Applied Probability Trust (11th March 2016)

SPACE-TIME MAX-STABLE MODELS WITH SPECTRAL SEPARABILITY

PAUL EMBRECHTS * AND
ERWAN KOCH,* *ETH Zürich (Department of Mathematics, RiskLab)*
CHRISTIAN ROBERT,** *ISFA Université Lyon 1*

Abstract

Natural disasters may have considerable impact on society as well as on the (re-)insurance industry. Max-stable processes are ideally suited for the modelling of the spatial extent of such extreme events, but it is often assumed that there is no temporal dependence. Only a few papers have introduced spatio-temporal max-stable models, extending the Smith, Schlather and Brown-Resnick spatial processes. These models suffer from two major drawbacks: time plays a similar role to space and the temporal dynamics are not explicit. In order to overcome these defects, we introduce spatio-temporal max-stable models where we partly decouple the influence of time and space in their spectral representations. We introduce both continuous- and discrete-time versions. We then consider particular Markovian cases with a max-autoregressive representation and discuss their properties. Finally, we briefly propose an inference methodology which is tested through a simulation study.

Keywords: extreme value theory; spatio-temporal max-stable processes; spectral separability; temporal dependence

2010 Mathematics Subject Classification: Primary 60G70
Secondary 60G60; 62M30

1. Introduction

In the context of climate change, some extreme events tend to be more and more frequent; see e.g. Swiss Re (2014). Meteorological and more generally environmental disasters have a considerable impact on society as well as on the (re-)insurance industry. Hence, the statistical modelling of extremes constitutes a crucial challenge. Extreme value theory (EVT) provides powerful statistical tools for this purpose.

EVT can basically be divided into three different streams closely linked to each other: the univariate case, the multivariate case and the theory of max-stable processes. For an introduction to the univariate theory, see e.g. Coles (2001) and for a detailed description, see e.g. Embrechts et al. (1997) or Beirlant et al. (2004). In the multivariate case, we refer to Resnick (1987), Beirlant et al. (2004) and de Haan and Ferreira (2007). Max-stable processes constitute an extension of EVT to the level of stochastic processes (de Haan, 1984; de Haan and Pickands, 1986) and are very well suited for the modelling of spatial extremes. Indeed, it can be shown that the distribution of

* Postal address: ETH Zürich, Department of Mathematics, Rämistrasse 101, 8092 Zürich, Switzerland
** Postal address: ISFA, 50 Avenue Tony Garnier, 69366 LYON CEDEX 07, France

the random field of the suitably normalized temporal maxima of independent and identically distributed (iid) random fields at each point of the space is necessarily max-stable when the number of temporal observations tends to infinity. For a detailed overview of max-stable processes, we refer to de Haan and Ferreira (2007).

In the literature about max-stable processes, measurements are often assumed to be independent in time and thus only the spatial structure is studied (see e.g. Padoan et al., 2010). Nevertheless, the temporal dimension should be taken into account in a proper way. To the best of our knowledge, only a few papers focus on such a question. The majority of the spatio-temporal models introduced is based on Schlather's spectral representation (Penrose, 1992; Schlather, 2002) which has given rise to the well-known Schlather (Schlather, 2002) and Brown-Resnick (Kabluchko et al., 2009) processes. This representation tells us that if $(U_i)_{i\geq 1}$ generates a Poisson point process on $(0,\infty)$ with intensity $u^{-2}\,\mathrm{d}u$ and $((Y_i(\mathbf{y}))_{\mathbf{y}\in\mathbb{R}^d})_{i\geq 1}$ are iid non-negative stationary stochastic processes such that $\mathbb{E}[Y_i(\mathbf{y})] = 1$ for each $\mathbf{y} \in \mathbb{R}^d$, then the process $(\bigvee_{i=1}^{\infty}\{U_i Y_i(\mathbf{y})\})_{\mathbf{y}\in\mathbb{R}^d}$ is stationary simple max-stable, where simple means that the margins are standard Fréchet. Here $\bigvee$ denotes the max-operator. In Davis et al. (2013a), Huser and Davison (2014) and Buhl and Klüppelberg (2015), the idea underlying the construction of the spatio-temporal model is to divide the dimension d into dimension $d-1$ for the spatial component and dimension 1 for time. Davis et al. (2013a) introduce the Brown–Resnick model in space and time by taking a log-normal process for Y_i while Buhl and Klüppelberg (2015) introduce an extension of this model to the anisotropic setting. Huser and Davison (2014) consider an extension of the Schlather model by using a truncated Gaussian process for the Y_i and a random set that allows the process to be mixing in space as well as to exhibit a spatial propagation. Advantages of these models lie in the facts that the Schlather and Brown-Resnick models have been widely studied and that the large literature about spatio-temporal correlation functions for Gaussian processes can be used, allowing for a considerable diversity of spatio-temporal behaviour. Davis et al. (2013a) also introduce the spatio-temporal version of the Smith model (Smith, 1990) that is based on de Haan's spectral representation (see de Haan, 1984). If $(U_i, \mathbf{C}_i)_{i\geq 1}$ are the points of a Poisson point process on $(0,\infty)\times\mathbb{R}^d$ with intensity $u^{-2}\,\mathrm{d}u \times \mathrm{d}\mathbf{c}$ and if $g_{\mathbf{y}}$ are measurable non-negative functions satisfying $\int_{\mathbb{R}^d} g_{\mathbf{y}}(\mathbf{c})\,\mathrm{d}\mathbf{c} = 1$ for each $\mathbf{y}\in\mathbb{R}^d$, then the process $(\bigvee_{i=1}^{\infty}\{U_i g_{\mathbf{y}}(\mathbf{C}_i)\})_{\mathbf{y}\in\mathbb{R}^d}$ is a simple max-stable process. However, they do not allow any interaction between the spatial components and the temporal one in the underlying covariance matrix. The previous spatio-temporal max-stable models suffer from some defects. First, they are all continuous-time processes whereas measurements in environmental science are often time-discrete. Second, time has no specific role but is equivalent to an additional spatial dimension. Especially, the spatial and temporal distributions belong to a similar class of models. This constitutes a serious drawback since such a similarity is not supported by any physical argument. Third, the temporal dynamics are not explicit and hence are difficult to identify and interpret. Finally, these models have in general no causal representation.

In this paper, we propose a class of models where we partly decouple the influence of time and space, but such that time influences space through a bijective operator on space. We present both continuous- and discrete-time versions. A first advantage of this class of models lies in their flexibility since they allow the distributions in time (when the location is fixed) to belong to a class different from the distributions in space (when time is fixed) and hence can be chosen with regard to their function in the application. Because of the spatial operator mentioned above, our models are able to account for physical processes such as propagations/contagions/diffusions. Furthermore, the estimation procedure can be simplified since the purely spatial parameters can be estimated independently of the purely temporal ones.

Then we study some particular sub-classes of our general class of models, where the function related to time in the spectral representation is the exponential density (in the continuous-time case) or takes as values the probabilities of a geometric random variable (in the discrete-time case). In this context, our models become Markovian and have a max-autoregressive representation. This makes the dynamics of these models explicit and easy to interpret physically.

The remainder of the paper is organized as follows. Section 2 presents our class of spectrally separable space-time max-stable models. In Section 3, we focus on the particular Markovian cases where the space is $\mathbb{R}^2$ and the unit sphere in $\mathbb{R}^3$, respectively. Section 4 briefly presents an estimation procedure as well as an application of the latter on simulated data. Some concluding remarks are given in Section 5. Throughout the paper, the elements belonging to $\mathbb{R}^d$ for some $d \geq 2$ are denoted using bold symbols whereas those in more general spaces are in light font.

2. A new class of space-time max-stable models

2.1. The models

The time index t and space index x will belong respectively to the sets $\mathcal{I}$ and $\mathcal{X}$. The models we introduce will be either continuous-time ($\mathcal{I} = \mathbb{R}$) or discrete-time ($\mathcal{I} = \mathbb{Z}$). In the following, we denote by δ the Lebesgue measure on $\mathbb{R}$ in the case $\mathcal{I} = \mathbb{R}$ and the counting measure $\sum_{z \in \mathbb{Z}} \partial_{\{z\}}$ when $\mathcal{I} = \mathbb{Z}$, where ∂ stands for the Dirac measure.

For the definition of the discrete-time models below, let $(N_k)_{k \in \mathbb{Z}}$ be iid Poisson(1) and define, for $A \subset \mathbb{Z}$, $N(A) = \sum_{k \in A} N_k$, a Poisson random measure on $\mathbb{Z}$ with intensity one, i.e. $N(A)$ is Poisson-distributed with parameter $\delta(A)$ and for any $l \geq 1$ and $A_1, \dots, A_l$ disjoint sets in $\mathbb{Z}$, the $N(A_i)$, $i = 1, \dots, l$, are independent random variables.

Space-time simple max-stable processes on $\mathcal{I} \times \mathcal{X}$ allow for a spectral representation of the following form (see e.g. de Haan (1984)):

$$X(t, x) = \bigvee_{i=1}^{\infty} \{U_i V_{(t,x)}(W_i)\}, \quad (t, x) \in \mathcal{I} \times \mathcal{X}, \tag{1}$$

where $(U_i, W_i)_{i \geq 1}$ are the points of a Poisson point process on $(0, \infty) \times E$ with intensity $u^{-2}\,\mathrm{d}u \times \mu(\mathrm{d}w)$ for some Polish measure space $(E, \mathcal{E}, \mu)$ and the functions $V_{(t,x)} : E \to (0, \infty)$ are measurable such that $\int_E V_{(t,x)}(w)\mu(\mathrm{d}w) = 1$ for each $(t, x) \in \mathcal{I} \times \mathcal{X}$. A class of space-time max-stable models avoiding the previously mentioned shortcomings is introduced below.

Definition 1. (*Space-time max-stable models with spectral separability.*) The class of space-time max-stable models with spectral separability is defined by inserting the following spectral decomposition in (1):

$$V_{(t,x)}(W_i) = V_t(B_i) V_{R_{(t,B_i)}x}(C_i),$$

where:

- $(U_i, B_i, C_i)_{i\geq 1}$ are the points of a Poisson point process on $(0,\infty)\times E_1\times E_2$ with intensity $u^{-2}\,\mathrm{d}u \times \mu_1(\mathrm{d}b)\times\mu_2(\mathrm{d}c)$ for some Polish measure spaces $(E_1,\mathcal{E}_1,\mu_1)$ and $(E_2,\mathcal{E}_2,\mu_2)$;
- the operators $R_{(t,b)}$ are bijective from $\mathcal{X}$ to $\mathcal{X}$ for each $(t,b)\in\mathcal{I}\times E_1$;
- the functions $V_t : E_1 \to (0,\infty)$ are measurable such that $\int_{E_1} V_t(b)\mu_1(\mathrm{d}b) = 1$ for each $t\in\mathcal{I}$ and the functions $V_x : E_2 \to (0,\infty)$ are measurable such that $\int_{E_2} V_x(c)\mu_2(\mathrm{d}c) = 1$ for each $x\in\mathcal{X}$.

We emphasize that the models belonging to this class are max-stable in space and time, since

$$\begin{aligned}\int_E V_{(t,x)}(w)\mu(\mathrm{d}w) &= \int_{E_1\times E_2} V_t(b)V_{R_{(t,b)}x}(c)\mu_1(\mathrm{d}b)\mu_2(\mathrm{d}c)\\ &= \int_{E_1} V_t(b)\int_{E_2} V_{R_{(t,b)}x}(c)\mu_2(\mathrm{d}c)\mu_1(\mathrm{d}b)\\ &= \int_{E_1} V_t(b)\mu_1(\mathrm{d}b) = 1,\end{aligned}$$

but of course also in space and in time only. A spectral decomposition in space, for example, is easily derived since, for a fixed t, $(U_iV_t(B_i), C_i)_{i\geq 1}$ defines a Poisson point process on $(0,\infty)\times E_2$ with intensity $u^{-2}\,\mathrm{d}u\times\mu_2(\mathrm{d}c)$, and $\int_{E_2} V_{R_{(t,b)}x}(c)\mu_2(\mathrm{d}c) = 1$ for each $x\in\mathcal{X}$ and $b\in E_1$.

