ARTICLE IN PRESS

Journal of Statistical Planning and Inference ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

A flexible dependence model for spatial extremes

Jean-Noel Bacro [a,*], Carlo Gaetan [b], Gwladys Toulemonde [a]

[a] IMAG, Université de Montpellier, Montpellier, France
[b] DAIS, Università Ca' Foscari - Venezia, Italy

ARTICLE INFO

Article history:
Received 18 August 2014
Received in revised form 30 November 2015
Accepted 9 December 2015
Available online xxxx

Keywords:
Spatial extremes
Asymptotic independence
Max-stable processes

ABSTRACT

Max-stable processes play a fundamental role in modeling the spatial dependence of extremes because they appear as a natural extension of multivariate extreme value distributions. In practice, a well-known restrictive assumption when using max-stable processes comes from the fact that the observed extremal dependence is assumed to be related to a particular max-stable dependence structure. As a consequence, the latter is imposed to all events which are more extreme than those that have been observed. Such an assumption is inappropriate in the case of asymptotic independence. Following recent advances in the literature, we exploit a max-mixture model to suggest a general spatial model which ensures extremal dependence at small distances, possible independence at large distances and asymptotic independence at intermediate distances. Parametric inference is carried out using a pairwise composite likelihood approach. Finally we apply our modeling framework to analyze daily precipitations over the East of Australia, using block maxima over the observation period and exceedances over a large threshold.

1. Introduction

The last decade has registered a considerable effort to model extremes of data collected through a collection of sites and the interested reader is referred to Bacro and Gaetan (2012) and Davison et al. (2012) for recent reviews.

If the main interest is producing return level maps, the modeling issue is mainly concentrated on relating the parameters of the marginal distributions in each site to geographical covariates. In case of a residual dependence, uncertainty of the estimates can be further adjusted (Fawcett and Walshaw, 2007). Additionally, this modeling approach can be extended to be hierarchical adding a layer for incorporating spatial dependence through a spatial random process (Casson and Coles, 1999; Cooley et al., 2007; Gaetan and Grigoletto, 2007; Sang and Gelfand, 2010).

If we are interested in modeling the joint occurrence of extremes over a region, then the dependence structure of a multivariate variable needs to be explicitly stated. In this case the usual modeling strategy consists in two steps (1) estimating the marginal distribution and (2) characterizing the dependence via a model issued by the multivariate extreme value (MEV) theory (see for example Beirlant et al., 2004, and the references therein). These two steps can be integrated in a proper inferential analysis (Padoan et al., 2010; Ribatet et al., 2012).

The MEV theory deals with the tail behavior of a multivariate distribution from which we pretend that a sample is drawn and distinguishes three different forms of extremal dependence: asymptotic dependence, asymptotic independence and exact independence.

* Corresponding author.
E-mail address: jean-noel.bacro@univ-montp2.fr (J.-N. Bacro).

http://dx.doi.org/10.1016/j.jspi.2015.12.002
0378-3758/

Please cite this article in press as: Bacro, J.-N., et al., A flexible dependence model for spatial extremes. J. Statist. Plann. Inference (2016), http://dx.doi.org/10.1016/j.jspi.2015.12.002

Asymptotic independence and dependence between a pair of random variables Z_1 and Z_2, with marginal distributions F_1 and F_2, can be defined in terms of

$$\chi = \lim_{u \to 1^-} \Pr(F_1(Z_1) > u | F_2(Z_2) > u), \tag{1}$$

where $\chi = 0$ and $\chi > 0$ represent asymptotic independence and dependence, respectively.

An example of a multivariate distribution which is asymptotically independent is given by the multivariate Gaussian distribution when the components are not perfectly correlated (Sibuya, 1960).

However the multivariate framework is inadequate for predicting or simulating values at unobserved sites. Therefore in the last years there was a general consensus in representing extreme spatial variability by max-stable processes (de Haan, 1984; Smith, 1990; Schlather, 2002; Kabluchko et al., 2009; Opitz, 2013) that are an infinite dimensional generalization of multivariate distributions for the maxima. The drawback of these processes is that they admit only two types of dependence structures in their finite dimensional distributions: asymptotic dependence or exact independence. This restriction is constraining when we model the tail behavior of the multivariate distribution of the data because it is difficult to assess in practice whether a data set should be modeled using an asymptotically dependent or asymptotically independent model (see Thibaud et al., 2013; Davison et al., 2013, for recent examples of these difficulties).

For coping with dependence structures that have not converged to their limiting form at observable levels, Wadsworth and Tawn (2012) introduced the hybrid spatial dependence model. The model originates from a max-mixture, namely $Z(s) = \max(\beta X(s), (1-\beta)Y(s))$ with $0 \le \beta \le 1$, of an asymptotically dependent process X (a max-stable process, for instance), and an asymptotically independent process Y.

In applications such as environmental ones different types of extremal dependencies could be present according the distance between two locations. As motivating example we shall consider winter daily cumulative rainfall, recorded at 31 monitoring sites in the East of Australia (Fig. 3). We quantify the strength of the asymptotic dependence of a pair of random variables $Z(s)$ and $Z(s+h)$, located at sites s and $s+h$, assuming the same marginal distribution F, by means of the dependence measure (Coles et al., 1999)

$$\chi(h,u) = 2 - \frac{\log \Pr(F(Z(s+h)) < u, F(Z(s)) < u)}{\log \Pr(F(Z(s)) < u)}, \quad 0 \le u \le 1. \tag{2}$$

In case of asymptotic independence, $\chi(u,h) \simeq 0$ for $u \simeq 1$ and $\chi(h,u)$ is zero for exactly independent variables for all u. Discriminating between asymptotic dependence and asymptotic independence by means of the estimates of $\chi(h,u)$ or $\chi(h)$, its limit version when $u \to 1^-$, is not easy, especially for rainfall extremes (Serinaldi et al., 2014). However the nonparametric estimates of $\chi(h,u)$ (see Fig. 5) suggest that asymptotic dependence is present up to a distance r_1 and asymptotic independence prevails for distances between r_1 and r_2 whereas for larger distances, exact independence could also be conjectured ($r_1 = 500$ km and $r_2 = 1000$ km, approximately, in Fig. 5).

In Wadsworth and Tawn (2012) the authors discuss the idea of having asymptotic dependence present up to a finite lag but in the reported examples asymptotic dependence or asymptotic independence are present for any distance, even if it is dimming with the distance. Following their idea, the contribution of the present paper is to consider examples of max-mixture between a max-stable process, that yields exact independence between maxima after a finite spatial lag and an asymptotically independent process that may or not yield exact independence between observations after that lag.

The max-stable process stems from the construction in Schlather (2002, p. 39) and, as example, we use the truncated Extremal Gaussian process (see also Davison and Gholamrezaee, 2012). For the asymptotically independent process we can consider stationary spatial processes with bivariate distributions satisfying only a general condition on the bivariate survivor functions (Ledford and Tawn, 1996, 1997). We exemplify our construction by means of a marginal transformed Gaussian process with possible finite range covariance function and of an inverse truncated Extremal Gaussian process.

The paper is organized as follows. In Section 2 we briefly introduce the max-stable and asymptotically independent processes and some classical extremal dependence measures. Our modeling proposal and its main properties are detailed in Section 3 and a pairwise likelihood approach is presented for the statistical inference. In Section 4 we show, by means of a simulation study, that this approach seems effective in order to identifying different max-mixture models. Section 5 is devoted to illustrate the modeling issues related to our motivating example. Concluding remarks and some perspectives are addressed in Section 6.

2. Spatial extremes modeling

2.1. Models for asymptotic dependence

Max-stable processes (de Haan, 1984) are an infinite-dimensional generalization of multivariate extreme value theory. The stochastic process $X = \{X(s), s \in \mathcal{D}\}$, where $\mathcal{D}$ is a spatial domain, is a max-stable process if and only if there exist functions $a_n(\cdot) > 0$ and $b_n(\cdot)$ on $\mathbb{R}$ such that

$$\max_{1 \le i \le n} \frac{X_i(s) - b_n(s)}{a_n(s)} \overset{d}{=} X(s)$$

Please cite this article in press as: Bacro, J.-N., et al., A flexible dependence model for spatial extremes. J. Statist. Plann. Inference (2016), http://dx.doi.org/10.1016/j.jspi.2015.12.002

where $X_1, X_2, \ldots$ are independent copies of X. In the sequel and without loss of generality, $\mathcal{D}$ is a subset of $\mathbb{R}^2$ and univariate margins of max-stable processes are assumed to be unit Fréchet, i.e. $\Pr(X(s) \le x) = \exp(-1/x), x > 0$.

A max-stable process has a spectral representation (de Haan, 1984; Schlather, 2002). Assume that $r_i, i \ge 1$, are points of a Poisson process on $(0, \infty)$ with intensity dr. Let $W_i, i \ge 1$, be independent and identically distributed (i.i.d.) copies of a real valued continuous random function $W = \{W(s), s \in \mathcal{D}\}$, independent of the $\{r_i\}$ and such that $\mathbb{E}[W^+(s)] = \mu \in (0, \infty)$, where $W^+(s) = \max(W(s), 0)$. Then

$$X(s) = \mu^{-1} \max_{i \ge 1} W_i^+(s)/r_i \tag{3}$$

is a max-stable process with unit Fréchet margins.

Choosing a particular expression for W_i leads to known examples of max-stable processes: the Gaussian extreme value process (Smith, 1990), the extremal Gaussian process (Schlather, 2002), the Brown–Resnick process (Kabluchko et al., 2009) and the extremal t process (Opitz, 2013).

In the sequel we focus on a particular instance of a max-stable process: the Truncated Extremal Gaussian (TEG) process. The TEG process has been introduced by Schlather (2002) and has been exemplified by Davison and Gholamrezaee (2012). As in the extremal Gaussian model the model derives from an underlying Gaussian process censored on a compact random set.

Let $r_i, i \ge 1$, be defined as previously and consider $W_i(s) = c \max(0, \varepsilon_i(s)) \mathbb{I}_{B_i}(s - U_i)$ with ε_i independent copies of a stationary Gaussian process $\varepsilon = \{\varepsilon(s), s \in \mathcal{D}\}$ with zero mean, unit variance and correlation function $\rho(\cdot)$, $\mathbb{I}_B$ is the indicator function of a compact random set $B \subset \mathcal{D}$, of which B_i are independent replicates and U_i are points of a homogeneous Poisson process of unit rate on $\mathcal{D}$, independent of the ε_i. Choosing the constant c such that $c^{-1} = \mathbb{E}\left(\max\{W_i(s), 0\} \mathbb{I}_{B_i}(x - U_i)\right)$, the TEG process is defined as

$$X(s) = \max_{i \ge 1} \frac{W_i(s)}{r_i}. \tag{4}$$

The marginal distribution of X is unit Fréchet and the bivariate one is given by

$$\begin{aligned} &\Pr\left(X(s) \le t_1, X(s+h) \le t_2\right) \\ &= \exp\left\{ -\left(\frac{1}{t_1} + \frac{1}{t_2}\right)\left[1 - \frac{\alpha(h)}{2}\left(1 - \left(1 - 2\frac{(\rho(h)+1)t_1 t_2}{(t_1+t_2)^2}\right)^{1/2}\right)\right]\right\} \end{aligned} \tag{5}$$

where $\alpha(h) = \mathbb{E}[|B \cap (h + B)|]/\mathbb{E}[|B|]$. If B is a disk of fixed radius r, $\alpha(h)$ can be approximated by

$$\alpha(h) \simeq (1 - \|h\|/(2r))\, \mathbb{I}_{[0,2r]}(\|h\|). \tag{6}$$

2.2. Models for asymptotic independence

A multivariate vector is asymptotically independent (AI) if and only if all its pairs of components are AI (de Oliveira, 1962). As a consequence, if all the bivariate distributions of a stochastic process are AI, the stochastic process is said to be AI.

