



HAL
open science

A SEMI-PARAMETRIC ESTIMATION FOR MAX-MIXTURE SPATIAL PROCESSES

Manaf Ahmed, Véronique Maume-Deschamps, Pierre Ribereau, Céline Vial

► **To cite this version:**

Manaf Ahmed, Véronique Maume-Deschamps, Pierre Ribereau, Céline Vial. A SEMI-PARAMETRIC ESTIMATION FOR MAX-MIXTURE SPATIAL PROCESSES. 2017. hal-01617358v2

HAL Id: hal-01617358

<https://hal.science/hal-01617358v2>

Preprint submitted on 4 Dec 2017

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.

A SEMI-PARAMETRIC ESTIMATION FOR MAX-MIXTURE SPATIAL PROCESSES

M. AHMED, V. MAUME-DESCHAMPS, P. RIBEREAU, AND C. VIAL

ABSTRACT. We propose a semi-parametric estimation procedure in order to estimate the parameters of a max-mixture model as an alternative to composite likelihood estimation. This procedure uses the *F-madogram*. We propose to minimize the square difference between the theoretical F-madogram and an empirical one. We evaluate the performance of this estimator through a simulation study and we compare our method to composite likelihood estimation. We apply our estimation procedure to daily rainfall data from East Australia.

1. INTRODUCTION

One of the main characteristic of environmental or climatic data is their spatial dependence. The dependencies may be strong even for large distances as observed for heat waves or they may be strong at short distances and weak at larger distances, as observed for cevenol rainfall events. Many dependence structures may arise, for example, asymptotic dependence, asymptotic independence, or both. Max-mixture models as defined in [36] are a mixture between a max-stable process and an Asymptotic Independent (AI) process. These kind of models may be useful to fit e.g. rainfall data (see [5]).

The estimation of the parameters of these processes remains challenging. The usual way to estimate parameters in spatial contexts is to maximize the composite likelihood. For example, in [27], [14] and many others, the composite likelihood maximization is used to estimate the parameters of max-stable processes. In [5] and [36], it is used to estimate the parameters of max-mixture processes. Nevertheless, the estimation remains unsatisfactory in some cases as it seems to have difficulties estimating the AI part. So that, an alternative should be welcomed.

We propose a semi-parametric estimation procedure as an alternative to composite likelihood maximization for max-mixture processes. Our procedure is a least square method on the *F-madogram*. That is, we minimize the squared difference between the theoretical *F-madogram* and the empirical one. Some of literature deals with semi-parametric estimation in modeling spatial extremes. For example, [26] and [1] provided semi-parametric estimators of extremal indexes. In [8], another semi-parametric procedure to estimate model parameters is introduced. It is based a non linear least square method based on the extremogram for isotropic space-time Brown-Resnick max-stable processes. The semi-parametric procedure that we propose in this study is close to that article.

Section 2 is dedicated to the main tools used in this study; it contains definitions of some spatial dependence measures and of different dependence structures (asymptotic dependence/independence and mixture of them). In Section 3, we calculate

Key words and phrases. Spatial dependence measures, asymptotic dependence/independence, max-stable process, max-mixture, madogram.

an expression for the F -madogram of max-mixture models. Section 4 is devoted to estimation procedures of the parameters of max-mixture processes. We prove that our least square-madogram estimation is consistent, provided that the parameters are identified by the F -madogram. In Section 5, simulation study is conducted, it allows us to evaluate the performance of the estimation procedure. Section 6 is devoted to an application on rainfall real data from East Australia. Finally, concluding remarks are discussed in Section 7.

2. MAIN TOOLS USED IN THE STUDY

Throughout this study, the spatial process $X := \{X(s), s \in \mathbb{S}\}$, $\mathbb{S} \in \mathbb{R}^d$ is assumed to be strongly stationary and isotopic. We recall some basic facts and definitions related to spatial extreme theory.

2.1. Dependence measures. The extremal dependence behavior of spatial processes may be described by several coefficient / measures.

The **upper tail dependence coefficient** χ (see [6, 30]) is defined for a stationary spatial process X on $\mathbb{S} \subset \mathbb{R}^2$ with margin F by:

$$(2.1) \quad \chi(h) = \lim_{u \rightarrow 1^-} \mathbb{P}(F(X(s+h)) > u | F(X(s)) > u).$$

The process is said AI (resp. AD) if for all $h \in \mathbb{S}$ $\chi(h) = 0$ (resp. $\chi(h) \neq 0$).

For any h the coefficient $\chi(h)$ can alternatively be expressed as the limit when $u \rightarrow 1^-$ of the function defined on $\mathbb{S} \times [0, 1]$ into $[0, 1]$, by

$$(2.2) \quad \chi(h, u) = 2 - \frac{\log \mathbb{P}(F(X(s)) < u, F(X(t)) < u)}{\log \mathbb{P}(F(X(s)) < u)}, \text{ for } h \in \mathbb{S}, u \in [0, 1[.$$

Such that, $\chi(h) = \lim_{u \rightarrow -1} \chi(h, u)$.

For AI processes, χ cannot reveal the strength of the dependence. This is why the authors in [11], introduced an alternative dependence coefficient called **lower tail dependence coefficient** $\bar{\chi}$. Consider for any $(s, s+h) \in \mathbb{S}^2$

$$(2.3) \quad \bar{\chi}(h, u) = \frac{2 \log \mathbb{P}(F(X(s)) > u)}{\log \mathbb{P}(F(X(s)) > u, F(X(s+h)) > u)} - 1, \quad 0 \leq u \leq 1$$

and $\bar{\chi}(h) = \lim_{u \rightarrow 1} \bar{\chi}(h, u)$.

If $\bar{\chi}(h) = 1$ for all h , the spatial process is asymptotically dependent. Otherwise, the process is asymptotically independent. Furthermore, if $\bar{\chi} \in]0, 1[$ (resp. $] -1, 0[$) the locations s and $s+h$ (for any s) are asymptotically positively associated (resp. asymptotically negatively associated).

Another important measure of dependence is the extremal coefficient which was introduced by [9, 29]. For any $s \in \mathbb{S}$ and $s+h \in \mathbb{S}$ and $x \in \mathbb{R}$, let

$$\theta_F(h, x) = \frac{\log(P(X(s) < x, X(s+h) < x))}{\log(P(X(s) < x))}.$$

θ_F is related with the upper tail dependence parameter; indeed if $\lim_{x \rightarrow x_F} \theta_F(h, x) = \theta(h)$ exists, we have the following relation (see [6]):

$$\chi(h) = 2 - \theta_F(h),$$

where $x_F = \sup\{x | F(x) < 1\}$. In that case, $P(X(s) < x, X(s+h) < x)$ may be approximated by $F(x)^{\theta_F(h)}$ for x large.

This coefficient is particularly useful when dealing with asymptotic dependence, but useless in case of asymptotic independence. To overcome this problem, [23] proposed a model allowing to gather all the different dependence behaviors. This

model has Fréchet marginal laws and for all $(s, s + h) \in \mathbb{S}^2$ the pairwise survivor function is given by:

$$(2.4) \quad \mathbb{P}(X(s) > x, X(s + h) > x) = \mathcal{L}_h(x)x^{-1/\eta(h)}, \quad \text{as } x \rightarrow \infty,$$

where \mathcal{L}_h is a slowly varying function and $\eta(h) \in (0, 1]$ is the **tail dependence coefficient**. This coefficient determines the decay rate of the bivariate tail probability for large x . The interest of this simple modelization, which appears to be quite general, is that the coefficient $\eta(h)$ provides a measure of the extremal dependence between $X(s)$ and $X(s + h)$.

$\mathcal{L}(x) \not\rightarrow 0$ (resp. $0 < \eta(h) < 1$ and $\mathcal{L}(x) \not\rightarrow 0$), corresponds to asymptotic dependence (resp. asymptotic independence); see [6, 23]. Finally, it is important to see the relation between η and $\bar{\chi}$. If equation (2.4) is satisfied, then $\bar{\chi}(h) = 2\eta(h) - 1$.

