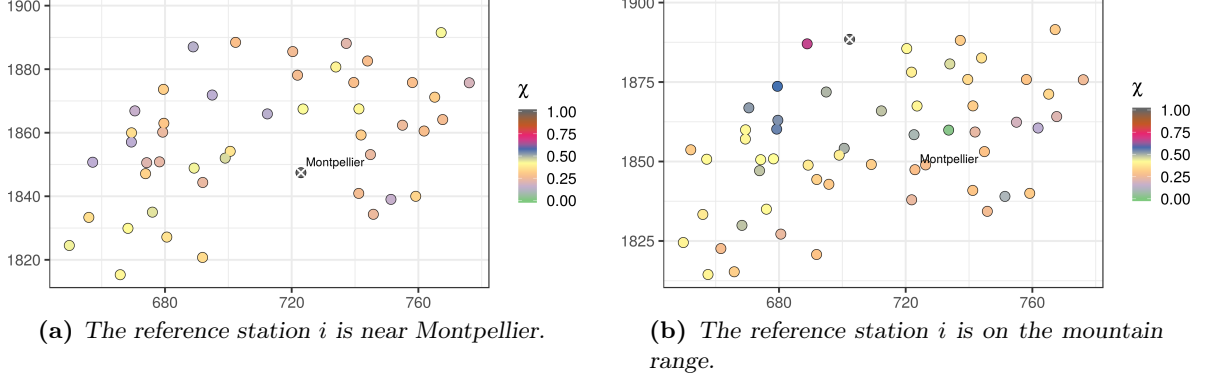
We rely on sample estimates of the upper tail dependence coefficient  $\chi$  introduced in Eq. (5) that summarizes the strength of the dependence between two sites  $i$  and  $j$ . Let  $U_i$  and  $U_j$  be random variables representing the annual maxima at each site transformed to the uniform scale. Then, the upper tail dependence coefficient  $\chi_{ij}$  between sites  $i$  and  $j$  can be written as

$$\chi_{ij} = 2 - \left( \frac{1 + \mathbb{E}[|U_i - U_j|]}{1 - \mathbb{E}[|U_i - U_j|]} \right) \quad (7)$$

where  $1/2\mathbb{E}[|U_i - U_j|]$  is the so-called madogram (Cooley *et al.*, 2006; Vannitsem & Naveau, 2007). Sample estimates  $\hat{\chi}_{ij}$  are obtained by replacing the expectation  $\mathbb{E}[|U_i - U_j|]$  in Eq. (7) by the sample average. To compute empirical estimates, observed annual maxima are rank-transformed to the uniform scale by applying empirical distribution functions. For a given pair of stations, we kept empirical estimates only when at least 30 years of observations are available.

To assess the assumption of stationarity in the strength of the dependence, we depicted maps of estimates  $\hat{\chi}_{ij}$ ,  $i$  being a fixed reference station and  $j \in \{1, \dots, 171\}$  being, in turn, each of the other calibration stations. In the left panel of Fig. 3, the reference station is the nearest one to the city of Montpellier which sits near the coastline. The strength of dependence is relatively low even for the closest stations ( $\hat{\chi}_{ij}$  is about 0.3). In the right panel of Fig. 3, the reference

1 station lies on the mountain range and the level of dependence is higher ( $\hat{\chi}_{ij}$  is about 0.75 for  
 2 the closest station). This change of dependence intensity with the location is an indication of  
 3 non-stationarity.



**Figure 3:** Maps of empirical upper tail dependence coefficient estimates  $\hat{\chi}_{ij}$  (see Eq. (7)) with respect to a given reference station  $i$  shown by a white cross.

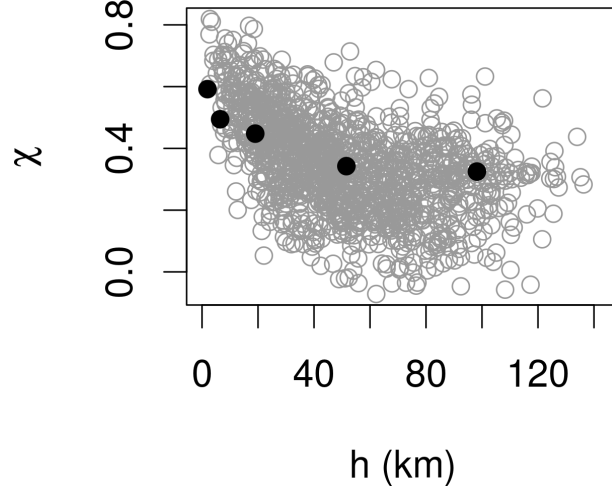
4 We also assess the spatial behavior of the strength of dependence by looking at plots of esti-  
 5 mates  $\hat{\chi}_{ij}$  with respect to  $h$ , the distance between stations  $i$  and  $j$ . To reduce variability, we also  
 6 computed estimates  $\hat{\chi}_{[h]}$  for five classes of distance  $[h] \in \{(0, 3], (3, 9], (9, 27], (27, 81], (81, 243]\}$   
 7 that follow a geometric progression. In Fig. 4, the pairwise estimates  $\hat{\chi}_{ij}$  are shown (in gray)  
 8 together with the distance class estimates  $\hat{\chi}_{[h]}$  (in black) for the 171 stations of the calibration set  
 9 over the 57 year period. Note that preliminary analyses performed by considering two orthogonal  
 10 directions detected no significant anisotropy. The strength of dependence, as shown in Fig. 4,  
 11 decreases with increasing distance, as is typical for extreme climatic spatial data (Blanchet &  
 12 Davison, 2011; Davison & Gholamrezaee, 2012). However, the level of dependence seems to  
 13 stabilize at a value clearly larger than zero, starting at a distance of about 40 km. This is an  
 14 indication of asymptotic dependence.

## 15 4 Spatial XGumbel

16 The spatial XGumbel model for maxima presented in § 4.3 combines a spatial regression model  
 17 for the univariate marginal distributions introduced in § 4.1 with the spatial extension of the  
 18 multivariate XGumbel copula (see § 2.1) in § 4.2. A two-stage inference scheme for the spatial  
 19 XGumbel model is described in § 4.4.

### 20 4.1 Response surfaces

For a given site  $i$ , we denote  $Y_i$  the random variable representing the annual maxima of daily precipitation. As is commonly done, we assume that  $G_i$ , the distribution function of  $Y_i$ , is the



**Figure 4:** Empirical upper tail dependence coefficient estimates  $\hat{\chi}_{ij}$  for pairs of stations with at least 30 years of observations (in gray) and distance class estimates  $\hat{\chi}_{[h]}$  (in black) with  $[h] \in \{(0, 3], (3, 9], (9, 27], (27, 81], (81, 243]\}$  (see Eq. (7)).

GEV distribution which has the form

$$G_i(y) = \exp \left[ - \left\{ 1 + \xi_i \left( \frac{y - \mu_i}{\sigma_i} \right) \right\}_+^{-1/\xi_i} \right], \quad (8)$$

1 where  $a_+ = \max(0, a)$ . The GEV distribution, which is theoretically justified by the univariate  
2 extreme value theory (Fisher & Tippett, 1928; Gnedenko, 1943; Gumbel, 1958; Coles, 2001),  
3 depends on three parameters, see Eq. (8) : the location parameter  $\mu_i \in \mathbb{R}$ , the scale parameter  
4  $\sigma_i > 0$  and the shape parameter  $\xi_i \in \mathbb{R}$ . The latter characterizes the behavior of the upper tail  
5 of the distribution : exponential decay when  $\xi_i = 0$ , polynomial decay when  $\xi_i > 0$  and finite  
6 endpoint for  $\xi_i < 0$ .

7 To obtain response surfaces that interpolate the GEV parameters over the study area, we  
8 rely on a vector generalized linear model (VGLM) approach, see Yee & Stephenson (2007); Yee  
9 (2015). This allows to fit the GEV distribution simultaneously at all the calibration stations.  
10 The three GEV parameters are defined as functions

$$\mu(\mathbf{x}; \boldsymbol{\alpha}_\mu) = \alpha_{\mu:0} + \alpha_{\mu:1}x_1 + \cdots + \alpha_{\mu:p}x_p \quad (9)$$

$$\log(\sigma(\mathbf{x}; \boldsymbol{\alpha}_\sigma)) = \alpha_{\sigma:0} + \alpha_{\sigma:1}x_1 + \cdots + \alpha_{\sigma:p}x_p \quad (10)$$

$$\log(\xi(\mathbf{x}; \boldsymbol{\alpha}_\xi) + 0.5) = \alpha_{\xi:0} + \alpha_{\xi:1}x_1 + \cdots + \alpha_{\xi:p}x_p, \quad (11)$$

11 where  $\mathbf{x} \in \mathbb{R}^p$  are geographical covariates known everywhere in the study area. For the shape  
12 parameter, an offset of 0.5, see Eq. (11), serves to enforce that  $\xi > -0.5$  thereby ensuring  
13 numerical stability (Yee, 2015).