For modeling AI we assume a specific model for bivariate joint tails as described in more detail in Ledford and Tawn (1996).

We assume that $\{Z(s), s \in \mathcal{D}\}$ is a stationary spatial process with unit Fréchet margins. Under weak conditions, Ledford and Tawn (1997, 1998) showed that the bivariate tail distribution of a pair of observations at s and $s + h$ admits an approximation such that

$$\Pr(Z(s) > z_1, Z(s+h) > z_2) \sim z_1^{-c_h^{(1)}} z_2^{-c_h^{(2)}} \mathcal{L}'_h(z_1, z_2)$$

for z_1 and z_2 simultaneously large, where $0 < 1/(c_h^{(1)} + c_h^{(2)}) \le 1$ and $\mathcal{L}'_h \ne 0$ a bivariate slowly varying function (Bingham et al., 1987, Appendix 1) with limit function g_h such that for all $x > 0, y > 0$ and $c > 0, g_h(x, y) = \lim_{t \to \infty} \mathcal{L}'_h(tx, ty)/\mathcal{L}'_h(x, y)$ and $g_h(cx, cy) = g_h(x, y)$. The homogeneity property of g_h implies that $g_h(x, y) = g_h^*(w)$ with $w = x/(x+y) \in [0, 1]$ where the function g_h^*, often called the ray dependence function, is assumed to be a slowly varying function at 0 and 1.

Assuming $z_1 = z_2 = z$ leads to the Ledford–Tawn (LT) model for bivariate joint tails (Ledford and Tawn, 1996):

$$\Pr(Z(s) > z, Z(s+h) > z) \sim z^{-1/\eta(h)} \mathcal{L}_h(z) \quad \text{for } z \to \infty \tag{7}$$

where $\mathcal{L}_h(\cdot)$ is a univariate slowly varying function. The coefficient $\eta(h)$ varies between 0 and 1 and determines the decay rate of the bivariate tail probability $\Pr(Z(s) > z, Z(s+h) > z)$ for large z. Despite its simplicity, Eq. (7) appears as a very general model for bivariate joint tails which can provide, as detailed below, a measure of the extremal dependence between $Z(s)$ and $Z(s+h)$ through the coefficient $\eta(h)$. Asymptotic independence corresponds to $\eta(h) < 1$ and in such a case, $\eta(h)$ measures the degree of dependence in the asymptotic independence at h, where $\eta(h) > 1/2$ and $\eta(h) < 1/2$ indicate

Please cite this article in press as: Bacro, J.-N., et al., A flexible dependence model for spatial extremes. J. Statist. Plann. Inference (2016), http://dx.doi.org/10.1016/j.jspi.2015.12.002

positive and negative association, respectively. When the variables $Z(s)$ and $Z(s+h)$ are independent $\eta(h) = 1/2$. There are few examples of AI processes in the literature. In the sequel three asymptotically independent processes are considered with explicit expressions of (7).

Example 1: *Stationary Gaussian process*

Let $\{Y(s), s \in \mathcal{D}\}$ be a stationary Gaussian process with zero mean, unit variance and correlation function $\rho(h)$. Because bivariate Gaussian variables are AI provided that they are not perfectly correlated (Sibuya, 1960), the spatial process $Z'(s) = -1/\log(\Phi(Y(s)))$ has unit Fréchet margins and verifies (Ledford and Tawn, 1996)

$$\Pr(Z'(s) > z, Z'(s+h) > z) \sim C_h z^{-2/\{1+\rho(h)\}} (\log z)^{-\rho(h)/\{1+\rho(h)\}}$$

with $C_h = (1+\rho(h))^{3/2}(1-\rho(h))^{-1/2}(4\pi)^{-\rho(h)/\{1+\rho(h)\}}$. So $\eta(h) = \{1+\rho(h)\}/2$.

Example 2: *Inverse max-stable process*

The inverse max-stable process (Wadsworth and Tawn, 2012) is obtained by simply inverting a max-stable process. More precisely, let $\{X(s), s \in \mathcal{D}\}$ be a max-stable process defined as in (3). Then the process

$$Z(s) = -1/\log[1 - \exp\{-1/X(s)\}]$$

is an AI process with Fréchet margins. For any fixed h, the tail dependence coefficient is $\eta(h) = 1/\theta(h)$ where $\theta(h)$ is the extremal coefficient of the max-stable process.

Example 3: *Max-Gaussian ratio process*

Recently Padoan (2013) introduced a new family of spatial processes whose univariate limit distributions are unit Fréchet and bivariate distributions are able to cope with different levels of dependence according to the magnitude of extreme events. Such processes, called max-Gaussian ratio processes, are obtained as pointwise maxima of samples from a ratio of Gaussian processes with common correlation function. For every $n \in \mathbb{N}$, let $\{U_n(s), s \in \mathcal{D}\}$ and $\{V_n(s), s \in \mathcal{D}\}$ be two independent Gaussian processes on $\mathcal{D}$ with mean zero, unit variance and common correlation function, $\rho_n(h)$, such that

$$\rho_n(h) = 1 - \frac{\lambda(h)^2}{2n^2} + o\left(n^{-2}\right), \quad \text{as } n \to \infty.$$

Here $\lambda(h) > 0$ for $\|h\| \neq 0$. Assume also that $Y_{i,n}(s)$ are independent copies of $Y_n(s) = U_n(s)/V_n(s)$ and define $M_n(s) = \max_{i=1,\ldots,n} Y_{i,n}(s)$. Then the normalized bivariate asymptotic distribution of $(M_n(s), M_n(s+h))$ is

$$\begin{aligned}\Pr(W(s) \le w_1, W(s+h) \le w_2) &\equiv \lim_{n\to\infty} \Pr\left(M_n(s) \le \frac{nw_1}{\pi}, M_n(s+h) \le \frac{nw_2}{\pi}\right) \\ &= \exp\left\{-V_{\lambda(h)}(w_1, w_2)\right\}\end{aligned}$$

where

$$V_{\lambda(h)}(w_1, w_2) = \frac{1}{2}\left(\frac{2}{w_1} + \frac{2}{w_2} + \lambda(h) + \sqrt{\left(\frac{1}{w_1} - \frac{1}{w_2}\right)^2 + \lambda(h)^2} - \sqrt{\frac{1}{w_1^2} + \lambda(h)^2} - \sqrt{\frac{1}{w_2^2} + \lambda(h)^2}\right).$$

For a given $\lambda(h)$,

$$\Pr(W(s) > w, W(s+h) > w) \sim \left(1 + \frac{1}{2\lambda(h)}\right) w^{-2} \quad \text{as } w \to \infty$$

leading to a constant tail dependence parameter $\eta(h) = 1/2$. As underlined by Padoan (2013), a general framework based on Eq. (7) allows for different speeds of convergence to the independence case and could be used for dependence structures with a slower convergence than that of max-Gaussian ratio processes.

2.3. Pairwise extremal dependence measures

We recall here some measures of extremal dependence for spatial processes. From a theoretical point of view, the dependence structure of any multivariate extreme distribution is characterized by the exponent measure (see Resnick, 1987, for example). Unfortunately, it is quite difficult to infer this measure from the data. That is why summaries of extremal dependence based on pairwise measures have been proposed (Coles et al., 1999). For a stationary spatial process $Z = \{Z(s), s \in \mathcal{D}\}$ with univariate cumulative distribution function F, the pairwise extremal dependence between two sites s and $s+h$ can be characterized using the function

$$\chi(h) = \lim_{u \to 1^-} \Pr(F(Z(s+h)) > u \mid F(Z(s)) > u)$$

since we have pairwise asymptotic independence or asymptotic dependence (AD) if and only if $\chi(h) = 0$ or $\chi(h) \neq 0$, respectively (Sibuya, 1960). Alternatively $\chi(h)$ can be expressed as limit for $u \to 1^-$ of $\chi(h,u)$, defined in (2). The function $\chi(h,\cdot)$ can be interpreted as a quantile-dependent measure of dependence between two sites separated by h, giving more

Please cite this article in press as: Bacro, J.-N., et al., A flexible dependence model for spatial extremes. J. Statist. Plann. Inference (2016), http://dx.doi.org/10.1016/j.jspi.2015.12.002

insight if $Z(s)$ and $Z(s+h)$ are positively or negatively associated (Coles et al., 1999). Note also that for a max-stable process any bivariate distribution is max-stable and then the function $\chi(h, u)$ is constant with respect to u for a fixed h.

The extremal coefficient function (Schlather and Tawn, 2003) is a specific measure of the dependence for a max-stable process X. Given a pair of sites s and $s+h$ the extremal coefficient function $\theta(h)$ is defined as

$$\Pr(X(s) \le x, X(s+h) \le x) = \Pr(X(s) \le x)^{\theta(h)}.$$

Here $1 \le \theta(h) \le 2$ and $\theta(h) = 1$ or $\theta(h) = 2$ corresponds to perfect dependence or exact independence, respectively. It is easy to show that $\theta(h) = 2 - \chi(h)$.

Special instances of the Gaussian extreme value process (Smith, 1990) or the Brown–Resnick process (Kabluchko et al., 2009) span the range of possible extremal dependencies from perfect dependence to exact independence provided that distance $\|h\|$ increases indefinitely. Instead the extremal Gaussian process (Schlather, 2002) cannot account for extremes that become independent after some distance. Note that the TEG process has the feature that its extremal coefficient function

$$\theta(h) = 2 - \alpha(h)\left\{1 - 2^{-1/2}[1-\rho(h)]^{1/2}\right\} \tag{8}$$

reaches the upper value ($\theta(h) = 2$) for $\|h\|$ large enough. This specific feature will be exploited later.

Under asymptotic independence, both $\chi(h)$ and $\theta(h)$ functions are uninformative and of limited interest. Assume again that Z is a stationary spatial process with univariate cumulative distribution function F, and define

$$\bar{\chi}(h, u) = \frac{2\log \Pr(F(Z(s)) > u)}{\log \Pr(F(Z(s)) > u, F(Z(s+h)) > u)} - 1, \quad 0 \le u \le 1. \tag{9}$$

The limit $\bar{\chi}(h) = \lim_{u\to 1^-} \bar{\chi}(h, u)$, with $-1 < \bar{\chi}(h) \le 1$, provides another measure that increases with the extremal dependence between $Z(s)$ and $Z(s+h)$ (Coles et al., 1999). It turns out that for AD process $\bar{\chi}(h) = 1$, for all h. Moreover under the condition (7), it is easy to show that $\bar{\chi}(h) = 2\eta(h) - 1$ and the tail dependence coefficient $\eta(h)$ appears as another dependence measure of interest (Ledford and Tawn, 1996, 1997; Ancona-Navarrete and Tawn, 2002).

