

# MODELING NON-STATIONARY TEMPERATURE MAXIMA BASED ON EXTREMAL DEPENDENCE CHANGING WITH EVENT MAGNITUDE

BY PENG ZHONG<sup>\*</sup>, RAPHAËL HUSER<sup>\*</sup> AND THOMAS OPITZ<sup>†</sup>

*Computer, Electrical and Mathematical Sciences and Engineering (CEMSE) Division,  
King Abdullah University of Science and Technology (KAUST)<sup>\*</sup>  
peng.zhong@kaust.edu.sa; raphael.huser@kaust.edu.sa*

*BioSP, INRAE, Avignon, 84914, France<sup>†</sup>  
thomas.opitz@inrae.fr*

## *Abstract*

The modeling of spatio-temporal trends in temperature extremes can help better understand the structure and frequency of heatwaves in a changing climate, and assess the environmental, societal, economic and health-related risks they entail. Here, we study annual temperature maxima over Southern Europe using a century-spanning dataset observed at 44 monitoring stations. Extending the spectral representation of max-stable processes, our modeling framework relies on a novel construction of max-infinitely divisible processes, which include covariates to capture spatio-temporal non-stationarities. Our new model keeps a popular max-stable process on the boundary of the parameter space, while flexibly capturing weakening extremal dependence at increasing quantile levels and asymptotic independence. This is achieved by linking the overall magnitude of a spatial event to its spatial correlation range, in such a way that more extreme events become less spatially dependent, thus more localized. Our model reveals salient features of the spatio-temporal variability of European temperature extremes, and it clearly outperforms natural alternative models. Results show that the spatial extent of heatwaves is smaller for more severe events at higher elevations, and that recent heatwaves are moderately wider. Our probabilistic assessment of the 2019 annual maxima confirms the severity of the 2019 heatwaves both spatially and at individual sites, especially when compared to climatic conditions prevailing in 1950–1975. Our results could be exploited in practice to understand the spatio-temporal dynamics, severity, and frequency of extreme heatwaves, and to design suitable region-specific mitigation measures.

**1. Introduction.** In the current era of climate change and ecological transitions, environmental risks such as heatwaves or floods are major threats that our society faces more than ever. High surface air temperature has indeed been shown to directly contribute to the increase in human mortality rate during the hot season (Ballester et al., 1997; Mitchell et al., 2016; Brown, 2016). Moreover, major ecosystem disruptions can be triggered in biodiversity-rich regions such as the Mediterranean basin or the Alps (Theurillat and Guisan, 2001; Klausmeyer and Shaw, 2009; Gordo and Sanz, 2010; Grabherr, Gottfried and Pauli, 2010). Available data are becoming increasingly rich and allow us to develop and implement mathematically sound statistical models to assess the risk associated with spatial environmental extreme events. There is now very broad scientific consensus that global warming is a fact, and climate science research strongly suggests that the nature and magnitude of extreme events undergo a strong evolution due to global change (Field et al., 2012). However, the regional responses may be quite varied, and it is of crucial importance to quantify region-specific

---

*Keywords and phrases:* Asymptotic dependence and independence, Max-infinitely divisible process, Max-stable process, Spatial extreme event, Spatio-temporal modeling

risks. We therefore need appropriate statistical models that shed light into the mechanisms leading to extreme episodes in environmental variables, and that can accurately describe their spatio-temporal dependence and variability.

