Bayesian Modeling of Air Pollution Extremes Using Nested Multivariate Max-Stable Processes

Huser Raphael   ORCID iD: 0000-0002-1228-2071

Bayesian Modeling of Air Pollution Extremes Using Nested Multivariate Max-Stable Processes

Sabrina Vettori

King Abdullah University of Science and Technology (KAUST),

Computer, Electrical and Mathematical Science and Engineering Division (CEMSE),

Thuwal 23955-6900, Saudi Arabia.

email: sabrina.vettori@kaust.edu.sa

and

Raphaël Huser

King Abdullah University of Science and Technology (KAUST),

Computer, Electrical and Mathematical Science and Engineering Division (CEMSE),

Thuwal 23955-6900, Saudi Arabia.

email: raphael.huser@kaust.edu.sa

and

Marc G. Genton

King Abdullah University of Science and Technology (KAUST),

Computer, Electrical and Mathematical Science and Engineering Division (CEMSE),

Thuwal 23955-6900, Saudi Arabia.

email: marc.genton@kaust.edu.sa

SUMMARY: Capturing the potentially strong dependence among the peak concentrations of multiple air pollutants across a spatial region is crucial for assessing the related public health risks. In order to investigate the multivariate spatial dependence properties of air pollution extremes, we introduce a new class of multivariate max-stable processes. Our proposed model admits a hierarchical tree-based formulation, in which the data are conditionally independent given some latent nested positive stable random factors. The hierarchical structure facilitates Bayesian inference and offers a convenient and interpretable characterization. We fit this nested multivariate max-stable model to the maxima of air pollution concentrations and temperatures recorded at a number of sites in the Los Angeles area, showing that the proposed model succeeds in capturing their complex tail dependence structure.

KEY WORDS: Air pollution; Bayesian hierarchical modeling; Extreme event; Multivariate max-stable process, Reich–Shaby model.

1. Introduction

Modeling the joint behavior of multivariate extreme events is of interest in a wide range of applications, ranging from finance to environmental sciences, such as hydrology and applications related to climate change or air pollution monitoring. Simultaneous exposure to multiple air pollutants seriously affects public health worldwide, causing loss of life and livelihood and requiring costly health care. Therefore, policymakers such as those at the US Environmental Protection Agency (EPA) are researching multivariate approaches to quantify air pollution risks (Dominici et al., 2010). The issue of air pollution is compounded by global warming and climate change, as increasingly high temperatures are suspected to contribute to raising ozone concentrations and aggravating their effect in the human body (Kahle et al., 2015). This situation urges a greater understanding and better monitoring of air pollution extremes, the complexity of which poses a challenge for standard statistical techniques. Several air pollutants are often recorded at multiple spatial locations and the modeling of peak exposures across a spatial region must transcend the assumption of independence in order to capture their spatial variability. In this paper, we propose a new methodological framework based on Extreme-Value Theory, for estimating the probability that various air pollutants and temperatures will be simultaneously extreme at multiple locations.

The statistical modeling of single extreme variables observed over space is usually based on spatial max-stable processes (see, e.g., the reviews by Davison et al., 2012, and Davison et al., 2019), which are the only possible limit models of properly renormalized block maxima from independent and identically distributed spatial processes. Notice that although pointwise maxima may or may not occur simultaneously, the dependence structure of maxima coincides asymptotically with the tail dependence structure from the original process. Therefore, except in the case of asymptotic independence, max-stable models can be used to capture the potentially strong spatial dependence that may exist among original variables at extreme levels. Several spatial models for asymptotically (in)dependent data were proposed by Wadsworth and Tawn (2012), Opitz (2016), Huser et al. (2017), Huser and Wadsworth (2019) and Krupskii et al. (2018). Here, we restrict ourselves to asymptotic dependence by modeling multivariate block maxima recorded over space using a suitable max-stable process. Genton et al. (2015) proposed multivariate versions of the Gaussian (Smith, 1990), the extremal-Gaussian (Schlather, 2002), extremal-t (Opitz, 2013) and the Brown–Resnick (Kabluchko et al., 2009) max-stable models, and Oesting et al. (2017) introduced a bivariate Brown–Resnick max-stable process to jointly model the spatial observations and forecasts of wind gusts

in Northern Germany. In this paper, we propose a new class of multivariate max-stable processes that extends the Reich–Shaby model (Reich and Shaby, 2012) to the multivariate setting, and that is suitable for studying the spatial and cross-dependence structures of multiple max-stable random fields within an intuitive and computationally convenient hierarchical tree-based framework. Our model is unique in its flexibility for capturing cross-dependence and it is applied below considering five variables over space. Recent related work includes Reich and Shaby (2018b), but their proposed model has a highly restrictive cross-dependence structure and has been applied in the bivariate case only.

In contrast to the standard spatial processes based on the Gaussian distribution, such as the Brown–Resnick, the computationally demanding nature of the likelihood function for max-stable processes has hampered their use in high-dimensional settings within both frequentist and Bayesian frameworks (Ribatet et al., 2012; Huser and Davison, 2013; Castruccio et al., 2016; Thibaud et al., 2016; Huser et al., 2019). In the Bayesian context, Thibaud et al. (2016) showed how the Brown–Resnick max-stable process may be fitted using a Markov chain Monte Carlo (MCMC) algorithm; however, this remains excessively expensive in high dimensions. From a computational perspective, it is convenient to relax the max-stable structure by assuming conditional independence of the extreme data given an unobserved latent process (Casson and Coles, 1999; Cooley et al., 2007; Davison et al., 2012; Opitz et al., 2018). This significantly facilitates Bayesian and likelihood-based inference and is helpful for estimating marginal distributions by borrowing strength across locations. Unfortunately, when the latent process is Gaussian, the resulting dependence structure lacks flexibility and cannot capture strong extremal dependence. The Reich–Shaby model (Reich and Shaby, 2012) on the other hand has a conditional independence representation given some latent positive stable random effects, and is jointly max-stable. Other popular max-stable processes do not possess such a convenient hierarchical characterization. In this paper, we generalize the Reich–Shaby process for modeling multivariate spatial extremes by assuming a nested, tree-based, positive stable latent structure, which maintains a low computational burden.

In our proposed model, the dependence structure may be represented by a tree framework, in which the "leaves" (i.e., the terminal nodes) correspond to different spatial Reich–Shaby processes representing different variables of interest (e.g., pollutants), and the tree "branches" describe the relationships among these processes, which are grouped into clusters. The intra-cluster cross-dependence is assumed to be exchangeable and stronger than the inter-cluster cross-dependence. In principle, the underlying tree struc-

ture can involve an arbitrary number of "layers" (i.e., node levels) in order to describe more complicated forms of cross-dependence among the spatial processes, although more complex trees necessarily imply an increased number of latent variables and parameters, thus complicating the inference procedure.

The remainder of this paper is organized as follows. In §2, we review the modeling of spatial extremes based on max-stable processes. In §3, we introduce our novel class of multivariate max-stable processes and we describe the inference procedure based on an MCMC algorithm. In §4, we use our model to study the dependence structure of concentration maxima of various air pollutants and temperature, observed at a number of sites across the Los Angeles area in California, US. §5 concludes with some final remarks and perspectives for future research.