Note that the empirical estimate of (9) provides a useful statistic for inspecting the tail behavior when $u < 1$. For the stationary Gaussian process with correlation function $\rho(h)$ we can show that $\bar{\chi}(h, u)$ varies with u (Coles et al., 1999, p. 348) with limit $\bar{\chi}(h) = \rho(h)$ and $\eta(h) = (1+\rho(h))/2$. For the inverse max-stable process, $\bar{\chi}(h, \cdot)$ is a constant function. In other words, bivariate survival distributions of inverse max-stable processes are uniquely linked to the marginal survival function of the process whatever the magnitude of the considered extreme events. Moreover we have $\bar{\chi}(h) = 2/\theta(h) - 1$.

Finally, the function $\bar{\chi}(h, u)$ of a max-Gaussian ratio process varies with u and tends to 0 as $u \to 1^-$ for a fixed value of $\lambda(h)$.

3. Max-mixture modeling of spatial extremal dependence

3.1. Model specification

Let $X = \{X(s), s \in \mathcal{D}\}$ and $Y = \{Y(s), s \in \mathcal{D}\}$ be two independent stationary spatial processes, such that X is a max-stable process and Y an AI process both with unit Fréchet univariate distributions. We define the max-mixture (MM) model as

$$Z(s) = \max(\beta X(s), (1-\beta)Y(s)), \quad 0 \le \beta \le 1. \tag{10}$$

The MM model has been introduced by Wadsworth and Tawn (2012) for modeling situations where the extremal dependence structure may vary with distance. Even if it is not max-stable process, the MM model allows a different order of decay towards an asymptotically dependent limit which inherits the same dependence structure of X. In Wadsworth and Tawn (2012) various instances of max-stable processes along with their inverted versions as AI processes have been considered and all fitted models had asymptotic dependence or asymptotic independence present at all spatial lags.

Owing to our motivating data set, we propose in the sequel to extend the set of examples by considering a max-mixture model that deals with asymptotic dependence at short lags, asymptotic independence at intermediate lags and possibly exact independence at larger lags. More precisely we choose as X a TEG process (4) with covariance function $\rho(\cdot)$. Moreover, with respect to the examples in Wadsworth and Tawn (2012), we broaden the class of considered AI processes by taking into account AI processes with unit Fréchet univariate distributions and bivariate distributions satisfying the LT model (7) for $\eta(h) < 1$.

Using the independence between the two processes X and Y it is straightforward to obtain the bivariate distribution for a pair of sites, namely

$$\begin{aligned}&\Pr(Z(s) \le z_1, Z(s+h) \le z_2)\\ &= \exp\left\{-\beta\left(\frac{1}{z_1}+\frac{1}{z_2}\right)\left[1 - \frac{\alpha(h)}{2}\left(1 - \left(1 - 2\frac{(\rho(h)+1)z_1z_2}{(z_1+z_2)^2}\right)^{1/2}\right)\right]\right\} F_Y^h\left(\frac{z_1}{1-\beta}, \frac{z_2}{1-\beta}\right)\end{aligned} \tag{11}$$

where $F_Y^h(y_1, y_2) = \Pr(Y(s) \le y_1, Y(s+h) \le y_2)$. Since $\Pr(Z(s) \le z) = \Pr(Z(s) \le z, Z(s+h) < \infty) = \exp(-1/z)$ the model has unit Fréchet univariate distribution.

Please cite this article in press as: Bacro, J.-N., et al., A flexible dependence model for spatial extremes. J. Statist. Plann. Inference (2016), http://dx.doi.org/10.1016/j.jspi.2015.12.002

3.2. Pairwise extremal dependence measures of the model

Exploiting characterization (7), the bivariate tail distribution of (10), for large z, can be expressed as:

$$\Pr(Z(s) > z, Z(s+h) > z) = \frac{\beta(2-\theta(h))}{z} + \left(\frac{z}{1-\beta}\right)^{-1/\eta(h)} \mathcal{L}_h\left(\frac{z}{1-\beta}\right) + O(z^{-2}).$$

So it is easy to deduce the $\chi(h)$ function using Eq. (8), namely

$$\chi(h) = \beta\,(2-\theta(h)) = \beta\,\alpha(h)\left(1-\sqrt{\frac{1-\rho(h)}{2}}\right).$$

If the approximation (6) holds, it turns out that pairs of sites separated by a distance $\|h\|$ are AD if this distance is smaller than $2r$ and AI otherwise.

For evaluating $\bar{\chi}(h)$, we need to evaluate the logarithm of the bivariate tail distribution. We obtain

$$\log \Pr(Z(s) > z, Z(s+h) > z) = \begin{cases} \log(\beta(2-\theta(h))) - \log z + o(\log(z)) & \text{if } 2-\theta(h) \neq 0 \\ -\eta(h)^{-1}\log\left(\frac{z}{1-\beta}\right) + \log \mathcal{L}_h\left(\frac{z}{1-\beta}\right) + o(1), & \text{otherwise.} \end{cases}$$

If $2-\theta(h) \neq 0$, we can conclude that $\bar{\chi}(h,z) \to 1$ as $z \to \infty$. On the other hand, if $2-\theta(h) = 0$, we have

$$\bar{\chi}(h,z) \sim \frac{\left(-2-\frac{2}{z\log z}\right)}{\left(-\frac{1}{\eta(h)}\left(1-\frac{\log(1-\beta)}{\log z}\right)+\frac{\log(\mathcal{L}_h(z/(1-\beta)))}{\log z}\right)} - 1,$$

i.e. $\bar{\chi}(h,z) \to 2\eta(h)-1$ as $z \to \infty$. Owing to (6) the results can be summarized into the formula

$$\bar{\chi}(h) = \mathbb{I}_{[0,2r]}(\|h\|) + (2\eta(h)-1)\mathbb{I}_{[2r,\infty)}(\|h\|),$$

that highlights the different behavior according to the distance between two sites. Let $R > 2r$ and assume that $\eta(h) = 1/2$ for $\|h\| > R$. Then pairs of sites separated by a distance $\|h\|$ are asymptotically dependent for $\|h\| < 2r$, asymptotically independent for $2r \leq \|h\| \leq R$ and near independent for $\|h\| > R$. For example, for the transformed stationary Gaussian process with unit Fréchet margins and correlation function $\rho_Y(h)$, we have:

$$\bar{\chi}(h) = \mathbb{I}_{[0,2r]}(\|h\|) + \rho_Y(h)\mathbb{I}_{[2r,\infty)}(\|h\|).$$

In that case, independence is achieved if the correlation function $\rho_Y(\cdot)$ is such that $\rho_Y(h) = 0$ when $\|h\| > R$.

3.3. Model inference

For the model (10) since the full likelihood is intractable to evaluate, a composite likelihood approach is used for parametric estimations using pairs. The composite likelihood is an inference function derived by multiplying likelihoods of marginal or conditional events (Lindsay, 1988; Varin, 2008). Such an approach has been applied in spatial extremes using bivariate densities of max-stable processes (Padoan et al., 2010) or bivariate density of exceedances over a large threshold (Jeon and Smith, 2012; Wadsworth and Tawn, 2012; Bacro and Gaetan, 2014; Huser and Davison, 2014). Recently, improvements have been obtained for the parameters estimations of some max-stable processes, e.g. Brown–Resnick processes: extremal increments of the process allow to work with a complete likelihood function (Engelke et al., 2015; Wadsworth and Tawn, 2014). A direct modeling of the exceedances of a max-stable process is also possible using a generalized Pareto process (Ferreira and de Haan, 2014) but such an approach is only of interest in the case of asymptotic dependence.

If z_{ik} is the site-wise block maximum, for instance seasonal maximum, observed at site s_i, $i = 1,\ldots,N$ and at time t_k, $k = 1,\ldots,M$, the pairwise (weighted) log-likelihood is defined by

$$\mathrm{pl}(\psi) = \sum_{k=1}^{M} \mathrm{pl}_k(\psi) = \sum_{k=1}^{M}\sum_{i=1}^{N-1}\sum_{j>i}^{N} w_{ij}\log L(z_{ik}, z_{jk}; \psi) \tag{12}$$

where $L(z_{ik}, z_{jk}; \psi)$ is the likelihood of a pair z_{ik}, z_{jk}, The weights w_{ij} are non negative and specify the contributions of each pairs. A simple weighting choice is to let $w_{ij} = 0$ for any pair whose distance exceeds a specified value δ, and let $w_{ij} = 1$, otherwise.

Please cite this article in press as: Bacro, J.-N., et al., A flexible dependence model for spatial extremes. J. Statist. Plann. Inference (2016), http://dx.doi.org/10.1016/j.jspi.2015.12.002

Recently Wadsworth and Tawn (2012) argued that, under asymptotic independence, it is more natural to model the original events provided that they exceed a large threshold, u. Following their proposal the pairwise likelihood contribution $L(z_{ik}, z_{jk}; \psi)$ becomes

$$L(z_{ik}, z_{jk}; \psi) = \begin{cases} \dfrac{\partial^2}{\partial z_{ik} \partial z_{jk}} G(z_{ik}, z_{jk}; \psi) & \text{if } \max(z_{ik}, z_{jk}) > u \\ G(u, u; \psi) & \text{if } \max(z_{ik}, z_{jk}) \leq u \end{cases} \tag{13}$$

where z_{ik} is the observed value and $G(\cdot, \cdot)$ designates the bivariate distribution (11).

When dealing with exceedances it is not reasonable to assume that the observations are independent over the time. Assuming that the space–time process is temporally α mixing, the function (12) is a contrast function and conditions in Guyon (1995, Theorem 3.4.7) are satisfied. Thus the maximum composite likelihood estimator $\widehat{\psi}$ is asymptotically Gaussian for large M and its asymptotic variance is given by the inverse of the Godambe information matrix $\mathcal{G}(\psi) = \mathcal{H}(\psi)[\mathcal{J}(\psi)]^{-1}\mathcal{H}(\psi)$. Standard error evaluation requires consistent estimation of the matrices $\mathcal{H}(\psi) = \mathbb{E}(-\nabla^2 \mathrm{pl}(\psi))$ and $\mathcal{J}(\psi) = \mathbb{V}ar(\nabla \mathrm{pl}(\psi))$.

It is worth noting that such results hold if the data are actually from the limit model and this fact can add a bias (for an accurate study see Jeon and Smith, 2012) and, consequently, further uncertainty in the estimates.

The matrix $\mathcal{H}(\psi)$ can be estimated by $\widehat{\mathcal{H}} = -\nabla^2 \mathrm{pl}(\widehat{\psi}_T)$. Estimation of the matrix

$$\mathcal{J}(\psi) = \sum_{k=1}^{M} \sum_{k'>k}^{M} \mathbb{C}ov \left\{ \nabla \mathrm{pl}_k(\psi) \nabla \mathrm{pl}_{k'}(\psi)' \right\}$$

requires some care when we deal with temporally dependent data. In this paper we estimate $\mathcal{J}$ by means of a subsampling technique (Carlstein, 1986). More precisely, we consider B overlapping blocks $D_b \subset \{1, \ldots, M\}$, $b = 1, \ldots, B$, of size d_b and the estimate

$$\widehat{\mathcal{J}} = \frac{M}{B} \sum_{b=1}^{B} \frac{1}{d_b} \nabla \mathrm{pl}_{D_b}(\widehat{\psi}) \nabla \mathrm{pl}_{D_b}(\widehat{\psi})'$$

where pl_{D_b} is the pairwise likelihood evaluated over the block D_b.