## 4.2 Spatialized XGumbel copula

The spatialized XGumbel copula is based on the definition of the extra-parameters as a mapping  $a : \mathbb{R}^2 \mapsto [0, 1]$ , with parameters  $\theta$ , of the x- and y-coordinates of the sites. Note that more general geographical covariates could be used as for the response surfaces. This mapping allows to extend the XGumbel copula from Eq. (1) to any set  $\mathcal{S}$  of sites by letting the extra-parameters be given by  $a_s = a(s_x, s_y; \theta)$ , for all sites  $s \in \mathcal{S}$  with x- and y-coordinates  $(s_x, s_y) \in \mathbb{R}^2$ . The vector of parameters  $\psi_{\text{spat}}$  of the spatialized XGumbel copula includes the global parameters  $\beta_A$  and  $\beta_B$ , as in Eq. (1), and  $\theta$  to define the extra-parameter mapping. The number of parameters is thus invariant to the dimension, i.e. the number of sites in a spatial application. However, the extra-parameter mapping must be designed so that the resulting spatialized XGumbel copula be able to reproduce the spatial dependence structure of the observations.

To this end, we rely on the properties of the upper tail dependence coefficient  $\chi$  of the XGumbel copula, see Fig. 1b. First, we note that the dependence between two sites is maximized when their extra-parameter values are both equal to zero. In such a case, the extremal coefficient  $\chi$  only depends on  $\beta_B$  (let  $a_1 = a_2 = 0$  in Eq. (6)). Second, two sites are independent when one has extra-parameter value zero and the other has value one (let  $a_1 = 0$  and  $a_2 = 1$  or vice-versa in Eq. (6)). To account for these two points, we designed the extra-parameter mapping shaped as a disk, as shown in Fig. 5a, with values approaching zero near the disk center indicating stronger dependence and values getting closer to one when moving away from the center implying independence between sites near the center and away from the center.

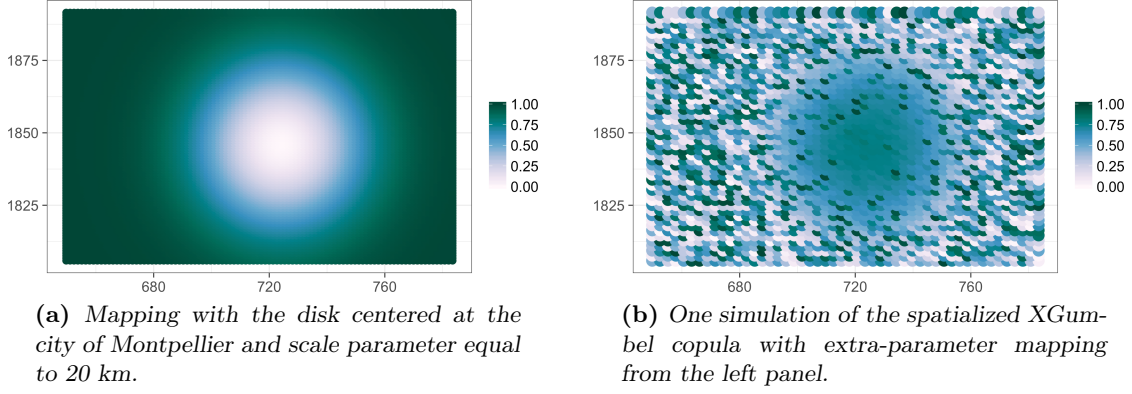
More precisely, for  $(s_x, s_y) \in \mathbb{R}^2$ , the extra-parameter mapping is parametrized as

$$a(s_x, s_y; \theta) = 1 - \exp \left\{ -\frac{(s_x - \mu_x)^2 + (s_y - \mu_y)^2}{2\delta^2} \right\} \quad (12)$$

where  $\delta > 0$  is a scale parameter and  $(\mu_x, \mu_y) \in \mathbb{R}^2$  is the center of the disk. The extra-parameter mapping has thus three parameters  $\theta = (\delta, \mu_x, \mu_y)$ . Note that any pair of sites located in the dark green area in Fig. 5a, whatever their distance, has the same dependence strength that only depends on  $\beta_A$  (let  $a_1 = a_2 = 1$  in Eq. (6)). In order to permit independence between pairs of sites with larger distances,  $\beta_A$  is fixed to 1, i.e. the independent copula. Therefore, the parameter vector of the spatialized XGumbel copula is  $\psi_{\text{spat}} = (\beta_B, \delta, \mu_x, \mu_y)$ . In addition, the Pickands function and the upper tail dependence coefficient from Eq. (4) and (6) are simplified as follows :

$$\begin{aligned} A(t) &= a_1(1-t) + a_2t + \left[ (1-a_1)^{\beta_B}(1-t)^{\beta_B} + (1-a_2)^{\beta_B}t^{\beta_B} \right]^{1/\beta_B} \\ \chi &= 2 - [(a_1 + a_2) + ((1-a_1)^{\beta_B} + (1-a_2)^{\beta_B})^{1/\beta_B}]. \end{aligned}$$

In Fig. 5b, a simulation of the spatialized XGumbel copula reveals the impact on the spatial dependence pattern of the shape of extra-parameter mapping shown in Fig. 5a. The area of



**Figure 5:** Effect of the shape of the extra-parameter mapping on the spatial pattern of a simulation of the spatialized XGumbel copula ( $\beta_B = 20$ ).

1 strong dependence is completely determined by the location of the disk center and the value of  
 2 the scale parameter  $\delta$  in Eq. (12). As areas of various degree of dependence may be defined,  
 3 the spatialized XGumbel copula allows the introduction of non-stationarity in the dependence  
 4 structure.

### 5 4.3 Proposed spatial model for maxima

6 The full spatial model for maxima combines the GEV distribution provided by the response  
 7 surfaces in § 4.1 and the spatialized XGumbel copula described in § 4.2. For any set of sites,  
 8 whether gauged or ungauged, this spatial model yields a well-defined MEV distribution. More  
 9 precisely, ungauged sites can be modeled in a consistent way such that the lower dimensional  
 10 distributions of sets of gauged and ungauged sites belong to the same class.

11 More precisely, let  $\mathcal{S} = \{s_1, \dots, s_K\}$  be any set of  $K$  sites in the study area, for any  
 12  $K \in \mathbb{N}$ . For all  $s \in \mathcal{S}$  with x- and y-coordinates  $(s_x, s_y) \in \mathbb{R}^2$ , let  $Y_s$  and  $\mathbf{x}_s$  be respec-  
 13 tively the random variate representing the annual maxima of daily precipitation and the ge-  
 14 ographical covariates at site  $s$ . The GEV distribution function  $G_s$ ,  $\forall s \in \mathcal{S}$ , has parameters  
 15  $(\mu(\mathbf{x}_s; \boldsymbol{\alpha}_\mu), \sigma(\mathbf{x}_s; \boldsymbol{\alpha}_\sigma), \xi(\mathbf{x}_s; \boldsymbol{\alpha}_\xi))$  as provided by Eqs (9)-(11). Moreover, the XGumbel copula  
 16 parameter vector  $\psi$  contains the Gumbel copula parameters  $\beta_A = 1$  and  $\beta_B$  that are shared for  
 17 all the sites and the extra-parameters given by  $a_s = a(s_x, s_y; \theta)$  from Eq. (12)  $\forall s \in \mathcal{S}$ . The  
 18 multivariate distribution of the maxima at the  $K$  sites is then given by

$$\mathbb{P}(Y_{s_1} \leq y_1, \dots, Y_{s_K} \leq y_K) = C_\psi(G_{s_1}(y_1), \dots, G_{s_K}(y_K)), \quad (13)$$

19 with  $C_\psi$  defined in Eq. (1). Thanks to Eq. (13), it is possible to simulate from the model  
 20 everywhere in the study area.



#### 4.4 Inference scheme

As the joint estimation of the marginal and the dependence structure parameters of the spatial XGumbel model would be too complex, we opted for a two-stage inference scheme as follows. The parameter vectors  $\alpha_\mu$ ,  $\alpha_\sigma$  and  $\alpha_\xi$  of the response surfaces of the GEV parameters in Eqs. (9)-(11) are estimated by maximizing the log-likelihood under the independence assumption (Yee & Stephenson, 2007; Yee, 2015). The parameter vector  $\psi_{\text{spat}}$  of the spatialized XGumbel copula is estimated with an Approximate Bayesian Computation (ABC) scheme on the rank-transformed observations (as recommended in Genest & Favre (2007)).

To constitute the reference table of the ABC scheme, we use as summary statistics sample upper tail dependence coefficient estimates for distance classes  $\hat{\chi}_{[h]}$  with  $[h] \in \{(0, 3], (3, 9], (9, 27], (27, 81], (81, 243]\}$  based on the madogram (see Eq. (7)). In ABC schemes for max-stable processes, related summary statistics containing information on the strength of the extremal dependence structure were proposed. In Erhardt & Smith (2012); Erhardt & Sisson (2016), summary statistics deduced from the madogram and the extremal coefficient, which is equivalent to the upper tail dependence coefficient for max-stable distributions, were evaluated and compared. In addition to pairwise information, information based on triplet of sites was considered. A smoothing procedure, similar in spirits to the use of distance classes, was based on either curve fitting or by grouping stations.