2. Max-stable processes

Owing to their asymptotic characterization, max-stable processes are widely used for modeling spatially-indexed block maxima. Here, we briefly summarize the theory and modeling of max-stable processes in the spatial context, and in §3 we extend these processes to the multivariate spatial framework. Let $\{Y_1(\mathbf{s}), \ldots, Y_n(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$ be independent and identically distributed processes defined over the region $\mathcal{S} \subset \mathbb{R}^2$. If there exist normalization functions $a_n(\mathbf{s}) > 0$, $b_n(\mathbf{s}) \in \mathbb{R}$ such that the renormalized process of the pointwise maxima, i.e., $M_n(\mathbf{s}) = a_n(\mathbf{s})^{-1}\{\max\{Y_1(\mathbf{s}), \ldots, Y_n(\mathbf{s})\} - b_n(\mathbf{s})\}$, $\mathbf{s} \in \mathcal{S}$, converges in the sense of finite-dimensional distributions to a process $Z^\star(\mathbf{s})$ with non-degenerate margins, as $n \to \infty$, then $Z^\star(\mathbf{s})$ is a max-stable process (see, e.g., de Haan, 1984). The max-stability of $Z^\star(\mathbf{s})$ implies that there exist functions $\alpha_n(\mathbf{s}) > 0$ and $\beta_n(\mathbf{s})$ for all $n \in \mathbb{N}$, such that, for each collection of sites $\mathbf{s}_1, \ldots, \mathbf{s}_D \in \mathcal{S}$, $D \in \mathbb{N}$, the finite-dimensional distribution $G(z_1, \ldots, z_D)$ of the variables $Z^\star(\mathbf{s}_1), \ldots, Z^\star(\mathbf{s}_D)$ satisfies $G^n\{\alpha_n(\mathbf{s}_1)z_1 + \beta_n(\mathbf{s}_1), \ldots, \alpha_n(\mathbf{s}_D)z_D + \beta_n(\mathbf{s}_D)\} = G(z_1, \ldots, z_D)$. The limit max-stable process $Z^\star(\mathbf{s})$ may be used to represent, e.g., monthly maxima of daily measurements for a specific air pollutant observed at various locations $\mathbf{s}$ within the study region $\mathcal{S}$. For each location $\mathbf{s} \in \mathcal{S}$, the Extremal Types Theorem (see, e.g., Coles, 2001, Chapter 3) implies that the random variable $Z^\star(\mathbf{s})$ follows the generalized extreme-value (GEV) distribution with location $\mu(\mathbf{s}) \in \mathbb{R}$, scale $\sigma(\mathbf{s}) > 0$, and shape $\xi(\mathbf{s}) \in \mathbb{R}$ parameters. To disentangle marginal and dependence effects, it is convenient to standardize $Z^\star(\mathbf{s})$ as

$$Z(\mathbf{s}) = \left\{ 1 + \xi(\mathbf{s}) \frac{Z^\star(\mathbf{s}) - \mu(\mathbf{s})}{\sigma(\mathbf{s})} \right\}^{1/\xi(\mathbf{s})}. \tag{1}$$

Thus, we obtain a residual, *simple*, max-stable process $Z(\mathbf{s})$, which is characterized by unit Fréchet marginal distributions, i.e., $\Pr\{Z(\mathbf{s}) \leqslant z\} = \exp(-1/z)$, $z > 0$, for all $\mathbf{s} \in \mathcal{S}$, corresponding to the case $\mu(\mathbf{s}) = \sigma(\mathbf{s}) = \xi(\mathbf{s}) = 1$. In practice, data are collected at a finite set of locations, and the joint distribution of $Z(\mathbf{s})$ at $\mathbf{s}_1, \ldots, \mathbf{s}_D \in \mathcal{S}$ is necessarily a multivariate extreme-value distribution that may be expressed as

$$\Pr\{Z(\mathbf{s}_1) \leqslant z_1, \ldots, Z(\mathbf{s}_D) \leqslant z_D\} = \exp\left\{-V(z_1, \ldots, z_D)\right\}, \quad z_1, \ldots, z_D > 0, \quad (2)$$

where $V(z_1, \ldots, z_D)$ is the associated *exponent function* containing information about the spatial dependence of the maxima. By max-stability, we can verify that V is homogeneous of order -1, i.e., $V(tz_1, \ldots, tz_D) = t^{-1}V(z_1, \ldots, z_D)$, $t > 0$; moreover, because of the unit Fréchet margins in (2), we have $V(z, \infty, \ldots, \infty) = 1/z$ for any permutation of the arguments. In particular, in the case of independence between $Z(\mathbf{s}_1), \ldots, Z(\mathbf{s}_D)$ we have $V(z_1, \ldots, z_D) = \sum_{d=1}^{D} z_d^{-1}$; in the case of perfect positive dependence, we have $V(z_1, \ldots, z_D) = \max_{1 \leqslant d \leqslant D} z_d^{-1}$. The pairwise extremal coefficient $\theta(\mathbf{s}_i, \mathbf{s}_j) = V(1,1) \in [1,2]$, where V is here restricted to sites $\mathbf{s}_i$ and $\mathbf{s}_j$, is a measure of the dependence between the variables $Z(\mathbf{s}_i)$ and $Z(\mathbf{s}_j)$. Perfect dependence corresponds to $\theta(\mathbf{s}_i, \mathbf{s}_j) = 1$, and $\theta(\mathbf{s}_i, \mathbf{s}_j) = 2$ leads to complete independence.

Thanks to the spectral characterization of max-stable processes (de Haan, 1984), various parametric models have been proposed in the literature. The most well-known in the spatial framework are the Smith (1990), Schlather (2002), Brown–Resnick (Kabluchko et al., 2009), extremal-t (Opitz, 2013), Reich–Shaby (Reich and Shaby, 2012), and Tukey (Xu and Genton, 2017) max-stable processes. For comprehensive reviews on spatial extremes, see, e.g., Davison et al. (2012) and Davison et al. (2019).

In this work, the Reich–Shaby max-stable model, which admits a hierarchical construction in terms of positive stable random effects, plays a key role. Its exponent function in (2) is

$$V(z_1, \ldots, z_D) = \sum_{l=1}^{L} \left[\sum_{d=1}^{D} \left\{ \frac{z_d}{\omega_l(\mathbf{s}_d)} \right\}^{-1/\alpha} \right]^{\alpha}, \qquad z_1, \ldots, z_D > 0, \qquad (3)$$

where $\omega_l(\mathbf{s}) \geqslant 0$, $l = 1, \ldots, L$, are deterministic spatial profiles (or kernels) such that $\sum_{l=1}^{L} \omega_l(\mathbf{s}) = 1$ for any location $\mathbf{s} \in \mathcal{S}$, and $\alpha > 0$. Although other kernels are possible, Reich and Shaby (2012) proposed using the isotropic Gaussian density function

$$g_l(\mathbf{s}) = \frac{1}{2\pi\tau^2} \exp\left\{ -\frac{1}{2\tau^2}(\mathbf{s} - \mathbf{v}_l)^\top(\mathbf{s} - \mathbf{v}_l) \right\}, \quad l = 1, \ldots, L, \qquad (4)$$

with a bandwidth (i.e., spatial range parameter) $\tau > 0$ and fixed spatial knots $\mathbf{v}_1, \ldots, \mathbf{v}_L$ distributed over the domain $\mathcal{S}$, which is then rescaled as $\omega_l(\mathbf{s}) = g_l(\mathbf{s})\left\{ \sum_{l=1}^{L} g_l(\mathbf{s}) \right\}^{-1}$.

The hierarchical construction of the Reich–Shaby model, detailed in Web Appendix A, allows for fast Bayesian inference in high dimensions. In Section 3, we extend the Reich–Shaby to the multivariate spatial setting.

3. Nested multivariate max-stable processes

3.1 *Tree-based construction of multivariate max-stable processes*

In our multivariate spatial process, the univariate spatial margins follow the Reich–Shaby model and interact with each other according to a nested tree-based structure for their latent positive stable random effects.