Finally we mention that an appropriate model selection criterion to the pairwise likelihood is the composite likelihood information criterion (Varin and Vidoni, 2005), namely

$$\mathrm{CLIC} = -2 \left[\mathrm{pl}(\hat{\psi}) - \mathrm{tr}\{\hat{\mathcal{H}}^{-1} \hat{\mathcal{J}}\} \right].$$

Lower values of CLIC indicate better fit.

4. Simulation study

To assess the quality of the estimation procedure in case of the MM model (10), a simulation study has been carried out. We have chosen for X a TEG process (4) where B is a disk with a fixed radius r and exponential correlation function $\rho(h) = \exp(-\|h\|/\rho_1)$, $\rho_1 > 0$. The asymptotically independent process, Y, is given by $Y(s) = -1/\log(\Phi(Y'(s)))$, where Φ is the cumulative distribution function of a standard normal distribution and $\{Y'(s), s \in \mathcal{D}\}$ is a Gaussian spatial process with spherical correlation function, i.e. $1 - 1.5(\|h\|/\rho_2) + 0.5(\|h\|/\rho_2)^3$, for $\|h\| \leq \rho_2$, zero otherwise, $\rho_2 > 0$.

Under this setup extreme observations at sites separated by a vector h are asymptotically dependent if $\|h\| < r$, asymptotically independent if $r \leq \|h\| < \rho_2$ and independent if $\|h\| \geq \rho_2$, provided that $r < \rho_2$.

Five simulated images of the MM model over the $[0, 1]^2$ square are shown in Fig. 1, according to different values of the mixing parameter β. Actually, in order to appreciate the role of the mixing parameter β, the values in the images are derived by considering the simulation when $\beta = 0$ (AI process) and $\beta = 1$ (AD process). Note that the degree of smoothness decreases with β.

In the simulation study we have considered a moderately sized data set with $N = 49$ sites and $M = 1000$ independent observations. To avoid too systematic distances between pairs of sites, a non regular spatial grid has been considered. The $[0, 1]^2$ square is divided into 49 equal sub-squares and within each sub-square a point is uniformly chosen at random. We set $\rho_1 = 0.2$, $\rho_2 = 0.8$ and $r = 0.25$ and different values of $\beta \in \{0, 0.25, 0.50, 0.75, 1\}$.

The parameters are estimated on 500 data replication using the composite likelihood approach detailed in Section 3.3. The threshold u is taken corresponding to the 0.9 empirical quantile at each site and the δ value is chosen as the 0.9 quantile of the distribution of the distances between pairs of sites.

For compactness we report only the results for 500 data replications with $\beta = 0, 0.25, 0.75, 1$. For $\beta = 0.5$ we have obtained similar results. The boxplots in Fig. 2 that, overall, the parameters are well estimated except ρ_2 the parameter of the spherical correlation for which the estimate is significantly biased. This inadequacy seems consistent with the difficulties in estimating the parameter of the spherical correlation function in Gaussian models (Mardia and Watkins, 1989). In our example a justification for choosing a spherical correlation function is to consider a potential extremal exact independence

Please cite this article in press as: Bacro, J.-N., et al., A flexible dependence model for spatial extremes. J. Statist. Plann. Inference (2016), http://dx.doi.org/10.1016/j.jspi.2015.12.002

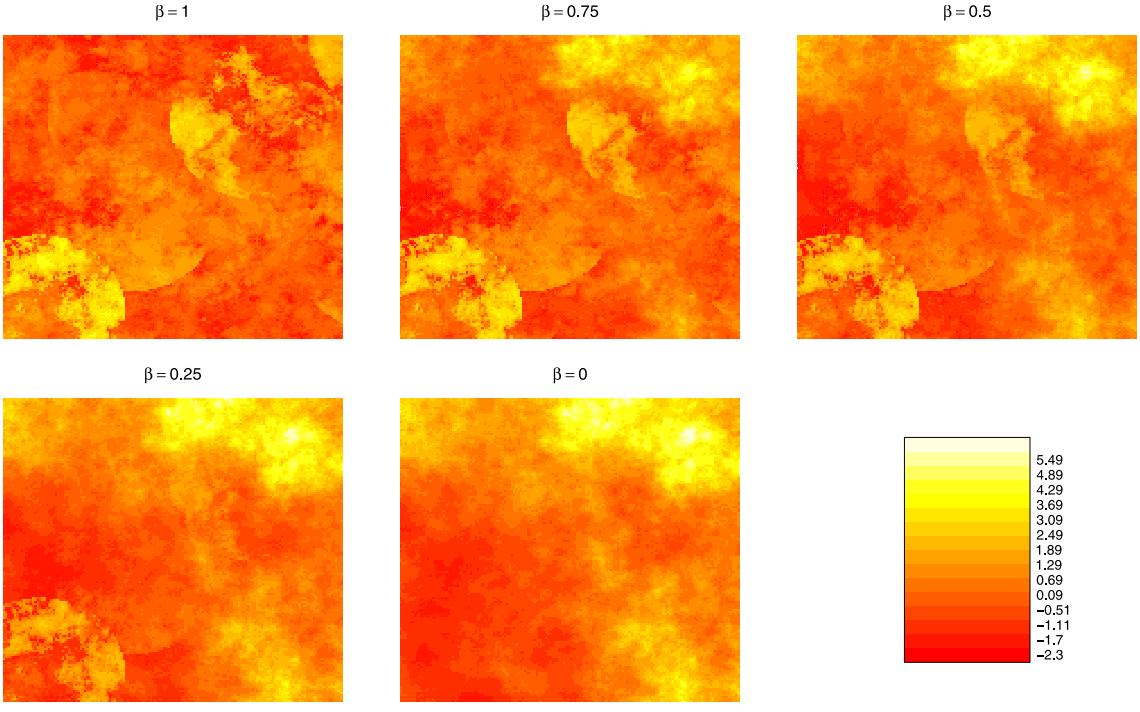

Fig. 1. Simulations of the MM_β model (10) on the logarithm scale according different values of $\beta \in \{0, 0.25, 0.50, 0.75, 1\}$. The compact set B is taken as a disk with a fixed radius $r = 0.25$. An exponential correlation function with parameter $\rho_1 = 0.2$ is chosen for the underlying Gaussian process. For the AI process a Gaussian random field is considered with a spherical correlation function with parameter $\rho_2 = 0.8$.

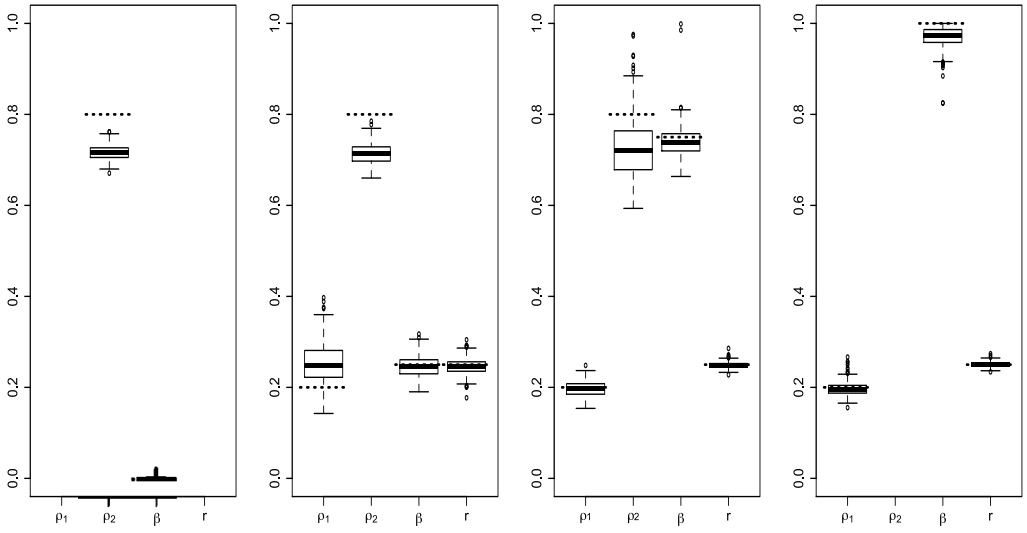

Fig. 2. Boxplots of 500 estimates from 1000 independent copies of the MM_β model (from left to right: $\beta = 0$, $\beta = 0.25$, $\beta = 0.75$ and $\beta = 1$) with $\rho_1 = 0.2$, $\rho_2 = 0.8$ and $r = 0.25$. For $\beta = 0$ and $\beta = 1$, only the results for the identifiable parameters are reported.

for distances larger than ρ_2. Simulations with an exponential correlation function not reported here lead to unbiased estimates of the range parameter.

Thereafter, we assessed whether CLIC is useful in identifying the true model, i.e. in our framework if we can use CLIC for discriminating between asymptotic independence, asymptotic dependence or a mixture of this. We have considered 500 simulations from mixture models with the same setting as before. In Table 1 we summarize our findings that are quite encouraging. We denote by MM_β, $\beta \in \{0, 0.25, 0.50, 0.75, 1\}$ the MM model according to different values of the mixing parameter. When the simulations come from MM_β, $\beta = 0.25, 0.5$ and 0.75, identification based on minimizing the CLIC value performs extremely well. Moreover the proportion of simulations in which the true model is detected is 68.6% if the true model is the AI process (MM_0). This proportion increases to 80% when the TEG process ($\beta = 1$) is the true model.

Please cite this article in press as: Bacro, J.-N., et al., A flexible dependence model for spatial extremes. J. Statist. Plann. Inference (2016), http://dx.doi.org/10.1016/j.jspi.2015.12.002

Table 1
Number of identified models according CLIC under different MM_β model, $\beta \in \{0, 0.25, 0.50, 0.75, 1\}$ with $\rho_1 = 0.2$, $\rho_2 = 0.8$, $r = 0.25$.

	Gaussian	MM	TEG
MM_0	346	154	0
$MM_{0.25}$	0	500	0
$MM_{0.50}$	0	500	0
$MM_{0.75}$	0	498	2
MM_1	0	100	400

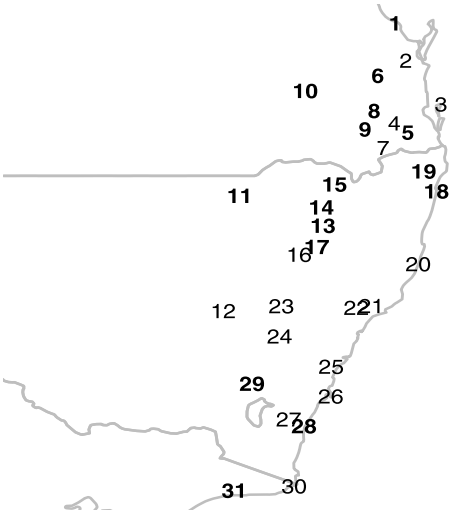

Fig. 3. Geographical locations of the 31 meteorological stations in the East Australia. Stations with a label in bold character are used for model inference and the other stations are put aside for validating the models.

5. Real data example

We analyze daily rainfall data from the 31 stations in the East of Australia whose locations are shown in Fig. 3. The values come from the daily rainfall data set of Lavery et al. (1992), available at time of writing at ftp.bom.gov.au/anon/home/ncc/www/change/HQdailyR.

Daily rainfall totals are for the 24-hours (measured at 9 am) and we consider days in the winter period (April–September) for 49 years ranging from 1955 to 2003.