The general scientific question that we aim to address in this work is to quantify the impact that climate change has had on European heatwaves in terms of their frequency, severity and spatial extent. Understanding the behavior of extreme temperature events over space and time is key for understanding the (past, present, and potential future) risk that heatwaves cause to society, assess their real impacts, and designing adequate regional mitigation measures of future extreme events. This point is well illustrated by the 2019 summer, during which two temporally distinct heatwaves occurred (one in June and one in July), severely affecting millions of people in major parts of Europe. As temperatures reached unprecedented levels over a wide area, it is natural to investigate from a risk assessment and mitigation standpoint whether such spatial extreme events could have been anticipated from historical records, both in terms of their severity and spatial extent. In the literature, some papers claim that heatwaves in different parts of the world are generally becoming wider due to climate change (see, e.g., [Osborn and Briffa, 2006](#); [Lyon et al., 2019](#)). However, these studies often do not use actual temperature measurements, but they rather rely on temperature-related proxy records or gridded climate model outputs, which may potentially be subject to strong biases due to lack of realism of models with respect to certain aspects (see [Jun, Knutti and Nychka, 2008](#), for instance). Moreover, the notion of “spatial extent” in these papers is often defined as the number of grid cells or sites, at which a certain climate index exceeds a high *marginal* level, while disregarding their actual *spatial dependence*. In this paper, we exploit real observations from 44 monitoring stations over the southern part of Europe from 1918 to 2018, in order to study the spatial dependence range and the severity of non-stationary temperature maxima, viewed as irregularly-spaced realizations from a spatial stochastic process. Specifically, we develop new statistical methodology based on extreme-value theory to understand whether the (local) spatial extent of heatwaves varies regionally and temporally. We then specifically analyze the 2019 European heatwaves and use our new modeling framework to understand whether such extreme events were particularly abnormal in the light of historical data.

In contrast to traditional statistical models that are appropriate for capturing the “average” behavior of such phenomena, spatial extreme-value models focus on modeling the joint tails of spatial processes. In this context, max-stable processes have played a central role, being the only possible non-degenerate limits of linearly rescaled pointwise maxima of random processes ([Davison, Padoan and Ribatet, 2012](#); [Davison and Huser, 2015](#); [Davison, Huser and Thibaud, 2019](#)). In practice, max-stable processes are commonly fitted to block maxima, which are often based on annual blocks ([Davison and Gholamrezaee, 2012](#)). By contrast to approaches based on high threshold exceedances ([Davison and Smith, 1990](#)), the block maximum approach focuses on the long-term behavior of extremes, avoiding the intricate treatment of seasonality and short-term temporal dependence. Moreover, it does not rely on the (sometimes arbitrary) choice of a threshold discriminating extremes from the bulk, which can be awkward under spatio-temporal non-stationarity ([Scarrott and MacDonald, 2012](#)). Popular choices for parametric max-stable models include the Brown–Resnick model ([Brown and Resnick, 1977](#); [Kabluchko, Schlather and de Haan, 2009](#)), or the extremal- $t$  model ([Opitz, 2013](#)), which comprises the Schlather model ([Schlather, 2002](#)) as a special case and the Brown–Resnick model as a limiting case. However, max-stable processes  $\{Z(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  always display a property known as *asymptotic dependence* (except in the trivial case of full independence), which means that the limit  $\chi = \lim_{u \rightarrow 1} \Pr\{Z(\mathbf{s}_1) > G_1^{-1}(u) \mid Z(\mathbf{s}_2) > G_2^{-1}(u)\}$ ,  $\mathbf{s}_1, \mathbf{s}_2 \in \mathcal{S}$ , where  $G_1$  and  $G_2$  denote the cumulative distribution function (CDF) of  $Z(\mathbf{s}_1)$  and  $Z(\mathbf{s}_2)$ , respectively, exists and is positive ( $\chi > 0$ ). This positive value implies that max-stable processes can only capture strong tail dependence, and are inappropriate when maxima—or

the original data from which maxima are extracted—are asymptotically independent ( $\chi = 0$ ), which corresponds to the situation where the extremal dependence strength eventually weakens and completely vanishes as the quantile level increases (i.e., as  $u \rightarrow 1$ ). Max-stability is in fact a strong theoretical property that arises asymptotically when considering blocks of increasing size, and which largely restricts the flexibility of extreme-value models derived from it. As the block size is often chosen to be one year (or less) in real data applications, *imposing* max-stability is often an overly restrictive simplification, which yields an artificially strong extremal dependence structure. This model misspecification is problematic, as it potentially leads to a significant overestimation of joint tail probabilities and thus impacts risk assessment of spatial extreme events, as well as the estimation of their spatial extent. However, while there is a wide body of literature developing peaks-over-threshold models for asymptotic independence or hybrid models bridging the two asymptotic dependence regimes (Wadsworth and Tawn, 2012; Opitz, 2016; Huser, Opitz and Thibaud, 2017; Huser and Wadsworth, 2019; Shooter et al., 2019; Wadsworth and Tawn, 2019), there are only a few papers so far where this problem has been rigorously tackled for block maxima data; see Bopp, Shaby and Huser (2020) and Huser, Opitz and Thibaud (2021). It is indeed difficult to develop principled sub-asymptotic models for block maxima, which reasonably depart from limiting max-stable processes, while keeping certain properties that reflect the specific type of positive dependence of maxima; see also the review paper Huser and Wadsworth (2020).