We first define the two-layer nested multivariate max-stable process, and then extend it to multilayer tree structures. Analogously to the hierarchical construction of the Reich–Shaby model detailed in Web Appendix A, for each $k = 1, \dots, K$, let $U_k(\mathbf{s}) \overset{\text{i.i.d.}}{\sim} \exp\{-z^{-1/(\alpha_k \alpha_0)}\}$, $z > 0$, denote a random Fréchet noise process controlled by the product of the two parameters $\alpha_k, \alpha_0 \in (0, 1]$; and let $\vartheta_k(\mathbf{s}) = \left\{ \sum_{l=1}^{L} A_{k;l} A_{0;l}^{1/\alpha_k} \omega_{k;l}(\mathbf{s})^{1/(\alpha_k \alpha_0)} \right\}^{\alpha_k \alpha_0}$ be a smooth spatial process, where the terms $\omega_{k;l}(\mathbf{s}) \geqslant 0$ are L kernel basis functions representing deterministic weights, such that $\sum_{l=1}^{L} \omega_{k;l}(\mathbf{s}) = 1$ for any $\mathbf{s} \in \mathcal{S}$, and the variables $A_{k;l}$ and $A_{0;l}$, $l = 1, \dots, L$, are mutually independent (across both $k = 1, \dots, K$ and $l = 1, \dots, L$) positive stable (PS) variables with parameters α_k and α_0, respectively. Thus, $A_{k;l} \overset{\text{i.i.d.}}{\sim} \mathrm{PS}(\alpha_k) \perp\!\!\!\perp A_{0;l} \overset{\text{i.i.d.}}{\sim} \mathrm{PS}(\alpha_0)$. Although the positive stable density is not available in closed form, the Laplace transform of $A \sim \mathrm{PS}(\alpha)$ is $\mathrm{E}(e^{-tA}) = \exp(-t^\alpha)$. Using this, we can show that $A_{k;l} A_{0;l}^{1/\alpha_k}$ follows a positive stable distribution with parameter $\alpha_k \alpha_0$; therefore, each process defined as $Z_k(\mathbf{s}) = U_k(\mathbf{s}) \vartheta_k(\mathbf{s})$ $(k = 1, \dots, K)$ is a univariate Reich–Shaby process defined over the region $\mathcal{S} \subset \mathbb{R}^2$, with the dependence parameter $\alpha_k \alpha_0$, kernels $\omega_{k;l}(\mathbf{s})$, $l = 1, \dots, L$, and unit Fréchet margins; see Web Appendix A for more details. Cross-dependence among these marginal processes is induced by their shared latent variables $A_{0;1}, \dots, A_{0;L}$. Then, we combine the residual univariate processes $Z_k(\mathbf{s})$, $k = 1, \dots, K$, into the multivariate max-stable process $\mathbf{Z}(\mathbf{s}) = \{Z_1(\mathbf{s}), \dots, Z_K(\mathbf{s})\}^\top$, whose finite-dimensional distributions at D locations $\mathbf{s}_1, \dots, \mathbf{s}_D$ are expressed as (2) in terms of the exponent function

$$V(\mathbf{z}_1, \dots, \mathbf{z}_K) = \sum_{l=1}^{L} \left(\sum_{k=1}^{K} \left[\sum_{d=1}^{D} \left\{ \frac{z_{k;d}}{\omega_{k;l}(\mathbf{s}_d)} \right\}^{-1/(\alpha_k \alpha_0)} \right]^{\alpha_k} \right)^{\alpha_0}, \quad z_{k;d} > 0 \quad \text{for all } k, d.$$

(5)

In (5), $\mathbf{z}_k = (z_{k;1}, \dots, z_{k;D})^\top$ denotes the vector containing the maxima of the kth variable observed at D locations, while the parameters $0 < \alpha_k, \alpha_0 \leqslant 1$ $(k = 1, \dots, K)$ control the

spatial and cross-dependence structures. Model (5) corresponds to a max-mixture of L independent *nested* logistic max-stable distributions (Tawn, 1990), with kernels $\omega_{k;l}(\mathbf{s})$ introducing spatial asymmetries. This nested cross-dependence structure, represented by a simple tree in the top panel of Figure 1, assumes that all the univariate processes are exchangeable. In the following, we refer to the max-stable process $\mathbf{Z}(\mathbf{s})$ with (5) as a two-layer nested multivariate max-stable process.

[Figure 1 about here.]

The exchangeability of such two-layer processes is not always realistic, but we overcome this limitation by generalizing the construction above to multilayer, partially exchangeable, tree structures based on nested α-stable random effects. We now illustrate the three-layer tree case. We define T exchangeable clusters, each comprised of K_t $(t = 1, \ldots, T)$ max-stable processes. Thus, we have a total of $K = \sum_{t=1}^{T} K_t$ spatial processes. For each cluster $t = 1, \ldots, T$ and variable $k = 1, \ldots, K_t$, let $U_{t;k}(\mathbf{s}) \overset{\text{i.i.d.}}{\sim}$ $\exp\{-z^{-1/(\alpha_{t;k}\alpha_t\alpha_0)}\}$, $z > 0$, be a random Fréchet noise process, and let $\vartheta_{t;k}(\mathbf{s}) = \left\{\sum_{l=1}^{L} A_{t;k;l} A_{t;l}^{1/\alpha_{t;k}} A_{0;l}^{1/(\alpha_{t;k}\alpha_t)} \omega_{t;k;l}(\mathbf{s})^{1/(\alpha_{t;k}\alpha_t\alpha_0)}\right\}^{\alpha_{t;k}\alpha_t\alpha_0}$ be a smooth spatial process, where $\alpha_{t;k}, \alpha_t, \alpha_0 \in (0,1]$ are dependence parameters, $\omega_{t;k;l}(\mathbf{s}) \geqslant 0$ are deterministic kernels such that $\sum_{l=1}^{L} \omega_{t;k;l}(\mathbf{s}) = 1$, for any $\mathbf{s} \in \mathcal{S}$, and $A_{t;k;l} \overset{\text{i.i.d.}}{\sim} \mathrm{PS}(\alpha_{t;k}) \perp\!\!\!\perp A_{t;l} \overset{\text{i.i.d.}}{\sim}$ $\mathrm{PS}(\alpha_t) \perp\!\!\!\perp A_{0;l} \overset{\text{i.i.d.}}{\sim} \mathrm{PS}(\alpha_0)$ are independent positive stable random amplitudes. Then, we set $Z_{t;k}(\mathbf{s}) = U_{t;k}(\mathbf{s})\vartheta_{t;k}(\mathbf{s})$, and define the nested multivariate max-stable process with unit Fréchet margins as $\mathbf{Z}(\mathbf{s}) = \{\mathbf{Z}_1^\top(\mathbf{s}), \ldots, \mathbf{Z}_T^\top(\mathbf{s})\}^\top$ with $\mathbf{Z}_t(\mathbf{s}) = \{Z_{t;1}(\mathbf{s}), \ldots, Z_{t;K_t}(\mathbf{s})\}^\top$, $t = 1, \ldots, T$. Analogously to (5), the finite-dimensional distributions of $\mathbf{Z}(\mathbf{s})$ observed at D locations $\mathbf{s}_1, \ldots, \mathbf{s}_D \in \mathcal{S}$ may be written as (2) with the exponent function

$$V(\mathbf{z}_1, \ldots, \mathbf{z}_T) = \sum_{l=1}^{L} \left\{ \sum_{t=1}^{T} \left(\sum_{k=1}^{K_t} \left[\sum_{d=1}^{D} \left\{ \frac{z_{t;k;d}}{\omega_{t;k;l}(\mathbf{s}_d)} \right\}^{-1/(\alpha_{t;k}\alpha_t\alpha_0)} \right]^{\alpha_{t;k}} \right)^{\alpha_t} \right\}^{\alpha_0}, \quad z_{t;k;d} > 0 \quad \text{for all } t, k, d$$