Another classical tool often used in geostatistics is the variogram. But for spatial processes with Fréchet marginal laws, the variogram does not exist. We shall use the **F -madogram** introduced in [12] which is defined for any spatial process X with univariate margin F and for all $(s, t) \in \mathbb{S}^2$

$$(2.5) \quad \nu^F(s - t) = \frac{1}{2} \mathbb{E}|F(X(s)) - F(X(t))|.$$

2.2. Spatial extreme models. For completeness, we recall definitions on max-stable, inverse max-stable and max-mixture processes.

2.3. Max-stable model. Max-stable processes are the extension of the multivariate extreme value theory to the infinite dimensional setting [8]. If there exist two sequences of continuous functions $(a_n(\cdot) > 0, n \in \mathbb{N})$ and $(b_n(\cdot) \in \mathbb{R}, n \in \mathbb{N})$ such that for all $n \in \mathbb{N}$ and n i.i.d. X_1, \dots, X_n and X a process, such that

$$(2.6) \quad \max_{i=1, \dots, n} \frac{X_i - b_n}{a_n} \xrightarrow{d} X, \quad n \rightarrow \infty,$$

then $X := \{X(s), s \in \mathbb{S}\}$ is a max-stable process [17]. When for all $n \in \mathbb{N}$, $a_n = 1$ and $b_n = 0$, the margin distribution of the process X is unit Fréchet, that is for any $s \in \mathbb{S}$ and $x > 0$,

$$F(x) := P(X(s) \leq x) = \exp[-1/x],$$

in that case, we say that X is a *simple max-stable process*.

In [16] it is proved that a max-stable process X can be constructed by using a random process and a Poisson process. This representation is named the **spectral representation**. Let X be a max-stable process on \mathbb{S} . Then there exists $\{\xi_i, i \geq 1\}$ i.i.d Poisson point process on $(0, \infty)$, with intensity $d\xi/\xi^2$ and a sequence $\{W_i, i \geq 1\}$ of i.i.d. copies of a positive process $W = (W(s), s \in \mathbb{S})$, such that $\mathbb{E}[W(s)] = 1$ for all $s \in \mathbb{S}$ such that

$$(2.7) \quad X =^d \max_{i \geq 1} \xi_i W_i.$$

The c.d.f. of a simple max-stable process satisfies: (see [7], Section 8.2.2.)

$$(2.8) \quad \mathbb{P}(X(s_1) \leq x_1, \dots, X(s_k) \leq x_k) = \exp\{-V(x_1, \dots, x_k)\},$$

where the function

$$(2.9) \quad V(x_1, \dots, x_k) = \mathbb{E} \left[\max_{\ell=1, \dots, k} \left(\frac{W(s_\ell)}{x_\ell} \right) \right].$$

is homogenous of order -1 and is named *the exponent measure*. One of the interests of the exponent measure is its interpretation in terms of dependence. In fact, the homogeneity of the exponent measure V implies

$$(2.10) \quad \max\{1/x_1, \dots, 1/x_k\} \leq V(x_1, \dots, x_k) \leq \{1/x_1 + \dots + 1/x_k\}.$$

In Inequalities (2.10), the lower (resp. upper) bound corresponds to complete dependence (resp. independence).

For simple max-stable process, **the extremal coefficient function** Θ , for any pairs of sites $(s, s+h) \in \mathbb{S}^2$ is the function Θ defined on \mathbb{S} (or in \mathbb{R}^+ in isotropic case) with values in $[1, 2]$ by

$$(2.11) \quad \mathbb{P}(X(s) \leq x, X(s+h) \leq x) = \exp\{-\Theta(h)/x\}, \quad x > 0$$

We have

$$(2.12) \quad \Theta(h) = \mathbb{E}[\max\{W(s), W(s+h)\}] = V(1, 1) \in [1, 2].$$

If for any $h \in \mathbb{S}$, $\Theta(h) = 1$ (resp. $\Theta(h) = 2$), then we have complete dependence (resp. complete independence). The case $1 < \Theta(h) < 2$, for all $h \in \mathbb{S}$ corresponds to asymptotic dependence. Remark that, for a simple max-stable process, the coefficients Θ and θ_F coincide.

Furthermore, it is easy to see the relationship between Θ and χ ; see [36] for any $h \in \mathbb{S}$

$$(2.13) \quad \Theta(h) = 2 - \chi(h).$$

In the max-stable case, [12] gives the relation for all $h \in \mathbb{S}$,

$$(2.14) \quad \Theta(h) = \frac{1 + 2\nu^F(h)}{1 - 2\nu^F(h)},$$

which appears to be helpful to estimate the extremal coefficient Θ from the empirical madogram. The max-stable process X with pairwise distribution function is given by the following equation, for all $(s, s+h) \in \mathbb{S}^2$,

$$(2.15) \quad \mathbb{P}(X(s) \leq x_1, X(s+h) \leq x_2) = \exp\{-V_h(x_1, x_2)\},$$

We provide below three examples of well-known max-stable models represented by different exponent measures V .

Smith Model (Gaussian extreme value model) [31] with unit Fréchet margin and exponent measure

$$(2.16) \quad V_h(x_1, x_2) = \frac{1}{x_1} \Phi\left(\frac{\tau(h)}{2} + \frac{1}{\tau(h)} \log \frac{x_2}{x_1}\right) + \frac{1}{x_2} \Phi\left(\frac{\tau(h)}{2} + \frac{1}{\tau(h)} \log \frac{x_1}{x_2}\right);$$

$\tau(h) = \sqrt{h^T \Sigma^{-1} h}$ and $\Phi(\cdot)$ the standard normal cumulative distribution function.

For isotropic case $\tau(h) = \sqrt{\frac{1}{\sigma} \|h\|^2}$.

The pairwise extremal coefficient equals

$$\Theta(h) = 2\Phi\left(\frac{\tau(h)}{2}\right).$$

Brown-Resnik Model [22] with unit Fréchet margin and exponent measure

$$V_h(x_1, x_2) = \frac{1}{x_1} \Phi\left(\frac{\sqrt{2\gamma(h)}}{2} + \frac{1}{\sqrt{2\gamma(h)}} \log \frac{x_2}{x_1}\right) + \frac{1}{x_2} \Phi\left(\frac{\sqrt{2\gamma(h)}}{2} + \frac{1}{\sqrt{2\gamma(h)}} \log \frac{x_1}{x_2}\right);$$

$2\gamma(h)$ is a variogram and $\Phi(\cdot)$ the standard normal cumulative distribution function. The pairwise extremal coefficient given by

$$\Theta(h) = 2\Phi\left(\frac{\sqrt{2\gamma(h)}}{2}\right).$$

Truncated extremal Gaussian Model (TEG) [28] with unit Fréchet margin and exponent measure

$$(2.17) \quad V_h(x_1, x_2) = \left(\frac{1}{x_1} + \frac{1}{x_2} \right) \left[1 - \frac{\alpha(h)}{2} \left(1 - \sqrt{1 - 2(\rho(h) + 1) \frac{x_1 x_2}{(x_1 + x_2)^2}} \right) \right].$$

The extremal coefficient is given by

$$(2.18) \quad \Theta(h) = 2 - \alpha(h) \left\{ 1 - \left(\frac{1 - \rho(h)}{2} \right)^{1/2} \right\}$$

where $\alpha(h) = \mathbb{E}\{|\mathcal{B} \cap (h + \mathcal{B})|\} / \mathbb{E}[|\mathcal{B}|]$ and \mathcal{B} is a random set which can consider it a disk with fixed radius r . In such a case, $\alpha(h) = \{1 - h/2r\}_+$ (see [15]) and $\chi(h) = 0, \forall h \geq 2r$.