The prior distribution in the ABC scheme of the spatialized XGumbel copula is meant to be vague. For the parameter vector  $\psi_{\text{spat}} = (\beta_B, \delta, \mu_x, \mu_y)$ , we set :  $\beta_B \sim U[10, 100]$ ,  $\delta \sim U[5, 100]$  and  $(\mu_x, \mu_y)$  is drawn uniformly from the locations of the 171 stations in the calibration set. The constitution of the reference table goes as follows, for all  $i \in \{1, \dots, 100000\}$  :

1. Draw candidate parameters  $\psi_{\text{spat}}^{(i)} = (\beta_B^{(i)}, \delta^{(i)}, \mu_x^{(i)}, \mu_y^{(i)})$  from the prior distribution ;
2. Simulate  $\mathbf{U}^{(i)} = (U_1^{(i)}, \dots, U_d^{(i)})$ , a sample of size  $n = 57$  from the spatialized XGumbel copula with parameters  $\psi_{\text{spat}}^{(i)}$  at the  $d = 171$  stations of the calibration set ;
3. Compute  $\hat{\chi}_{[h]}$ , the sample upper tail dependence coefficients for all  $[h] \in \{(0, 3], (3, 9], (9, 27], (27, 81], (81, 243]\}$  on the simulated sample  $\mathbf{U}^{(i)}$ .

We apply a simple version of ABC called rejection-ABC in which the posterior distribution consists of a subset of candidate parameters such that the distance in terms of summary statistics to the observations is small. More specifically, let  $\{\psi_{\text{spat}}^{(i_j)}\}_{j=1}^{100}$  with  $1 \leq i_j \leq 100000$  be the subset of 100 candidate parameters such that Euclidean distances in terms of summary statistics are the smallest. This corresponds to 0.1% of the simulations from the prior distribution.

## 5 Assessment of spatial models for maxima

We evaluate and compare spatial models for maxima on the annual maxima of the daily precipitation data described in Section 3. In § 5.1, a single spatial regression model for the univariate margins (see § 4.1) is considered. In § 5.2, the dependence structure as modeled by the spatialized XGumbel copula is compared with the one from a Brown-Resnick process (Brown & Resnick, 1977). The Brown-Resnick process is fitted by pairwise log-likelihood on the annual maxima rank-transformed to the Fréchet scale (this is performed with the R package from Ribatet (2018)). Uncertainty assessment is based on non-parametric bootstrap : 100 sets of Brown-Resnick parameters are estimated on bootstrap samples obtained by sampling with replacement the years of the calibration period. Finally, two complete spatial models, i.e. GEV margins combined with either the spatialized XGumbel copula or the Brown-Resnick process, are compared in § 5.3 in terms of simulated fields of maxima and in terms of their ability to reproduce conditional trivariate probabilities involving the validation stations. These probabilities may be of interest for hazard assessment at a regional scale.

### 5.1 Response surfaces

In addition to the x- and y-coordinates along with the altitude, we considered as covariates for the response surfaces in Eqs. (9)-(11) ten landscape features (Benichou & Le Breton, 1987). Based on a digital elevation model, these features are deduced from a principal component (PC) analysis applied to the relative elevation of a square neighborhood centered on each cell of the digital elevation grid. The first ten components are retained.

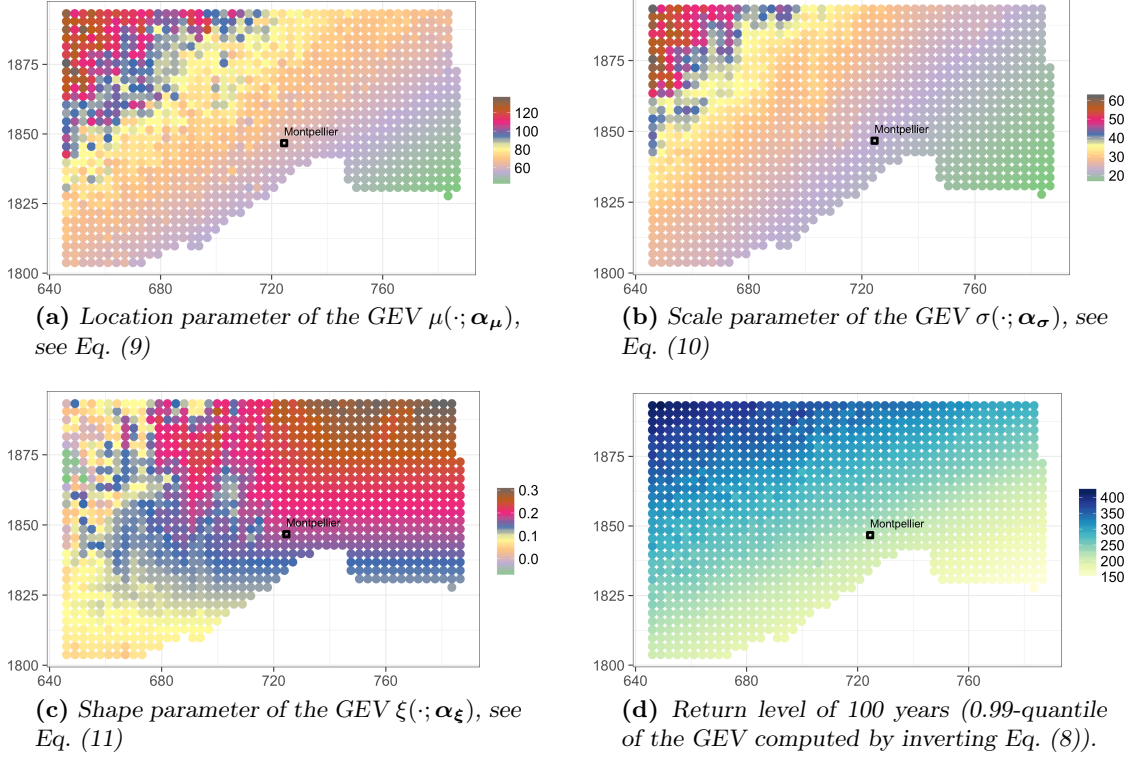
Covariate selection is performed in two stages. First, a screening is performed by applying LASSO regression to the natural logarithm of the annual maxima with the initial 13 covariates (Friedman *et al.*, 2010). Six covariates are selected : the x- and y- coordinates, the altitude and three landscape features resulting from the 1st, 4th and 9th PC. This selection is further refined by constraining the coefficients of the VGLM to be null when not sufficiently significant for a subset of the GEV parameters. The Bayesian Information Criterion (BIC) is used to ensure that the exclusion of covariates does not deteriorate the fit (Schwarz, 1978). The final covariate selection is summarized in Table 1.

**Table 1:** Selected covariates for the response surfaces of the GEV parameters (Eqs. (9)-(11)). In addition to the x- and y-coordinates and the altitude z, three landscape features (PC1, PC4 and PC9) are obtained from a principal component analysis of the digital elevation grid.

	x	y	z	PC1	PC4	PC9
$\mu(\cdot; \alpha_\mu)$	✓	✓	✓	✓	✓	✓
$\sigma(\cdot; \alpha_\sigma)$	✓	✓	✓			
$\xi(\cdot; \alpha_\xi)$		✓	✓			

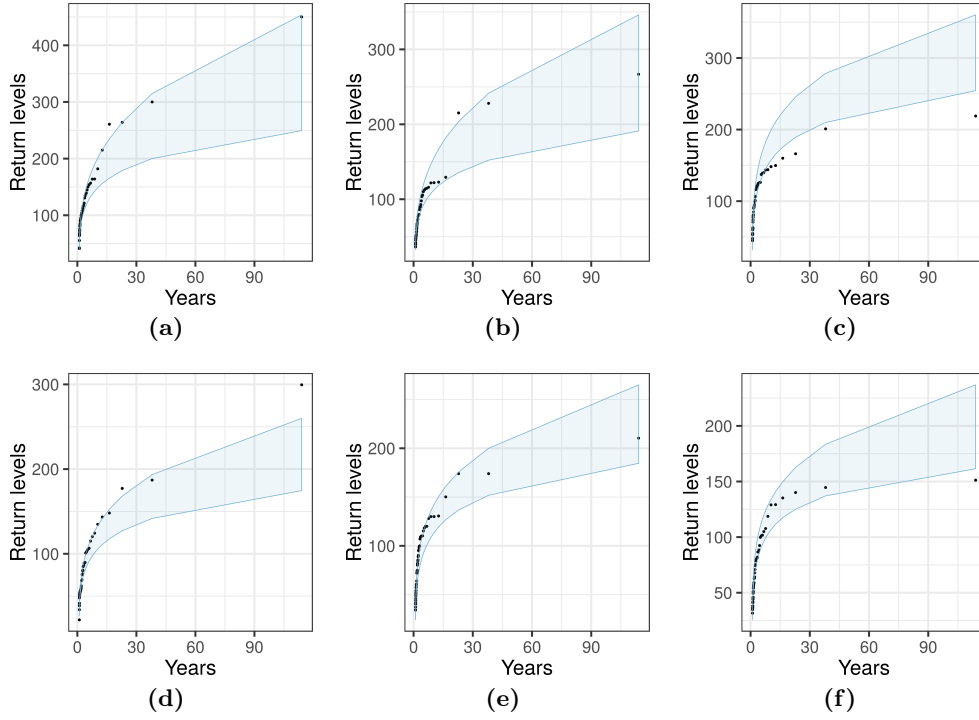
The response surfaces of the GEV parameters as provided by the fitted VGLM by interpo-

1 lating over a grid covering our study area are shown in Figs. 6a-6c. While the spatial patterns of  
 2 the location and scale parameters are strongly influenced by the altitude, the shape parameter  
 3 displays a different pattern with higher values in the Rhône river valley. The map of the 100-year  
 4 return levels, i.e. quantiles of probability 0.99 computed by inverting Eq. (8), is shown in Fig. 6d.  
 5 As is typical for this area, values ranging from 150 mm near the coastline to 400 mm on the  
 6 mountain range are observed (Carreau *et al.*, 2017).