The crucial point in the previous definition lies in the fact that we have decoupled the spectral functions with respect to time and the spectral functions with respect to space given time. This allows one to deal with the temporal and the spatial aspects separately. Moreover, the latter depend on time through a bijective transformation which typically may account for an underlying physical process.

The spectral function with respect to space basically drives the shape of the main spatial patterns whereas the bijective transformation describes how these spatial patterns move in space. Thus, the transformation contributes to the temporal dynamics of the process. Finally, the spectral function with respect to time also contributes to the temporal dynamics of the process. This interpretation will become clearer with the illustrations of Section 3.1 (see Figure 1).

The finite dimensional distributions of X defined above are given, for $M\in\mathbb{N}\backslash\{0\}$, $t_1,\dots,t_M\in\mathcal{I}$, $x_1,\dots,x_M\in\mathcal{X}$ and $z_1,\dots,z_M>0$, by

$$\begin{aligned}&-\log\mathbb{P}(X(t_1,x_1)\leq z_1,\dots,X(t_M,x_M)\leq z_M)\\ &\qquad= \int_{E_1\times E_2}\bigvee_{m=1}^{M}\left\{\frac{V_{t_m}(b)V_{R_{(t_m,b)}x_m}(c)}{z_m}\right\}\mu_1(\mathrm{d}b)\mu_2(\mathrm{d}c). \quad (2)\end{aligned}$$

We now provide some examples of sub-classes of the general class of space-time max-stable processes given in Definition 1.

2.1.1. *Models of type 1: de Haan's representation with* $\mathcal{X}=\mathbb{R}^2$. We take $E_1=\mathcal{I}$ with $\mu_1=\delta$ and $E_2=\mathcal{X}=\mathbb{R}^2$ with $\mu_2=\lambda_2$, where λ_2 is the Lebesgue measure on $\mathbb{R}^2$. Let

g be a probability density function (case $\mathcal{I} = \mathbb{R}$) or a discrete probability distribution (case $\mathcal{I} = \mathbb{Z}$), and f a probability density function on $\mathbb{R}^2$. We then assume that

$$V_t(b) = g(t-b) \quad \text{and} \quad V_{\mathbf{x}}(\mathbf{c}) = f(\mathbf{x} - \mathbf{c}),$$

and that the operators $R_{(t,b)}$ are translations: for all $t, b \in \mathcal{I}$ and $\mathbf{x} \in \mathbb{R}^2$, $R_{(t,b)}\mathbf{x} = \mathbf{x} - (t-b)\boldsymbol{\tau}$, where $\boldsymbol{\tau} \in \mathbb{R}^2$.

The class of moving-maxima max-stable processes with general spectral representation (1) assumes the existence of a probability density function h on $\mathcal{I} \times \mathbb{R}^2$ such that $V_{(t,\mathbf{x})}(w) = h(t-b, \mathbf{x} - \mathbf{c})$. The density function h can always be decomposed as $h(t, \mathbf{x}) = g(t)h_1(\mathbf{x} \mid t)$, where $h_1(\mathbf{x} \mid t)$ is the conditional probability density function on $\mathbb{R}^2$ given t. For models of type 1, we have implicitly assumed that this density function satisfies the equality $h_1(\mathbf{x} \mid t) = f(\mathbf{x} - t\boldsymbol{\tau})$.

Note that the translation operator allows one to model physical processes such as propagation and diffusion.

2.1.2. *Models of type 2: de Haan's representation with $\mathcal{X}$ the unit sphere in* $\mathbb{R}^3$*.* We denote by $\mathbb{S}^2 = \{\mathbf{x} \in \mathbb{R}^3 : \|\mathbf{x}\| = 1\}$ the unit sphere in $\mathbb{R}^3$. We choose $E_1 = \mathcal{I}$ with $\mu_1 = \delta$, and $E_2 = \mathcal{X} = \mathbb{S}^2$ with $\mu_2 = \lambda_{\mathbb{S}^2}$, where $\lambda_{\mathbb{S}^2}$ is Lebesgue measure on $\mathbb{S}^2$. Let g be a probability density function (case $\mathcal{I} = \mathbb{R}$) or a discrete probability distribution (case $\mathcal{I} = \mathbb{Z}$) and let f be the von Mises-Fisher probability density function (see e.g. Mardia and Jupp, 1999, Section 9.3.2) on $\mathbb{S}^2$ with parameters $\boldsymbol{\mu} \in \mathbb{S}^2$ and $\kappa \geq 0$:

$$f(\mathbf{x}; \boldsymbol{\mu}, \kappa) = \frac{\kappa}{4\pi \sinh \kappa} \exp(\kappa \boldsymbol{\mu}' \mathbf{x}), \quad \mathbf{x} \in \mathbb{S}^2,$$

where $'$ denotes transposition. The parameters $\boldsymbol{\mu}$ and κ are called the mean direction and concentration parameter, respectively. The greater the value of κ, the higher the concentration of the distribution around the mean direction $\boldsymbol{\mu}$. The distribution is uniform on the sphere for $\kappa = 0$ and unimodal for $\kappa > 0$. We assume that

$$V_t(b) = g(t-b) \quad \text{and} \quad V_{\mathbf{x}}(\mathbf{c}) = f(\mathbf{x}; \mathbf{c}, \kappa)$$

and that, for $\mathbf{u} = (u_x, u_y, u_z)' \in \mathbb{S}^2$, $R_{(t,b)} = R_{\theta(t-b),\mathbf{u}}$, where $R_{\theta,\mathbf{u}}$ is the rotation matrix of angle θ around an axis in the direction of $\mathbf{u}$. We have that

$$R_{\theta,\mathbf{u}} = \cos\theta \, I_3 + \sin\theta \, [\mathbf{u}]_\times + (1 - \cos\theta)\mathbf{u}\mathbf{u}',$$

where I_3 is the identity matrix of $\mathbb{R}^3$ and $[\mathbf{u}]_\times$ the cross-product matrix of $\mathbf{u}$, defined by

$$[\mathbf{u}]_\times = \begin{pmatrix} 0 & -u_z & u_y \\ u_z & 0 & -u_x \\ -u_y & u_x & 0 \end{pmatrix}.$$

To the best of our knowledge, the resulting models are the first max-stable models on a sphere. Such models can of course be relevant in practice due to the natural spherical shape of planets and stars.

2.1.3. *Models of type 3: Schlather's representation with $\mathcal{X} = \mathbb{R}^2$.* For $d \in \mathbb{N}\backslash\{0\}$, let $\mathcal{C}_d = \mathcal{C}(\mathbb{R}^d, \mathbb{R}_+\backslash\{0\})$ be the space of continuous functions from $\mathbb{R}^d$ to $\mathbb{R}_+\backslash\{0\}$. For this sub-class of models, $(E_1, \mathcal{E}_1, \mu_1)$ and $(E_2, \mathcal{E}_2, \mu_2)$ are probability spaces with $E_1 = \mathcal{C}_1$,

$E_2 = \mathcal{C}_2$, and μ_1 and μ_2 are probability measures on E_1 and E_2 respectively. The function V_t (respectively $V_{\mathbf{x}}$) is defined as the natural projection from $\mathcal{C}_1$ (respectively $\mathcal{C}_2$) to $\mathbb{R}_+$ such that

$$V_t(b) = b(t) \quad \text{and} \quad V_{\mathbf{x}}(c) = c(\mathbf{x}).$$

Moreover, we assume that $\mathbb{E}[b(t)] = 1$ for all $t \in \mathcal{I}$ and $\mathbb{E}[c(\mathbf{x})] = 1$ for all $\mathbf{x} \in \mathbb{R}^2$. Note that for notational consistency, we use small letters for the stochastic processes b and c. The spectral process c is assumed to be either stationary and in this case $R_{(t,b)}\mathbf{x} = \mathbf{x} - t\boldsymbol{\tau}$ where $\boldsymbol{\tau} \in \mathbb{R}^2$, or to be isotropic and in this case $R_{(t,b)}\mathbf{x} = A^t\mathbf{x}$ where A is an orthogonal matrix ($R_{(t,b)}$ corresponds to a rotation).

2.1.4. *Models of type 4: Mixed representation with $\mathcal{X} = \mathbb{R}^2$.* We choose $E_1 = \mathcal{I}$, $\mu_1 = \delta$, $\mathcal{X} = \mathbb{R}^2$, $E_2 = \mathcal{C}_2$. Let g be a probability density function (case $\mathcal{I} = \mathbb{R}$) or a discrete probability distribution (case $\mathcal{I} = \mathbb{Z}$) and μ_2 a probability measure on $\mathcal{C}_2$. We take

$$V_t(b) = g(t-b) \quad \text{and} \quad V_{\mathbf{x}}(c) = c(\mathbf{x}).$$

Moreover, we assume that $\mathbb{E}[c(\mathbf{x})] = 1$ for all $\mathbf{x} \in \mathbb{R}^2$. As in the previous case, $V_{\mathbf{x}}$ is the natural projection from $\mathcal{C}_2$ to $\mathbb{R}_+$. Once again, note that we use a small letter for the stochastic process c. The spectral process c is assumed to be stationary and $R_{(t,b)}\mathbf{x} = \mathbf{x} - (t-b)\boldsymbol{\tau}$, where $\boldsymbol{\tau} \in \mathbb{R}^2$.

2.2. Spectral separability and marginal distributions

We are interested in conditions such that the marginal distributions of the process of Definition 1 are characterized by spectral representations involving only one spectral function (in time or in space).

We first consider the case of a fixed $t \in \mathcal{I}$. We define the process $(X_t(x))_{x\in\mathcal{X}} = (X(t,x))_{x\in\mathcal{X}}$. For two processes, $\stackrel{\mathrm{D}}{=}$ denotes equality in distribution for any finite dimensional vectors of the two processes.

Theorem 1. *Assume that for each $b \in E_1$, $M \in \mathbb{N}\backslash\{0\}$, $x_1, \dots, x_M \in \mathcal{X}$ and $z_1, \dots, z_M > 0$,*

$$\int_{E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_{R_{(t,b)}x_m}(c)}{z_m} \right\} \mu_2(\mathrm{d}c) = \int_{E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_{x_m}(c)}{z_m} \right\} \mu_2(\mathrm{d}c). \tag{3}$$

Then

$$X_t(x) \stackrel{\mathrm{D}}{=} \bigvee_{i=1}^{\infty} \{U_i V_x(C_i)\}, \quad x \in \mathcal{X},$$

where $(U_i, C_i)_{i\geq 1}$ are the points of a Poisson point process on $(0,\infty) \times E_2$ with intensity $u^{-2}\,\mathrm{d}u \times \mu_2(\mathrm{d}c)$. Moreover, Assumption (3) is satisfied for models of types 1, 2, 3 and 4.