Empirical estimates of the functions $\chi(h, u)$ and $\bar{\chi}(h, u)$ can be constructed on the basis of observed data by using the empirical estimates of univariate and bivariate distributions. In order to explore possible anisotropy of the dependence we have plotted (Fig. 4) the loess smoothing of $\hat{\chi}(h, u)$ and $\hat{\bar{\chi}}(h, u)$ at $u = 0.975$ with respect to the distances in different directional sectors, namely $(-\pi/8, \pi/8]$, $(\pi/8, 3\pi/8]$, $(3\pi/8, 5\pi/8]$, $(5\pi/8, 7\pi/8]$, where 0 represents the northing direction. Based on these estimates there is no clear evidence of anisotropy even if a stronger spatial dependence appears in the northing direction.

Moreover, as we mentioned in the introduction, the isotropic estimates (Fig. 5) of the functions $\hat{\chi}(h, u)$ and $\hat{\bar{\chi}}(h, u)$ at different values of the threshold u suggest that asymptotic dependence between sites seems to be present up to a distance of 500 km, and asymptotic independence could be conjectured between 500 and 1000 km distances. Therefore a max-mixture model seems a good candidate for interpreting the extreme value dependence. However the strength of dependence decreases when considering exceedances of increasing thresholds. This fact highlights the difficulty in a proper modeling of the asymptotic dependence for short distances.

In the sequel, we shall consider seven models that belong to three classes: max-mixture, max-stable and asymptotically independent. Each model is fitted using a subset of 16 sites and the remaining sites are used to perform model validation. We shall consider three MM models, namely

A_1 a MM model (10) specification in which X is a TEG process with exponential correlation function $\exp\{-\|h\|/\rho_1\}$, $\rho_1 > 0$ and B_1 is a disk of fixed and unknown radius r_1. The asymptotically independent process is given by $Y(s) = -1/\log((\Phi(Y'(s))))$, where Φ is the cumulative distribution function of a normalized Gaussian random variable and Y' is a Gaussian spatial process with spherical correlation function $1 - 1.5(\|h\|/\rho_2) + 0.5(\|h\|/\rho_2)^3$, for $\|h\| \leq \rho_2$, zero otherwise, $\rho_2 > 0$;

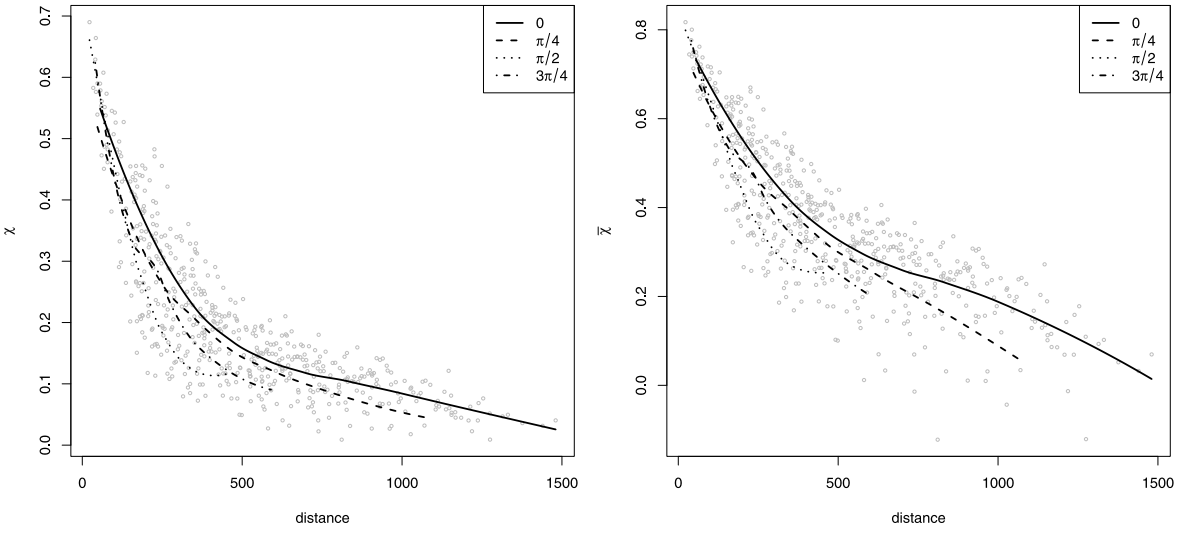

Fig. 4. Empirical evaluation of the functions $\widehat{\chi}(h,u)$ (left) and $\widehat{\bar{\chi}}(h,u)$ (right) at $u = 0.975$. Gray circles give empirical value between all available pairs. Lines represent smoothed values of the empirical estimates using the pairs in the directional sectors $(-\pi/8, \pi/8]$, $(\pi/8, 3\pi/8]$, $(3\pi/8, 5\pi/8]$ and $(5\pi/8, 7\pi/8]$.

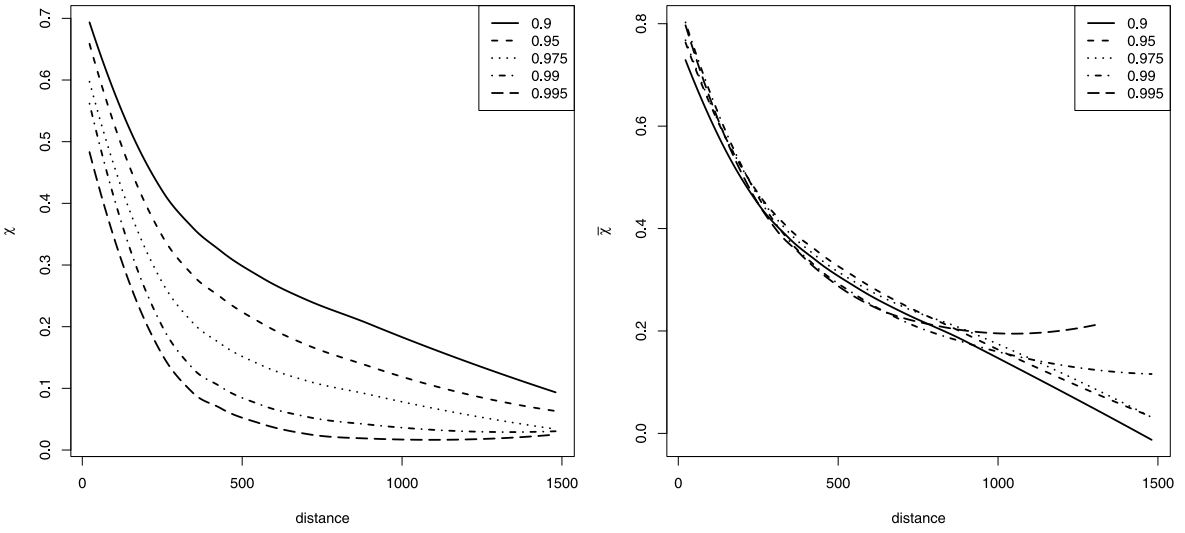

Fig. 5. Smoothed values of the empirical estimates of the functions $\hat{\chi}(h,u)$ (left) and $\hat{\bar{\chi}}(h,u)$ (right) at different values of the threshold u.

A_2 a MM model (10) where X is a TEG process as in A_1 and Y' is a Gaussian spatial process with exponential correlation function $\exp\{-\|h\|/\rho_2\}$;

A_3 a MM model with the same X as specified in A_1 and A_2 and in which Y is an inverse TEG process with exponential correlation function $\exp\{-\|h\|/\rho_2\}$, $\rho_2 > 0$. The B_2 disk has a fixed and unknown radius r_2.

As max-stable model candidate, we consider a max-stable model that entails exact independence between sites after a distance greater than $2r_1$, i.e.

B the TEG process specified in A_1.

Finally we take into account three asymptotically independent models, namely

C_1 the asymptotically independent process specified as Y in A_1;
C_2 the asymptotically independent process specified as Y in A_2;
C_3 the asymptotically independent process specified as Y in A_3.

Note that models C_1 and C_3 result in exact independence after distances greater than ρ_2 and r_2, respectively.

Please cite this article in press as: Bacro, J.-N., et al., A flexible dependence model for spatial extremes. J. Statist. Plann. Inference (2016), http://dx.doi.org/10.1016/j.jspi.2015.12.002

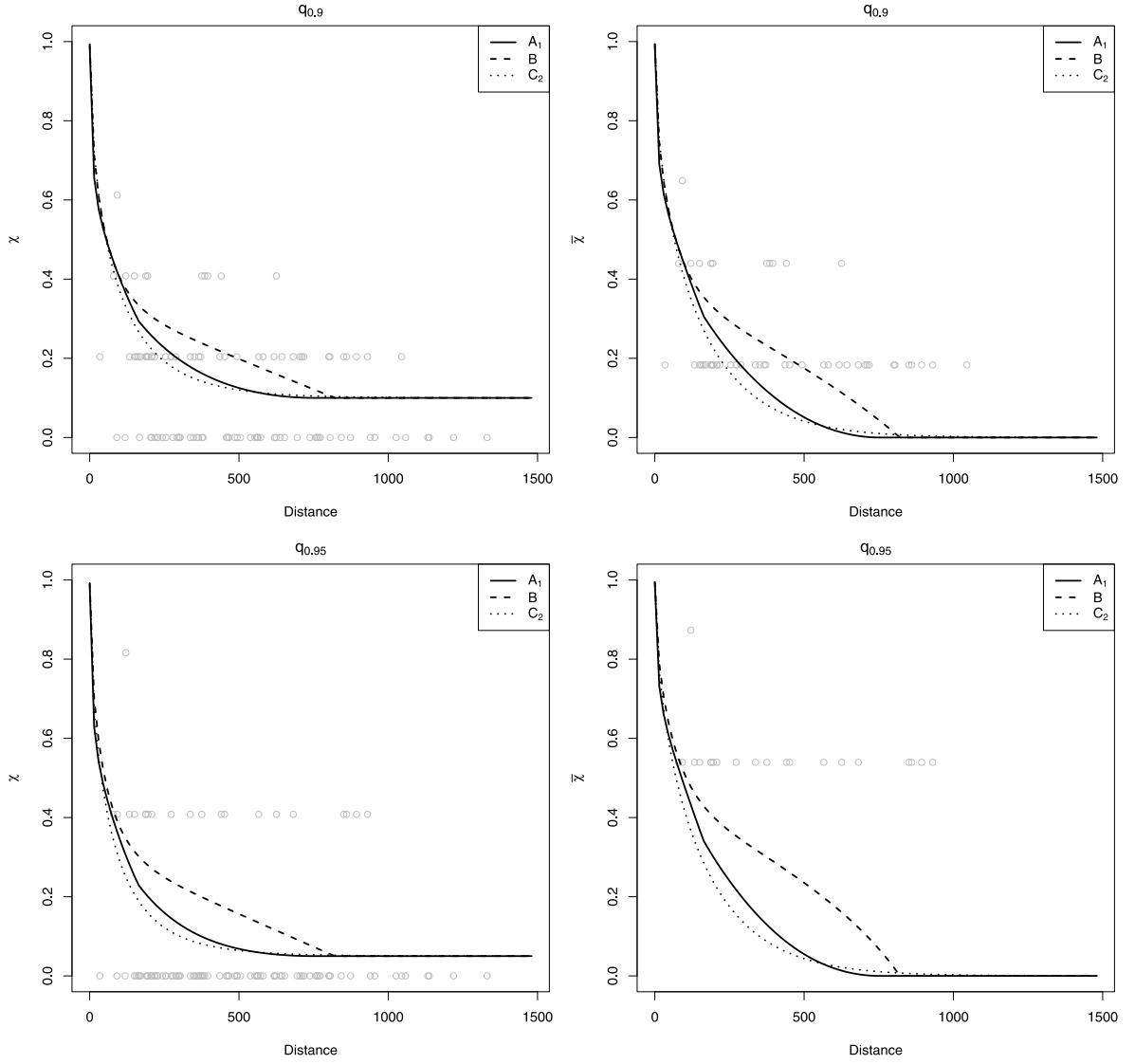

Fig. 6. Site-wise winter maxima: empirical and fitted values for $\hat{\chi}(h, u)$ and $\hat{\bar{\chi}}(h, u)$. Empirical values are computed using the validation data set. Top row: $u = 0.9$; bottom row: $u = 0.95$.