**Figure 6:** Interpolation of the GEV parameters over a grid covering the study area with a vector generalized linear model approach, see Eqs. (9)-(11), and geographical covariates (see Table 1).

7 In Fig. 7, the goodness-of-fit of the response surfaces is evaluated in terms of return levels at  
 8 the six validation stations. Each validation station, depicted in red filled circles in Fig. 2, wears  
 9 a letter that is related to a panel in Fig. 7. Empirical return levels are depicted as black dots.  
 10 The light blue bands are 99% non-parametric bootstrap confidence bands (10000 replications  
 11 obtained by sampling with replacement the 57 years of annual maxima) for the return levels  
 12 computed from the GEV parameters interpolated by the fitted VGLM. At the third station  
 13 which is located in the mountain area (corresponding to the red filled circle wearing the letter c  
 14 in Fig. 2), the VGLM interpolation tends to overestimate the larger empirical return levels (see  
 15 Fig. 7c). Nevertheless, the fit is overall quite satisfactory.

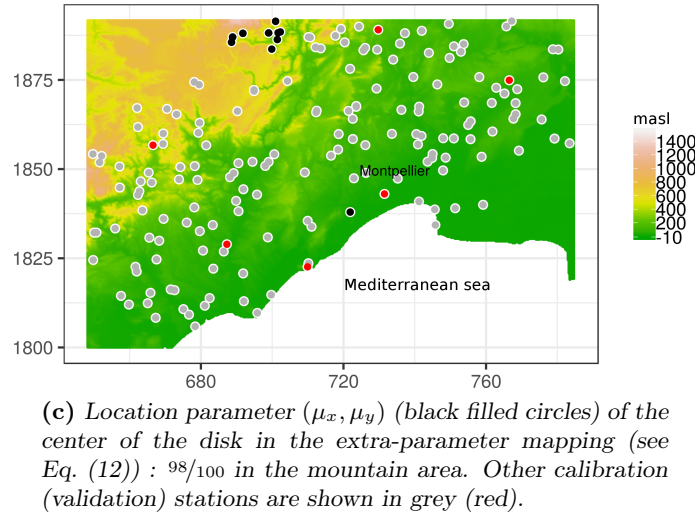
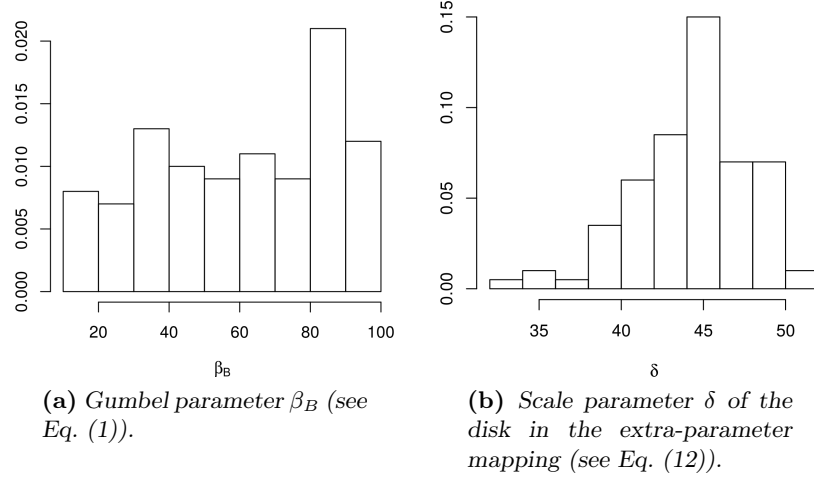


**Figure 7:** Return levels at the six validation stations, each panel corresponding to a red filled circle wearing the same letter in Fig. 2 : empirical estimates are depicted as black dots and 99% non-parametric bootstrap confidence bands for the return levels computed from GEV parameters interpolated by the fitted VGLM are shown in light blue.

## 5.2 Spatial dependence structures

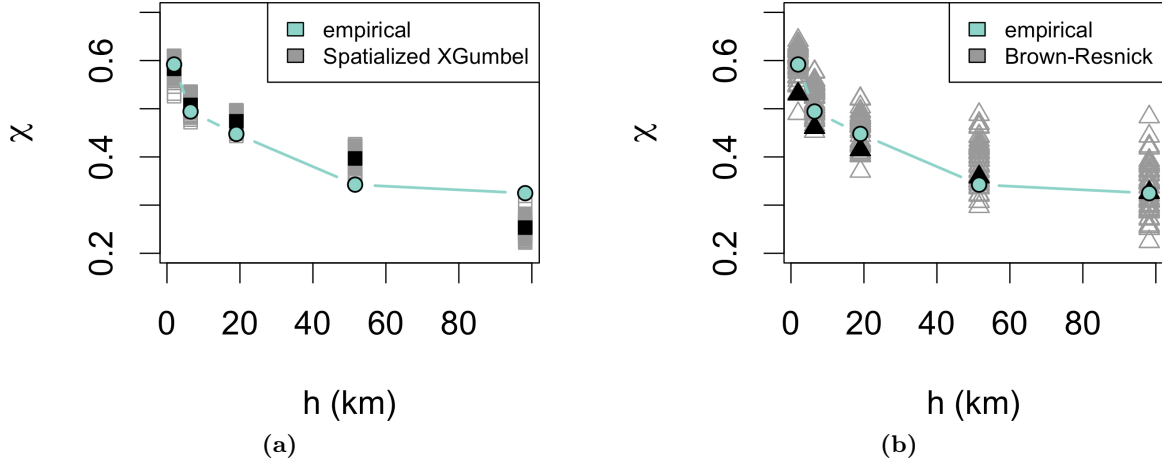
The posterior distribution of the spatialized XGumbel copula parameter vector resulting from the ABC scheme, that is the subset  $\{\psi_{\text{spat}}^{(i_j)}\}_{j=1}^{100}$  with  $1 \leq i_j \leq 100000$  of candidate parameters leading to the summary statistics closest to the observed ones, is illustrated in Fig. 8. For the Gumbel parameter  $\beta_B$ , in Fig. 8a, the posterior distribution is similar to the prior distribution  $U[10, 100]$ . This might indicate that the designed ABC scheme is not able to infer properly this parameter. In contrast, the posterior distribution of  $\delta$ , the scale parameter of the disk in the extra-parameter mapping, has a clear mode at about 45 km, see Fig. 8b. The posterior distribution of the location of the disk center in the extra-parameter mapping is represented as black filled circles in Fig. 8c. The selected disk centers are located preferentially, i.e. 98 times out of 100, over the mountain range, in a very specific area which might be explained by orographic effects.

The spatialized XGumbel copula and the Brown-Resnick process are compared in Fig. 9, left and right panel respectively, in terms of the statistics  $\hat{\chi}_{[h]}$ , i.e. the sample upper tail dependence coefficients for distance classes  $[h]$ , with  $[h] \in \{(0, 3], (3, 9], (9, 27], (27, 81], (81, 243]\}$ . The empirical estimates computed from the observations are shown in light blue in both panels. For each model, there are 100 statistics  $\hat{\chi}_{[h]}$  depicted in gray. For the spatialized XGumbel copula, these statistics, retrieved directly from the reference table, correspond to the 100 sets of parameters



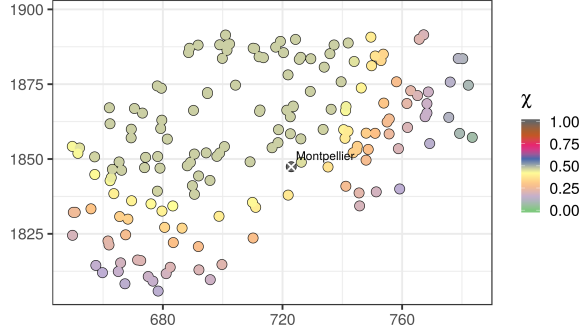
**Figure 8:** Posterior distribution of the spatialized XGumbel copula parameters  $\{\psi_{\text{spat}}^{(i_j)}\}_{j=1}^{100}$  with  $1 \leq i_j \leq 100000$  from the rejection-ABC inference scheme described in § 4.4.