2.4. Inverse Max-stable processes. Max-stable processes are either asymptotically Dependent (AS) or independent. This means that they are not useful to model non trivial AI processes. In [36] a class of asymptotically independent processes is obtained by inverting max-stable processes. These processes are called **inverse max-stable processes**; their survivor function satisfy Equation (2.4). Let $X' := \{X'(s), s \in \mathbb{S}\}$ be a simple max-stable process with exponent measure V . Let $g : (0, \infty) \mapsto (0, \infty)$ be defined by $g(x) = -1/\log\{1 - e^{-1/x}\}$ and $X(s) = g(X'(s))$. Then, $X := \{X(s), s \in \mathbb{S}\}$ is an asymptotic independent spatial process with unit Fréchet margin. The d -dimensional joint survivor function satisfies

$$(2.19) \quad \mathbb{P}(X(s_1) > x_1, \dots, X(s_d) > x_d) = \exp\{-V(g(x_1), \dots, g(x_d))\}.$$

The tail dependent coefficient is given by $\eta(h) = 1/\Theta(h)$, where $\Theta(h)$ is the extremal coefficient of the max-stable process X' . Moreover, we have $\tilde{\chi}(h) = 2/\Theta(h) - 1$. With a slight abuse of notations, we shall say that V is the exponent measure of X .

2.5. Max-mixture model. In spatial contexts, specifically in an environmental domain, many scenarios of dependence could arise and AD and AI might cohabite. The work by [36] provides a flexible model called max-mixture.

Let $X := \{X(s), s \in \mathbb{S}\}$ be a simple max-stable process with extremal coefficient $\Theta(h)$ and bivariate distribution function F_X , and let $Y := \{Y(s), s \in \mathbb{S}\}$ be an inverse max-stable process whose tail dependence coefficient is $\eta(h)$. Its bivariate distribution function is denoted F_Y . Assume that X and Y are independent. Let $a \in [0, 1]$ and define

$$(2.20) \quad Z(s) = \max\{aX(s), (1 - a)Y(s)\}, \quad s \in \mathbb{S},$$

then Z unit Fréchet marginals and its pairwise survivor function satisfies

$$(2.21) \quad \mathbb{P}(Z(s) > z, Z(t) > z) \sim \frac{a\{2 - \Theta(h)\}}{z} + \frac{(1 - a)^{1/\eta(h)}}{z^{1/\eta(h)}} + O(z^{-2}), \quad z \rightarrow \infty.$$

The process $(Z(s))_{s \in \mathbb{S}}$ is called a max-mixture process. Assume there exists finite $h^* = \inf\{h : \Theta(h) \neq 0\}$; then,

$$(2.22) \quad \chi(h) = a(2 - \Theta(h))$$

and

$$(2.23) \quad \bar{\chi}(h) = \mathbf{1}_{[h^* < h]}(h) + (2\eta(h) - 1)\mathbf{1}_{[h^* \geq h]}(h).$$

Remark 2.1. *If there exists finite $h^* = \inf\{h : \Theta(h) \neq 0\}$, then Z is asymptotically dependent up to distance h^* and asymptotically independent for larger distances.*

Of course, if $a = 0$ then Z is AI and if $a = 1$ then Z is a simple max-stable process.

In [5] max-mixture processes are studied. The authors emphasize the fact that these models allow asymptotic dependence and independence to be present at a short and intermediate distances. Furthermore, the process may be independent at long distances (using e.g. TEG processes).

3. F -MADOGRAM FOR MAX-MIXTURE SPATIAL PROCESS

In extreme value theory and therefore for spatial extremes, one of the main concerns is to find a dependence measure that can quantify the dependences between locations. The χ and $\bar{\chi}$ dependence measures are designed to quantify asymptotic dependence and asymptotic independence respectively (see equations (2.22) and (2.23)). Max-mixture processes have been introduced in order to provide both behaviors. We are then faced with the question of finding an adapted tool which would give information on more than one dependence structure.

In [12], the F -madogram has been introduced for max-stable processes. There exists several definitions of madograms. For example, in [25], the λ -madogram is considered in order to take into account the dependence information from the exponent measure $V_h(u, v)$ when $u \neq v$. This λ -madogram has been extended in [19] to evaluate the dependence between two observations located in two disjoint regions in \mathbb{R}^2 . [20] adopted a F -madogram suitable for asymptotic independence instead of asymptotic dependence only. Finally, [3] used F -madogram as a test statistic for asymptotic independence bivariate maxima.

Below, we calculate $\nu^F(h)$ for a max-mixture process. It appears that contrary to χ and $\bar{\chi}$, it combines the parameters coming from the AD and the AI parts.

Proposition 3.1. *Let Z be a max-mixture process, with mixing coefficient $a \in [0, 1]$. Let X be its max-stable part with extremal coefficient $\Theta(h)$. Let Y be its inverse max-stable part with tail dependence coefficient $\eta(h)$. Then, the F -madogram of Z is given by*

$$(3.1) \quad \nu^F(h) = \frac{a(\Theta(h) - 1)}{a(\Theta(h) - 1) + 2} - \frac{a\Theta(h) - 1}{2a\Theta(h) + 2} - \frac{1/\eta(h)}{a\Theta(h) + (1-a)/\eta(h) + 1} \beta\left(\frac{a\Theta(h) + 1}{(1-a)}, 1/\eta(h)\right),$$

where β is beta function.

Proof. We have

$$(3.2) \quad \nu^F(h) = \frac{1}{2} \mathbb{E}|F(Z(s)) - F(Z(s+h))|.$$

The equality $|a - b|/2 = \max(a, b) - (a + b)/2$ leads to

$$(3.3) \quad \nu^F(h) = \mathbb{E}[\max(F(Z(s)), F(Z(s+h)))] - \mathbb{E}[F(Z(s))].$$

Let $M(h) = \max(F(Z(s)), F(Z(s+h)))$, we have:

$$(3.4) \quad \mathbb{P}(M(h) \leq u) = \mathbb{P}(Z(s) \leq F^{-1}(u), Z(s+h) \leq F^{-1}(u)).$$

From definition of the max-mixture spatial process Z , we have

$$(3.5) \quad \mathbb{P}(Z(s) \leq z_1, Z(s+h) \leq z_2) = e^{-aV_X^h(z_1, z_2)} \left[e^{-\frac{-(1-a)}{z_1}} + e^{-\frac{-(1-a)}{z_2}} - 1 + e^{-V_Y^h(g(z_1), g(z_2))} \right],$$

where V_X (resp. V_Y) corresponding to the exponent measures of X (resp. Y) and $g(z) = -1/\log(1 - e^{-\frac{(1-a)}{z}})$. That leads to

$$\mathbb{P}(M(h) \leq u) = u^{a\Theta(h)} [2u^{(1-a)} - 1 + (1 - u^{(1-a)})^{1/\eta(h)}], \quad u \in [0, 1].$$

We deduce that

$$(3.6) \quad \begin{aligned} \mathbb{E}[M(h)] &= \int_0^1 u f_M(h)(u) du \\ &= \frac{2a(\Theta(h) - 1) + 2}{a(\Theta(h) - 1) + 2} - \frac{a\Theta(h)}{a\Theta(h) + 1} - \frac{\beta\left(\frac{a\Theta(h)+1}{(1-a)}, 1/\eta(h)\right)}{\eta(h)(1-a)\left[\frac{a\Theta(h)+1}{(1-a)} + (1/\eta(h))\right]}. \end{aligned}$$

where $f_M(h)$ is the density of $M(h)$. Recall that $\mathbb{E}(F(Z(s))) = \frac{1}{2}$ because $F(Z(s)) \sim \mathcal{U}([0, 1])$ and return to equation (3.3) to get equation (3.1). \square

In the particular cases where $a = 1$ or $a = 0$, Proposition 3.1 reduces to known results for max-stable processes (see [12]) and inverse max-stable processes (see [20]). That is, the F -madogram for a max-stable spatial process is given by

$$(3.7) \quad 2\nu^F(h) = \frac{\Theta(h) - 1}{\Theta(h) + 1}.$$

and the F -madogram of an asymptotically independent spatial process is given by

$$(3.8) \quad 2\nu^F(h) = \frac{1 - \eta(h)}{1 + \eta(h)}$$

In order to have a comprehensive view of the behavior of $\nu^F(h)$, we have plotted in Figure 1 below $h \rightsquigarrow \nu^F(h)$. We have considered two max-mixture models MM1 and MM2 described as follows:

- **MM1** max-mixture between a TEG max-stable process X and an inverse Smith max-stable process Y ;
- **MM2** max-mixture between X as in MM1 and an inverse TEG max-stable process Y .