In this paper, we build upon Bopp, Shaby and Huser (2020) and Huser, Opitz and Thibaud (2021), and develop flexible spatial models that pertain to the wider class of max-infinitely divisible (max-id) processes. Max-id processes naturally extend max-stable processes and relax their rigid dependence structure. While the theory behind max-id processes has been well established for decades (Resnick, 1987; Giné, Hahn and Vatan, 1990; Dombry and Eyi-Minko, 2013), Padoan (2013) was the first to propose a max-id model that has a magnitude-dependent extremal dependence structure. This parametric model stems from taking the limit of block maxima over independent and identically distributed (i.i.d.) Gaussian process ratios, with correlation strength increasing to one as the block size tends to infinity. However, while this model captures asymptotic independence, it has quite an artificial construction and is rather inflexible in its ability to capture weakening but strong spatial dependence. In fact, it was found to be the worst-performing model fitted in the application of Huser, Opitz and Thibaud (2021), and was even largely outperformed by a relatively simple Gaussian process. More importantly, this model does not have a max-stable model as a special case, which makes it unsuitable for maxima defined over moderate or large blocks. Alternatively, Huser, Opitz and Thibaud (2021) proposed general construction principles for building relatively flexible max-id models that remain in the “neighborhood” of the extremal- $t$  max-stable process. In particular, they adapted the spectral representation of max-stable processes to construct flexible max-id models that have a smooth transition between asymptotic dependence classes on the boundary of the parameter space. However, the dependence structure of those max-id models remains quite rigid for describing the central part of the distribution of component-wise maxima. Alternatively, Bopp, Shaby and Huser (2020) recently developed a Bayesian hierarchical max-id model that scales well with large datasets and keeps the Reich and Shaby (2012) max-stable model as a special case, but whose tail properties are less flexible than the models proposed by Huser, Opitz and Thibaud (2021). In this paper, we extend the max-id models of Huser, Opitz and Thibaud (2021) even further, in order to retain their appealing tail dependence properties and gain significant flexibility in the bulk of the max-id distribution with just one additional parameter. The novel approach that we develop here is to construct max-id processes by taking maxima over random fields whose spatial correlation range depends on a random variable representing the overall event magnitude. Furthermore, the max-id models of Padoan (2013), Huser, Opitz and Thibaud (2021) and Bopp, Shaby and



FIGURE 1. Study region, with the 44 monitoring stations (dots) distributed over Europe. The color of each dot represents the empirical quantile level of the annual maximum temperature observed in 2019, calculated from the sample of annual temperature maxima available at each site from 1918–2018, with yellow corresponding approximately to the median and dark red corresponding to higher quantiles. White dots show stations where no data were available in 2019.

Huser (2020) have stationary and isotropic dependence structures, both in space and time, which is not realistic when modeling environmental data (e.g., temperature) over relatively large areas and long time periods. In this paper, we develop non-stationary max-id models that have a rather parsimonious construction and include spatial and temporal covariates in their dependence structure, in order to flexibly capture spatio-temporal variations.

The remainder of the paper is organized as follows. In §2, we introduce the dataset used in our application, and an exploratory analysis illustrates the main applied scientific questions that we aim to address in this paper. In §3, we develop our general modeling framework. More precisely, after giving some background theory about max-id processes, we build a new non-stationary max-id model with a spatial dependence structure that varies according to the three dimensions of (i) space, (ii) time, and (iii) event magnitude. In §4, we develop our inference approach based on a pairwise likelihood, and demonstrate its good performance with a simulation study. In §5, we further detail our parametric max-id models, and we detail a substantial application of these models to study European temperature extremes, in order to assess the risk of spatial extreme temperatures, such as the 2019 European Heatwaves. Concluding remarks are enclosed in §6.