1  $\{\psi_{\text{spat}}^{(i_j)}\}_{j=1}^{100}$  with  $1 \leq i_j \leq 100000$  from the posterior distribution of the rejection-ABC inference  
2 scheme. The median of the 100  $\hat{\chi}_{[h]}$  is also shown in black. For the Brown-Resnick process, the  
3 100 statistics are estimated by simulating samples of the same size as the observations' from  
4 the 100 sets of Brown-Resnick parameters obtained by non-parametric bootstrap. The statistics  
5 estimated from the fit on the original calibration data are shown in black. The patterns of de-  
6 crease in extremal dependence with the distance produced by both models of spatial dependence  
7 structure are comparable to the one obtained from the observed annual maxima. However, the  
8 spread and thus the uncertainty of the Brown-Resnick estimates is larger.

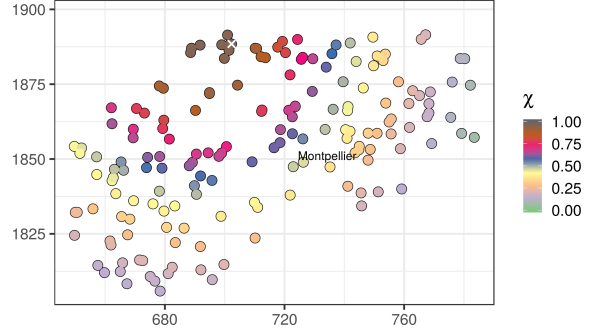


**Figure 9:** Upper tail dependence coefficient estimates  $\hat{\chi}_{[h]}$  for five classes of distance  $[h] \in \{(0, 3], (3, 9], (9, 27], (27, 81], (81, 243]\}$  (see Eq. (7)). For the spatialized XGumbel (left panel), the best 100 estimates (in gray, with the median in black) are retrieved from the reference table. For the Brown-Resnick process (right panel), estimates are computed on samples of the same size as the observations' (57 years); from 100 models fitted on bootstrap samples (in gray) and from the model on the original sample (in black).

9 In Fig. 10 and Fig. 11, the two models are compared in terms of non-stationarity patterns in  
10 the dependence structure. These patterns are produced when drawing the maps of the upper tail  
11 dependence coefficient estimates  $\hat{\chi}_{ij}$  with respect to two different reference sites  $i$ . In Fig. 10,  
12 these patterns are depicted for the spatialized XGumbel copula, with the  $\hat{\chi}_{ij}$  obtained by re-  
13 placing the parameters in Eq. (6) by the best set of parameters from the posterior distribution  
14 of the ABC scheme. For the Brown-Resnick process, the maps are shown in Fig. 11 with  $\hat{\chi}_{ij}$   
15 computed with the madogram, as in Eq. (7), on a sample of size 1000 simulated from the fitted  
16 model. Although the values are a bit too high with respect to the empirical estimates in Fig. 3,  
17 the non-stationary pattern of the spatialized XGumbel copula in Fig. 10 is generally reasonable.  
18 In contrast, the Brown-Resnick process in Fig. 11 not only fails to exhibit any non-stationarity,  
19 as expected since it is not designed to account for it, but it yields rather low values with little  
20 spatial variability compared to the empirical estimates in Fig. 3.

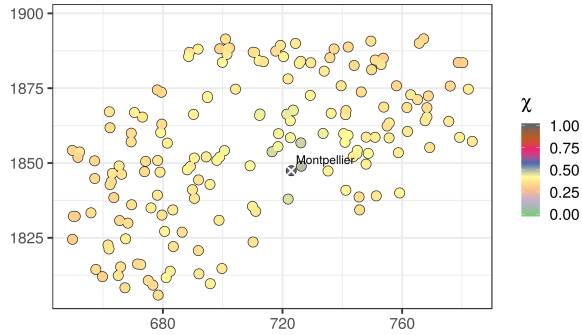


(a) The reference station  $i$  is near Montpellier.

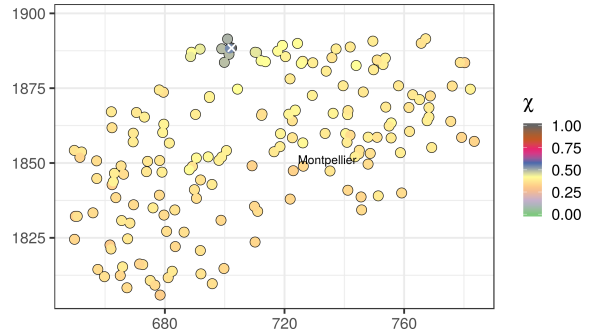


(b) The reference station  $i$  is on the mountain range.

**Figure 10:** Maps of spatialized XGumbel copula upper tail dependence coefficient estimates  $\hat{\chi}_{ij}$ , computed from Eq. (6). The reference station  $i$  is shown by a white cross.



(a) The reference station  $i$  is near Montpellier.

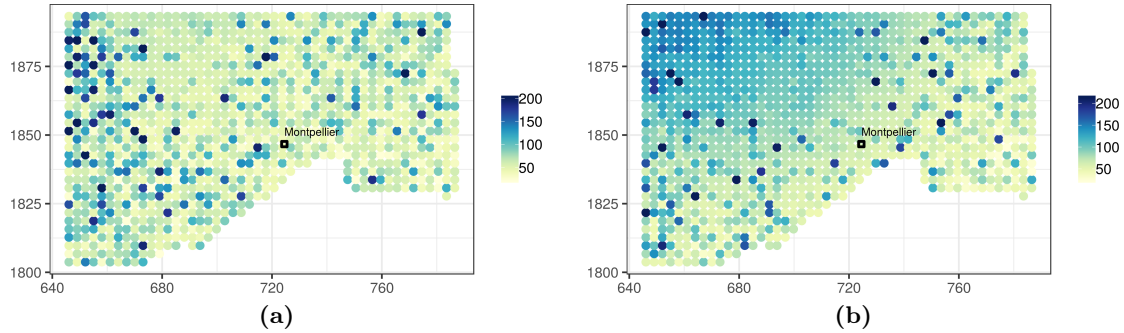


(b) The reference station  $i$  is on the mountain range.

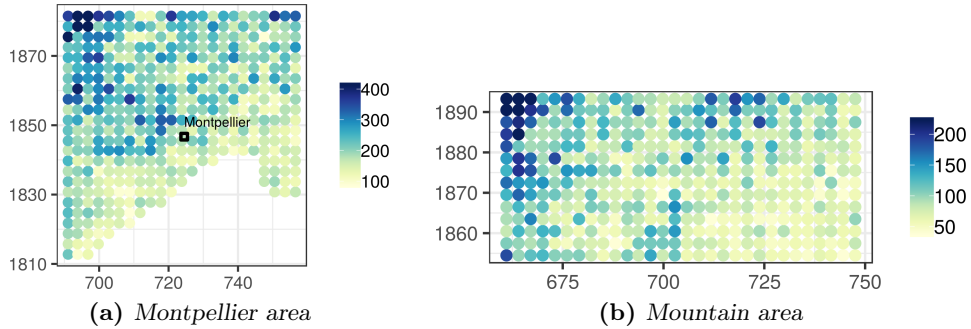
**Figure 11:** Maps of Brown-Resnick upper tail dependence coefficient estimates  $\hat{\chi}_{ij}$ , obtained by estimating the madogram with a sample of size 1000 (see Eq. (7)). The reference station  $i$  is shown by a white cross.

### 5.3 Complete spatial models

Simulations from the two complete fitted spatial models for maxima are illustrated in Fig. 12 and 13. In the former case, the dependence structure is modeled by the spatialized XGumbel copula whereas in the latter case, it is modeled by the Brown-Resnick process. In both cases, univariate marginal distributions are provided by the response surfaces for the GEV parameters from § 5.1. In the spatialized XGumbel copula case, the best set of parameters from the posterior distribution of the ABC scheme is used. The location of the disk center of the extra-parameter mapping, see Eq. (12), on the mountain range can easily be detected in Fig. 12. In the Brown-Resnick case, the grid for the simulation is restricted to two sub-areas (a first one encompassing the city of Montpellier and a second one in the mountain area) owing to computing limitations (Ribatet, 2018).



**Figure 12:** Two data-scale simulations of the spatialized XGumbel copula combined with the response surfaces for the GEV over the study area. The color scale is capped at the 99% quantile of the simulated values.