In this Figure and for the two models MM1 and MM2, $\nu^F(h)$ has two sill one corresponding to X and the second corresponding to Y . This is completely in accordance with the nested variogram concept as presented in [35]. In data analysis, these two levels of the sill gives the researcher a hint about whether there is more than one spatial dependence structure in the data.

Since the F -madogram expresses with all the model parameters it should be useful for the parameter estimation.

4. MODEL INFERENCE

This section is devoted to the parametric inference for max-mixture processes. We begin with the presentation of the maximum composite likelihood estimation, then we present the least squares madogram estimation.

4.1. Parametric Estimation using Composite Likelihood.

Consider $(Z^k(s_1), \dots, Z^k(s_D))$, $k = 1, \dots, N$, be N independent copies of a spatial process $(Z(s))_{s \in \mathbb{S}}$, observed at D locations s_1, \dots, s_D . Composite likelihood inference is an appropriate approach in estimating the parameter models of a spatial process X ([24, 33]). Asymptotic properties of this estimator has been proved in [13]. This approach has been applied successfully to spatial max-stable processes by [14] and [27] and is also used to identify the parameters of data exceedances over

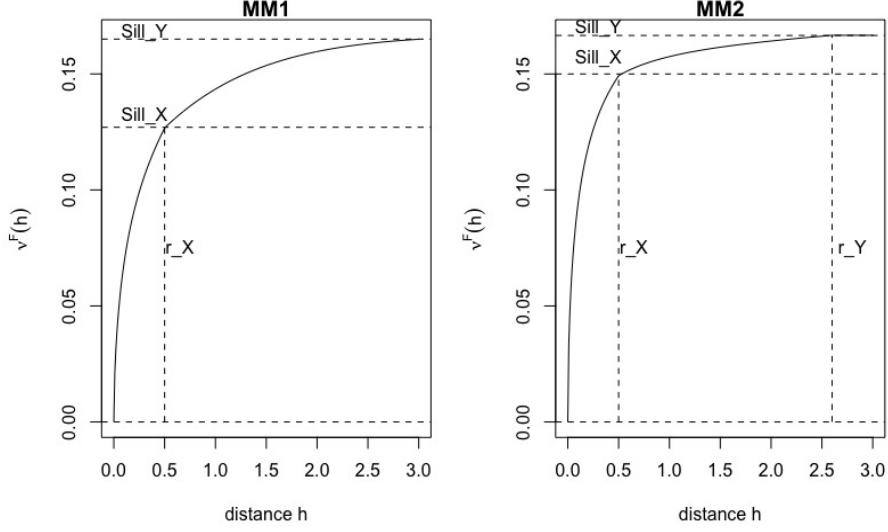


FIGURE 1. $h \rightsquigarrow \nu^F(h)$ for the max-mixture processes models MM1 and MM2. In MM1, X has correlation function $\rho(h) = \exp(-h/\theta_X)$, $r_X = 0.25$ and fixed radius $\theta_X = 0.2$ and Y has $\Sigma = \sigma_Y \mathbf{I}_d$, $\sigma_Y = 0.6$. In MM2, X has the same setting as in MM1 and Y has $\rho_Y(h) = \exp(-h/\theta_Y)$, $\theta_Y = 0.8$ and fixed radius $r_Y = 1.35$. For the two models, we set $a = 0.5$.

a large threshold, for example, [4] and [32]. In this article, we focus on max-mixture models. Composite likelihood inference for max-mixture processes has been studied in [5] and [36]. We will compare our madogram based estimation to the composite likelihood estimation described in [5].

If the pairwise density of Z can be computed and its parameter ψ is identifiable, then it is possible to estimate ψ by maximizing the pairwise weighted log likelihood. For simplicity, we denote Z_i^k for $Z^k(s_i)$. Let

$$\hat{\psi}_L = \max_{\psi} \mathcal{P}(\psi),$$

where

$$(4.1) \quad \mathcal{P}(\psi) = \sum_{k=1}^N \sum_{i=1}^{D-1} \sum_{j>i}^D w_{ij} \log \mathcal{L}(Z_i^k, Z_j^k; \psi) =: \sum_{k=1}^N \mathcal{P}_k(\psi).$$

where \mathcal{L} is the likelihood of the pair (Z_i^k, Z_j^k) and $w_{i,j} \geq 0$ is the weight that specifies the contribution for each pair. In [27], it is suggested to

In [10], it is suggested to consider a censor approach of the likelihood, taking into account a threshold. This approach has been adopted in this study. Let $G(\cdot, \cdot)$ be a pairwise distribution function and consider the thresholds u_1 and u_2 ; the likelihood contribution is

$$\mathcal{L}(z_1, z_2; \psi) = \begin{cases} \partial_{12}^2 G(z_1, z_2; \psi) & \text{if } z_1 > u_1, z_2 > u_2, \\ G(z_1, z_2; \psi) & \text{if } z_1 \leq u_1, z_2 \leq u_2, \end{cases}$$

In [36], the censored likelihood is used in order to improve the estimation of the parameters related to asymptotic independence. This censored approach was also applied by [5] for the estimation of parameters of max-mixture processes. In that paper, the replications Z^1, \dots, Z^N of Z are assumed to be α -mixing rather than

independent. We denote generically by ψ the parameters of the model. In [5], it is proved, under some smoothness assumptions on the composite likelihood, that the composite maximum likelihood estimator $\widehat{\psi}_L$ for max-mixture processes is asymptotically normal as N goes to infinity with asymptotic variance

$$\mathcal{G}(\psi) = \mathcal{J}(\psi)[\mathcal{K}(\psi)]^{-1}\mathcal{J}(\psi),$$

where $\mathcal{J}(\psi) = \mathbb{E}[-\nabla^2\mathcal{P}(\psi)]$, $\mathcal{K}(\psi) = \text{var}(\nabla\mathcal{P}(\psi))$. The matrix $\mathcal{G}(\psi)$ is called the Godambe information matrix (see [5] and Theorem 3.4.7 in [21]).

An estimator $\widehat{\mathcal{J}}$ of $\mathcal{J}(\psi)$ is obtained from the Hessian matrix computed in the optimization algorithm. The variability matrix $\mathcal{K}(\psi)$ has to be estimated too. In our context, we have independent replications of Z and N is large compared with respect to the dimension of ψ . Then, we can use the outer product of the estimation of $\widehat{\psi}$. Let

$$\widehat{\mathcal{K}}(\psi) = N^{-1} \sum_{k=1}^N \nabla\mathcal{P}_k(\widehat{\psi})\nabla\mathcal{P}_k(\widehat{\psi})'$$

or by Monte Carlo simulation with explicit formula of $\mathcal{P}_k(\psi)$ (see section 5. in [33]). In the case of samples of Z satisfying the α -mixing property, the estimation of $\mathcal{K}(\psi)$ can be done using a subsampling technique introduced by [18]; this was used in [5]. Finally, model selection can be done by using the composite likelihood information criterion [34]:

$$\text{CLIC} = -2 \left[\mathcal{P}(\widehat{\psi}) - \text{tr}(\widehat{\mathcal{J}}^{-1}\widehat{\mathcal{K}}) \right].$$

Considering several max-stable models, the one that has the smallest CLIC will be chosen. In [32], the criterion $\text{CLIC}^* = (D-1)^{-1}\text{CLIC}$ is proposed. It is close to Akaike information criterion (AIC).