(6)

where $\mathbf{z}_t = (\mathbf{z}_{t;1}^\top, \ldots, \mathbf{z}_{t;K_t}^\top)^\top$ and $\mathbf{z}_{t;k} = (z_{t;k;1}, \ldots, z_{t;k;D})^\top$, $t = 1, \ldots, T$, $k = 1, \ldots, K_t$. The proof of (6) is provided in Web Appendix B. As illustrated in the bottom panel of Figure 1, the multivariate process $\mathbf{Z}(\mathbf{s})$ with (6) may be represented graphically using a tree, whereby the terminal nodes represent the marginal max-stable processes and the upper nodes describe the cross-dependence relationships among the processes. When $T = 1$ and $\alpha_{1;k}\alpha_1 \equiv \alpha_k$, (6) reduces to the two-layer case in (5). Hence, (6) provides more flexibility than (5) for representing complex cross-dependence structures, and we call it a three-layer nested multivariate max-stable process. Using the inverse of the mapping (1), we get a multivariate max-stable process with arbitrary GEV margins and

the hierarchical formulation:

$$Z_{t;k}^{\star}(\mathbf{s}) \mid \{A_{0;l}, A_{t;l}, A_{t;k;l}\}_{l=1}^{L} \stackrel{\text{ind}}{\sim} \text{GEV}\{\mu_{t;k}^{\star}(\mathbf{s}), \sigma_{t;k}^{\star}(\mathbf{s}), \xi_{t;k}^{\star}(\mathbf{s})\}, \tag{7}$$

$$A_{t;k;l} \stackrel{\text{i.i.d.}}{\sim} \text{PS}(\alpha_{t;k}) \perp\!\!\!\perp A_{t;l} \stackrel{\text{i.i.d.}}{\sim} \text{PS}(\alpha_t) \perp\!\!\!\perp A_{0;l} \stackrel{\text{i.i.d.}}{\sim} \text{PS}(\alpha_0), \quad l = 1, \ldots, L,$$

where $\mu_{t;k}^{\star}(\mathbf{s}) = \mu_{t;k}(\mathbf{s}) + \sigma_{t;k}(\mathbf{s})/\xi_{t;k}(\mathbf{s})\{\vartheta_{t;k}(\mathbf{s})^{\xi_{t;k}(\mathbf{s})} - 1\}$, $\sigma_{t;k}^{\star}(\mathbf{s}) = \alpha_0 \alpha_t \alpha_{t;k} \sigma_{t;k}(\mathbf{s}) \vartheta_{t;k}(\mathbf{s})^{\xi_{t;k}(\mathbf{s})}$, and $\xi_{t;k}^{\star}(\mathbf{s}) = \alpha_0 \alpha_t \alpha_{t;k} \xi_{t;k}(\mathbf{s})$. By integrating out the positive stable random effects, the process $Z_{t;k}^{\star}(\mathbf{s})$ is max-stable and has GEV margins with parameters $\mu_{t;k}(\mathbf{s})$, $\sigma_{t;k}(\mathbf{s})$, $\xi_{t;k}(\mathbf{s})$.

3.2 Spatial and cross-dependence properties

In a three-layer multivariate max-stable model (see §3.1), each marginal process $Z_{t;k}(\mathbf{s})$, $t = 1, \ldots, T$, $k = 1, \ldots, K_t$, is a Reich–Shaby process with unit Fréchet margins; therefore, it inherits its spatial dependence properties, which were studied in depth by Reich and Shaby (2012). The product $\alpha_{t;k} \alpha_t \alpha_0$ plays the role of the dependence parameter α in (3) and acts as a mediator between the noise component $U_{t;k}(\mathbf{s})$ and the smooth spatial process $\vartheta_{t;k}(\mathbf{s})$. Below, we describe the cross-dependence properties of our new multivariate spatial model.

Similar to Reich and Shaby (2012), we propose using the Gaussian kernel (4) with a different bandwidth $\tau_{t;k} > 0$ for each marginal process, and the spatial knots $\mathbf{v}_1, \ldots, \mathbf{v}_L \in \mathcal{S}$ are fixed on a regular grid. Again, we rescale the kernels to ensure that they sum to one at each location, i.e., $\omega_{t;k;l}(\mathbf{s}) = g_{t;k;l}(\mathbf{s})\{\sum_{l=1}^{L} g_{t;k;l}(\mathbf{s})\}^{-1}$, $t = 1, \ldots, T$, $k = 1, \ldots, K_t$, with $g_{t;k;l}$ defined similarly to g_l in (4). In practice, we must choose a number of knots L that balances computational feasibility with modeling accuracy. A too small L might not be realistic and could affect subsequent inferences by artificially creating a non-stationary process (Castruccio et al., 2016), whereas a too large L would significantly increase the computational burden. Reich and Shaby (2012) suggested fixing the number of knots, L, such that the grid spacing is approximately equal to or smaller than the kernel bandwidth $\tau_{t;k}$.

To understand the cross-dependence structure of (6) and the meaning of its parameters, we consider the product of the dependence parameters along a specific path through the underlying tree. The amount of noise assigned to the marginal (univariate) Reich–Shaby process in the $(t; k)$th terminal node is governed by $\alpha_{t;k} \alpha_t \alpha_0$; in contrast, $\alpha_t \alpha_0$ controls the cross-dependence among the variables belonging to the same cluster t. The cross-dependence among variables belonging to distinct clusters is controlled by α_0. Since

$\alpha_t \alpha_0 \leqslant \alpha_0$, the intra-cluster cross-dependence is always stronger than the inter-cluster cross-dependence.

For model (6), the pairwise extremal coefficient $\theta\{\mathbf{s}_i, \mathbf{s}_j; (t_1; k_1), (t_2; k_2)\} \in [1, 2]$ (see §2) summarizes the strength of dependence between each pair of variables $\{Z_{t_1;k_1}(\mathbf{s}_i), Z_{t_2;k_2}(\mathbf{s}_j)\}^{\top}$, with $t_1, t_2 = 1, \ldots, T$, $k_1 = 1, \ldots, K_{t_1}$, $k_2 = 1, \ldots, K_{t_2}$. The variables $Z_{t_1;k_1}(\mathbf{s}_i)$ (process k_1 in cluster t_1 observed at location $\mathbf{s}_i$) and $Z_{t_2;k_2}(\mathbf{s}_j)$ (process k_2 in cluster t_2 observed at location $\mathbf{s}_j$) are perfectly dependent when $\theta\{\mathbf{s}_i, \mathbf{s}_j; (t_1; k_1), (t_2; k_2)\} = 1$, and completely independent when $\theta\{\mathbf{s}_i, \mathbf{s}_j; (t_1; k_1), (t_2; k_2)\} = 2$. The dependence strength increases monotonically as the value of the extremal coefficient approaches unity. Writing $\theta\{\mathbf{s}_i, \mathbf{s}_j; (t_1; k_1), (t_2; k_2)\} \equiv \theta(\mathbf{s}_i, \mathbf{s}_j)$ for simplicity, we distinguish three cases from (6):

$$\theta(\mathbf{s}_i, \mathbf{s}_j) = \begin{cases} \sum_{l=1}^{L} \left\{ \omega_{t;k;l}(\mathbf{s}_i)^{1/(\alpha_{t;k}\alpha_t\alpha_0)} + \omega_{t;k;l}(\mathbf{s}_j)^{1/(\alpha_{t;k}\alpha_t\alpha_0)} \right\}^{\alpha_{t;k}\alpha_t\alpha_0}, & t_1 = t_2 = t, \ k_1 = k_2 = k, \\ \sum_{l=1}^{L} \left\{ \omega_{t;k_1;l}(\mathbf{s}_i)^{1/(\alpha_t\alpha_0)} + \omega_{t;k_2;l}(\mathbf{s}_j)^{1/(\alpha_t\alpha_0)} \right\}^{\alpha_t\alpha_0}, & t_1 = t_2 = t, \ k_1 \neq k_2, \\ \sum_{l=1}^{L} \left\{ \omega_{t_1;k_1;l}(\mathbf{s}_i)^{1/\alpha_0} + \omega_{t_2;k_2;l}(\mathbf{s}_j)^{1/\alpha_0} \right\}^{\alpha_0}, & t_1 \neq t_2, \ k_1 \neq k_2. \end{cases}$$
(8)

Hence, as the two sites get closer to each other (i.e., as $\mathbf{s}_i \to \mathbf{s}_j$), the cross-extremal coefficient reduces to $\theta(\mathbf{s}_i, \mathbf{s}_j) \to 2^{\alpha_{t;k}\alpha_t\alpha_0}$ if $t_1 = t_2 = t, k_1 = k_2 = k$ (same cluster, same variable), $\theta(\mathbf{s}_i, \mathbf{s}_j) \to 2^{\alpha_t\alpha_0}$ if $t_1 = t_2 = t, k_1 \neq k_2$ (same cluster, different variable), or $\theta(\mathbf{s}_i, \mathbf{s}_j) \to 2^{\alpha_0}$ if $t_1 \neq t_2, k_1 \neq k_2$ (different cluster, different variable). This clearly confirms that intra-cluster cross-dependence is stronger than inter-cluster cross-dependence. Moreover, the nugget effect is evident when we notice that $2^{\alpha_0} > 1$ for all values of $\alpha_0 \in (0, 1]$. Figure 2 illustrates these pairwise dependence properties and shows realizations of the three-layer nested max-stable model with exponent function (6) and the underlying tree structure displayed in the bottom panel of Figure 1, for specific values of the dependence parameters.