**Figure 13:** Two data-scale simulations of the Brown-Resnick process combined with the response surfaces for the GEV over the study area. The color scale is capped at the 99% quantile of the simulated values. Two sub-areas are selected as the implementation of the Brown-Resnick process we used did not allow simulation on the full area (Ribatet, 2018)

We then compare the two complete fitted spatial models for maxima in terms of a quantity that could be useful for regional hazard analysis. This quantity is related to the multivariate extension of the upper tail dependence coefficient termed  $m$ -dimensional joint tail dependence coefficients (Wadsworth & Tawn, 2013). Higher dimensional properties of the models can be



investigated as these coefficients involve  $m$ -dimensional distributions instead of being limited to bivariate marginals as is the case for the upper tail dependence coefficient.

More precisely, we focus on trivariate properties, i.e.  $m = 3$ . Let  $Y_k$ ,  $Y_i$  and  $Y_j$  represent annual maxima at three sites  $k$ ,  $i$  and  $j$  respectively. Moreover, let  $R_k^T$ ,  $R_i^T$  and  $R_j^T$  be the  $T$ -year return level at each site. The quantity of interest for our regional hazard analysis is the 3-dimensional joint tail dependence coefficient that is defined as follows :

$$\mathbb{P}(Y_i > R_i^T, Y_j > R_j^T | Y_k > R_k^T). \quad (14)$$

Note that, given that the univariate marginals are the same in both spatial models, differences in terms of the coefficient in Eq. (14) are only caused by differences in the spatial dependence structure. The interpolation ability of the spatial models is evaluated by setting the conditioning site  $k$  in Eq. (14) as one of the six validation stations not used for model inference (see the stations depicted with red filled circles in Fig. 2). For the other two sites  $i$  and  $j$  in Eq. (14), we selected two nearby sites from the calibration stations within a 20 km radius with the most complete observation record. These calibration stations wear numbers from 1 to 11 in Fig. 2.

In Fig. 14 and 15, empirical and theoretical estimates of the 3-dimensional joint tail dependence coefficient from Eq. (14) are compared, with each of the six validation stations taken as the conditioning site  $k$  in turn. Empirical estimates, colored in light blue in both cases, are obtained by computing the sample proportions from the observed annual maxima with return levels determined from empirical quantiles. As there is no closed-form expression for Eq. (14), theoretical estimates are also deduced from proportions of samples of size 10000 simulated from each of the spatial models (GEV margins combined with either the spatialized XGumbel copula, in Fig. 14, or the Brown-Resnick process, in Fig. 15), with the return levels provided by the response surfaces for the GEV parameters (see § 4.1, Eqs (9)-(11)). For each return level, there are 100 theoretical estimates corresponding to different sets of parameters (from the posterior distribution resulting from the ABC scheme for the spatialized XGumbel copula or from the non-parametric bootstrap for the Brown-Resnick process). In addition, for the spatialized XGumbel copula, the median of the theoretical estimates of the 3-dimensional joint tail dependence coefficient is shown in black in Fig. 14 while, for the Brown-Resnick process, the theoretical estimates of the fit on the original calibration data is shown in black in Fig. 15.

As the dependence structure in both spatial models is max-stable, both theoretical coefficient estimates stabilize at longer return periods (greater than five years). Being a stationary model, the Brown-Resnick process always yields estimates at about the same level, wherever is located the conditioning site. In contrast, the spatialized XGumbel copula, thanks to its non-stationarity, can adapt to the location of the conditioning site. For instance, the estimates stabilize at about 0.5 for the validation station labeled “a” in Fig. 14a whereas they stabilize at about 0.3 for the validation station labeled “b” in Fig. 14b. The empirical estimates are mostly contained within the spread of the theoretical estimates for both models, although it happens in a few instances

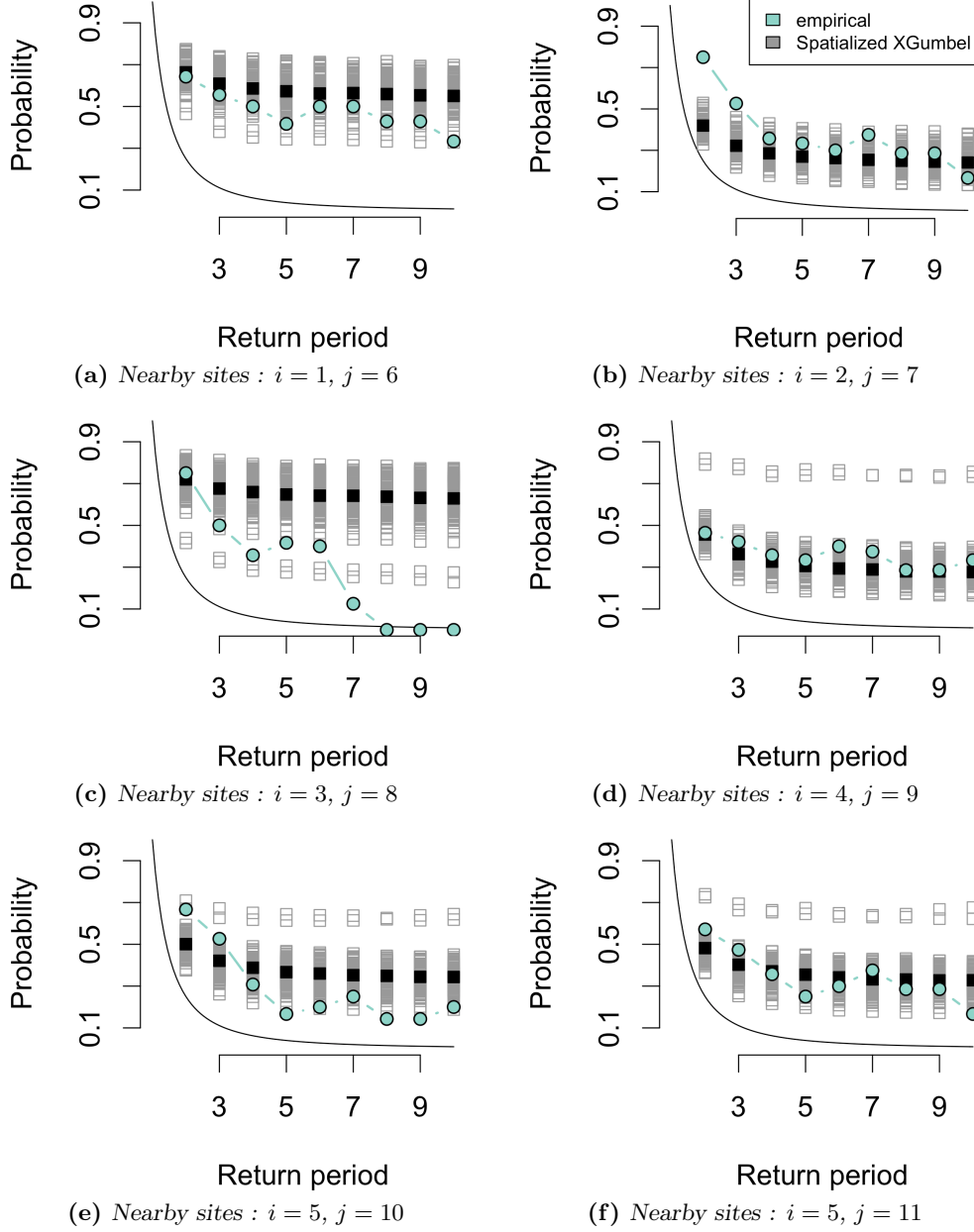
that they fall outside, e.g. in Fig. 14c or Fig. 15c and 15e. For some conditioning sites, e.g. Fig. 14d, the spatialized XGumbel copula yielded two estimates that are far away from the others. These correspond to parameter vectors for which the disk centers are located near the coastline, see Fig. 8c.