4.2. Semi-parametric estimation using NLS of F-madogram. In this section, we shall define the non-linear least square estimation procedure of the parameters set ψ corresponding to the max-mixture model Z using the F -madogram. This procedure can be considered as an alternative method to the composite likelihood method.

Consider Z^k , $k = 1, \dots, N$ copies of an isotropic max-mixture process Z with unit Fréchet marginal laws (F denotes the distribution function of a unit Fréchet law). It may be independent copies for example, if the data is recorded yearly (see [25]) or we shall consider that $(Z^k)_{k=1, \dots, N}$ satisfies a α -mixing property ([5]). Let \mathcal{H} be a finite subset of \mathbb{S} , $J(x, y) = \frac{1}{2}|x - y|$ and $Y_{h,k} = J(F(Z^k(s)), F(Z^k(s+h)))$, $k = 1, \dots, N$ and $h \in \mathcal{H}$. Therefore, for $k = 1, \dots, N$, the vectors $(Y_{h,k})_{h \in \mathcal{H}}$ have the same law and are considered either independent or α -mixing (in k). The main motivation for using the F-madogram in estimation is that it contains the dependence structure information for a fixed h of $Y_{h,k}$ (see Section 3.2 in [3]).

In what follows, we make the assumption that the vectors $(Y_{h,k})_{h \in \mathcal{H}}$ are i.i.d. Note that from the definition of the F -madogram, we have $\mathbb{E}[Y_{h,k}] = \nu^F(h, \psi)$ where $\nu^F(h, \psi)$ is the F -madogram of Z with parameters ψ defined in (3.1). If Z has an unknown true parameter ψ_0 on a compact set $\Psi \subset \mathbb{R}^d$, we rewrite

$$(4.2) \quad Y_{h,k} = \nu^F(h, \psi_0) + \varepsilon_{h,k}.$$

The vectors $(\varepsilon_{h,k})_{h \in \mathcal{H}}$ are i.i.d errors with $\mathbb{E}[\varepsilon_{h,k}] = 0$ and $\text{Var}(\varepsilon_{h,k}) = \sigma_h^2 > 0$ is finite and unknown.

Let

$$(4.3) \quad \mathcal{L}(\psi) = \sum_{h \in \mathcal{H}} \frac{1}{N} \sum_{k=1, \dots, N} (Y_{h,k} - \nu^F(h, \psi))^2$$

Any vector $\widehat{\psi}_M$ in Ψ which minimizes $\mathcal{L}(\psi)$ will be called a least square estimate of ψ_0 :

$$(4.4) \quad \widehat{\psi}_M \in \underset{\psi \in \Psi}{\operatorname{argmin}} \mathcal{L}(\psi).$$

Theorem 4.1. *Assume that $\Psi \subset \mathbb{R}^d$ is compact and that $\psi \mapsto \nu^F(h, \psi)$ is continuous for all $h \in \mathcal{H}$. We assume that the vectors $(Y_{h,k})_{h \in \mathcal{H}}$ are i.i.d. Let $(\widehat{\psi}_M^N)_{N \in \mathbb{N}}$ be least square estimators of ψ_0 . Then, any limit point (as N goes to infinity) ψ of $(\widehat{\psi}_M^N)_{N \in \mathbb{N}}$ satisfies $\nu(h, \psi) = \nu(h, \psi_0)$ for all $h \in \mathcal{H}$.*

Remark 4.1. *Of course, if $\psi \rightsquigarrow (\nu(h, \psi))_{h \in \mathcal{H}}$ is injective, then Theorem 4.1 implies that the least square estimation is consistent, i.e. $\widehat{\psi}_M^N \rightarrow \psi_0$ a.s. as T goes to infinity. In the examples considered below, it seems that the injectivity is satisfied provided $|\mathcal{H}| \geq d$, but we were unable to prove it.*

Proof. We follow the proof of Theorem II.5.1 in [2]. From (4.2), we have, for all $\psi \in \Psi$

$$\begin{aligned} \mathcal{L}(\psi) &= \sum_{h \in \mathcal{H}} \frac{1}{N} \sum_{k=1, \dots, N} (\nu^F(h, \psi_0) + \varepsilon_{h,k} - \nu^F(h, \psi))^2 \\ &= \sum_{h \in \mathcal{H}} (\nu^F(h, \psi_0) - \nu^F(h, \psi))^2 \\ &\quad + \frac{2}{N} \sum_{h \in \mathcal{H}} (\nu^F(h, \psi_0) - \nu^F(h, \psi)) \left(\sum_{k=1, \dots, N} \varepsilon_{h,k} \right) + \sum_{h \in \mathcal{H}} \frac{1}{N} \sum_{k=1, \dots, N} \varepsilon_{h,k}^2. \end{aligned}$$

From the law of large numbers, we have

$$\frac{1}{N} \sum_{h \in \mathcal{H}} \sum_{k=1, \dots, N} \varepsilon_{h,k}^2 \rightarrow \sum_{h \in \mathcal{H}} \sigma_h^2 \quad \text{a.s. as } N \rightarrow \infty$$

and for any $h \in \mathcal{H}$,

$$\frac{1}{N} \sum_{k=1, \dots, N} \varepsilon_{h,k} \rightarrow 0 \quad \text{a.s.}$$

Therefore,

$$\mathcal{L}(\psi) \rightarrow \sum_{h \in \mathcal{H}} \sigma_h^2 + \sum_{h \in \mathcal{H}} (\nu^F(h, \psi_0) - \nu^F(h, \psi))^2 \quad \text{a.s. as } N \rightarrow \infty.$$

Take a sequence $(\widehat{\psi}_M^N)_{N \in \mathbb{N}}$ of least square estimators, taking if necessary a subsequence, we may assume that it converges to some $\psi^* \in \Psi$. Using the continuity of $\psi \rightsquigarrow \nu^F(h, \psi)$, we have

$$\mathcal{L}(\widehat{\psi}_M^N) \rightarrow \sum_{h \in \mathcal{H}} \sigma_h^2 + \sum_{h \in \mathcal{H}} (\nu^F(h, \psi_0) - \nu^F(h, \psi^*))^2 \quad \text{a.s. as } N \rightarrow \infty.$$

Since $\widehat{\psi}_M^N$ is a least square estimator, $\mathcal{L}(\widehat{\psi}_M^N) \leq \mathcal{L}(\psi_0) \rightarrow \sum_{h \in \mathcal{H}} \sigma_h^2$. It follows that

$$\sum_{h \in \mathcal{H}} (\nu^F(h, \psi_0) - \nu^F(h, \psi^*))^2 = 0$$

and thus $\nu(h, \psi^*) = \nu(h, \psi_0)$ for all $h \in \mathcal{H}$. \square

The asymptotic normality of the least square estimators should also be obtained by following, e.g., [8] and using the asymptotic normality of the F -madogram obtained in [12]. Nevertheless, the calculation of the asymptotic variance will require to calculate the covariances between $\nu^F(h_1, \psi)$ and $\nu^F(h_2, \psi)$, which is not straightforward.

5. SIMULATION STUDY

This section is devoted to some simulations in order to evaluate the performance of the least square estimator and to compare it with the maximum composite likelihood estimator. Recall that $\widehat{\psi}_M$ denotes the least square estimator of the parameter vector ψ and $\widehat{\psi}_L$ denotes the composite likelihood estimator.