[Figure 2 about here.]

The pairwise extremal coefficients shown in the top row confirm that spatial dependence for each individual process is stronger than cross-dependence, and that intra-cluster cross-dependence is stronger than inter-cluster cross-dependence. The realizations displayed in the bottom row show that strong cross-dependence between two distinct variables may result in co-localized spatial extremes. For example, by comparing the plots of $Z_{1;1}(\boldsymbol{s})$ and $Z_{1;2}(\boldsymbol{s})$, which belong to the same cluster, it can be noticed that spatial extreme events (red/yellow regions) tend to occur in the same area of the figure. This figure appears in color in the electronic version of this article, and color refers to that version.

3.3 *Inference*

Parameter estimation for this type of model may be performed within a Bayesian framework by implementing a standard Metropolis–Hastings Markov chain Monte Carlo (MH-MCMC) algorithm, which takes advantage of the hierarchical formulation (7); see, e.g., Reich and Shaby (2012). In Web Appendix C, we detail the implementation of the MH-MCMC algorithm, which draws approximate samples from the posterior distributions of the dependence parameters $\alpha_{t;k}, \alpha_t, \alpha_0$, and $\tau_{t;k}$, where $t = 1, \ldots, T$, $k = 1, \ldots, K_t$, for the nested max-stable model (6). For simplicity, we assume here that the prior distributions of the parameters $\alpha_{t;k}, \alpha_t$, and α_0 are non-informative Unif$(0, 1)$, and that the range parameters $\tau_{t;k}$ have prior distribution equal to $0.5 h_{\max} \times$ Beta$(2, 5)$ as suggested by Sebille (2016), p. 97, where $h_{\max}$ denotes the maximum distance between stations. This slightly informative prior distribution stabilizes the estimation of the range parameters, whose posterior distribution is usually right-skewed, but it should not have an important impact on the fitted model when the spatial dependence is not too strong. While our model inherits the same computational benefits as the (univariate) Reich–Shaby model, the computational time increases by considering additional variables or adding layers to the latent tree structure.

We conducted a simulation study reported in Web Appendix D, based on the two-layer and three-layer multivariate models, which suggests that the MH-MCMC works correctly and that the parameters can be properly estimated in a reasonable amount of time.

4. Multivariate spatial analysis of air pollution extremes

4.1 *Motivation*

High concentrations of pollution in the air can harm the human body. Current methods for assessing air pollution dangers typically consider each pollutant separately, ignoring the heightened threat of exposure to multiple air pollutants. In order to inform the public and government administrations, the US Environmental Protection Agency (EPA) and other international organizations are moving towards a multi-pollutant approach for quantifying health risks of air pollution. In this work, we investigate the extremal dependence among air pollutants and temperature jointly across space. Here, we use our new methodology based on nested multivariate max-stable processes to characterize the spatial and cross-dependence structures among these variables of interest. In this paper, we focus on the estimation of the dependence structure, while the spatial modeling of margins is carried out separately using the Bayesian hierarchical model proposed by

Davison et al. (2012). More details on the marginal estimation and model assessment are provided in Web Appendix E.

4.2 *Data, model fitting and diagnostics*

Following up on the study in Vettori et al. (2018), we here investigate the extremal dependence between air pollutants and meteorological data jointly across space. We select daily measurements of Carbon Monoxide (CO), Nitric Oxide (NO), Nitrogen Dioxide (NO$_2$), Ozone (O$_3$) and temperature (T) at a number of sites in California from January 2006 to December 2015. Vettori et al. (2018) developed the Tree Mixture MCMC (TM-MCMC) algorithm for investigating plausible multivariate cross-dependence structures for each site separately. This algorithm exploits reversible jump MCMC and the tree-structure of the nested logistic distribution to sample from the posterior distribution of the parameters and the tree itself. In Vettori et al. (2018), the multivariate extremal cross-dependence between air pollutants and meteorological parameters was analyzed separately at each site, thus ignoring spatial dependence. Since the clusters CO-NO and NO$_2$-O$_3$-T appeared to be characterized by stronger dependence, it makes sense to assume two-layer dependence structures, with one single tree summarizing the dependence of CO-NO and another tree summarizing the dependence of NO$_2$-O$_3$-T, over the 21 sites under study. The results of the fits based on the MH-MCMC algorithm are reported in Web Appendix F; see Web Figures S11 and S12.

We now extend this study by fitting a more complex three-layer nested max-stable model to the five variables (CO, NO, NO$_2$, O$_3$, T), but at a smaller number (nine) of sites, see Figure 3, for which stationarity and a single tree structure over space is a reasonable assumption.

[Figure 3 about here.]

To choose the tree structure, we run again the TM-MCMC algorithm of Vettori et al. (2018) (see details therein) to the stationary, standardized, monthly maxima data for all variables separately for each site. The most likely dependence structures identified by the TM-MCMC algorithm and the associated posterior probabilities are represented in Figure 4.

[Figure 4 about here.]

The tree structures that appear more often across the chains are trees A, B and C. In these three trees, the extreme concentrations of CO and NO are grouped together. Moreover, extreme concentrations of NO$_2$ are grouped with extreme concentrations of

O_3 and high temperatures in Tree B and with extreme concentrations of CO and NO in Tree C.

We then proceed and fit the three-layer nested multivariate max-stable model with the tree dependence structures A, B, and C; the posterior medians are reported in Figure 5.

[Figure 5 about here.]

The estimated parameters α_0 take values close to 0.9 in all three trees, indicating that the extreme concentrations of CO and NO are weakly related to the extreme concentrations of NO_2 and O_3 and to high temperatures across the Los Angeles area. The cross-dependence is strongest between the pollutants CO and NO in Tree A and between CO, NO and NO_2 in the case of Tree C, whereas Tree B also identifies the variables NO_2, O_3, and high temperature with fairly strong cross-dependence. Web Figure S13 (in Web Appendix G) compares empirical and model-based estimates of the pairwise extremal coefficients; recall §2 and (8). Overall, model-based estimates are similar to their empirical counterparts (taking the high variability of empirical estimates into account), which suggests that the fitted model is reasonable and adequately captures the complex spatial cross-dependence structure of extremes in our dataset. As expected, the pairwise extremal coefficient estimates computed for individual variables seem to increase with the distance between sites. Moreover, both the empirical and model-based pairwise cross-extremal coefficient estimates indicate a moderate dependence strength between variables belonging to the same cluster, such as CO and NO, O_3 and NO_2, or O_3 and T, regardless of the distance between sites, whereas the variables belonging to different clusters, such as O_3 and CO or O_3 and NO, are weakly dependent at any distances.

To verify that the nested multivariate max-stable model provides a good marginal fit for each of the processes under study, Web Figure S14 (in Web Appendix G) compares the MCMC output obtained from the joint fit of our three-layer model to the posterior medians of the dependence parameters α and τ obtained from the (univariate) Reich–Shaby model fitted to each process independently. Overall, the joint and individual models provide similar values for the marginal parameters, confirming that our approach yields sensible marginal fits, while simultaneously providing information about the cross-dependence structure. However, the processes characterized by a large dependence parameter α are quite noisy, and thus harder to estimate. The computational time for the three-layer model fits was less than a few hours with $R = 50000$ MCMC iterations.

Plots and results are produced after removing an initial burn-in of $R/5$ iterations, and thinning the Markov chain by a factor 500.