5.1. Site-wise maxima

First of all we have considered model site-wise winter maxima. Model (10) assumes common marginal Fréchet distributions and a proper inferential approach requires to fit marginal and dependence parameters. However, because we are interested in the appropriateness of different degrees of spatial asymptotic dependence, we prefer to follow a more simple and pragmatic approach. Specifically, we fit separately a GEV distribution in each site and we use the estimates to transform the marginals to unit Fréchet. The dependence parameters are estimated using the pairwise likelihood approach. Padoan et al. (2010) found in their simulation study that relatively small values of the distance δ in (12) produce gains in computation efficiency as well as in statistical efficiency of the estimates. However in our case we prefer to set $\delta \simeq 1000$ km, which entails to consider about 90% of all distinct observational pairs. For evaluating the CLIC and the standard errors we assume that the seasonal maxima are independent. In that case the estimation of the matrix $\mathcal{J}$ is greatly simplified in the subsampling procedure and we have considered $M = 49$ non overlapping blocks D_b corresponding to a single year, i.e. $d_b = 1$.

Our findings are summarized in Table 2. The rather wide standard-error of the spatial parameters, in particular for the max-mixture models, probably can be justified by the small number of independent replications over the years, pointing out that it is hard to separate the contribution of the components in the max-mixture. As suggested by the CLIC, the best-fitting model is A_1, for which pairs of sites separated by a distance d smaller than 160 km or greater than 750 km are

Please cite this article in press as: Bacro, J.-N., et al., A flexible dependence model for spatial extremes. J. Statist. Plann. Inference (2016), http://dx.doi.org/10.1016/j.jspi.2015.12.002

Table 2
Summary of the fitted models based on the site-wise winter maxima from the Australian data. Standard errors are reported between parentheses.

Model	$\widehat{\rho}_1$	$\widehat{r}_1$	$\widehat{\rho}_2$	$\widehat{r}_2$	$\widehat{\beta}$	CLIC
A_1	10.82 (14.10)	81.22 (301.93)	752.42 (278.52)	–	0.38 (0.22)	22 623.54
A_2	29.05 (37.56)	177.91 (32.16)	1451.92 (187.49)	–	0.72 (0.04)	22 661.76
A_3	5.47 (5.63)	311.22 (81.88)	367.48 (129.36)	707.73 (217.12)	0.43 (0.07)	22 644.5
B	78.09 (18.32)	410.34 (86.77)	–	–	–	22 692.3
C_1	–	–	359.51 (42.96)	–	–	22 689.73
C_2	–	–	179.34 (21.59)	–	–	22 642.23
C_3	–	–	71.84 (19.11)	440.48 (63.99)	–	22 679.72

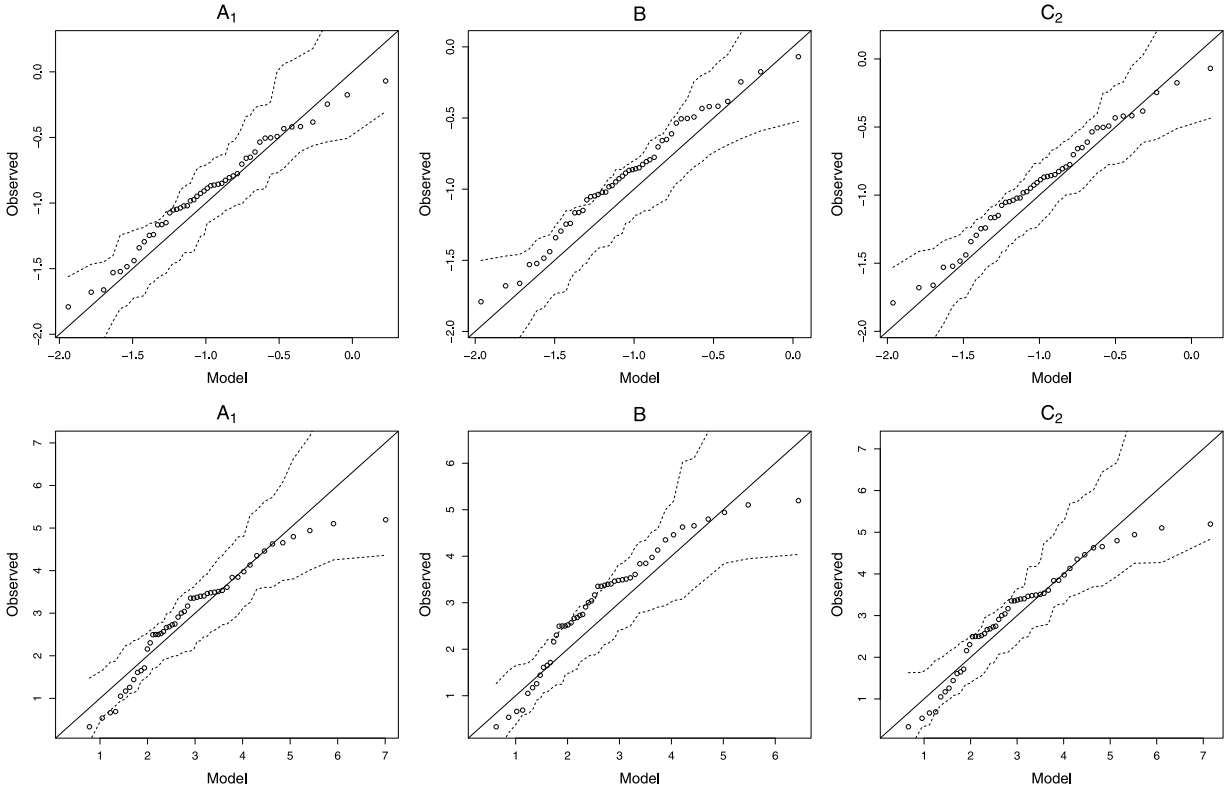

Fig. 7. Site-wise winter maxima: quantile–quantile plots for the minimum and maximum values on the validation data set (15 sites). The three columns correspond to fitted models A_1, B and C_2, respectively. The top row compares the minimum of the validation data set with its corresponding value under the fitted models. The bottom row compares in the same way the maximum values on the validation data set.

asymptotically max-stable dependent or exactly independent, respectively. At intermediate distances the seasonal maxima exhibit asymptotic independence. Moreover according to the CLIC values the MM models and the asymptotic independence models appear superior to the max-stable model B. So the max-stable model seems to overestimate the level of dependence in the data.

The goodness of fit has been also assessed in two different ways. Fig. 6 shows the empirical values for $\chi(h, u)$ and $\bar{\chi}(h, u)$, with $u = 0.9$ and 0.95 and their model-based counterparts of the three best models in each class according to the CLIC. Empirical estimates are calculated on the validation data set. The fits at finite thresholds are similar for A_1 and C_2 and the max stable model B entails stronger dependence for any distance. Considering the general patterns and owing to the small

Please cite this article in press as: Bacro, J.-N., et al., A flexible dependence model for spatial extremes. J. Statist. Plann. Inference (2016), http://dx.doi.org/10.1016/j.jspi.2015.12.002

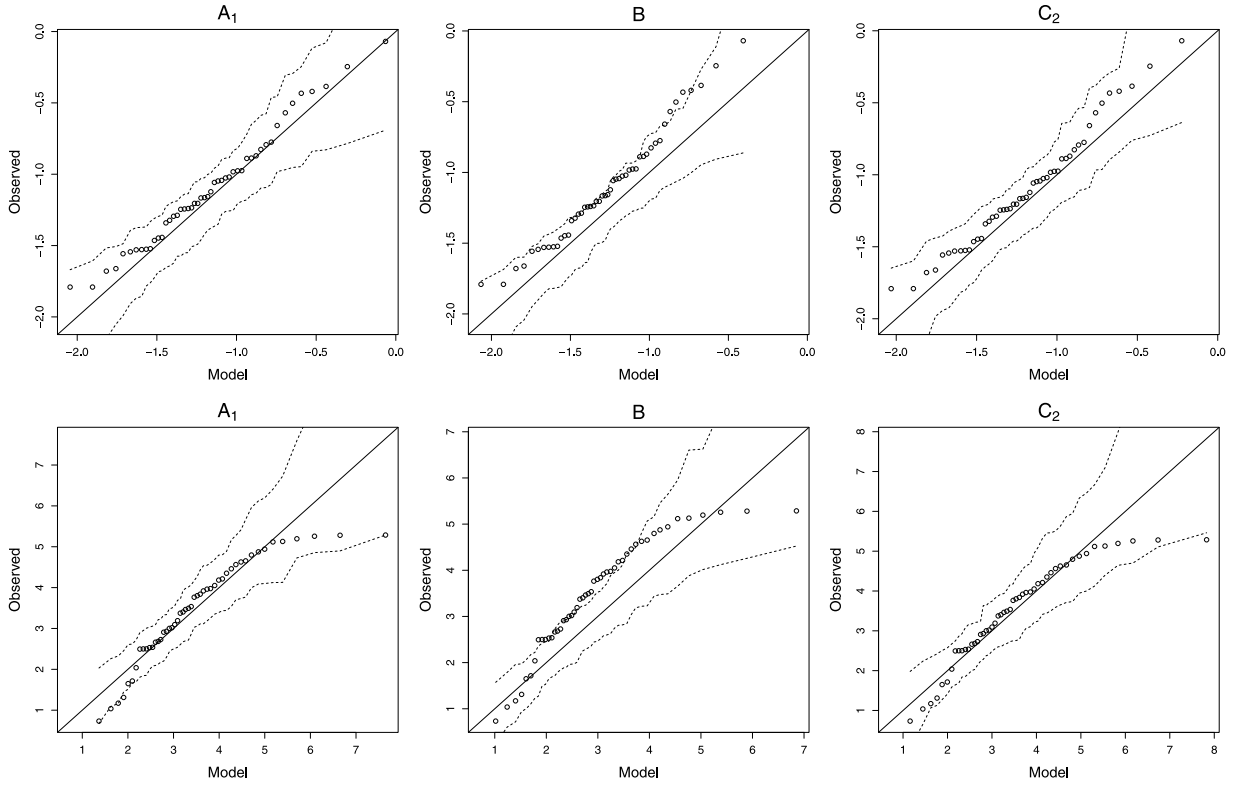

Fig. 8. Site-wise winter maxima: quantile–quantile plots for the minimum and maximum values on the complete data set (31 sites). The three columns correspond to fitted models A_1, B and C_2, respectively. The top row compares the minimum of the validation data set with its corresponding value under the fitted models. The bottom row compares in the same way the maximum values on the validation data set.

number of repeated observations for each site it is difficult to see which model catches better the bulk of the empirical values.