**2. Dataset and explanatory analysis.** In our study, we use the dataset of Klein Tank et al. (2002), from which we extract annual maximum temperatures for the period 1918–2018 at  $D = 44$  (irregularly-spaced) monitoring stations in Europe covering a belt between latitudes  $40^\circ$  and  $50^\circ$  from Western to Eastern Europe, with the mountain range of the Alps in its central part. The distance between stations, calculated using the great-circle metric (i.e., geodesic distance), ranges between 22.6km and 2227.42km, while elevations vary moderately from 48m (Debrecen airport, Hungary) to 893m (Hohenpeissenberg, Germany) above mean sea level. The study region and the monitoring stations are displayed in Figure 1. This dataset contains 22 stations with complete records (i.e., without any missing data over the whole century-long study period), but overall, missing values account for about 14.7% of observations across the 44 stations. Europe is climatologically very diverse: while Western

Europe is mostly characterized by an oceanic climate, Southern Europe has a Mediterranean climate, and Central-Eastern Europe has a continental climate. Therefore, when modeling temperature data over such a large region, it is important to account for such regional variations that affect marginal distributions, and potentially also the spatial dependence structure. Here, the only available covariates that we have at our disposal are the geographical information about the monitoring stations including longitude, latitude and elevation.

During the summer in 2019, major parts of Europe were hit hard by two consecutive heatwaves in late June and late July, with new temperature records observed at many locations. The color of the dots plotted in Figure 1 represents the severity of the 2019 temperature maxima at all the stations, compared to the sample of maxima from the years 1918–2018. For many stations, the 2019 temperature maxima were very extreme, and in some cases more extreme than any past observations. This raises the natural question of how likely the temperature levels reached during the 2019 summer were compared to historical observations. A major applied objective of this work is thus to build a flexible non-stationary statistical model that reliably quantifies the severity of the 2019 heatwaves, in terms of the return period of marginal and spatial events. Such an assessment requires a specialized extreme-value model that properly accounts for spatio-temporal trends and dependence. In §5.1, we detail how we flexibly model margins using a semi-parametric model that uses a penalized tensor product of splines in terms of covariates (longitude, latitude, elevation, and time) to capture complex local temperature variations. In §5.2, we further detail our non-stationary modeling of the dependence structure to capture local patterns in spatial extreme events, while in §5.3, we precisely assess the severity of the 2019 heatwaves.

To further explore the data's temporal and spatial properties, Figure 2 displays the daily maximum temperatures for the months of June, July and August from 2013 to 2019 at the 23 stations with the most complete temperature records. The red dots represent the annual temperature maxima that are used in our analysis. From these data plotted over a much shorter period, the time series appear roughly stationary from year to year and it is (visually) hard to detect any time trend. However, as mentioned above, we shall nevertheless include a flexible time trend component in our spatial extremes model (see §§5.1–5.2), both in the margins and the dependence structure, in order to precisely quantify the effect of climate change on the local distribution of temperatures, which should be evident when considering our entire century-long dataset. Furthermore, from Figure 2, it appears that annual maxima often occur around the same time of year at most stations, even for stations that are far away from each other. Within the same country, annual maxima even often occur on the exact same day. This is an indication that spatial dependence among temperature extremes over Europe is quite strong, and that heatwaves are large-scale natural processes that typically affect a wide area simultaneously. However, from these simple exploratory plots, it is unclear what the actual size of heatwaves is, and whether the spatial extent of heatwaves changes with time, or whether certain regions are more prone than others to spatially widespread extreme temperature events. [Osborn and Briffa \(2006\)](#) and [Lyon et al. \(2019\)](#) studied the spatial extent of heatwaves in different parts of the world in terms of their affected areas using indirect and potentially biased data (i.e., temperature-related proxy records or gridded climate model outputs, respectively). They claimed that the spatial extent of heatwaves has been increasing and might further increase in the future. Another major applied objective of our work is to verify this claim based on real temperature measurements and sound statistical methodology for the spatial and temporal window considered in this study, and to precisely quantify by how much the spatial extent of heatwaves might have increased—if it has. However, since our dataset comprises real observations at irregularly-spaced monitoring stations, directly measuring the size of heatwaves is not feasible. Therefore, we here instead define the spatial extent of heatwaves through the effective (extremal) dependence range of the continuous-space stochastic process fitted to extreme temperatures. This definition intuitively describes