4.3 *Return level projections and air pollution risk assessment*

The US EPA typically uses the Air Quality Index (AQI) to communicate air pollution risks to the public. In order to illustrate the impacts of neglecting spatial and cross-dependence structures on the AQI return level estimates, Figure 6 shows high p-quantiles with probabilities ranging from $p = 0.5$ to $p = 0.996$, considering April 2009 as the baseline, computed for the maximum AQI over the pollutants CO, O_3, and NO_2 and across space, using different dependence models (but the same marginal model).

[Figure 6 about here.]

Under stationary conditions, the return levels for the return periods of 1 and 20 years roughly correspond to $p = 1 - 1/12 \approx 0.917$ and $p = 1 - 1/(12 \times 20) \approx 0.996$, respectively. The AQI categories, representing different levels of health concern, are represented by different colors. We compare the fits of the full nested multivariate max-stable model, the Reich–Shaby model fitted to each individual process separately (ignoring cross-dependence), the nested logistic distribution fitted at each site separately using the TM-MCMC algorithm (ignoring spatial dependence), and the GEV distribution fitted to each site and pollutant independently (ignoring both spatial and cross-dependence). The AQI quantiles obtained from the posterior predictive distribution of the Reich–Shaby fits are much smaller than the ones obtained from the multivariate distribution fitted to each site separately or the GEV distribution fitted to each site and pollutants separately. Furthermore, the high quantile projections based on the nested multivariate max-stable model are generally smaller than the high quantiles based on the Reich–Shaby model. Therefore, when neglecting spatial dependence or the multivariate cross-dependence among processes, the return levels calculated for the maximum of several extreme observations might be strongly overestimated. Such a large difference in return levels for spatial risk measures was also similarly noticed by Huser and Genton (2016) when they explored the effect of model misspecification in non-stationary max-stable processes. Using our proposed multivariate max-stable process based on Tree A, the high quantiles z_p (for the maximum AQI across all sites and pollutants) lie within the very unhealthy category for probabilities $p \geqslant 0.92$, indicating that at least one of the criteria pollutants under study exceeds this critical threshold at one or more of the monitoring sites approximately once every year. We obtained similar results from trees B and C.

5. Conclusion

We introduced a novel class of hierarchical multivariate max-stable processes that have the Reich–Shaby model as univariate margins, and that can capture the spatial and cross-dependence structures among extremes of multiple variables, based on latent nested positive stable random effects. These hierarchical models may be conveniently represented by a tree structure, and the complexity of the dependence relations among the various spatial variables might be increased by adding an arbitrary number of nesting layers. Parameter estimation can be carried out within a Bayesian framework using a standard Metropolis–Hastings MCMC algorithm. As shown in our simulation experiments, the dependence parameters governing the spatial dependence of individual variables and the cross-dependence among different variables can be satisfactorily identified using our proposed algorithm.

We fitted the nested multivariate max-stable process to air pollution extremes collected in the Los Angeles area. In addition to providing good spatial marginal fits for each of the air pollutants under study, our model detects their extremal spatial cross-dependence, and takes into account temperature extremes. Extreme concentrations of toxic air pollutants, such as CO, NO, NO_2, and O_3 and high temperatures are weakly related across the area of Los Angeles. Furthermore, a strong cross-dependence is detected between the maxima of CO and NO, which are both pollutants released by fossil fuel combustion. Also, high concentrations of O_3 and high temperatures often occur simultaneously, which leads to a heightened health threat according to Kahle et al. (2015). Modeling air pollution extremes using the proposed nested multivariate max-stable model allows us to provide sensible multi-pollutant return level estimates based on the Air Quality Index (AQI); thus, our new methodology is useful for assessing the risks associated with simultaneous exposure to several air pollutants over space, and might be used to develop future air pollution monitoring regulations.

In order to fit the nested multivariate max-stable process, we must assume a single fixed tree structure across space. To generalize our approach to spatially-varying tree structures, one could define a partition of space with homogeneous sub-regions governed by possibly different cross-dependence tree structures. For more flexibility, this partition could also be treated as random, similarly to Reich and Shaby (2018a).

Another open problem is the modeling and characterization of multivariate spatial extremes defined as high threshold exceedances. While Eastoe and Tawn (2009) discussed the modeling of non-stationary extreme ozone data in the univariate context, and Thibaud and Opitz (2015) and de Fondeville and Davison (2018) showed how to

model spatial threshold exceedances using suitable risk functionals, it is less clear how to properly capture the joint spatial and cross-dependence structures of *multivariate* high-threshold exceedances.

References

Casson, E. and Coles, S. G. (1999). Spatial regression models for extremes. *Extremes* **1,** 449–468.

Castruccio, S., Huser, R., and Genton, M. G. (2016). High-order composite likelihood inference for max-stable distributions and processes. *Journal of Computational and Graphical Statistics* **25,** 1212–1229.

Coles, S. G. (2001). *An Introduction to Statistical Modeling of Extreme Values.* Springer, London.

Cooley, D., Nychka, D., and Naveau, P. (2007). Bayesian spatial modeling of extreme precipitation return levels. *Journal of the American Statistical Association* **102,** 824–840.

Davison, A. C., Huser, R., and Thibaud, E. (2019). Spatial Extremes. In Gelfand, A. E., Fuentes, M., Hoeting, J. A., and Smith, R. L., editors, *Handbook of Environmental and Ecological Statistics.* CRC Press.

Davison, A. C., Padoan, S. A., and Ribatet, M. (2012). Statistical modeling of spatial extremes. *Statistical Science* **27,** 161–186.

de Fondeville, R. and Davison, A. C. (2018). High-dimensional peaks-over-threshold inference. *Biometrika* **105,** 575–592.

de Haan, L. (1984). A spectral representation for max-stable processes. *The Annals of Probability* **12,** 1194–1204.

Dominici, F., Peng, R. D., Barr, C. D., and Bell, M. L. (2010). Protecting human health

from air pollution: shifting from a single-pollutant to a multi-pollutant approach. *Epidemiology* **21,** 187–194.

Eastoe, E. F. and Tawn, J. A. (2009). Modelling non-stationary extremes with application to surface level ozone. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **58,** 25–45.

Genton, M. G., Padoan, S. A., and Sang, H. (2015). Multivariate max-stable spatial processes. *Biometrika* **102,** 215–230.

Huser, R. and Davison, A. (2013). Composite likelihood estimation for the Brown-Resnick process. *Biometrika* **100,** 511–518.

Huser, R., Dombry, C., Ribatet, M., and Genton, M. G. (2019). Full likelihood inference for max-stable data. *Stat* **8,** e218.

Huser, R. and Genton, M. G. (2016). Non-stationary dependence structures for spatial extremes. *Journal of Agricultural, Biological, and Environmental Statistics* **21,** 470–491.

Huser, R., Opitz, T., and Thibaud, E. (2017). Bridging asymptotic independence and dependence in spatial extremes using Gaussian scale mixtures. *Spatial Statistics.* **21,** 166–186.

Huser, R. and Wadsworth, J. L. (2019). Modeling spatial processes with unknown extremal dependence class. *Journal of the American Statistical Association.* To appear.

Kabluchko, Z., Schlather, M., and de Haan, L. (2009). Stationary max-stable fields associated to negative definite functions. *Annals of Probability* **37,** 2042–2065.

Kahle, J. J., Neas, L. M., Devlin, R. B., Case, M. W., Schmitt, M. T., Madden, M. C., and Diaz-Sanchez, D. (2015). Interaction effects of temperature and ozone on lung function and markers of systemic inflammation, coagulation, and fibrinolysis: A crossover study of healthy young volunteers. *Environmental Health Perspectives* **123,** 310–316.

Krupskii, P., Huser, R., and Genton, M. G. (2018). Factor copula models for replicated spatial data. *Journal of the American Statistical Association* **113,** 467–479.

Oesting, M., Schlather, M., and Friederichs, P. (2017). Statistical post-processing of forecasts for extremes using bivariate Brown–Resnick processes with an application to wind gusts. *Extremes* **20,** 309–332.