5.1. Outline the simulation experiment. In order to evaluate the performance of the non-linear least square estimator $\widehat{\psi}_M$ as defined in (3.1), we have generated data from the model MM1 above. The least square madogram estimator $\widehat{\psi}_M$ has been compared with true one ψ_0 and also with parameters estimated by maximum composite likelihood $\widehat{\psi}_L$ proposed in [5, 36], on the same simulated data. We considered 50 sites randomly and uniformly distributed in the square $\mathcal{A} = [0, 1]^2$. We have generated $N = 1.000$ i.i.d observations for each site and replicated this experiment $J = 100$ times. We have considered several mixing parameters: $a := \{0, 0.25, 0.5, 0.75, 1\}$. For the composite likelihood estimator $\widehat{\psi}_L$, we used the censored procedure with the 0.9 empirical quantile of data at each site as threshold u . The fitting of $\widehat{\psi}_L$ was done using the code which was used in [5] with some appropriated modifications.

5.2. Results on the parameters estimate. In Figure 2, we represented the boxplots of the errors, that is $(\widehat{\psi}_M - \psi_0)$ and $(\widehat{\psi}_L - \psi_0)$ for model MM1. Generally, the estimators above worked well, although the variability in some estimates were relatively large, especially for the asymptotic independence parameters. It also shows some bias in the estimation of the asymptotic independence model parameters.

It is well known that asymptotic independence is difficult to estimate (see [15]). Therefore, the estimation accuracy of the parameters is very sensitive. On one other hand, the fitting of $\alpha(h)$ which appears in TEG models in (2.17), is delicate and might lead to quite different estimates with different data [14]. Another comparison indicator is the root mean square error (RMSE) ([37, 38]): let $\widehat{\psi}_j$ denote the j th estimation (either least square or composite likelihood estimation),

$$(5.1) \quad \text{RMSE} = \left[J^{-1} \sum_{j=1}^J (\widehat{\psi}_j - \psi)^2 \right]^{1/2},$$

The barplot in Figure 3 displays the RMSE for each parameter of MM1 model. We see on these barplots that when a is close to 0 ($a = 0; 0.25$), the estimator $\widehat{\psi}_M$ over-performs the estimator $\widehat{\psi}_L$ and vice versa when $a \in \{0.75, 1\}$. For $a = 0.5$ the performance of the two estimators seem equivalent.

5.3. Asymptotic normality. We may figure out whether the least square estimator $\widehat{\psi}_M$ is asymptotically normal through the graphe of the errors $\widehat{\psi}_L - \psi_0$. In Figure 4, the graphs represent the distributions of the errors of each estimated parameters of MM1 model for $a := \{0, 0.5, 1\}$. We did $J = 500$ simulation experiments. We can see on Figure 4 that the densities of the errors of the parameters seem close to the shape of centered normal distribution.

6. REAL DATA EXAMPLE

In this section, we analyze a real data and fit it to some models considered in this study by composite likelihood and LS-madogram procedures.

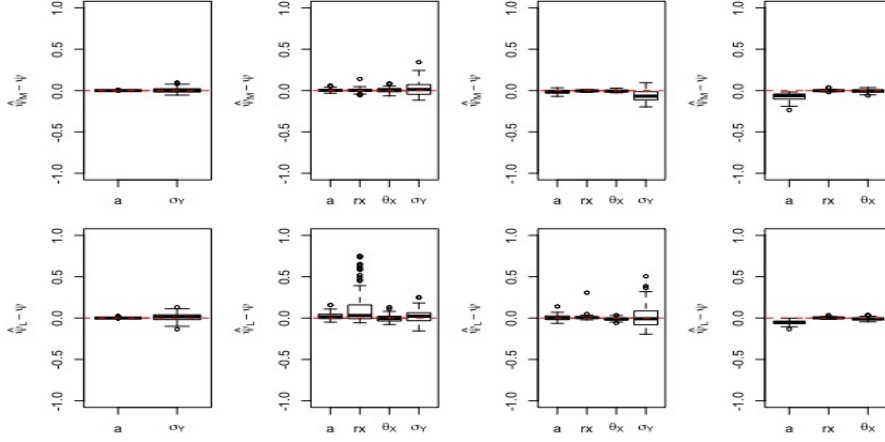


FIGURE 2. Boxplots display $(\hat{\psi} - \psi)$ of estimated parameters vector $\hat{\psi} = (\hat{a}, \hat{r}_X, \hat{\theta}_X, \hat{\sigma}_Y)^T$ for the MM1 model by the two estimators $\hat{\psi}_M$ and $\hat{\psi}_L$. The figures in the first row and from left to right concern the estimator $\hat{\psi}_M$ for $a \in \{0, 0.25, 0.75, 1\}$, the second row concerns $\hat{\psi}_L$. We have set, $r_X = 0.25$, $\theta_X = 0.20$ and $\sigma_Y = 0.6$ over a square $\mathcal{A} = [0, 1]^2$.

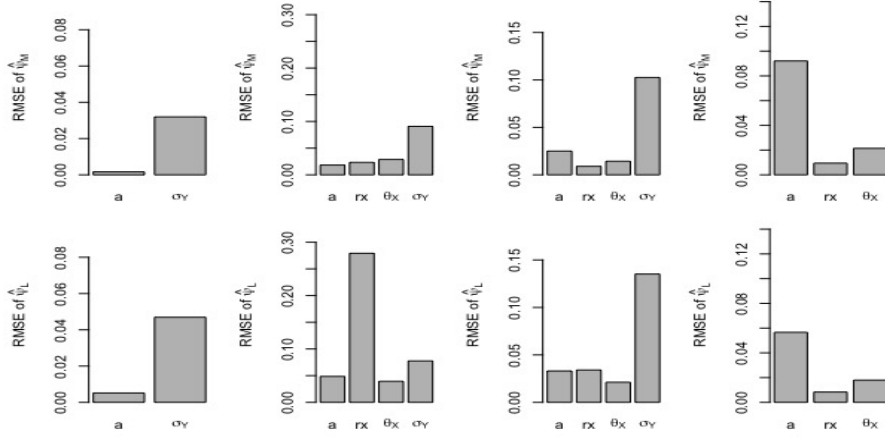


FIGURE 3. Barplots display the RMSE of $\hat{\psi}$ for each estimated parameters $\hat{\psi} = (\hat{a}, \hat{r}_X, \hat{\theta}_X, \hat{\sigma}_Y)^T$ for MM1 and the corresponding two estimators $\hat{\psi}_M$ and $\hat{\psi}_L$. The bars in the first row and from left to right represent the RMSE of the estimator $\hat{\psi}_M$ when $a := \{0, 0.25, 0.75, 1\}$, respectively and the same for the second row for $\hat{\psi}_L$. We set $r_X = 0.25$, $\theta_X = 0.20$ and $\sigma_Y = 0.6$ over a square $\mathcal{A} = [0, 1]^2$.

6.1. Data analysis. We analyzed daily rainfall along the east coast of Australia. We selected 39 locations randomly from this region. The data is daily measured in the period from April to September for 35 years from 1982-2016. The data is available from the Australian Bureau of Meteorology (<http://www.bom.gov.au/climate/data/>).