## 6 Conclusion

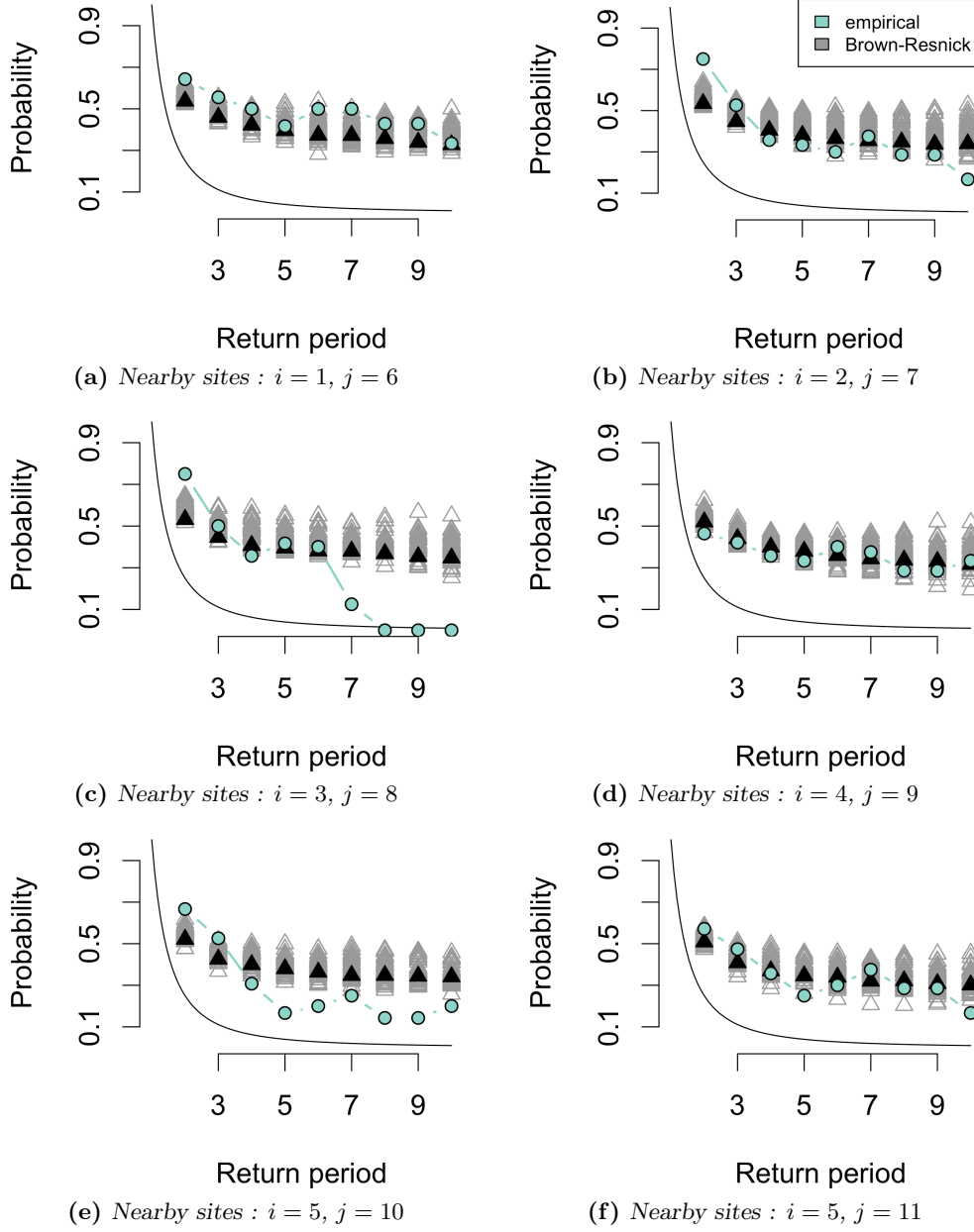
We proposed a spatial extension of the XGumbel copula that relies on the definition of the extra-parameters as a mapping of geographical covariates. Although the XGumbel copula could in principle be fitted in high dimension, the large number of extra-parameters, corresponding to the number of sites in a spatial application, might hamper inference. The spatialized XGumbel copula is more parsimonious as it requires only a four parameter vector  $\psi_{\text{spat}} = (\beta_B, \delta, \mu_x, \mu_y)$ , independently of the number of sites. We designed the extra-parameter mapping shaped as a disk by relating the behavior of the strength of dependence between two sites, as characterized by the upper tail dependence coefficient  $\chi$ , to desirable spatial properties. In particular, we focused on the pattern of decrease of the dependence with the distance by using  $\chi$  estimates for five distance classes. These distance class  $\chi$  estimates also serve as summary statistics in an ABC scheme to infer the parameters of the spatialized XGumbel copula. The spatialized XGumbel copula, when combined with a spatial regression model for the GEV marginal distributions, yields well-defined MEV distributions for any set of sites. Therefore, simulation is possible everywhere in the study area.

The proposed spatialized XGumbel copula is evaluated and compared with a Brown-Resnick process on annual maxima of daily precipitation totals in a region of the French Mediterranean with 177 gauged stations, six of which are kept for validation purposes. A vector generalized linear (VGLM) model is considered for the interpolation of the GEV parameters to model the univariate marginal distributions. The goodness-of-fit of the VGLM model is evaluated in terms of return levels at the validation stations. We analyzed the posterior distribution of the spatialized XGumbel copula parameters resulting from the rejection ABC scheme. Except for the parameter  $\beta_B$ , a global parameter inherited from one of the Gumbel copulas of the XGumbel, the ABC scheme inferred interpretable parameters. The Brown-Resnick process is fitted by pairwise log-likelihood minimization and uncertainty estimates are obtained by performing the fit on bootstrap resamples.

Comparison between the spatialized XGumbel copula and the Brown-Resnick process shows the following. The pattern of decrease of the strength of dependence is well reproduced in both cases. Owing to asymptotic dependence, the strength of dependence remains constant at extreme levels. However, strong non-stationarity patterns in the strength of dependence are present for the spatialized XGumbel copula whereas the Brown-Resnick process, by construction, has none. Simulations from both complete spatial models for maxima, GEV marginals together with spatial dependence structure, were provided for illustrations. A further downside of the Brown-Resnick



**Figure 14:** 3-dimensional joint tail dependence coefficient estimates, see Eq. (14), with respect to return periods  $T$  on the x-axis. The Spatialized XGumbel estimates (gray squares) are proportions of simulated samples of size 10000 for each of the 100 sets of parameters of the posterior distribution. The median estimates are shown as black squares. The conditioning site  $k$  is one of the six validation stations, red filled circles in Fig. 2 wearing the letter corresponding to the panel. The other two sites  $i$  and  $j$  are calibration stations wearing numbers in Fig. 2 that are reported under each panel. The black line corresponds to the perfect independence case.



**Figure 15:** 3-dimensional joint tail dependence coefficient estimates, see Eq. (14), with respect to return periods  $T$  on the x-axis. The Brown-Resnick estimates (gray triangles) are proportions of simulated samples of size 10000 for each of the 100 sets of parameters of the non-parametric bootstrap. The estimates from the fit on the original data are shown as black triangles. The conditioning site  $k$  is one of the six validation stations, red filled circles in Fig. 2 wearing the letter corresponding to the panel. The other two sites  $i$  and  $j$  are calibration stations wearing numbers in Fig. 2 that are reported under each panel. The black line corresponds to the perfect independence case.

process is that simulation on the full grid covering the study area was not possible due to computing limitations. We proposed a regional hazard analysis based on 3-dimensional joint tail dependence coefficients. These involve the trivariate distributions at three stations one of which is taken as a validation station and the other two are neighbor calibration stations. The simulations and the regional hazard analysis also highlight differences due to the presence or absence of non-stationarity in the dependence structures.

Earlier propositions to extend copulas to the spatial framework are based on a parametrization in terms of distance but are not especially targeting spatial maxima (Bárdossy & Li, 2008; Gräler, 2014; Krupskii *et al.*, 2018). The construction of the XGumbel copula as the maximum between two weighted random variables is directly related to the max-mixture model (Wadsworth & Tawn, 2012; Bacro *et al.*, 2016). Instead of relying on processes with well-defined spatial dependence structures, the spatial dependence of the spatialized XGumbel copula is driven by the mapping of extra-parameters. The shape of the mapping determines the non-stationarity pattern of the dependence structure. A completely different proposition to introduce non-stationarity in the dependence structure for spatial maxima was put forward in Huser & Genton (2016) concerning max-stable processes.

Further analyses are needed to develop and test different shapes for the extra-parameter mapping. An interesting development, that was already considered in preliminary work, would be to let the shape of the mapping change from year to year, leading to a conditionally max-stable model. This would allow, in particular, to let the areas of stronger and weaker dependence vary from one year to another. Another way to achieve this, while keeping the max-stable property, would be to iterate the extra-parametrization (or maximization) operation in § 2.1, e.g. by assuming that the phenomenon of interest can be modeled as :

$$\max\{[\max(\mathbf{U}^{1/\mathbf{b}}, \mathbf{W}^{1/(1-\mathbf{b})})]^{1/\mathbf{a}}, \mathbf{V}^{1/(1-\mathbf{a})}\},$$

with  $\mathbf{U} \sim C_{\beta_A}$ ,  $\mathbf{V} \sim C_{\beta_B}$ ,  $\mathbf{W} \sim C_{\beta_C}$ ,  $\beta_A, \beta_B, \beta_C \geq 1$  are Gumbel copula parameters and  $\mathbf{a}, \mathbf{b} \in [0, 1]^d$  two extra-parameter vectors. Although pairwise log-likelihood inference is widely used, ABC inference scheme yields promising results. For complex dependence structure models, even pairwise log-likelihood might be intractable. We have used summary statistics that convey information on the strength of extremal dependence. Other statistics, for instance, conveying information on asymmetry or non-stationarity, as well as other ways to compute distances, such as the Wasserstein distance, could be considered (Arbel *et al.*, 2019).