Opitz, T. (2013). Extremal t processes: Elliptical domain of attraction and a spectral representation. *Journal of Multivariate Analysis* **122,** 409–413.

Opitz, T. (2016). Modeling asymptotically independent spatial extremes based on

Laplace random fields. *Spatial Statistics* **16,** 1–18.

Opitz, T., Huser, R., Bakka, H., and Rue, H. (2018). INLA goes extreme: Bayesian tail regression for the estimation of high spatio-temporal quantiles. *Extremes* **21,** 441–462.

Reich, B. J. and Shaby, B. A. (2012). A hierarchical max-stable spatial model for extreme precipitation. *Annals of Applied Statistics* **6,** 1430–1451.

Reich, B. J. and Shaby, B. A. (2018a). A spatial Markov model for climate extremes. *Journal of Computational and Graphical Statistics.* To appear.

Reich, B. J. and Shaby, B. A. (2018b). Modeling of multivariate spatial extremes. *RESEARCHERS.ONE* https://www.researchers.one/article/2018-09-12. https://www.researchers.one/article/2018-09-12.

Ribatet, M., Cooley, D., and Davison, A. C. (2012). Bayesian inference from composite likelihoods, with an application to spatial extremes. *Statistica Sinica* **22,** 813–845.

Schlather, M. (2002). Models for stationary max-stable random fields. *Extremes* **5,** 33–34.

Sebille, Q. (2016). *Modélisation spatiale de valeurs extrêmes: application à l'étude de précipitations en France.* PhD thesis, Université Claude Bernard Lyon 1, Lyon, France.

Smith, R. L. (1990). Max-stable processes and spatial extremes. Unpublished manuscript. University of North Carolina, Chapel Hill, U.S.

Tawn, J. A. (1990). Modelling multivariate extreme value distributions. *Biometrika* **77,** 245–253.

Thibaud, E., J. Aalto, D. S. C., Davison, A. C., and Heikkinen, J. (2016). Bayesian inference for the Brown–Resnick process, with an application to extreme low temperatures. *Annals of Applied Statistics.* **10,** 2303–2324.

Thibaud, E. and Opitz, T. (2015). Efficient inference and simulation for elliptical Pareto processes. *Biometrika* **102,** 855–870.

Vettori, S., Huser, R., Segers, J., and Genton, M. G. (2018). Bayesian model averaging over tree-based dependence structures for multivariate extremes. arXiv preprint 1705.10488.

Wadsworth, J. L. and Tawn, J. A. (2012). Dependence modelling for spatial extremes. *Biometrika* **99,** 253–272.

Xu, G. and Genton, M. G. (2017). Tukey max-stable processes for spatial extremes. *Spatial Statistics* **18,** 431–443.
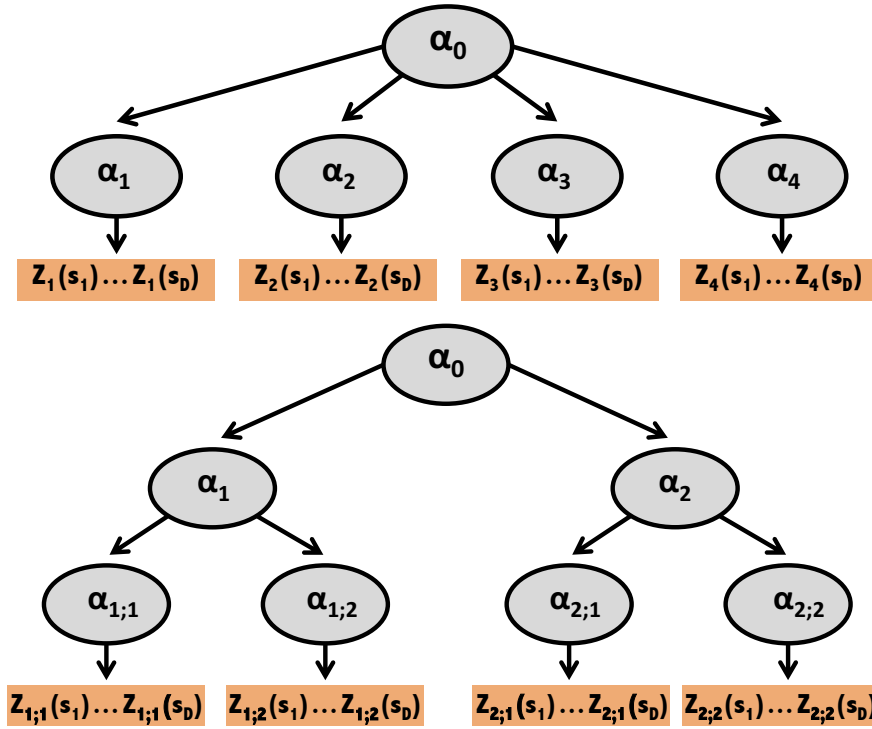
Figure 1. Example of simple two-layer (top) and three-layer (bottom) tree structures, summarizing the extremal dependence of a four-dimensional process written as $\mathbf{Z}(\mathbf{s}) = \{Z_1(\mathbf{s}), Z_2(\mathbf{s}), Z_3(\mathbf{s}), Z_4(\mathbf{s})\}^\top$ (top) and $\mathbf{Z}(\mathbf{s}) = \{Z_{1;1}(\mathbf{s}), Z_{1;2}(\mathbf{s}), Z_{2;1}(\mathbf{s}), Z_{2;2}(\mathbf{s})\}^\top$ (bottom). The cross-dependence strength among variables is controlled by the product of dependent parameters assigned the nodes that they share; e.g., in the three-layer tree (bottom), the intra-cluster cross-dependence, i.e., between processes $Z_{t;1}(\mathbf{s})$ and $Z_{t;2}(\mathbf{s})$, is summarized by the product $\alpha_t \alpha_0$, whereas the inter-cluster cross-dependence, i.e., between the variables $Z_{t_1;k_1}(\mathbf{s}), Z_{t_2;k_2}(\mathbf{s})$, with $t_1 \neq t_2$ and $k_1, k_2 = 1, 2$, is summarized by the parameter α_0. The number of latent α-stable random variables involved in the model is equal to the number of upper tree nodes (excluding the terminal nodes) multiplied by the number of basis functions, L; here, there are $5L$ (top) and $7L$ (bottom) latent variables. This figure appears in color in the electronic version of this article.
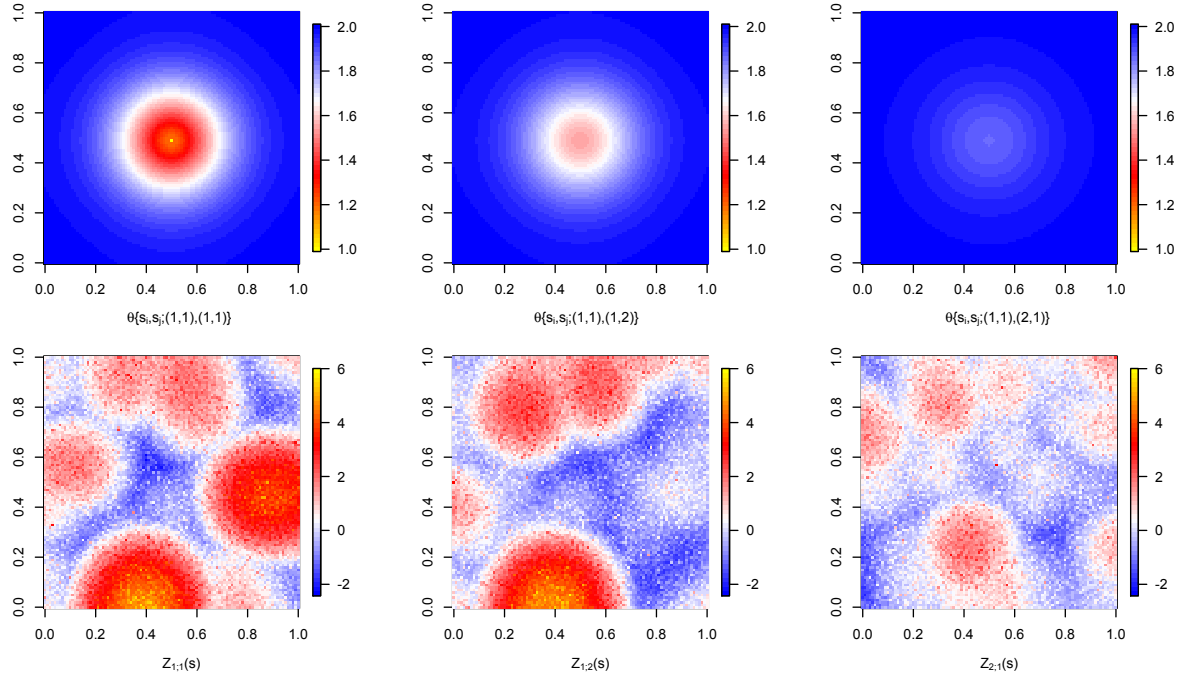
Figure 2. *Top row:* Pairwise extremal coefficient $\theta\{\mathbf{s}_i, \mathbf{s}_j; (t_1, k_1), (t_2, k_2)\}$, see (8), for the three-layer multivariate max-stable model with exponent function (6) and underlying tree structure displayed in the bottom panel of Figure 1, for fixed reference location $\mathbf{s}_i = (0.5, 0.5)^\top$ and $\mathbf{s}_j \in [0,1]^2$. Here, $\alpha_0 = 0.9$, $\alpha_1 = \alpha_2 = 0.7$ and $\alpha_{1;1} = \alpha_{1;2} = \alpha_{2;1} = \alpha_{2;2} = 0.4$, while the kernels are Gaussian densities as in (3) with bandwidths $\tau_{1;1} = \tau_{1;2} = \tau_{2;1} = \tau_{2;2} = 0.1$, with knots taken on a 100×100 regular grid. The panels summarize the spatial dependence of each individual process (left), the intra-cluster cross-dependence (middle) and the inter-cluster cross-dependence (right). *Bottom row:* Realizations of $Z_{1;1}(\mathbf{s})$ (left), $Z_{1;2}(\mathbf{s})$ (middle) and $Z_{2;1}(\mathbf{s})$ (right). The color scale on the right-hand side of each bottom panel indicates the level of each process simulated on the standard Gumbel scale. This figure appears in color in the electronic version of this article, and color refers to that version.