Proof. For $M \in \mathbb{N}\backslash\{0\}$, let $t \in \mathbb{Z}$, $x_1, \dots, x_M \in \mathcal{X}$ and $z_1, \dots, z_M > 0$. We deduce by (2) and (3) that

$$\begin{aligned}
&-\log \mathbb{P}(X(t,x_1)\le z_1,\dots,X(t,x_M)\le z_M)\\
&= \int_{E_1\times E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_t(b)V_{R_{(t,b)}x_m}(c)}{z_m} \right\} \mu_1(\mathrm{d}b)\mu_2(\mathrm{d}c)\\
&= \int_{E_1} V_t(b) \int_{E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_{R_{(t,b)}x_m}(c)}{z_m} \right\} \mu_2(\mathrm{d}c)\mu_1(\mathrm{d}b)\\
&= \int_{E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_{x_m}(c)}{z_m} \right\} \mu_2(\mathrm{d}c) \int_{E_1} V_t(b)\mu_1(\mathrm{d}b)\\
&= \int_{E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_{x_m}(c)}{z_m} \right\} \mu_2(\mathrm{d}c).
\end{aligned}$$

We now show that Assumption (3) is satisfied for models of types 1, 2, 3 and 4. For models of type 1 we have $E_2=\mathbb{R}^2$,

$$V_{R_{(t,b)}\mathbf{x}_m}(\mathbf{c}) = f(R_{(t,b)}\mathbf{x}_m - \mathbf{c}) = f\big(\mathbf{x}_m - (\mathbf{c}+(t-b)\boldsymbol{\tau})\big) = V_{\mathbf{x}_m}(\mathbf{c}+(t-b)\boldsymbol{\tau})$$

and $\mu_2=\lambda_2$. Since λ_2 is invariant under translation, we derive by a change of variable that

$$\begin{aligned}
\int_{E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_{R_{(t,b)}\mathbf{x}_m}(\mathbf{c})}{z_m} \right\} \mu_2(\mathrm{d}\mathbf{c}) &= \int_{E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_{\mathbf{x}_m}(\mathbf{c}+(t-b)\boldsymbol{\tau})}{z_m} \right\} \mu_2(\mathrm{d}\mathbf{c})\\
&= \int_{E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_{\mathbf{x}_m}(\mathbf{c})}{z_m} \right\} \mu_2(\mathrm{d}\mathbf{c}).
\end{aligned}$$

For models of type 2 we have $E_2=\mathbb{S}^2$ and

$$\begin{aligned}
V_{R_{(t,b)}\mathbf{x}_m}(\mathbf{c}) &= f(R_{\theta(t-b),\mathbf{u}}\mathbf{x}_m;\mathbf{c},\kappa)\\
&= \frac{\kappa}{4\pi\sinh\kappa}\exp\big(\kappa(R_{-\theta(t-b),\mathbf{u}}\mathbf{c})'\mathbf{x}\big) = V_{\mathbf{x}_m}(R_{-\theta(t-b),\mathbf{u}}\mathbf{c}),
\end{aligned}$$

and it follows, since $\mu_2=\lambda_{\mathbb{S}^2}$ is invariant under rotation, that

$$\begin{aligned}
\int_{E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_{R_{(t,b)}\mathbf{x}_m}(\mathbf{c})}{z_m} \right\} \mu_2(\mathrm{d}\mathbf{c}) &= \int_{E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_{\mathbf{x}_m}(R_{-\theta(t-b),\mathbf{u}}\mathbf{c})}{z_m} \right\} \mu_2(\mathrm{d}\mathbf{c})\\
&= \int_{E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_{\mathbf{x}_m}(\mathbf{c})}{z_m} \right\} \mu_2(\mathrm{d}\mathbf{c}).
\end{aligned}$$

For models of type 3 we have $V_{\mathbf{x}}(c)=c(\mathbf{x})$. Thus if $R_{(t,b)}\mathbf{x}=\mathbf{x}-t\boldsymbol{\tau}$, we have $V_{R_{(t,b)}\mathbf{x}_m}(c)=c(\mathbf{x}_m-t\boldsymbol{\tau})$, and deduce by stationarity that

$$\int_{E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_{R_{(t,b)}\mathbf{x}_m}(c)}{z_m} \right\} \mu_2(\mathrm{d}c) = \int_{E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_{\mathbf{x}_m}(c)}{z_m} \right\} \mu_2(\mathrm{d}c).$$

If $R_{(t,b)}\mathbf{x} = A^t\mathbf{x}$, we have $V_{R_{(t,b)}\mathbf{x}_m}(c) = c(A^t\mathbf{x}_m)$, and obtain by isotropy that

$$\int_{E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_{R_{(t,b)}\mathbf{x}_m}(c)}{z_m} \right\} \mu_2(\mathrm{d}c) = \int_{E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_{\mathbf{x}_m}(c)}{z_m} \right\} \mu_2(\mathrm{d}c).$$

For models of type 4, $V_{\mathbf{x}}(c) = c(\mathbf{x})$. Thus, if $R_{(t,b)}\mathbf{x} = \mathbf{x} - (t-b)\boldsymbol{\tau}$, we have $V_{R_{(t,b)}\mathbf{x}_m}(c) = c(\mathbf{x}_m - (t-b)\boldsymbol{\tau})$. Hence we deduce by stationarity that

$$\int_{E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_{R_{(t,b)}\mathbf{x}_m}(c)}{z_m} \right\} \mu_2(\mathrm{d}c) = \int_{E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_{\mathbf{x}_m}(c)}{z_m} \right\} \mu_2(\mathrm{d}c). \qquad \square$$

Note that the distribution of $X_t(x)$ does not depend on t. Therefore the process $((X_t(x))_{x\in\mathcal{X}})_{t\in\mathcal{I}}$ is stationary in time, and the distribution in space can be referred to as the stationary spatial distribution.

We also see that the spectral separability and the use of specific operators R make the spectral function V_x (with its associated point process $(C_i)_{i\geq 1}$) the function which appears in the spatial spectral representation. The stationary spatial distribution depends only on the spatial parameters of the model. This property is interesting from a statistical point of view since any estimation procedure can be simplified by considering as a first step the spatial parameters only, without taking into account the temporal ones (see Section 4). Note that the idea of using a transformation of space in (3) can also be found in Strokorb et al. (2015), in a different context.

We now consider the case of a fixed site $x \in \mathcal{X}$. As previously, we define the process $(X_x(t))_{t\in\mathcal{I}} = (X(t,x))_{t\in\mathcal{I}}$.

Theorem 2. *Assume that there exist two operators S and G from $\mathcal{X}$ to $\mathcal{X}$ such that*

$$R_{(t,b)}S_{(t)}x = G_{(b)}x, \quad (t,b) \in \mathcal{I}\times E_1. \tag{4}$$

Then

$$X_{S_{(t)}x}(t) \overset{\mathrm{D}}{=} \bigvee_{i=1}^{\infty} \{U_i V_t(B_i)\}, \quad t \in \mathcal{I},$$

where $(U_i, B_i)_{i\geq 1}$ are the points of a Poisson point process on $(0,\infty)\times E_1$ with intensity $u^{-2}\,\mathrm{d}u \times \mu_1(\mathrm{d}b)$. Assumption (4) is satisfied for models of types 1 and 4 with $S_{(t)}\mathbf{x} = \mathbf{x} + t\boldsymbol{\tau}$, for models of type 2 with $S_{(t)}\mathbf{x} = R_{-\theta t,\mathbf{u}}\mathbf{x}$ and for models of type 3 with $S_{(t)}\mathbf{x} = \mathbf{x} + t\boldsymbol{\tau}$ or $S_{(t)}\mathbf{x} = A^{-t}\mathbf{x}$, where $\boldsymbol{\tau} \in \mathbb{R}^2$ and A is an orthogonal matrix.

Proof. For $M \in \mathbb{N}\backslash\{0\}$, $t_1,\dots,t_M \in \mathbb{Z}$, $x \in \mathcal{X}$ and $z_1,\dots,z_M > 0$,

$$\begin{aligned}
&-\log \mathbb{P}\big(X(t_1, S_{(t_1)}x) \leq z_1, \dots, X(t_M, S_{(t_M)}x) \leq z_M\big) \\
&\quad= \int_{E_1\times E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_{t_m}(b) V_{R_{(t_m,b)}S_{(t_m)}x}(c)}{z_m} \right\} \mu_1(\mathrm{d}b)\mu_2(\mathrm{d}c) \\
&\quad= \int_{E_1\times E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_{t_m}(b) V_{G_{(b)}x}(c)}{z_m} \right\} \mu_1(\mathrm{d}b)\mu_2(\mathrm{d}c) \\
&\quad= \int_{E_1} \bigvee_{m=1}^{M} \left\{ \frac{V_{t_m}(b)}{z_m} \right\} \int_{E_2} V_{G_{(b)}x}(c)\mu_2(\mathrm{d}c)\mu_1(\mathrm{d}b) = \int_{E_1} \bigvee_{m=1}^{M} \left\{ \frac{V_{t_m}(b)}{z_m} \right\} \mu_1(\mathrm{d}b).
\end{aligned}$$

Moreover, it is easy to show that Assumption (4) is satisfied for models of types 1, 2, 3 and 4 with the operators $S_{(t)}$ that are given. □

Contrary to Theorem 1, it is not possible to say that the marginal distributions when x is fixed are those given by the temporal spectral representation with the spectral function V_t and its associated point process $(B_i)_{i\geq 1}$. In order to obtain such a representation, it is necessary to apply a time transformation $S_{(t)}$ on x. As a consequence, it is difficult to estimate the temporal parameters separately since this transformation is not necessarily known in practice. The transformation indeed depends on the type of model and the parameters we want to estimate. Note that if $R_{t,b}$ does not depend on t (for instance under translation with $\boldsymbol{\tau} = \mathbf{0}$), i.e. if space and time are fully separated in the spectral representation, then $S_{(t)}$ is equal to the identity.

3. Markovian cases

In this section, if $\mathcal{I} = \mathbb{R}$, g is the density of a standard exponential random variable whereas if $\mathcal{I} = \mathbb{Z}$, g corresponds to the probability weights of a geometric random variable:

$$g(t) = \begin{cases} \nu \mathrm{e}^{-\nu t}\mathbb{I}_{\{t\geq 0\}} & \text{if } \mathcal{I} = \mathbb{R}, \\ (1-\phi)\phi^t \mathbb{I}_{\{t\geq 0\}} & \text{if } \mathcal{I} = \mathbb{Z}, \end{cases} \tag{5}$$

where $\nu > 0$ and $\phi \in (0,1)$. We first consider models of type 1 and type 4 and then models of type 2. The choice of the function g in (5) makes these spatio-temporal max-stable models time-Markovian.

3.1. Markovian models of type 1 and type 4

Recall that we assume the transformations $R_{(t,b)}$ to be translations: $R_{(t,b)}(\mathbf{x}) = \mathbf{x} - (t-b)\boldsymbol{\tau}$, where $\boldsymbol{\tau} \in \mathbb{R}^2$. In this context, we obtain

$$X(t,\mathbf{x}) = \begin{cases} \bigvee_{i\geq 1} \left\{ U_i \nu \mathrm{e}^{-\nu(t-B_i)} \mathbb{I}_{\{t-B_i\geq 0\}} V_{\mathbf{x}-(t-B_i)\boldsymbol{\tau}}(C_i) \right\} & \text{if } \mathcal{I} = \mathbb{R}, \\ \bigvee_{i\geq 1} \left\{ U_i \phi(1-\phi)^{t-B_i} \mathbb{I}_{\{t-B_i\geq 0\}} V_{\mathbf{x}-(t-B_i)\boldsymbol{\tau}}(C_i) \right\} & \text{if } \mathcal{I} = \mathbb{Z}. \end{cases} \tag{6}$$

Note that for $\mathcal{I} = \mathbb{R}$ the function g has been introduced by Dombry and Eyi-Minko (2014), Section 4, in the form $g(t) = -\log(a)a^t\mathbb{I}_{\{t\geq 0\}}$ for $a \in (0,1)$, in order to build the continuous-time version of the real-valued max-autoregressive process. Let us denote by a the constant $\mathrm{e}^{-\nu}$ if $\mathcal{I} = \mathbb{R}$ and the constant ϕ if $\mathcal{I} = \mathbb{Z}$.