# References

- Arbel, J., Crispino, M., & Girard, S. 2019. Dependence properties and Bayesian inference for asymmetric multivariate copulas. *Journal of Multivariate Analysis*, **174**(Nov.), 104530:1–20.
- Bacro, J.-N., Gaetan, C., & Toulemonde, G. 2016. A flexible dependence model for spatial extremes. *Journal of Statistical Planning and Inference*, **172**, 36–52.
- Bárdossy, A., & Li, J. 2008. Geostatistical interpolation using copulas. *Water Resources Research*, **44**, 1–15.
- Beaumont, M. A. 2010. Approximate Bayesian computation in evolution and ecology. *Annual review of ecology, evolution, and systematics*, **41**, 379–406.
- Beirlant, J., Goegebeur, Y., Teugels, J., Segers, J., De Waal, D., & Ferro, C. 2004. *Statistics of Extremes Theory and Applications*. John Wiley & Sons.
- Benichou, P., & Le Breton, O. 1987. Prise en compte de la topographie pour la cartographie des champs pluviométriques statistiques. Une application de la méthode Aurelhy: la cartographie nationale de champs de normales pluviométriques. *Météorologie*.
- Blanchet, J., & Creutin, J.-D. 2017. Co-Occurrence of Extreme Daily Rainfall in the French Mediterranean Region. *Water Resources Research*, **53**, 9330–9349.
- Blanchet, J., & Davison, A. C. 2011. Spatial modeling of extreme snow depth. *The Annals of Applied Statistics*, **5**, 1699–1725.
- Braud, I., Ayral, P.-A., Bouvier, C., Branger, F., Delrieu, G., Le Coz, J., Nord, G., Vandervaere, J.-P., Anquetin, S., Adamovic, M., Andrieu, J., Batiot, C., Boudevillain, B., Brunet, P., Carreau, J., Confoland, A., Didon-Lescot, J.-F., Domergue, J.-M., Douvinet, J., Dramais, G., Freydier, R., Gérard, S., Huza, J., Leblois, E., Le Bourgeois, O., Le Boursicaud, R., Marchand, P., Martin, P., Nottale, L., Patris, N., Renard, B., Seidel, J.-L., Taupin, J.-D., Vannier, O., Vincendon, B., & Wijbrans, A. 2014. Multi-scale hydrometeorological observation and modelling for flash flood understanding. *Hydrology and Earth System Sciences*, **18**, 3733–3761.
- Brown, B., & Resnick, S. 1977. Extremes values of independent stochastic processes. *Journal of Applied Probability*, **14**, 732–739.
- Brunet, P., Bouvier, C., & Neppel, L. 2018. Retour d’expérience sur les crues des 6 et 7 octobre 2014 à Montpellier-Grabels (Hérault, France) : caractéristiques hydro-météorologiques et contexte historique de l’épisode. *Géographie physique et environnement*, **12**, 43–59.
- Carreau, J., Naveau, P., & Neppel, L. 2017. Partitioning into hazard subregions for regional peaks-over-threshold modeling of heavy precipitation. *Water Resources Research*, **53**, 4407–4426.

- 1 Coles, S. 2001. *An Introduction to Statistical Modeling of Extreme Values*. Springer Series in  
2 Statistics. Springer.
- 3 Coles, S. G., Heffernan, J. E., & Tawn, J. A. 1999. Dependence measures for extremes value  
4 analyses. *Extremes*, **2**, 339–365.
- 5 Cooley, D., Naveau, P., & Poncet, P. 2006. Variograms for spatial max-stable random fields.  
6 *Pages 373–390 of: Dependence in probability and statistics*. Lect. Notes Stat., vol. 187.  
7 Springer, New York.
- 8 Davison, A. C., & Gholamrezaee, M. M. 2012. Geostatistics of extremes. *Proceedings of the*  
9 *Royal Society London, Series A*, **468**, 581–608.
- 10 Davison, A. C., Padoan, S. A., & Ribatet, M. 2012. Statistical modeling of spatial extremes.  
11 *Statistical science*, **27**, 161–186.
- 12 de Haan, L. 1984. A spectral representation for max-stable processes. *Annals of Probability*, **12**,  
13 1194–1204.
- 14 Delrieu, G., Nicol, J., Yates, E., Kirstetter, P.-E., Creutin, J.-D., Anquetin, S., Obled, C.,  
15 Saulnier, G.-M., Ducrocq, V., Gaume, E., Payrastre, O., Andrieu, H., Ayrat, P.-A., Bouvier,  
16 C., Neppel, L., Livet, M., Lang, M., du Châtelet, J. P., Walpersdorf, A., & Wobrock, W. 2005.  
17 The Catastrophic Flash-Flood Event of 8-9 September 2002 in the Gard Region, France: A  
18 First Case Study for the Cévennes-Vivarais Mediterranean Hydrometeorological Observatory.  
19 *Journal of Hydrometeorology*, **6**, 34–52.
- 20 Durante, F., & Salvadori, G. 2010. On the construction of multivariate extreme value models  
21 via copulas. *Environmetrics*, **21**(2), 143–161.
- 22 Erhardt, R. J., & Sisson, S. A. 2016. Modelling extremes using approximate Bayesian computa-  
23 tion. *Pages 281–306 of: Extreme value modeling and risk analysis*. CRC Press, Boca Raton,  
24 FL.
- 25 Erhardt, R. J., & Smith, R. L. 2012. Approximate Bayesian computing for spatial extremes.  
26 *Computational Statistics & Data Analysis*, **56**, 1468–1481.
- 27 Fisher, R.A., & Tippett, L.H.C. 1928. Limiting forms of the frequency of the largest or smallest  
28 member of a sample. *Math. Proc. Cambridge Philos. Soc.*, **24**, 180–190.
- 29 Friedman, J., Hastie, T., & Tibshirani, R. 2010. Regularization Paths for Generalized Linear  
30 Models via Coordinate Descent. *Journal of Statistical Software*, **33**, 1–22.
- 31 Genest, C., & Favre, A.-C. 2007. Everything you always wanted to know about copula modeling  
32 but were afraid to ask. *Journal of hydrologic engineering*, **12**, 347–368.
- 33 Gnedenko, B.V. 1943. Sur la distribution limite du terme maximum d’une série aléatoire. *Ann.*  
34 *of Math.*, **44**, 423–453.

- 1 Gräler, B. 2014. Modelling skewed spatial random fields through the spatial vine copula. *Spatial*  
2 *Statistics*, **10**, 87–102.
- 3 Gumbel, E.J. 1958. *Statistics of Extremes*. New York: Columbia Univ. Press.
- 4 Huser, R., & Genton, M. G. 2016. Non-Stationary Dependence Structures for Spatial Extremes.  
5 *Journal of Agricultural, Biological, and Environmental Statistics*, **21**(3), 470–491.
- 6 Krupskii, P., Huser, R., & Genton, M. 2018. Factor copula models for replicated spatial data.  
7 *Journal of the American Statistical Association*, **113**, 467–479.
- 8 Lee, Xing Ju, Hainy, Markus, McKeone, James P., Drovandi, Christopher C., & Pettitt, An-  
9 thony N. 2018. ABC model selection for spatial extremes models applied to South Australian  
10 maximum temperature data. *Comput. Statist. Data Anal.*, **128**, 128–144.
- 11 Liebscher, E. 2008. Construction of asymmetric multivariate copulas. *J. Multivariate Anal.*,  
12 **99**(10), 2234–2250.
- 13 Marcon, G., Padoan, S. A., Naveau, P., Muliere, P., & Segers, J. 2017. Multivariate nonpara-  
14 metric estimation of the Pickands dependence function using Bernstein polynomials. *Journal*  
15 *of Statistical Planning and Inference*, **183**, 1–17.
- 16 Pickands, J. 1981. Multivariate extreme value distributions. *Proceedings of the 43rd session of*  
17 *the International Statistical Institute, Bulletin de l'Institut International de Statistique*, **49**,  
18 859–878, 894–902.
- 19 Ribatet, M. 2018. *SpatialExtremes: Modelling Spatial Extremes*. R package version 2.0-7.
- 20 Salvadori, G., & De Michele, C. 2010. Multivariate multiparameter extreme value models and  
21 return periods: A copula approach. *Water Resources Research*, **46**, 1–11.
- 22 Schwarz, G. 1978. Estimating the dimension of a model. *Annals of Statistics*, **6**, 461–464.
- 23 Sibuya, M. 1960. Bivariate extreme statistics. *Annals of the Institute of Statistical Mathematics*,  
24 **11**, 195–210.
- 25 Thibaud, E., Mutzner, R., & Davison, A. C. 2013. Threshold modeling of extreme spatial rainfall.  
26 *Water Resources Research*, **49**, 4633–4644.
- 27 Vannitsem, S., & Naveau, P. 2007. Spatial dependences among precipitation maxima over Bel-  
28 gium. *Nonlinear Processes in geophysics*, **14**, 621–630.
- 29 Wadsworth, J. L., & Tawn, J. A. 2012. Dependence modelling for spatial extremes. *Biometrika*,  
30 **99**, 253–272.
- 31 Wadsworth, J. L., & Tawn, J. A. 2013. A new representation for multivariate tail probabilities.  
32 *Bernoulli*, **19**, 2689–2714.



- 1 Yee, T. W. 2015. *Vector Generalized Linear and Additive Models: With an Implementation in*  
2 *R*. New York, USA: Springer.
- 3 Yee, T. W., & Stephenson, A. G. 2007. Vector generalized linear and additive extreme value  
4 models. *Extremes*, **10**, 1–19.