Model checking is also performed through QQ-plots for the logarithm of different groupwise minima and maxima on the validation set (Fig. 7) and the complete data (Fig. 8). Such plots provide some insight into whether the dependence models inferred using pairwise likelihood are capturing higher order dependence structures (Wadsworth and Tawn, 2012). Inspecting these plots, it appears that the multivariate distribution of the seasonal maxima is poorly modeled by the max-stable model B. Instead models A_1 and C_2 lead to quite similar results with an overall agreement to the data. Considering these plots and the CLIC values there is an overall evidence in favor of the max-mixture model.

5.2. Threshold exceedances

Now we deal with daily precipitations in the winter period and we use exceedances above a threshold corresponding to the 0.975 quantile in the empirical distribution at each site. We transform the observations to a unit Fréchet variable using the empirical distribution for data below the threshold and a site-wise fitted Generalized Pareto Distribution for data above the threshold. We estimate the spatial dependence parameters using the pairwise likelihood contribution (13). Because the original event data appear temporal dependent, the estimates $\hat{\mathcal{H}}$ and $\hat{\mathcal{J}}$ are carried out using a sliding temporal window of $d_b = 30$ days.

According to the CLIC value (Table 3), the preferred model is the *MM* model A_3. Nevertheless, the results for this model, here reported for completeness, have to be carefully considered because the estimate of β is virtually indistinguishable from zero, pointing out there is no mixture between the max-stable process and the asymptotically independent one. For $\beta = 0$ the parameters of the max-stable component are not identifiable and this fact affects the values of the estimates, their standard errors and finally the CLIC value. Moreover model A_3 reduces to model C_3 for $\beta = 0$ which corresponds the second best CLIC value, indicating some evidence for asymptotic independence for all distances.

Setting aside A_3 we reconsider the empirical and fitted values for $\hat{\chi}(h, u)$ and $\hat{\bar{\chi}}(h, u)$, $u = 0.9$ and $u = 0.95$, for the three best models in each class, namely A_2, B and C_3 (see Fig. 9). Model B seems to overestimate the asymptotic dependence at large distances. Again the fits of A_2 and C_3 look overall similar with a slight preference for A_2.

Finally, in order to illustrate the behavior of the models and check the fitting, we consider empirical and simulation based model estimates of few conditional probabilities. We choose the site s_1 in the top-right corner of the map (see

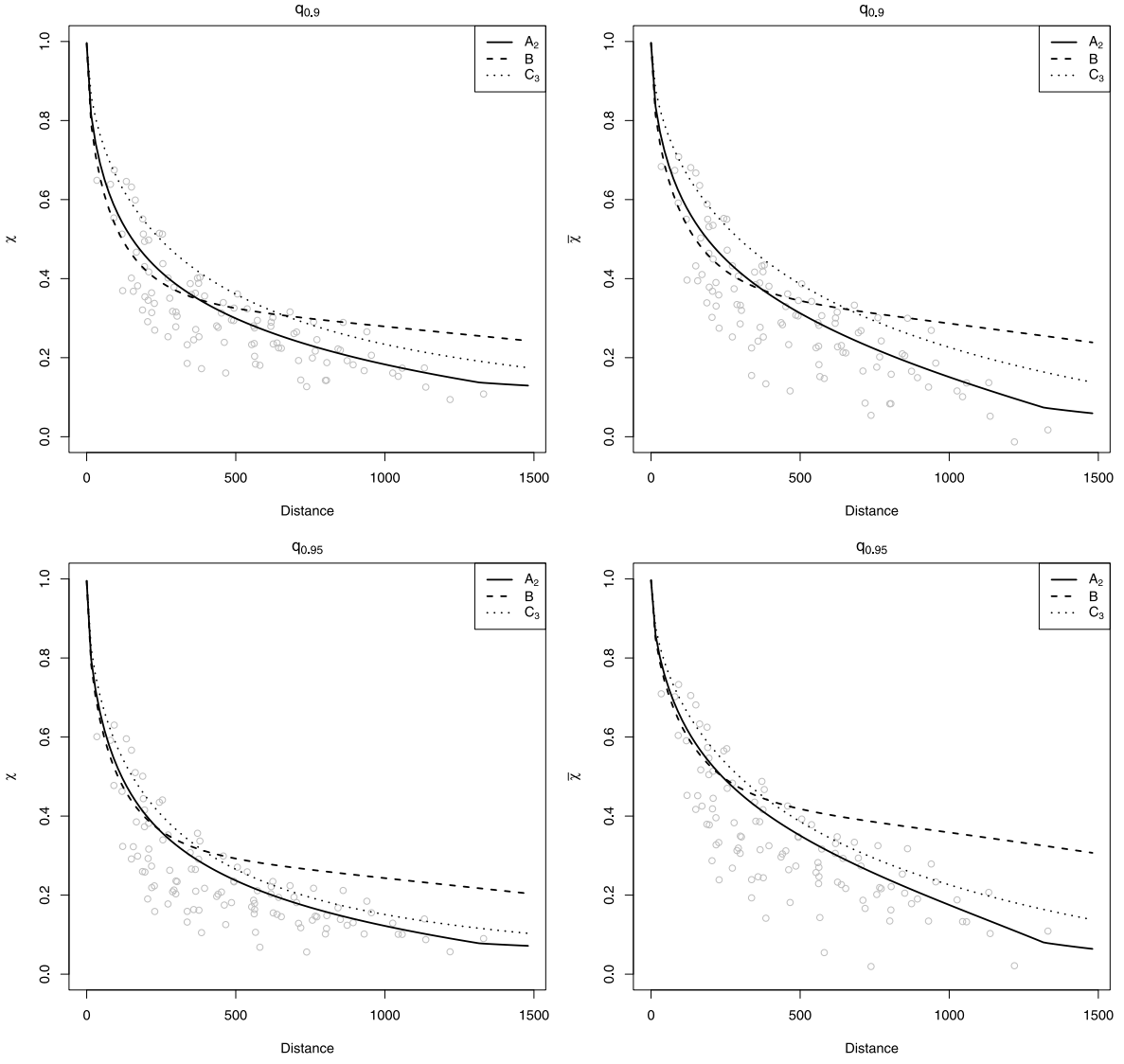

Fig. 9. Winter daily data: empirical and fitted values for $\hat{\chi}(h, u)$ and $\hat{\bar{\chi}}(h, u)$. Empirical values are computed using the validation data set and models are fitted using the q_u quantile exceedances. Top row: $u = 0.9$; bottom row: $u = 0.95$.

Fig. 3) as a reference location and we consider three subsets of sites $\mathcal{S}_1 = \{s_2, s_3, s_6, s_8, s_{10}\}$, $\mathcal{S}_2 = \{s_{11}, s_{13}, s_{14}, s_{15}, s_{18}\}$ and $\mathcal{S}_3 = \{s_{25}, s_{26}, s_{27}, s_{28}, s_{29}\}$ that roughly correspond to three different classes of distances from s_1. Then we compute the conditional probabilities $\Pr(Z(s) > z, s \in \mathcal{S}_i \mid Z(s_1) > z)$, $i = 1, 2, 3$ for different large values of p such that $\Pr(Z(s_1) \leq z) = p$. The confidence intervals in Fig. 10 are based on block bootstrapping of simulated daily data. The overall impression is that the max-stable model B is not able to describe the extremal dependence at medium ($\mathcal{S}_2$) and large distances ($\mathcal{S}_3$). Model C_3 basically overestimates the empirical probabilities for different thresholds and exhibits a lack of fitting for relative small distances ($\mathcal{S}_1$). On the other hand the fitting of model A_2 is more consistent at different thresholds and distances. Lastly note that both models agree for very large thresholds. These findings indicate that the max-mixture models we propose add modeling flexibility to spatial extreme analysis and seem able to encompass different degrees of spatial extremal dependence.

6. Conclusion

In this paper we have proposed a unifying spatial model which combines different degrees of extremal dependence depending on the distance between pairs of sites. Our approach exploits the max-mixture model proposed by Wadsworth and Tawn (2012) and focuses on the possible detection of pairwise max-stable dependence at short distances, asymptotic independence at intermediate ones and possibly exact independence at large distances. At short distances the extremal

Please cite this article in press as: Bacro, J.-N., et al., A flexible dependence model for spatial extremes. J. Statist. Plann. Inference (2016), http://dx.doi.org/10.1016/j.jspi.2015.12.002