Theorem 3. (i) *For all $t, s \in \mathcal{I}$ such that $s > 0$,*

$$X(t,\mathbf{x}) = \max(a^s X(t-s, \mathbf{x} - s\boldsymbol{\tau}), (1-a^s)Z(t,\mathbf{x})), \tag{7}$$

where the process $(Z(t,\mathbf{x}))_{\mathbf{x}\in\mathbb{R}^2}$ is independent of $(X(t-s,\mathbf{x}))_{\mathbf{x}\in\mathbb{R}^2}$ and

$$Z(t,\mathbf{x}) \overset{\mathrm{D}}{=} \bigvee_{i=1}^{\infty} \{U_i V_{\mathbf{x}}(C_i)\}, \quad (t,\mathbf{x}) \in \mathcal{I} \times \mathbb{R}^2, \tag{8}$$

with $(U_i, C_i)_{i\geq 1}$ the points of a Poisson point process on $(0,\infty) \times E_2$ of intensity $u^{-2}\,\mathrm{d}u \times \mu_2(\mathrm{d}c)$. Therefore the process (6) is time-Markovian.

(ii) *Let* $\mathcal{I} = \mathbb{Z}$ *and* $((Z(t,\mathbf{x}))_{\mathbf{x}\in\mathbb{R}^2})_{t\in\mathcal{I}}$ *be a family of iid max-stable processes with spectral representation* (8). *Then*

$$X(t,\mathbf{x}) \overset{\mathrm{D}}{=} \bigvee_{j=0}^{\infty} \left\{a^j(1-a)Z(t-j,\mathbf{x}-j\boldsymbol{\tau})\right\}, \quad (t,\mathbf{x}) \in \mathcal{I}\times\mathbb{R}^2. \tag{9}$$

Proof. (i) Consider the case $\mathcal{I} = \mathbb{R}$ (the case $\mathcal{I} = \mathbb{Z}$ is similar). We have that

$$\begin{aligned} X(t,\mathbf{x}) &= \bigvee_{i=1}^{\infty} \left\{U_i\nu \mathrm{e}^{-\nu(t-B_i)}\mathbb{I}_{\{t-B_i\geq 0\}}V_{\mathbf{x}-(t-B_i)\boldsymbol{\tau}}(C_i)\right\} \\ &= \bigvee_{i=1}^{\infty} \left\{U_i\nu \mathrm{e}^{-\nu(s+t-s-B_i)}\mathbb{I}_{\{s+t-s-B_i\geq 0\}}V_{\mathbf{x}-(s+t-s-B_i)\boldsymbol{\tau}}(C_i)\right\} \\ &= \max\left(\mathrm{e}^{-\nu s}\bigvee_{i=1}^{\infty} \left\{U_i\nu \mathrm{e}^{-\nu(t-s-B_i)}\mathbb{I}_{\{t-s-B_i\geq 0\}}V_{\mathbf{x}-s\boldsymbol{\tau}-(t-s-B_i)\boldsymbol{\tau}}(C_i)\right\},\right. \\ &\qquad \left.\bigvee_{i=1}^{\infty} \left\{U_i\nu \mathrm{e}^{-\nu(t-B_i)}\mathbb{I}_{\{t\geq B_i>t-s\}}V_{\mathbf{x}-(t-B_i)\boldsymbol{\tau}}(C_i)\right\}\right) \\ &= \max\left(\mathrm{e}^{-\nu s}X(t-s,\mathbf{x}-s\boldsymbol{\tau}),(1-\mathrm{e}^{-\nu s})Z(t,\mathbf{x})\right), \end{aligned}$$

where

$$Z(t,\mathbf{x}) = \frac{1}{1-\mathrm{e}^{-\nu s}}\bigvee_{i\geq 1}\left\{U_i\nu \mathrm{e}^{-\nu(t-B_i)}\mathbb{I}_{\{t\geq B_i>t-s\}}V_{\mathbf{x}-(t-B_i)\boldsymbol{\tau}}(C_i)\right\}.$$

Since the sets $\{t \geq B > t-s\}$ and $\{t-s \geq B\}$ are disjoint, the Poisson point processes $\{(U_i,B_i,C_i), i : t \geq B_i > t-s\}$ and $\{(U_i,B_i,C_i), i : t-s \geq B_i\}$ are independent and it follows that $(X(t-s,\mathbf{x}))_{\mathbf{x}\in\mathbb{R}^2}$ and $(Z(t,\mathbf{x}))_{\mathbf{x}\in\mathbb{R}^2}$ are also independent.

We now show that $Z(t,\mathbf{x}) \overset{\mathrm{D}}{=} \bigvee_{i=1}^{\infty}\{U_iV_{\mathbf{x}}(C_i)\}$ for all $\mathbf{x}\in\mathbb{R}^2$, where $(U_i,C_i)_{i\geq 1}$ are the points of a Poisson point process on $(0,\infty)\times E_2$ of intensity $u^{-2}\,\mathrm{d}u\times\mu_2(\mathrm{d}c)$. Let $(U_i,B_i,C_i)_{i\geq 1}$ be the points of a Poisson point process on $(0,\infty)\times\mathbb{R}\times E_2$ with intensity $u^{-2}\,\mathrm{d}u\times\mathrm{d}b\times\mu_2(\mathrm{d}c)$. For $M\in\mathbb{N}\backslash\{0\}$, let $\mathbf{x}_1,\ldots,\mathbf{x}_M\in\mathbb{R}^2$ and $z_1,\ldots,z_M>0$. We consider the set

$$\begin{aligned} B_{z_1,\ldots,z_M} &= \{(u,b,c) : \\ &\quad u\nu \mathrm{e}^{-\nu(t-b)}\mathbb{I}_{\{t\geq b>t-s\}}V_{\mathbf{x}_m-(t-b)\boldsymbol{\tau}}(c) > z_m \text{ for at least one } m=1,\ldots,M\}. \end{aligned}$$

Denoting by $\bigwedge$ the min-operator, the Poisson measure of $B_{z_1,\ldots,z_M}$ is

$$\begin{aligned} &\Lambda(B_{z_1,\ldots,z_M}) \\ &= \int_{E_2}\int_{\mathbb{R}}\int_0^{\infty}\mathbb{I}\left\{u > \bigwedge_{m=1}^{M}\left\{\frac{z_m}{\nu \mathrm{e}^{-\nu(t-b)}\mathbb{I}_{\{t\geq b>t-s\}}V_{\mathbf{x}_m-(t-b)\boldsymbol{\tau}}(c)}\right\}\right\}\frac{\mathrm{d}u}{u^2}\,\mathrm{d}b\,\mu_2(\mathrm{d}c) \\ &= \int_{E_2}\int_{\mathbb{R}}\bigvee_{m=1}^{M}\left\{\frac{\nu \mathrm{e}^{-\nu(t-b)}\mathbb{I}_{\{t\geq b>t-s\}}V_{\mathbf{x}_m-(t-b)\boldsymbol{\tau}}(c)}{z_m}\right\}\mathrm{d}b\,\mu_2(\mathrm{d}c) \end{aligned}$$

$$= \nu \mathrm{e}^{-\nu t} \int_{\mathbb{R}} \mathbb{I}_{\{t \geq b > t-s\}} \mathrm{e}^{\nu b} \int_{E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_{\mathbf{x}_m - (t-b)\boldsymbol{\tau}}(c)}{z_m} \right\} \mu_2(\mathrm{d}c)\, \mathrm{d}b.$$

Since

$$\int_{E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_{\mathbf{x}_m - (t-b)\boldsymbol{\tau}}(c)}{z_m} \right\} \mu_2(\mathrm{d}c) = \int_{E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_{\mathbf{x}_m}(c)}{z_m} \right\} \mu_2(\mathrm{d}c),$$

we deduce that

$$\begin{aligned} \Lambda(B_{z_1,\dots,z_M}) &= \nu \mathrm{e}^{-\nu t} \int_{E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_{\mathbf{x}_m}(c)}{z_m} \right\} \mu_2(\mathrm{d}c) \int_{\mathbb{R}} \mathbb{I}_{\{t \geq b > t-s\}} \mathrm{e}^{\nu b}\, \mathrm{d}b \\ &= (1 - \mathrm{e}^{-\nu s}) \int_{E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_{\mathbf{x}_m}(c)}{z_m} \right\} \mu_2(\mathrm{d}c). \end{aligned}$$

It follows that

$$\begin{aligned} &-\log \mathbb{P}(Z(t,\mathbf{x}_1) \leq z_1, \dots, Z(t,\mathbf{x}_M) \leq z_M) \\ &\qquad = \Lambda(B_{(1-\mathrm{e}^{-\nu s})z_1,\dots,(1-\mathrm{e}^{-\nu s})z_M}) = \int_{E_2} \bigvee_{m=1}^{M} \left\{ \frac{V_{\mathbf{x}_m}(c)}{z_m} \right\} \mu_2(\mathrm{d}c). \end{aligned}$$

(ii) It is easily shown that the right-hand side of (9) is a solution of (7). Moreover, as in Davis and Resnick (1989), this solution is unique, yielding (9). □

Given Theorem 3, it is natural to consider the spatio-temporal max-stable process satisfying the following stochastic recurrence equation:

$$X(t,\mathbf{x}) = \max\bigl(aX(t-1,\mathbf{x}-\boldsymbol{\tau}), (1-a)Z(t,\mathbf{x})\bigr), \quad (t,\mathbf{x}) \in \mathcal{J} \times \mathbb{R}^2, \tag{10}$$

where $\mathcal{J} = \mathbb{Z}$ or $\mathbb{N}$, $a \in (0,1)$, $\boldsymbol{\tau} \in \mathbb{R}^2$ are both fixed and $\bigl((Z(t,\mathbf{x}))_{\mathbf{x}\in\mathbb{R}^2}\bigr)_{t\in\mathcal{J}}$ is a sequence of iid spatial max-stable processes with spectral representation (8). In the case $\mathcal{J} = \mathbb{Z}$, the process X defined by $X(t,\mathbf{x}) = \bigvee_{j=0}^{\infty} \bigl\{a^j(1-a)Z(t-j,\mathbf{x}-j\boldsymbol{\tau})\bigr\}$, for $(t,\mathbf{x}) \in \mathcal{I} \times \mathbb{R}^2$, clearly satisfies (10) and is stationary in time. In the case $\mathcal{J} = \mathbb{N}$, we must initialize the recurrence with a process $(X_0(\mathbf{x}))_{\mathbf{x}\in\mathbb{R}^2}$ having the spatial distribution given by (8). This makes the process defined by (10) stationary in time.

Remark. A real-valued process $(R(t))_{t\in\mathbb{Z}}$ follows the max-autoregressive moving-average process of orders p and q (MARMA(p,q)), introduced by Davis and Resnick (1989), if it satisfies the recursion

$$R(t) = \max\bigl(\phi_1 R(t-1), \dots, \phi_p R(t-p), S(t), \theta_1 S(t-1), \dots, \theta_q S(t-q)\bigr), \quad t \in \mathbb{Z},$$

where $\phi_i, \theta_j \geq 0$ for $i = 1,\dots,p$ and $j = 1,\dots,q$, and the max-stable random variables $S(t)$ for $t \in \mathbb{Z}$ are iid. It is a time series model which is max-stable in time. However, the spatial aspect is absent. From (10), it can be seen that our model X extends the real-valued MARMA$(1,0)$ process to the spatial setting.