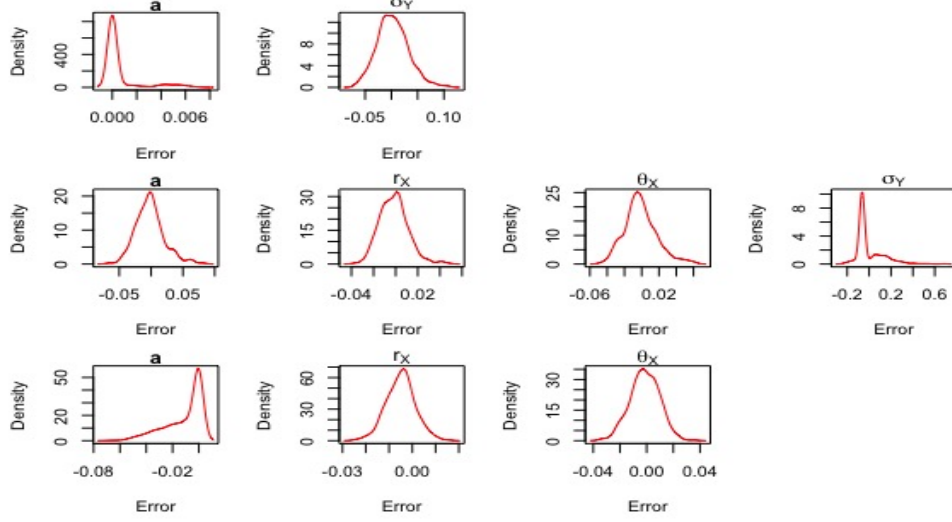


FIGURE 4. Graphs display the densities of the errors between $\hat{\psi}_M$ and ψ_0 for each estimated parameters in the set $\psi = (a, r_X, \theta_X, \sigma_Y)^T$ of MM1 model. The graphs from first row and last one represent the densities of the errors $\hat{\psi}_M - \psi_0$ when $a := \{0, 0.5, 1\}$. We set $r_X = 0.25$, $\theta_X = 0.20$ and $\sigma_Y = 0.6$ over a square $\mathcal{A} = [0, 1]^2$.

In order to explore the possibility of anisotropy of the spatial dependence, we used the same test as in [5]. We divided all data set according to directional sectors $(-\pi/8, \pi/8]$, $(\pi/8, 3\pi/8]$, $(3\pi/8, 5\pi/8]$ and $(5\pi/8, 7\pi/8]$, where 0 indicate to north direction. We use the empirical F-madogram $\hat{\nu}^F(h)$. The directional loss smoothing of such empirical measure in the Figure 5. (A), shows no evidence of anisotropy.

In Figure 5. (B), the empirical F-madogram is plotted for the whose data set. It seems that asymptotic dependence between the locations is present up to distance 500 km and asymptotic independence could be present at the remaining distance.

6.2. Data fitting. Our interest in this section is to chose a reasonable model for the data. We considered 7 models described below, for each model, the parameters are estimated by LS-madogram and maximum composite likelihood. The selection criteria for LS-madogram estimator $\hat{\psi}_M$ is computed as the following:

$$MIC := \log \mathcal{L}(\hat{\psi}) + (2k(k+1)/(T-k-1)),$$

where k is the number of parameters in the model and T is the number of the observations, that is: $T = N \times |\mathcal{H}|$, where $|\mathcal{H}|$ is the number of observed pairs. With respect to censored composite likelihood estimators we adopted the CLIC selection criteria. The two criteria selected the model MM1 as the best model.

We consider the following models:

MM1: max-mixture between asymptotic dependence process represented by TEG max-stable process X with exponential correlation function $\rho(h) = \exp\{-(h/\theta_X)\}$, $\theta_x > 0$ and \mathcal{B}_X is a disk of fixed and unknown radius r_X . The asymptotic independence is represented by an inverse Brown-Resnik max-stable process Y with variogram $2\gamma(h) = 2\sigma^2(1 - \exp\{-(h/\theta_Y)\})$, $\theta_Y, \sigma > 0$; σ^2 is the sill of the variogram.

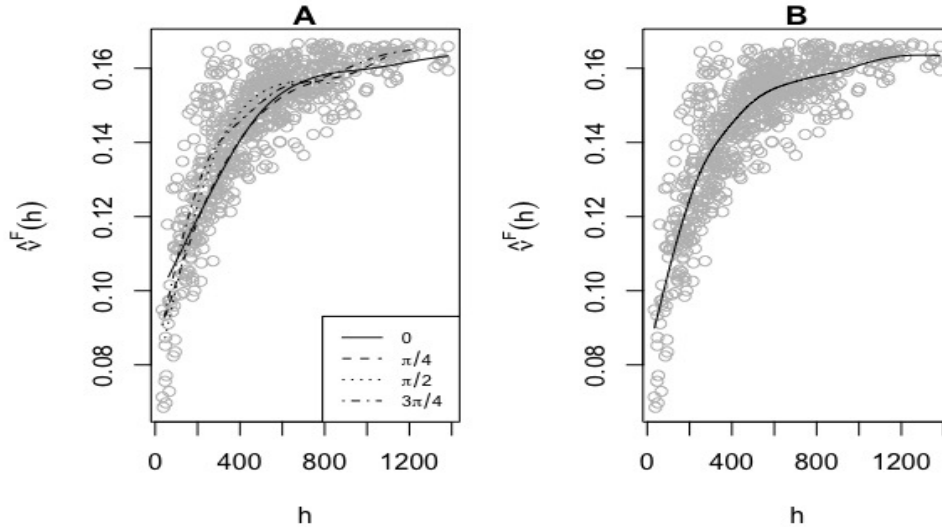


FIGURE 5. Empirical evaluation of $\hat{\nu}^F(h)$. The Grey circles represented the empirical value between all pairs. The lines in (A) represent the smoothed value of the empirical of $\hat{\nu}^F(h)$ according to directional sectors $(-\pi/8, \pi/8]$, $(\pi/8, 3\pi/8]$, $(3\pi/8, 5\pi/8]$ and $(5\pi/8, 7\pi/8]$. The line in (B) represent the smoothed value of the empirical of $\hat{\nu}^F(h)$ for all directions.

MM2: max-mixture between X as in MM1 and an inverse inverse Smith max-stable process Y with $\tau(h) = h/\sqrt{\sigma_Y}$.

M1: A TEG max-stable process X specified in MM1.

M2: A Brown-Resnik max-stable process X as specified in MM1.

M3: An inverse Brown-Resnik max-stable process Y as specified in MM1.

M4: A Smith max-stable process X as specified in M2.

M5: A inverse Smith max-stable process Y as specified in M2.

For all the considered models, the margin distribution are assumed to be unit Fréchet. Therefore it requires to transform the data set to Fréchet. Most of papers (see for example [5] and [36]) use parametric transformations: they fit GEV parameters for each location separately and then transform the data to Fréchet. In this study, we adopted a non-parametric transformation by using the empirical c.d.f.. For censored composite likelihood procedures, we set $u = 0.9$ and $\delta = \infty$.

We remark that the two Decision Criteria (CLIC and MIC) choose the same model. We would like to emphasize that CLIC and MIC are not comparable, one is related to the composite likelihood while the other one is related to the least squared madogram difference. For the tested models, we would keep the model with the smallest CLIC or the smallest MIC, depending on the used estimation method. Also, the least squared madogram estimation is involves less computations, indeed, the maximum composite likelihood estimation requires to estimate the Godambe matrix in order to compute the CLIC.

7. CONCLUSIONS

The calculation of the F-madogram for max-mixture processes show that it writes with both the AD and the AI parameters. This leads us to propose a semi-parametric estimation procedure using F-madogram $\nu^F(h)$ as an alternative

TABLE 1. Summary of the fitted models. the distance scale is kilometres. Composite likelihood procedure indicated by CL and LS-madogram procedure indicated by LS with selection criteria SC are CLIC and MIC, respectively.

Model		a	θ_X	r_X	θ_Y	σ_Y	SC
MM1	CL	0.262	1217.3	1364.5	3102.4	3.457	6807406
	LS	0.259	1285.7	1390.0	5794.8	2.013	1.917034
MM2	CL	0.248	31.16	70.15	998.84		7924609
	LS	0.185	35.51	48.14	871.19		1.917234
			θ_X	r_X			
M1	CL	931	307.86				7926261
	LS	1270	255.64				1.945177
			θ_X	σ_X	θ_Y	σ_Y	
M2	CL	931.02	3.078663				7926261
	LS	361.36	1.90816				1.96165
M3	CL			1644.76	2.702282		7918643
	LS			1383.08	1.394928		1.924574
M4	CL		85.34				8016633
	LS		193.43				1.988753
M5	CL					256.39	7988838
	LS					334.60	1.929235

to composite likelihood. The simulation study showed that the estimation procedure based on $\nu^F(h)$ performs better than the composite likelihood procedure when the model is near to asymptotic independence. We applied these estimator procedures to real data example. On the considered example, the results obtained by composite likelihood maximization and least squared madogram difference are similar.