Figure 3. Study region around Los Angeles, with the nine selected sites indicated by numbers. This figure appears in color in the electronic version of this article.
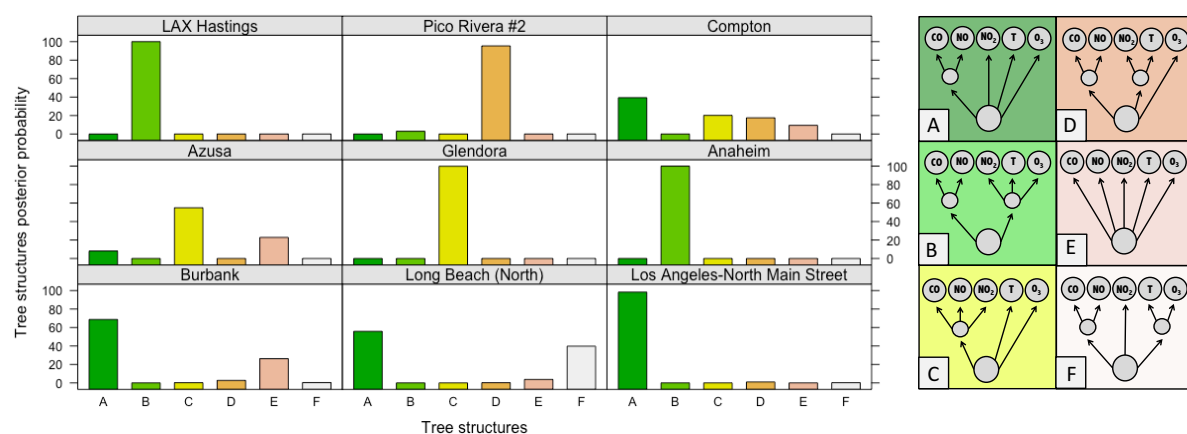
Figure 4. The most frequent tree structures (right), indicated by letters A–F, identified by the TM-MCMC algorithm of Vettori et al. (2018) after $R = 5000$ iterations, burn-in $R/4$ and thinning factor 5. The histograms (left) report the posterior probability associated with each tree for each site in Figure 3 calculated as the number of times each tree appears in the algorithm chain. This figure appears in color in the electronic version of this article.
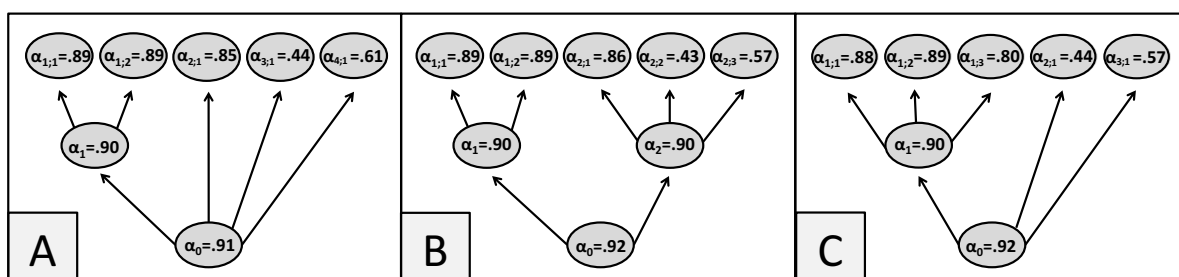
Figure 5. Posterior medians of the dependence parameters $\alpha_{t;k}, \alpha_t, \alpha_0$ obtained by fitting our new three-layer nested multivariate max-stable model (6) to the concentration maxima of CO, NO, NO_2, O_3, and temperature using the MCMC algorithm and assuming the tree structures A, B, or C from Figure 4.
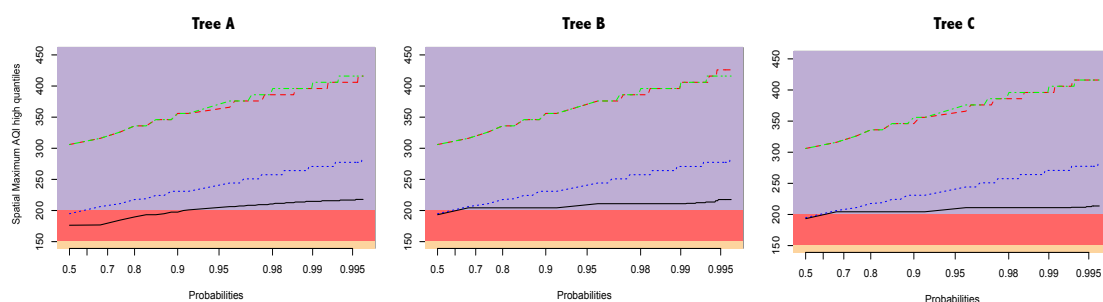
Figure 6.   High p-quantiles z_p computed for the spatial maximum of the largest Air Quality Index (AQI) for CO, O_3, and NO_2, setting April 2009 as baseline, obtained by fitting the nested multivariate max-stable model (black lines); by fitting the Reich–Shaby model to each individual process separately, neglecting cross-dependence (blue lines); by running the TM-MCMC algorithm to estimate the cross-dependence between variables at each site separately, neglecting spatial dependence (red lines); and by fitting the GEV distribution to each site and pollutants independently, neglecting spatial and cross-dependence structures (green lines). Results are shown underlying trees A (left), B (middle), and C (right). Probabilities are displayed on a Gumbel scale, i.e., z_p is plotted against $-\log\{-\log(p)\}$. AQI categories: 0-50 satisfactory; 51-100 acceptable; 101-150 unhealthy for sensitive groups (orange); 151-200 unhealthy (red); >200 very unhealthy (purple). This figure appears in color in the electronic version of this article, and color refers to that version.