The parameter a measures the influence of the past, whereas the parameter $\boldsymbol{\tau}$ represents some kind of specific direction of propagation (contagion) in space. The value at location $\mathbf{x}$ and time t is either related to the value at location $\mathbf{x}-\boldsymbol{\tau}$ at time $t-1$ or to the value of another process (the innovation), Z, that characterizes a new event happening at location $\mathbf{x}$. If a and the value at location $\mathbf{x}-\boldsymbol{\tau}$ are large, it is likely that there will be a propagation from location $\mathbf{x}-\boldsymbol{\tau}$ to location $\mathbf{x}$, i.e. contagion of the extremes, with an attenuation effect. Contrary to the existing spatio-temporal max-stable models, the dynamics are described by an equation that can be physically interpreted.

Moreover, the combination of Theorems 1 and 3 shows that the stationary spatial distribution of the Markov process/chain $((X(t,\mathbf{x}))_{\mathbf{x}\in\mathbb{R}^2})_{t\in\mathcal{J}}$ is the same as that of Z.

The space-time exponent measure of the process (10) is given in the next proposition.

Proposition 1. *For* $M\in\mathbb{N}\backslash\{0\}, t_1\leq\cdots\leq t_M\in\mathbb{Z}$, $\mathbf{x}_1,\ldots,\mathbf{x}_M\in\mathbb{R}^2$ *and* $z_1,\ldots,z_M>0$, *we have that*

$$\begin{aligned}
&-\log\mathbb{P}(X(t_1,\mathbf{x}_1)\leq z_1,\ldots,X(t_M,\mathbf{x}_M)\leq z_M)\\
&=\mathcal{V}_{\mathbf{x}_1,\mathbf{x}_2-(t_2-t_1)\boldsymbol{\tau},\ldots,\mathbf{x}_M-(t_M-t_1)\boldsymbol{\tau}}\left(z_1,\frac{z_2}{a^{t_2-t_1}},\ldots,\frac{z_M}{a^{t_M-t_1}}\right)\\
&+\sum_{m=2}^{M-1}(1-a^{t_m-t_{m-1}})\mathcal{V}_{\mathbf{x}_m,\mathbf{x}_{m+1}-(t_{m+1}-t_m)\boldsymbol{\tau},\ldots,\mathbf{x}_M-(t_M-t_m)\boldsymbol{\tau}}\left(z_m,\frac{z_{m+1}}{a^{t_{m+1}-t_m}},\ldots,\frac{z_M}{a^{t_M-t_m}}\right)\\
&+\frac{1-a^{t_M-t_{M-1}}}{z_M},
\end{aligned}\tag{11}$$

where $\mathcal{V}$ *is the exponent function characterizing the spatial distribution, defined by*

$$\mathcal{V}_{\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_M}(z_1,z_2,\ldots,z_M)=\int_{\mathbb{R}^2}\bigvee_{m=1}^{M}\left\{\frac{V_{\mathbf{x}_m}(c)}{z_m}\right\}\mu_2(\mathrm{d}c).$$

Proof. For the sake of notational simplicity, we give the proof only in the case $M=3$; this proof can easily be extended. Using the independence of the replications $(Z(t,\mathbf{x}))_{\mathbf{x}\in\mathbb{R}^2}$, and changes of indices, we obtain

$$\begin{aligned}
&\mathbb{P}\left(\bigvee_{j=0}^{J}\left\{a^j(1-a)Z(t_i-j,\mathbf{x}_i-j\boldsymbol{\tau})\right\}\leq z_i\text{ for }i=1,2,3\right)\\
&=\mathbb{P}\Bigg(\bigvee_{j=0}^{J}\left\{a^j(1-a)Z(t_1-j,\mathbf{x}_1-j\boldsymbol{\tau})\right\}\leq z_1,\\
&\qquad\bigvee_{j=t_1-t_2}^{J+t_1-t_2}\left\{a^{j+t_2-t_1}(1-a)Z(t_1-j,\mathbf{x}_2-(j+t_2-t_1)\boldsymbol{\tau}\right\}\leq z_2,\\
&\qquad\bigvee_{j=t_1-t_3}^{J+t_1-t_3}\left\{a^{j+t_3-t_1}(1-a)Z(t_1-j,\mathbf{x}_3-(j+t_3-t_1)\boldsymbol{\tau}\right\}\leq z_3\Bigg)\\
&=\mathbb{P}\Bigg(\bigvee_{j=0}^{J+t_1-t_3}\left\{a^j(1-a)Z(t_1-j,\mathbf{x}_1-j\boldsymbol{\tau})\right\}\leq z_1,
\end{aligned}$$

$$
\begin{aligned}
&\bigvee_{j=0}^{J+t_1-t_3} \left\{a^{j+t_2-t_1}(1-a)Z(t_1-j,\mathbf{x}_2-(j+t_2-t_1)\boldsymbol{\tau}\right\} \leq z_2,\\
&\bigvee_{j=0}^{J+t_1-t_3} \left\{a^{j+t_3-t_1}(1-a)Z(t_1-j,\mathbf{x}_3-(j+t_3-t_1)\boldsymbol{\tau}\right\} \leq z_3\Bigg)\\
&\times \mathbb{P}\Bigg(\bigvee_{j=t_1-t_2}^{-1} \left\{a^{j+t_2-t_1}(1-a)Z(t_1-j,\mathbf{x}_2-(j+t_2-t_1)\boldsymbol{\tau})\right\} \leq z_2,\\
&\bigvee_{j=t_1-t_2}^{-1} \left\{a^{j+t_3-t_1}(1-a)Z(t_1-j,\mathbf{x}_3-(j+t_3-t_1)\boldsymbol{\tau})\right\} \leq z_3\Bigg)\\
&\times \mathbb{P}\left(\bigvee_{j=t_1-t_3}^{t_1-t_2-1} \left\{a^{j+t_3-t_1}(1-a)Z(t_1-j,\mathbf{x}_3-(j+t_3-t_1)\boldsymbol{\tau})\right\} \leq z_3\right)\\
&\times \mathbb{P}\left(\bigvee_{J+t_1-t_2+1}^{J} \left\{a^{j}(1-a)Z(t_1-j,\mathbf{x}_1-j\boldsymbol{\tau})\right\} \leq z_1\right)\\
&\times \mathbb{P}\Bigg(\bigvee_{j=J+t_1-t_3+1}^{J+t_1-t_2} \left\{a^{j}(1-a)Z(t_1-j,\mathbf{x}_1-j\boldsymbol{\tau})\right\} \leq z_1),\\
&\bigvee_{j=J+t_1-t_3+1}^{J+t_1-t_2} \left\{a^{j+t_2-t_1}(1-a)Z(t_1-j,\mathbf{x}_2-(j+t_2-t_1)\boldsymbol{\tau})\right\} \leq z_2\Bigg). \qquad (12)
\end{aligned}
$$

Using the independence of the replications $(Z(t,\mathbf{x}))_{\mathbf{x}\in\mathbb{R}^2}$, the stationarity of the processes $(Z(t,\mathbf{x}))_{\mathbf{x}\in\mathbb{R}^2}$ and the homogeneity of order -1 of $\mathcal{V}$, we obtain

$$
\begin{aligned}
&\mathbb{P}\Bigg(\bigvee_{j=0}^{J+t_1-t_3} \left\{a^{j}(1-a)Z(t_1-j,\mathbf{x}_1-j\boldsymbol{\tau})\right\} \leq z_1,\\
&\qquad \bigvee_{j=0}^{J+t_1-t_3} \left\{a^{j+t_2-t_1}(1-a)Z(t_1-j,\mathbf{x}_2-(j+t_2-t_1)\boldsymbol{\tau})\right\} \leq z_2,\\
&\qquad \bigvee_{j=0}^{J+t_1-t_3} \left\{a^{j+t_3-t_1}(1-a)Z(t_1-j,\mathbf{x}_3-(j+t_3-t_1)\boldsymbol{\tau})\right\} \leq z_3\Bigg)\\
&= \prod_{j=0}^{J+t_1-t_3} \mathbb{P}\Bigg(Z(t_1-j,\mathbf{x}_1-j\boldsymbol{\tau}) \leq \frac{z_1}{a^j(1-a)},\\
&\qquad Z(t_1-j,\mathbf{x}_2-(j+t_2-t_1)\boldsymbol{\tau}) \leq \frac{z_2}{a^{j+t_2-t_1}(1-a)},\\
&\qquad Z(t_1-j,\mathbf{x}_3-(j+t_3-t_1)\boldsymbol{\tau} \leq \frac{z_3}{a^{j+t_3-t_1}(1-a)}\Bigg)\\
&= \prod_{j=0}^{J+t_1-t_3} \exp\Bigg(-\mathcal{V}_{\mathbf{x}_1-j\boldsymbol{\tau},\mathbf{x}_2-(t_2-t_1)\boldsymbol{\tau}-j\boldsymbol{\tau},\mathbf{x}_3-(t_3-t_1)\boldsymbol{\tau}-j\boldsymbol{\tau}}
\end{aligned}
$$

$$\left(\frac{z_1}{a^j(1-a)}, \frac{z_2}{a^j a^{t_2-t_1}(1-a)}, \frac{z_3}{a^j a^{t_3-t_1}(1-a)}\right)\Bigg)$$

$$= \exp\left(-\mathcal{V}_{\mathbf{x}_1,\mathbf{x}_2-(t_2-t_1)\boldsymbol{\tau},\mathbf{x}_3-(t_3-t_1)\boldsymbol{\tau}}\left(z_1, \frac{z_2}{a^{t_2-t_1}}, \frac{z_3}{a^{t_3-t_1}}\right) \times (1-a) \sum_{j=0}^{J+t_1-t_3} a^j\right)$$

$$= \exp\left(-\mathcal{V}_{\mathbf{x}_1,\mathbf{x}_2-(t_2-t_1)\boldsymbol{\tau},\mathbf{x}_3-(t_3-t_1)\boldsymbol{\tau}}\left(z_1, \frac{z_2}{a^{t_2-t_1}}, \frac{z_3}{a^{t_3-t_1}}\right) \times \left(1 - a^{J+t_1-t_3+1}\right)\right).$$

The other terms in (12) are calculated similarly. Finally, since $\lim_{J\to\infty} a^J = 0$,

$$\mathbb{P}(X(t_1,\mathbf{x}_1) \le z_1, X(t_2,\mathbf{x}_2) \le z_2, X(t_3,\mathbf{x}_3) \le z_3)$$
$$= \lim_{J\to\infty} \mathbb{P}\left(\bigvee_{j=0}^{J} \left\{a^j(1-a)Z(t_i-j, \mathbf{x}_i - j\boldsymbol{\tau})\right\} \le z_i \text{ for } i=1,2,3\right)$$
$$= \exp\left(-\mathcal{V}_{\mathbf{x}_1,\mathbf{x}_2-(t_2-t_1)\boldsymbol{\tau},\mathbf{x}_3-(t_3-t_1)\boldsymbol{\tau}}\left(z_1, \frac{z_2}{a^{t_2-t_1}}, \frac{z_3}{a^{t_3-t_1}}\right)\right)$$
$$\times \exp\left(-\mathcal{V}_{\mathbf{x}_2,\mathbf{x}_3-(t_3-t_2)\boldsymbol{\tau}}\left(\frac{z_2}{a^{t_2-t_1}}, \frac{z_3}{a^{t_3-t_1}}\right) \times (1-a^{t_2-t_1})\right) \exp\left(-\frac{1-a^{t_3-t_2}}{z_3}\right).$$

□

By using the approach developed by Bienvenüe and Robert (2014), the right-hand term of (11) can easily be computed provided that the distribution of $(V_{\mathbf{x}_m}(c))_{m=1,\dots,M}$ with $c \sim \mu_2$ is absolutely continuous with respect to Lebesgue measure. This is the case for example for the spatial Schlather and Brown-Resnick processes.