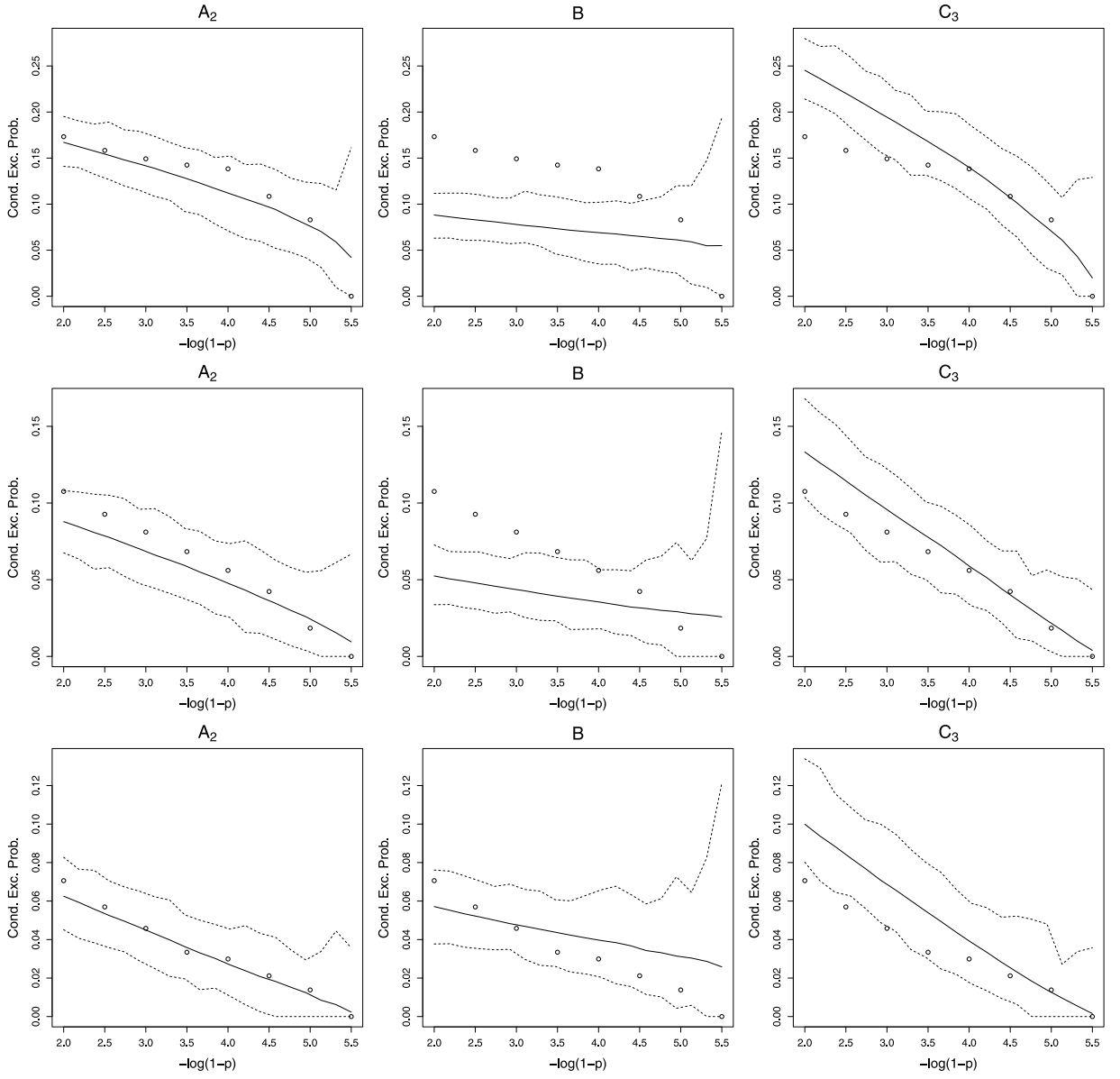

Fig. 10. Winter daily data: empirical and fitted values for the conditional probabilities $\Pr(Z(s) > z, s \in \mathcal{S} \mid Z(s_1) > z)$. The three columns correspond to models A_2, B and C_3, respectively. Top row: $\mathcal{S} = \{s_2, s_3, s_6, s_8, s_{10}\}$ (near sites data set); middle row $\mathcal{S} = \{s_{11}, s_{13}, s_{14}, s_{15}, s_{18}\}$ (medium sites data set); bottom row: $\mathcal{S} = \{s_{25}, s_{26}, s_{27}, s_{28}, s_{29}\}$ (far sites data set). The $1 - p$ values are such that $\Pr(Z(s_1) > z) = 1 - p$.

dependence is driven by a truncated extremal Gaussian max-stable process (Schlather, 2002) whereas at larger distances asymptotic independence is induced by any stochastic process with bivariate distributions satisfying a general condition proposed by Ledford and Tawn (1996). In this respect the hybrid models in Wadsworth and Tawn (2012) are particular instances.

Due to the intractability of the multivariate likelihoods parametric inference is carried out using a composite likelihood approach. A small and preliminary simulation study has shown that the inference procedure performs well, even when we have considered the boundary values for the mixture parameter.

In our real example we have highlighted that the max-mixture approach appears of interest for modeling environmental data. In particular it has the merit to overcome the limits of the max-stable models in which only asymptotic dependence or exact independence can be modeled.

Our attention has been concentrated on modeling the spatial dependence. In the future, we plan to consider spatio-temporal extensions that have fundamental interest in practice. Currently space–time models are still taking up little space in the literature and the major emphasis is in modeling asymptotic dependence treating the time just as additional

Please cite this article in press as: Bacro, J.-N., et al., A flexible dependence model for spatial extremes. J. Statist. Plann. Inference (2016), http://dx.doi.org/10.1016/j.jspi.2015.12.002

Table 3
Summary of the fitted models based on the daily exceedances from the Australian data. Standard errors are reported in parentheses.

Model	$\widehat{\rho}_1$	$\widehat{r}_1$	$\widehat{\rho}_2$	$\widehat{r}_2$	$\widehat{\beta}$	CLIC
A_1	78.71 (9.80)	833.76 (77.70)	1448.52 (57.72)	–	0.38 (0.02)	575 518.3
A_2	101.03 (13.93)	658.94 (54.26)	841.08 (51.23)	–	0.38 (0.02)	575 515.9
A_3	210.07 (10^{-13})	211.15 (10^{-13})	2164.57 (140.85)	1400.11 (95.08)	0 (10^{-13})	575 183.7
B	147.09 (6.17)	1706.55 (213.31)	–	–	–	580 455.
C_1	–	–	814.81 (19.34)	–	–	580 351.3
C_2	–	–	429.68 12.38	–	–	578 445.3
C_3	–	–	2084.84 (139.76)	1447.33 (106.76)	–	575 188.3

dimension of the space (Davis et al., 2013; Huser and Davison, 2014). However it seems reasonable to suppose that the spatial and temporal components behave asymptotically in a different way.

Acknowledgments

The research was partially supported by ANR-McSim and GICC-Miracle projects and by the Labex NUMEV. 137478, 2004. We are also indebted with Simone Padoan and the referees for comments that have led to improvements in the article.

References

Ancona-Navarrete, M., Tawn, J., 2002. Diagnostics for pairwise extremal dependence in spatial processes. Extremes 5, 271–285.
Bacro, J.N., Gaetan, C., 2012. A review on spatial extreme modelling. In: Porcu, E., Montero, J.M., Schlather, M. (Eds.), Advances and Challenges in Space–Time Modelling of Natural Events. Springer, New York, pp. 103–124.
Bacro, J.N., Gaetan, C., 2014. Estimation of spatial max-stable models using threshold exceedances. Stat. Comput. 24, 651–662.
Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J., 2004. Statistics of Extremes: Theory and Applications. John Wiley & Sons, New York.
Bingham, N.H., Goldie, C.M., Teugels, J.L., 1987. Regular Variation. In: Encyclopedia of Mathematics and its Applications, vol. 27. Cambridge University Press, Cambridge.
Carlstein, A., 1986. The use of subseries values for estimating the variance of a general statistic from a stationary sequence. Ann. Statist. 14, 1171–1179.
Casson, E., Coles, S.G., 1999. Spatial regression models for extremes. Extremes 1, 449–468.
Coles, S., Heffernan, J., Tawn, J., 1999. Dependence measures for extremes value analyses. Extremes 2, 339–365.
Cooley, D., Nychka, D., Naveau, P., 2007. Bayesian spatial modeling of extreme precipitation return levels. J. Amer. Statist. Assoc. 102, 824–840.
Davis, R.A., Klüppelberg, C., Steinkohl, C., 2013. Max-stable processes for modeling extremes observed in space and time. J. Korean Stat. Soc. 42, 399–414.
Davison, A.C., Gholamrezaee, M.M., 2012. Geostatistics of extremes. Proc. R. Soc. Lond. Ser. A 468, 581–608.
Davison, A.C., Huser, R., Thibaud, A., 2013. Geostatistics of dependent and asymptotically independent extremes. Math. Geosci. 45, 511–529.
Davison, A.C., Padoan, S.A., Ribatet, M., 2012. Statistical modelling of spatial extremes. Statist. Sci. 27, 161–186.
de Haan, L., 1984. A spectral representation for max-stable processes. Ann. Probab. 12, 1194–1204.
de Oliveira, T., 1962. Structure theory of bivariate extremes: extensions. Estud. Math. Estat. Econometrica 7, 165–195.
Engelke, S., Malinowski, A., Kabluchko, Z., Schlather, M., 2015. Estimation of Hüsler–Reiss distributions and Brown–Resnick processes. J. R. Stat. Soc. Ser. B Stat. Methodol 77, 239–265.
Fawcett, L., Walshaw, D., 2007. Improved estimation for temporally clustered extremes. Environmetrics 18, 173–188.
Ferreira, A., de Haan, L., 2014. The generalized Pareto process; with a view towards application and simulation. Bernoulli 20, 1717–1737.
Gaetan, C., Grigoletto, M., 2007. A hierarchical model for the analysis of spatial rainfall extremes. J. Agric. Biol. Environ. Stat. 12, 434–449.
Guyon, X., 1995. Random Fields on a Network. Springer, New York.
Huser, R., Davison, A.C., 2014. Space–time modeling of extreme events. J. R. Stat. Soc. Ser. B 76, 439–461.
Jeon, S., Smith, R., 2012. Dependence structure of spatial extremes using threshold approach, Technical report. arXiv:1209.6344.
Kabluchko, Z., Schlather, M., de Haan, L., 2009. Stationary max-stable fields associated to negative definite functions. Ann. Probab. 37, 2042–2065.
Lavery, B., Kariko, A., Nicholls, N., 1992. A historical rainfall data set for Australia. Aust. Meteorol. Mag. 40, 33–39.
Ledford, A.W., Tawn, J.A., 1996. Statistics for near independence in multivariate extreme values. Biometrika 83, 169–187.
Ledford, A., Tawn, J., 1997. Modelling dependence within joint tail regions. J. R. Stat. Soc. Ser. B 59, 475–499.
Ledford, A., Tawn, J., 1998. Concomitant tail behaviour for extremes. Adv. Appl. Probab. 30, 197–215.
Lindsay, B., 1988. Composite likelihood methods. Contemp. Math. 80, 221–239.
Mardia, K.V., Watkins, A.J., 1989. On multimodality of the likelihood in the spatial linear model. Biometrika 76, 289–295.
Opitz, T., 2013. Extremal t processes: elliptical domain of attraction and a spectral representation. J. Multivariate Anal. 122, 409–413.
Padoan, S.A., 2013. Extreme dependence model based on event magnitude. J. Multivariate Anal. 122, 1–19.
Padoan, S.A., Ribatet, M., Sisson, S., 2010. Likelihood-based inference for max-stable processes. J. Amer. Statist. Assoc. 105, 263–277.
Resnick, S., 1987. Extreme Values, Regular Variation and Point Processes. Springer, New York.
Ribatet, M., Cooley, D., Davison, A., 2012. Bayesian inference for composite likelihood models and an application to spatial extremes. Statist. Sinica 22, 813–845.
Sang, H., Gelfand, A., 2010. Continuous spatial process models for spatial extreme values. J. Agric. Biol. Environ. Stat. 15, 49–65.
Schlather, M., 2002. Models for stationary max-stable random fields. Extremes 5, 33–44.

Please cite this article in press as: Bacro, J.-N., et al., A flexible dependence model for spatial extremes. J. Statist. Plann. Inference (2016), http://dx.doi.org/10.1016/j.jspi.2015.12.002

Schlather, M., Tawn, J.A., 2003. A dependence measure for multivariate and spatial extreme values: properties and inference. Biometrika 90, 139–156.
Serinaldi, F., Bárdossy, A., Kilsby, C.G., 2014. Upper tail dependence in rainfall extremes: would we know it if we saw it? Stoch. Environ. Res. Risk Assess. 29, 1211–1233.
Sibuya, M., 1960. Bivariate extreme statistics. Ann. Inst. Statist. Math. 11, 195–210.
Smith, R.L., 1990. Max-stable processes and spatial extremes, Preprint. University of Surrey.
Thibaud, E., Mutzner, R., Davison, A.C., 2013. Threshold modeling of extreme spatial rainfall. Water Resour. Res. 49, 4633–4644.
Varin, C., 2008. On composite marginal likelihoods. Adv. Stat. Anal. 92, 1–28.
Varin, C., Vidoni, P., 2005. A note on composite likelihood inference and model selection. Biometrika 92, 519–528.
Wadsworth, J., Tawn, J., 2012. Dependence modelling for spatial extremes. Biometrika 99, 253–272.
Wadsworth, J., Tawn, J., 2014. Efficient inference for spatial extreme value processes associated to log-Gaussian random functions. Biometrika 101, 1–15.

Please cite this article in press as: Bacro, J.-N., et al., A flexible dependence model for spatial extremes. J. Statist. Plann. Inference (2016), http://dx.doi.org/10.1016/j.jspi.2015.12.002