Acknowledgements: This work was supported by the LABEX MILYON (ANR-10-LABX-0070) of Université de Lyon, within the program “Investissements d’Avenir” (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR). We also acknowledge the projet LEFE CERISE.

REFERENCES

[1] A. AghaKouchak and N. Nasrollahi, *Semi-parametric and parametric inference of extreme value models for rainfall data*, Water resources management **24** (2010), no. 6, 1229–1249.

[2] A. Antoniadis, J. Berruyer, and C. René, *Régression non linéaire et applications*, Economica, 1992.

[3] J-N. Bacro, L. Bel, and C. Lantuéjoul, *Testing the independence of maxima: from bivariate vectors to spatial extreme fields*, Extremes **13** (2010), no. 2, 155–175.

[4] J-N. Bacro and C. Gaetan, *Estimation of spatial max-stable models using threshold exceedances*, Statistics and Computing **24** (2014), no. 4, 651–662.

[5] J-N. Bacro, C. Gaetan, and G. Toulemonde, *A flexible dependence model for spatial extremes*, Journal of Statistical Planning and Inference **172** (2016), 36–52.

[6] J-N. Bacro and G. Toulemonde, *Measuring and modelling multivariate and spatial dependence of extremes*, Journal de la Société Française de Statistique **154** (2013), no. 2, 139–155.

[7] J. Beirlant, Y. Goegebeur, J. Segers, and J. Teugels, *Statistics of extremes: theory and applications*, John Wiley & Sons, 2006.

[8] S. Buhl, R. A. Davis, C. Klüppelberg, and C. Steinkohl, *Semiparametric estimation for isotropic max-stable space-time processes*, submitted (2016).

- [9] T.A. Buishand, *Bivariate extreme-value data and the station-year method*, Journal of Hydrology **69** (1984), no. 1, 77–95.
- [10] S. Coles, J. Bawa, L. Trenner, and P. Dorazio, *An introduction to statistical modeling of extreme values*, Vol. 208, Springer, 2001.
- [11] S. Coles, J. Heffernan, and J.A. Tawn, *Dependence measures for extreme value analyses*, Extremes **2** (1999), no. 4, 339–365.
- [12] D. Cooley, P. Naveau, and P. Poncet, *Variograms for spatial max-stable random fields*, Dependence in probability and statistics, 2006, pp. 373–390.
- [13] R. A. Davis and C. Y. Yau, *Comments on pairwise likelihood in time series models*, Statistica Sinica (2011), 255–277.
- [14] A. C. Davison and M. Gholamrezaee, *Geostatistics of extremes*, Proceedings of the royal society of london a: Mathematical, physical and engineering sciences, 2012, pp. 581–608.
- [15] A.C. Davison, R. Huser, and E. Thibaud, *Geostatistics of dependent and asymptotically independent extremes*, Mathematical Geosciences **45** (2013), no. 5, 511–529.
- [16] L. De Haan, *A spectral representation for max-stable processes*, The annals of probability (1984), 1194–1204.
- [17] L. De Haan and T. T. Pereira, *Spatial extremes: Models for the stationary case*, The annals of statistics (2006), 146–168.
- [18] C. Edward, *The use of subseries values for estimating the variance of a general statistic from a stationary sequence*, The Annals of Statistics (1986), 1171–1179.
- [19] C. Fonseca, L. Pereira, H. Ferreira, and A.P. Martins, *Generalized madogram and pairwise dependence of maxima over two regions of a random field*, KYBERNETIKE **51** (2015), no. 2, 193–211.
- [20] A. Guillo, P. Naveau, and A. Schorgen, *Madogram and asymptotic independence among maxima*, REVSTAT—Statistical Journal **12** (2014), no. 2, 119–134.
- [21] X. Guyon, *Random fields on a network: modeling, statistics, and applications*, Springer Science & Business Media, 1995.
- [22] Z. Kabluchko, M. Schlather, and L. De Haan, *Stationary max-stable fields associated to negative definite functions*, The Annals of Probability (2009), 2042–2065.
- [23] A.W. Ledford and J.A. Tawn, *Statistics for near independence in multivariate extreme values*, Biometrika **83** (1996), no. 1, 169–187.
- [24] B.G. Lindsay, *Composite likelihood methods*, Contemporary Math. (1988), 221–239.
- [25] P. Naveau, A. Guillo, D. Cooley, and J. Diebolt, *Modelling pairwise dependence of maxima in space*, Biometrika **96** (2009), no. 1, 1–17.
- [26] P. J. Northrop, *An efficient semiparametric maxima estimator of the extremal index*, Extremes **18** (2015), no. 4, 585–603.
- [27] S. A. Padoan, M. Ribatet, and S. A. Sisson, *Likelihood-based inference for max-stable processes*, Journal of the American Statistical Association **105** (2010), no. 489, 263–277.
- [28] M. Schlather, *Models for stationary max-stable random fields*, Extremes **5** (2002), no. 1, 33–44.
- [29] M. Schlather and J.A. Tawn, *A dependence measure for multivariate and spatial extreme values: Properties and inference*, Biometrika **90** (2003), no. 1, 139–156.
- [30] M. Sibuya, *Bivariate extreme statistics, i*, Annals of the Institute of Statistical Mathematics **11** (1959), no. 2, 195–210.
- [31] R. L. Smith, *Max-stable processes and spatial extremes*, Unpublished manuscript, Univer (1990).
- [32] E. Thibaud, R. Mutzner, and A.C. Davison, *Threshold modeling of extreme spatial rainfall*, Water resources research **49** (2013), no. 8, 4633–4644.
- [33] C. Varin, N. Reid, and D. Firth, *An overview of composite likelihood methods*, Statistica Sinica (2011), 5–42.
- [34] C. Varin and P. Vidoni, *A note on composite likelihood inference and model selection*, Biometrika (2005), 519–528.
- [35] H. Wackernagel, *Multivariate nested variogram*, Multivariate geostatistics, 1998, pp. 172–180.
- [36] J.L. Wadsworth and J.A. Tawn, *Dependence modelling for spatial extremes*, Biometrika **99** (2012), no. 2, 253–272.
- [37] F. Zheng, E. Thibaud, M. Leonard, and S. Westra, *Assessing the performance of the independence method in modeling spatial extreme rainfall*, Water Resources Research **51** (2015), no. 9, 7744–7758.
- [38] F. Zheng, S. Westra, M. Leonard, and S. A. Sisson, *Modeling dependence between extreme rainfall and storm surge to estimate coastal flooding risk*, Water Resources Research **50** (2014), no. 3, 2050–2071.

UNIVERSITÉ DE LYON, UNIVERSITÉ LYON 1, INSTITUT CAMILLE JORDAN ICJ UMR 5208 CNRS,
FRANCE, DEPARTMENT OF STATISTICS, UNIVERSITY OF MOSUL, IRAQ

UNIVERSITÉ DE LYON, UNIVERSITÉ LYON 1, INSTITUT CAMILLE JORDAN ICJ UMR 5208 CNRS,
FRANCE

UNIVERSITÉ DE LYON, UNIVERSITÉ LYON 1, INSTITUT CAMILLE JORDAN ICJ UMR 5208 CNRS,
FRANCE

UNIVERSITÉ DE LYON, UNIVERSITÉ LYON 1, INSTITUT CAMILLE JORDAN ICJ UMR 5208 CNRS,
FRANCE