It is easily shown that the models of types 1 and 4 are stationary in time (see Theorem 1) and space (in the case of models of type 4, the process c is stationary). In order to measure the spatio-temporal dependence, we propose extensions to the spatio-temporal setting of quantities that have been introduced in the spatial context. The first is the spatio-temporal extremal coefficient function, stemming from the spatial version by Schlather and Tawn (2003), which is defined for all $t_1, t_2 \in \mathbb{Z}$ and $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^2$ by

$$\mathbb{P}(X(t_1,\mathbf{x}_1) \le u, X(t_2,\mathbf{x}_2) \le u) = \exp\left(-\frac{\theta(t_2-t_1, \mathbf{x}_2-\mathbf{x}_1)}{u}\right), \quad u > 0.$$

The second is the spatio-temporal Φ_1-madogram, coming from the spatial version introduced by Cooley et al. (2006), where Φ_1 is the standard Fréchet probability distribution function. It is defined by

$$\nu_{\Phi_1}(t_2-t_1, \mathbf{x}_2-\mathbf{x}_1) = \frac{1}{2}\mathbb{E}\left[\left|\Phi_1(X(t_2,\mathbf{x}_2)) - \Phi_1(X(t_1,\mathbf{x}_1))\right|\right].$$

Proposition 2. *In the case of* (10)*, for $l \in \mathbb{Z}$ and $\mathbf{h} \in \mathbb{R}^2$, the spatio-temporal extremal coefficient is given by*

$$\theta(l,\mathbf{h}) = \mathcal{V}_{\mathbf{0},\mathbf{h}-l\boldsymbol{\tau}}\left(1, a^{-l}\right) + 1 - a^l \tag{13}$$

and the spatio-temporal Φ_1-madogram of X by

$$\nu_{\Phi_1}(l,\mathbf{h}) = \frac{1}{2}\frac{\theta(l,\mathbf{h})-1}{\theta(l,\mathbf{h})+1} = \frac{1}{2} - \frac{1}{\mathcal{V}_{\mathbf{0},\mathbf{h}-l\boldsymbol{\tau}}\left(1,a^{-l}\right) + 2 - a^l}. \tag{14}$$

Proof. Applying (11) with $M = 2$ and setting $z_1 = z_2 = u$ for $u > 0$, we obtain

$$\mathbb{P}(X(0,\mathbf{0})) \leq u, X(l,\mathbf{h}) \leq u) = \exp\left(-\mathcal{V}_{\mathbf{0},\mathbf{h}-l\boldsymbol{\tau}}\left(u, \frac{u}{a^l}\right)\right)\exp\left(-\frac{1-a^l}{u}\right)$$
$$= \exp\left(-\frac{\mathcal{V}_{\mathbf{0},\mathbf{h}-l\boldsymbol{\tau}}\left(1, a^{-l}\right) + 1 - a^l}{u}\right),$$

yielding (13) by definition of the spatio-temporal extremal coefficient.

In the same way as in the purely spatial case (see e.g. Cooley et al., 2006, p.379), it is easy to show the following link between the spatio-temporal Φ_1-madogram and the spatio-temporal extremal coefficient:

$$\nu_F(l,\mathbf{h}) = \frac{1}{2}\frac{\theta(l,\mathbf{h}) - 1}{\theta(l,\mathbf{h}) + 1} = \frac{1}{2} - \frac{1}{\theta(l,\mathbf{h}) + 1}. \tag{15}$$

Inserting (13) into (15) gives (14). □

Similarly, it would also be possible to extend the λ-madogram, introduced by Naveau et al. (2009), to the spatio-temporal setting.

Proposition 2 shows that we do not fully separate space and time in the extremal dependence measure given by the extremal coefficient, even if $\boldsymbol{\tau} = 0$. On the other hand, in the latter case, space and time are entirely separated in the spectral representation: V_t depends only on time and $V_{R(t,b)\mathbf{x}} = V_{\mathbf{x}}$ depends only on space.

Furthermore, we have $\lim_{l\to\infty} \theta(l,\mathbf{h}) = 2$, showing asymptotic time-independence. Moreover, from Theorem 3.1 in Kabluchko and Schlather (2010), we deduce that, for a fixed $\mathbf{x}$, the process $(X_{\mathbf{x}}(t))_{t\in\mathcal{I}}$ is strongly mixing in time. Finally, $\lim_{\|\mathbf{h}\|\to\infty} \theta(0,\mathbf{h}) = 2$ if and only if X is strongly mixing in space.

Before showing some simulations, let us define the spatial Smith and Schlather models. Let $(U_i, \mathbf{C}_i)_{i\geq 1}$ be the points of a Poisson point process on $(0,\infty)\times\mathbb{R}^2$ with intensity $u^{-2}\,\mathrm{d}u \times \lambda_2(\mathrm{d}\mathbf{c})$ and let h_Σ denote the bivariate Gaussian density with mean $\mathbf{0}$ and covariance matrix Σ. Then the spatial Smith model (Smith, 1990) is defined as $Z(\mathbf{x}) = \bigvee_{i=1}^{\infty}\{U_i h_\Sigma(\mathbf{x} - \mathbf{C}_i)\}$, for $\mathbf{x} \in \mathbb{R}^2$. Let $(U_i)_{i\geq 1}$ be the points of a Poisson point process on $(0,\infty)$ with intensity $u^{-2}\,\mathrm{d}u$ and $Y_1, Y_2, \ldots$ independent replications of the stochastic process $Y(\mathbf{x}) = \sqrt{2\pi}\varepsilon(\mathbf{x})$, for $\mathbf{x} \in \mathbb{R}^2$, where ε is a stationary standard Gaussian process with correlation function $\rho(\cdot)$. Then the spatial Schlather process (Schlather, 2002) is defined as $Z(\mathbf{x}) = \bigvee_{i=1}^{\infty}\{U_i Y_i(\mathbf{x})\}$, for $\mathbf{x} \in \mathbb{R}^2$.

In the left panel of Figure 1, we show the evolution of the process (10) when $((Z(t,\mathbf{x}))_{\mathbf{x}\in\mathbb{R}^2})_{t\in\mathbb{N}}$ is a sequence of iid spatial Smith processes with covariance matrix

$$\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$a = 0.7$ and $\boldsymbol{\tau} = (-1,-1)'$ (translation to the bottom left). In the right panel of Figure 1, we show the evolution of the process (10) when $((Z(t,\mathbf{x}))_{\mathbf{x}\in\mathbb{R}^2})_{t\in\mathbb{N}}$ is a sequence of iid spatial Schlather processes with correlation function of type 'powered exponential', defined for all $h \geq 0$ by $\rho(h) = \exp\left[-(h/c_1)^{c_2}\right]$ for $c_1 > 0$ and $0 < c_2 < 2$, where c_1 and c_2 are the range and the smoothing parameters, respectively. We take $c_1 = 3$, $c_2 = 1$ and, as previously, $a = 0.7$ and $\boldsymbol{\tau} = (-1,-1)'$. The simulations have been carried out using the function rmaxstab of the R package SpatialExtremes (Ribatet, 2015). The interpretations drawn below are independent of the values of parameters that are chosen. Note that the processes represented correspond to models of types 1 and 4, respectively. They are respectively spatial Smith and Schlather processes which evolve dynamically in time.

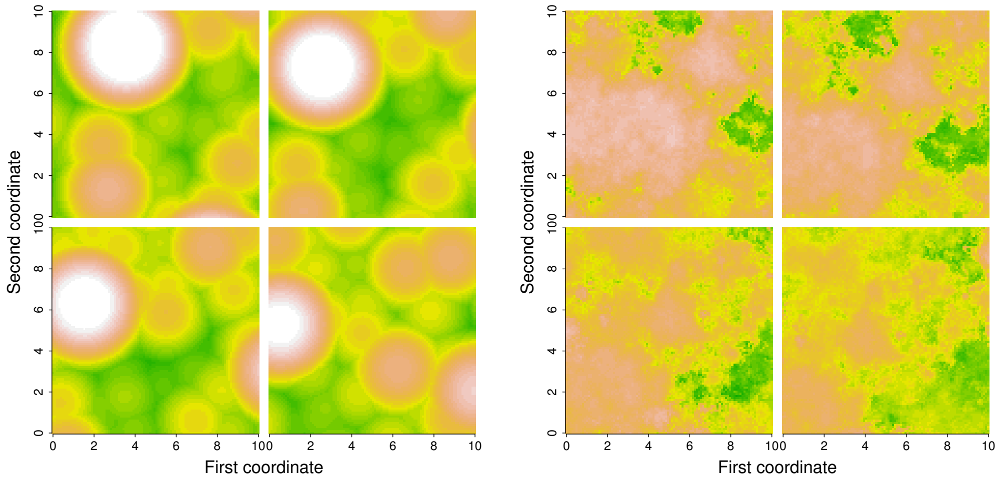

FIGURE 1: Simulation of the process (10) when $((Z(t,\mathbf{x}))_{\mathbf{x}\in\mathbb{R}^2})_{t\in\mathbb{N}}$ is a sequence of iid spatial Smith processes (left panel) and spatial Schlather processes (right panel). We depict the logarithm of the value of the process in space. In both cases, the evolution over four periods is represented (top left: $t=1$, top right: $t=2$, bottom left: $t=3$, bottom right: $t=4$).

In both cases, we observe a translation of the main spatial structures (the "storms" in the case of the spatial Smith model) to the bottom left, hence highlighting the usefulness of models like (10) for phenomena that propagate in space.

3.2. Markovian models of type 2

As in the previous section, we deduce that the models of type 2 satisfy the following stochastic recurrence equation:

$$X(t,\mathbf{x}) = \max\big(aX(t-1, R_{\theta,\mathbf{u}}\mathbf{x}), (1-a)Z(t,\mathbf{x})\big), \quad (t,\mathbf{x}) \in \mathcal{J}\times\mathbb{S}^2,$$

where $((Z(t,\mathbf{x}))_{\mathbf{x}\in\mathbb{S}^2})_{t\in\mathcal{J}}$ is a sequence of iid spatial max-stable processes with spectral representation

$$Z(t,\mathbf{x}) \overset{\mathrm{D}}{=} \bigvee_{i=1}^{\infty}\left\{U_i f(\mathbf{x};\boldsymbol{\mu}_i,\kappa)\right\}, \quad (t,\mathbf{x}) \in \mathcal{J}\times\mathbb{S}^2,$$

with $(U_i,\boldsymbol{\mu}_i)_{i\geq 1}$ the points of a Poisson point process on $(0,\infty)\times\mathbb{S}^2$ with intensity $u^{-2}\,\mathrm{d}u \times \mathrm{d}\lambda_{\mathbb{S}^2}$.

4. Estimation on simulated data

In this section, we briefly discuss statistical inference for the process (10). Although far from being exhaustive, this study illustrates the fact that this model can be used at an operational level. Of course, the compatibility of this model with real observations still has to be shown. We denote by $\boldsymbol{\theta}$ the vector gathering the parameters to be estimated. One possible method of estimation consists in using the pairwise likelihood (see e.g. Davis et al., 2013b), which requires the knowledge of the bivariate density function for each $t_1, t_2 \in \mathbb{R}$ and $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^2$. The latter is given in the following proposition.

Proposition 3. *For $t_1, t_2 \in \mathbb{R}$, $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^2$ and $z_1, z_2 > 0$, the bivariate density of the process* (10) *is given by*

$$
\begin{aligned}
&f_{(t_1,\mathbf{x}_1),(t_2,\mathbf{x}_2)}(z_1, z_2, \boldsymbol{\theta}) \\
&\quad = \exp\left(-\mathcal{V}_{\mathbf{x}_1,\mathbf{x}_2-(t_2-t_1)\boldsymbol{\tau}}\left(z_1, \frac{z_2}{a^{t_2-t_1}}\right) - \frac{1-a^{t_2-t_1}}{z_2}\right) \\
&\quad \times \Bigg[-\frac{\partial}{\partial z_1}\mathcal{V}_{\mathbf{x}_1,\mathbf{x}_2-(t_2-t_1)\boldsymbol{\tau}}\left(z_1, \frac{z_2}{a^{t_2-t_1}}\right) \\
&\qquad \times \left(-\frac{\partial}{\partial z_2}\mathcal{V}_{\mathbf{x}_1,\mathbf{x}_2-(t_2-t_1)\boldsymbol{\tau}}\left(z_1, \frac{z_2}{a^{t_2-t_1}}\right) + \frac{1-a^{t_2-t_1}}{z_2^2}\right) \\
&\qquad - \frac{\partial}{\partial z_1}\frac{\partial}{\partial z_2}\mathcal{V}_{\mathbf{x}_1,\mathbf{x}_2-(t_2-t_1)\boldsymbol{\tau}}\left(z_1, \frac{z_2}{a^{t_2-t_1}}\right)\Bigg].
\end{aligned} \tag{16}
$$

Proof. We have that

$$
\begin{aligned}
&f_{(t_1,\mathbf{x}_1),(t_2,\mathbf{x}_2)}(z_1, z_2, \boldsymbol{\theta}) \\
&= \frac{\partial}{\partial z_1}\frac{\partial}{\partial z_2}\exp\left(-\mathcal{V}_{\mathbf{x}_1,\mathbf{x}_2-(t_2-t_1)\boldsymbol{\tau}}\left(z_1, \frac{z_2}{a^{t_2-t_1}}\right) - \frac{1-a^{t_2-t_1}}{z_2}\right) \\
&= \frac{\partial}{\partial z_1}\Bigg(\exp\left(-\mathcal{V}_{\mathbf{x}_1,\mathbf{x}_2-(t_2-t_1)\boldsymbol{\tau}}\left(z_1, \frac{z_2}{a^{t_2-t_1}}\right) - \frac{1-a^{t_2-t_1}}{z_2}\right) \\
&\quad \times \left(-\frac{\partial}{\partial z_2}\mathcal{V}_{\mathbf{x}_1,\mathbf{x}_2-(t_2-t_1)\boldsymbol{\tau}}\left(z_1, \frac{z_2}{a^{t_2-t_1}}\right) + \frac{1-a^{t_2-t_1}}{z_2^2}\right)\Bigg) \\
&= \exp\left(-\mathcal{V}_{\mathbf{x}_1,\mathbf{x}_2-(t_2-t_1)\boldsymbol{\tau}}\left(z_1, \frac{z_2}{a^{t_2-t_1}}\right) - \frac{1-a^{t_2-t_1}}{z_2}\right) \\
&\quad \times \left(-\frac{\partial}{\partial z_1}\mathcal{V}_{\mathbf{x}_1,\mathbf{x}_2-(t_2-t_1)\boldsymbol{\tau}}\left(z_1, \frac{z_2}{a^{t_2-t_1}}\right)\right) \\
&\quad \times \left(-\frac{\partial}{\partial z_2}\mathcal{V}_{\mathbf{x}_1,\mathbf{x}_2-(t_2-t_1)\boldsymbol{\tau}}\left(z_1, \frac{z_2}{a^{t_2-t_1}}\right) + \frac{1-a^{t_2-t_1}}{z_2^2}\right) \\
&\quad + \exp\left(-\mathcal{V}_{\mathbf{x}_1,\mathbf{x}_2-(t_2-t_1)\boldsymbol{\tau}}\left(z_1, \frac{z_2}{a^{t_2-t_1}}\right) - \frac{1-a^{t_2-t_1}}{z_2}\right) \\
&\quad \times \left(-\frac{\partial}{\partial z_1}\frac{\partial}{\partial z_2}\mathcal{V}_{\mathbf{x}_1,\mathbf{x}_2-(t_2-t_1)\boldsymbol{\tau}}\left(z_1, \frac{z_2}{a^{t_2-t_1}}\right)\right),
\end{aligned}
$$

yielding the result. □

We now consider the case of the spatial Smith model. Its covariance matrix is

$$
\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}
$$

and $\boldsymbol{\theta}$ is now given by $(\sigma_{11}, \sigma_{12}, \sigma_{22}, a, \boldsymbol{\tau}')'$. The bivariate density function is given below.

Corollary 1. *We set*

$$w_1 = \frac{h_1}{2} + \frac{1}{h_1} \log\left(\frac{z_2}{a^{t_2-t_1} z_1}\right)$$

and $v_1 = h_1 - w_1$, *where*

$$h_1 = \sqrt{(\mathbf{x}_2 - (t_2 - t_1)\boldsymbol{\tau} - \mathbf{x}_1)'\Sigma^{-1}(\mathbf{x}_2 - (t_2 - t_1)\boldsymbol{\tau} - \mathbf{x}_1)}.$$

Let $t_1, t_2 \in \mathbb{R}$, $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^2$ *and* $z_1, z_2 > 0$. *The bivariate density function of the process* (10) *when* $((Z(t,\mathbf{x}))_{\mathbf{x}\in\mathbb{R}^2})_{t\in\mathbb{N}}$ *is a sequence of iid spatial Smith processes is given by*

$$\begin{aligned} f_{(t_1,\mathbf{x}_1),(t_2,\mathbf{x}_2)}&(z_1, z_2, \boldsymbol{\theta}) \\ &= \exp\left(-\frac{\Phi(w_1)}{z_1} - \frac{a^{t_2-t_1}\Phi(v_1)}{z_2} - \frac{1-a^{t_2-t_1}}{z_2}\right) \\ &\times \Bigg[\left(\frac{\Phi(w_1)}{z_1^2} + \frac{\phi(w_1)}{h_1 z_1^2} - \frac{a^{t_2-t_1}\phi(v_1)}{h_1 z_1 z_2}\right) \\ &\quad \times \left(\frac{a^{t_2-t_1}\Phi(v_1)}{z_2^2} + \frac{a^{t_2-t_1}\phi(v_1)}{h_1 z_2^2} - \frac{\phi(w_1)}{h_1 z_1 z_2} + \frac{1-a^{t_2-t_1}}{z_2^2}\right) \\ &\quad + \frac{v_1\phi(w_1)}{h_1^2 z_1^2 z_2} + \frac{a^{t_2-t_1} w_1 \phi(v_1)}{h_1^2 z_1 z_2^2}\Bigg], \end{aligned}$$

where Φ *and* ϕ *are respectively the probability distribution function and probability density function of a standard Gaussian random variable.*

Proof. Set

$$w = \frac{h}{2} + \frac{1}{h}\log\left(\frac{z_2}{z_1}\right)$$

and $v = h - w$, where $h = \sqrt{(\mathbf{x}_2 - \mathbf{x}_1)'\Sigma^{-1}(\mathbf{x}_2 - \mathbf{x}_1)}$. From Padoan et al. (2010), p. 275, we know that

$$\begin{aligned} -\frac{\partial}{\partial z_1}\mathcal{V}_{\mathbf{x}_1,\mathbf{x}_2}(z_1, z_2) &= \frac{\Phi(w)}{z_1^2} + \frac{\phi(w)}{h z_1^2} - \frac{\phi(v)}{h z_1 z_2}, \\ -\frac{\partial}{\partial z_2}\mathcal{V}_{\mathbf{x}_1,\mathbf{x}_2}(z_1, z_2) &= \frac{\Phi(v)}{z_2^2} + \frac{\phi(v)}{h z_2^2} - \frac{\phi(w)}{h z_1 z_2} \end{aligned}$$

and

$$-\frac{\partial^2}{\partial z_1 \partial z_2}\mathcal{V}_{\mathbf{x}_1,\mathbf{x}_2}(z_1, z_2) = \frac{v\phi(w)}{h^2 z_1^2 z_2} + \frac{w\phi(v)}{h^2 z_1 z_2^2}.$$

Hence we obtain

$$-\frac{\partial}{\partial z_1}\mathcal{V}_{\mathbf{x}_1,\mathbf{x}_2-(t_2-t_1)\boldsymbol{\tau}}\left(z_1,\frac{z_2}{a^{t_2-t_1}}\right)=\frac{\Phi(w_1)}{z_1^2}+\frac{\phi(w_1)}{h_1z_1^2}-\frac{a^{t_2-t_1}\phi(v_1)}{h_1z_1z_2}, \tag{17}$$

$$\begin{aligned}-\frac{\partial}{\partial z_2}&\mathcal{V}_{\mathbf{x}_1,\mathbf{x}_2-(t_2-t_1)\boldsymbol{\tau}}\left(z_1,\frac{z_2}{a^{t_2-t_1}}\right)\\&=\frac{1}{a^{t_2-t_1}}\left(\frac{a^{2(t_2-t_1)}\Phi(v_1)}{z_2^2}+\frac{a^{2(t_2-t_1)}\phi(v_1)}{h_1z_2^2}-\frac{a^{t_2-t_1}\phi(w_1)}{h_1z_1z_2}\right)\\&=\frac{a^{t_2-t_1}\Phi(v_1)}{z_2^2}+\frac{a^{t_2-t_1}\phi(v_1)}{h_1z_2^2}-\frac{\phi(w_1)}{h_1z_1z_2}\end{aligned} \tag{18}$$

and

$$\begin{aligned}-\frac{\partial^2}{\partial z_1\partial z_2}\mathcal{V}_{\mathbf{x}_1,\mathbf{x}_2-(t_2-t_1)\boldsymbol{\tau}}\left(z_1,\frac{z_2}{a^{t_2-t_1}}\right)&=\frac{1}{a^{t_2-t_1}}\left(\frac{a^{t_2-t_1}v_1\phi(w_1)}{h_1^2z_1^2z_2}+\frac{a^{2(t_2-t_1)}w_1\phi(v_1)}{h_1^2z_1z_2^2}\right)\\&=\frac{v_1\phi(w_1)}{h_1^2z_1^2z_2}+\frac{a^{t_2-t_1}w_1\phi(v_1)}{h_1^2z_1z_2^2}.\end{aligned} \tag{19}$$

Finally,

$$\mathcal{V}_{\mathbf{x}_1,\mathbf{x}_2-(t_2-t_1)\boldsymbol{\tau}}\left(z_1,\frac{z_2}{a^{t_2-t_1}}\right)=\frac{\Phi(w_1)}{z_1}+\frac{a^{t_2-t_1}\Phi(v_1)}{z_2}. \tag{20}$$

Inserting (17), (18), (19) and (20) in (16), we obtain the result. □

Assume that we observe the process at M locations $\mathbf{x}_1,\ldots,\mathbf{x}_M$ and N dates $t_1,\ldots,t_N$. Then the spatio-temporal pairwise log-likelihood is defined by (see e.g. Davis et al., 2013b, Section 3.1)

$$L_P^{ST}(\boldsymbol{\theta})=\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\sum_{k=1}^{M-1}\sum_{l=k+1}^{M}\omega_{i,j}\omega_{k,l}\log f_{(t_i,\mathbf{x}_k),(t_j,\mathbf{x}_l)}(z_{i,k},z_{j,l},\boldsymbol{\theta}),$$

where the $\omega_{i,j}$ and $\omega_{k,l}$ are temporal and spatial weights, respectively, and $z_{n,m}$ denotes the observation of the process at date n and site m. Then the maximum pairwise likelihood estimator is given by $\hat{\boldsymbol{\theta}}=\operatorname{argmax} L_P^{ST}(\boldsymbol{\theta})$.

We consider two different estimation schemes:

- **Scheme 1**: As previously explained, because of Theorem 1 it is possible to separate the estimation of $\boldsymbol{\theta}_1=(\sigma_{11},\sigma_{12},\sigma_{22})'$ and $\boldsymbol{\theta}_2=(a,\boldsymbol{\tau}')'$. As a first step, the estimation of $\boldsymbol{\theta}_1$ is carried out by maximizing the spatial pairwise log-likelihood (see Padoan et al., 2010, Section 3.2). Once $\boldsymbol{\theta}_1$ is known, it is held fixed and we estimate $\boldsymbol{\theta}_2$ by maximizing $L_P^{ST}(\boldsymbol{\theta}_2)$ with respect to $\boldsymbol{\theta}_2$.

- **Scheme 2**: We optimize $L_P^{ST}(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$, meaning that we estimate all parameters in a single step.

As an illustration of the above, we simulate the process considered 100 times, with parameter $\boldsymbol{\theta}=(1,0,1,0.7,-1,-1)$, at M sites and N dates. We compute statistical summaries from the 100 estimates obtained. In both schemes, we optimize L_P^{ST} with $\omega_{i,j}=1$ and $\omega_{k,l}=1$ for all $i=1,\ldots,N-1$, $j=i+1,\ldots,N$, $k=1,\ldots,M-1$ and

Table 1: Performance of the estimation in the case of Scheme 1. The mean estimate, the mean bias and the standard deviation are displayed.

True	**Pairwise likelihood (M=N=20)**			**Pairwise likelihood (M=N=30)**		
	Mean estimate	Mean bias	Stdev	Mean estimate	Mean bias	Stdev
$\sigma_{11} = 1$	1.139	0.139	0.421	1.105	0.105	0.232
$\sigma_{12} = 0$	0.040	0.040	0.286	−0.024	−0.024	0.162
$\sigma_{22} = 1$	1.185	0.185	0.325	1.066	0.066	0.254
$a = 0.7$	0.707	0.007	0.059	0.701	0.001	0.026
$\boldsymbol{\tau}_1 = -1$	−0.990	0.010	0.123	−0.999	0.001	0.032
$\boldsymbol{\tau}_2 = -1$	−0.990	0.010	0.101	−0.998	0.002	0.043

Table 2: Performance of the estimation in the case of Scheme 2. The mean estimate, the mean bias and the standard deviation are displayed.

True	**Pairwise likelihood (M=N=20)**			**Pairwise likelihood (M=N=30)**		
	Mean estimate	Mean bias	Stdev	Mean estimate	Mean bias	Stdev
$\sigma_{11} = 1$	1.288	0.288	0.678	1.239	0.239	0.483
$\sigma_{12} = 0$	0.043	0.043	0.621	0.057	0.057	0.314
$\sigma_{22} = 1$	1.453	0.453	1.159	1.264	0.264	0.574
$a = 0.7$	0.706	0.006	0.050	0.700	0.000	0.016
$\boldsymbol{\tau}_1 = -1$	−0.998	0.002	0.115	−1.002	−0.002	0.034
$\boldsymbol{\tau}_2 = -1$	−0.982	0.018	0.111	−1.003	−0.003	0.035

$l = k+1, \dots, M$. Tables 1 and 2 display the results for different values of M and N, in the cases of Scheme 1 and Scheme 2 respectively.

For both schemes, the estimation is more accurate (the mean bias and the standard deviation decrease) as M and N increase. Moreover, we observe that the estimation of the spatio-temporal parameters a and $\boldsymbol{\tau}$ is satisfactory and clearly more accurate than that of the purely spatial parameters σ_{11}, σ_{12} and σ_{22} (the mean bias and the standard deviation are lower). Finally, the estimation of the purely spatial parameters is more accurate when using Scheme 1 (the mean bias and the standard deviation are lower). This stems probably from the fact that in Scheme 2, the number of pairs used is higher than in Scheme 1, introducing more variability. Indeed, contrary to what is assumed in the pairwise log-likelihood, the pairs considered are not independent. This dependence generates instability. For a discussion about the impact of the choice of pairs on estimation efficiency, see Padoan et al. (2010), pp. 266, 268. This finding shows that from a statistical point of view, spatio-temporal max-stable models that allow a separate estimation of the purely statistical parameters can be preferable; needless to say that a more extensive analysis would be needed at this point.

5. Concluding remarks

In summary, in order to overcome the defects of the spatio-temporal max-stable models introduced in the literature, we propose a class of models where we partly decouple the influence of time and space in the spectral representations. Time has an influence on space through a bijective operator in space. Then, we propose several sub-classes of models where our operator is either a translation or a rotation. An advantage of the class of models we propose lies in the fact that it allows the roles of time and space to be distinct. Especially, the distribution in space (when time is fixed) can differ from the distributions in time (when the location is fixed). Moreover, the space operator allows us to account for physical processes. Our models have both a continuous-time and a discrete-time version.

Then we consider a special case of some of our models where the function related to time in the spectral representation is the exponential density (continuous-time case) or takes as values the probabilities of a geometric random variable (discrete-time case). In this context, the corresponding models become Markovian and have a useful max-autoregressive representation. They appear as an extension to a spatial setting of the real-valued MARMA$(1, 0)$ process introduced by Davis and Resnick (1989). The main advantage of these models lies in the fact that the temporal dynamics are explicit and easy to interpret. Finally, we briefly describe an inference method and show that it works well on simulated data, especially in the case of the parameters related to time. A detailed study of possible estimation methodologies for our class of models will be considered in a subsequent paper. In particular, we are able to show, using Hairer (2010), Theorem 3.6, that for instance the models of type 2 in Section 3.2 are geometrically ergodic.

Finally, note that it could be interesting to consider the following generalization of the model (10):

$$X(t, \mathbf{x}) = \max\big(\Psi(X(t-1, \cdot))(\mathbf{x}), Z(t, \mathbf{x})\big), \quad (t, \mathbf{x}) \in \mathbb{Z} \times \mathbb{R}^2,$$

where $\big((Z(t, \mathbf{x}))_{\mathbf{x} \in \mathbb{R}^2}\big)_{t \in \mathbb{Z}}$ is a sequence of iid spatial max-stable processes and Ψ is an operator from the space of continuous functions on $\mathbb{R}^2$ to itself such that, if $(X(t-1, \mathbf{x}))_{\mathbf{x} \in \mathbb{R}^2}$ is max-stable in space, then $\big(\Psi(X(t-1, \cdot))(\mathbf{x})\big)_{\mathbf{x} \in \mathbb{R}^2}$ is also max-stable in space. Such an operator could for instance be a "moving-maxima" operator

$$\Psi(X(t-1, \cdot))(\mathbf{x}) = \bigvee_{\mathbf{s} \in \mathbb{R}^2} \{K(\mathbf{s}, \mathbf{x}) X(t-1, \mathbf{s})\}, \quad \mathbf{x} \in \mathbb{R}^2,$$

where K is a kernel (see Meinguet (2012) for a similar idea).

Acknowledgements

Paul Embrechts acknowledges financial support by the Swiss Finance Institute (SFI). Erwan Koch would like to thank Andrea Gabrielli for interesting discussions. He also acknowledges RiskLab at ETH Zurich and the SFI for financial support. Christian Robert would like to thank Mathieu Ribatet and Johan Segers for fruitful discussions on a related topic when Johan Segers visited ISFA in October 2013. Finally, the three authors acknowledge the referee for useful comments and suggestions.

References

Beirlant, J., Goegebeur, Y., Segers, J. and Teugels, J. (2004). *Statistics of Extremes: Theory and Applications.* John Wiley & Sons, New York.

Bienvenüe, A. and Robert, C. Y. (2014). Likelihood based inference for high-dimensional extreme value distributions. Preprint, `http://arxiv.org/abs/1403.0065`.

Buhl, S. and Klüppelberg, C. (2015). Anisotropic Brown-Resnick space-time processes: estimation and model assessment. Preprint, `http://arxiv.org/abs/1503.06049`.

COLES, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag, London.

COOLEY, D., NAVEAU, P. AND PONCET, P. (2006). Variograms for spatial max-stable random fields. In *Dependence in Probability and Statistics*, ed. P. Bertail, P. Doukhan, P. Soulier. Lecture Notes in Statist. **187**, Springer-Verlag, New York, pp. 373–390.

DAVIS, R. A., KLÜPPELBERG, C. AND STEINKOHL, C. (2013a). Max-stable processes for modeling extremes observed in space and time. *J. Korean Statist. Soc.* **42**, 399–414.

DAVIS, R. A., KLÜPPELBERG, C. AND STEINKOHL, C. (2013b). Statistical inference for max-stable processes in space and time. *J. R. Statist. Soc. Ser. B* **75**, 791–819.

DAVIS, R. A. AND RESNICK, S. I. (1989). Basic properties and prediction of max-ARMA processes. *Adv. Appl. Prob.* **21**, 781–803.

DOMBRY, C. AND EYI-MINKO, F. (2014). Stationary max-stable processes with the Markov property. *Stoch. Proc. Appl.* **124**, 2266–2279.

EMBRECHTS, P., KLÜPPELBERG, C. AND MIKOSCH, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer-Verlag, Berlin.

HAAN, L. DE (1984). A spectral representation for max-stable processes. *Ann. Prob.* **12**, 1194–1204.

HAAN, L. DE AND FERREIRA, A. (2007). *Extreme Value Theory: An Introduction*. Springer-Verlag, New York.

HAAN, L. DE AND PICKANDS, J. (1986). Stationary min-stable stochastic processes. *Prob. Theory Rel. Fields* **72**, 477–492.

HAIRER, M. (2010). Convergence of Markov processes. Probability at Warwick course, `http://www.hairer.org/notes/Convergence.pdf`.

HUSER, R. AND DAVISON, A. (2014). Space-time modelling of extreme events. *J. R. Statist. Soc. Ser. B* **76**, 439–461.

KABLUCHKO, Z. AND SCHLATHER, M. (2010). Ergodic properties of max-infinitely divisible processes. *Stoch. Proc. Appl.* **120**, 281–295.

KABLUCHKO, Z., SCHLATHER, M. AND HAAN, L. DE (2009). Stationary max-stable fields associated to negative definite functions. *Ann. Prob.* **37**, 2042–2065.

MARDIA, K. V. AND JUPP, P. E. (1999). *Directional Statistics*. John Wiley & Sons, New York.

MEINGUET, T. (2012). Maxima of moving maxima of continuous functions. Extremes **15**, 267–297.

NAVEAU, P., GUILLOU, A., COOLEY, D. AND DIEBOLT, J. (2009). Modelling pairwise dependence of maxima in space. *Biometrika* **96**, 1–17.

PADOAN, S. A., RIBATET, M. AND SISSON, S. A. (2010). Likelihood-based inference for max-stable processes. *J. Amer. Statist. Assoc.* **105**, 263–277.

PENROSE, M. D. (1992). Semi-min-stable processes. *Ann. Prob.* **20**, 1450–1463.

RESNICK, S. I. (1987). *Extreme Values, Regular Variation, and Point Processes*. Springer-Verlag, New York.

RIBATET, M. (2015). SpatialExtremes. *R* Package, version 2.0–2: `http://spatialextremes.r-forge.r-project.org/`.

SCHLATHER, M. (2002). Models for stationary max-stable random fields. *Extremes* **5**, 33–44.

SCHLATHER, M. AND TAWN, J. A. (2003). A dependence measure for multivariate and spatial extreme values: properties and inference. *Biometrika* **90**, 139–156.

SMITH, R. L. (1990). Max-stable processes and spatial extremes. Unpublished manuscript, University of North Carolina.

STROKORB, K., BALLANI, F. AND SCHLATHER, M. (2015). Tail correlation functions of max-stable processes: construction principles, recovery and diversity of some mixing max-stable processes with identical TCF. *Extremes* **18**, 241–271.

SWISS RE. (2014). Natural catastrophes and man-made disasters in 2013: large losses from floods and hail; Haiyan hits the Philippines. *Sigma* **1/2014**, Swiss Re, Zürich.