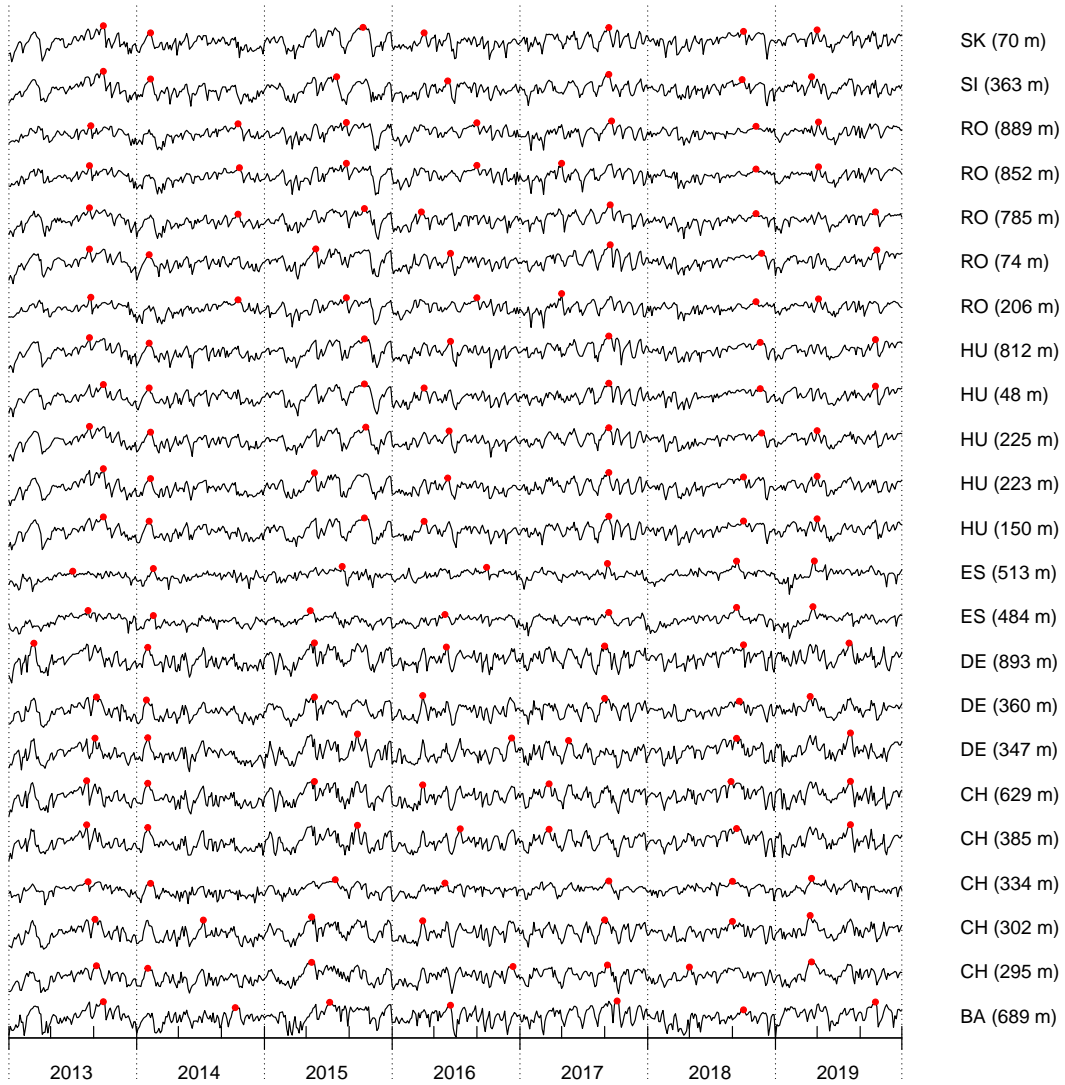


FIGURE 2. Daily temperature maxima (solid black curves) for the months of June, July and August during the period 2013–2019 at the 23 stations with the most complete records. Annual temperature maxima are indicated by red dots. The corresponding country code (in alphabetical order from bottom to top) and elevation of each station are provided on the right-hand side. In our data analysis, we used only the annual maxima until 2018 for fitting, and left the 2019 maximum out for risk assessment.

the spatial extent through the “expected size” of individual extreme events. To the best of our knowledge, our approach to measure the spatial extent of heatwaves and its potential temporal changes from irregularly-spaced data, based on a rigorous spatial extreme-value theory framework, is novel in the literature; see §5 for details on the application and the results of our study.

### 3. Modeling based on max-infinitely divisible processes.

3.1. *Max-infinitely divisible processes.* A random process  $\{Z(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  is called max-infinitely divisible (max-id) if, for any finite collection of sites  $\mathcal{D} = \{\mathbf{s}_1, \dots, \mathbf{s}_D\} \subset \mathcal{S}$ , the joint distribution  $G$  of the random vector  $\{Z(\mathbf{s}_1), \dots, Z(\mathbf{s}_D)\}^T$  is such that  $G^t$  defines a valid CDF for any positive real  $t > 0$ . While this property is always true in the univariate case

( $D = 1$ ) or when  $t$  is a positive integer, it may not be true for  $D \geq 2$  with non-integer, e.g., fractional, values of  $t$ . Moreover,  $G^t$  does not necessarily stay within the same location-scale family as  $G$ ; this property is only satisfied for the subclass of max-stable distributions  $\tilde{G}$ , for which  $\tilde{G}^m(\mathbf{a}_m \mathbf{z} + \mathbf{b}_m) = \tilde{G}(\mathbf{z})$ , for any integer  $m = 1, 2, \dots$ , and some normalizing vectors  $\mathbf{a}_m \in (0, \infty)^D$  and  $\mathbf{b}_m \in \mathbb{R}^D$ . In this work, we build new max-id models for spatial extremes, which nest a certain family of max-stable processes that have been used extensively for modeling extremes in the literature, while accommodating more flexible forms of dependence structures that remain natural for spatially-indexed block maxima.

Giné, Hahn and Vatan (1990), Resnick (1987) and Balkema, de Haan and Karandikar (1993) showed that any max-id process can be constructed by taking pointwise maxima over a Poisson point process (PPP) defined on a suitable functions space. Let  $\{X_i(\mathbf{s}); i = 1, 2, \dots, N\}_{\mathbf{s} \in \mathcal{S}}$  be the points of a Poisson point process with mean measure  $\Lambda$  on the space  $\mathcal{C}$  of continuous functions defined on a compact support  $\mathcal{S}$ , where the measure  $\Lambda$  must satisfy certain regularity constraints such as being finite on compact sets; see the above references for details. When  $\Lambda(\mathcal{C}) = \infty$  (such that  $N = \infty$  almost surely), we get a max-id process on  $\mathcal{S}$  by setting

$$(1) \quad Z(\mathbf{s}) = \sup_{i=1,2,\dots} X_i(\mathbf{s}), \quad \mathbf{s} \in \mathcal{S}.$$

Therefore, max-id processes can be constructed as pointwise maxima over an infinite number of continuous functions from the space  $\mathcal{C}$ , and the Poisson process weights the functions through its deterministic mean measure  $\Lambda$  when sampling from  $\mathcal{C}$ .

The mean measure  $\Lambda$  is also called the exponent measure of the max-id process, and it determines joint probabilities. Specifically, for a finite number of sites  $\mathcal{D} = \{\mathbf{s}_1, \dots, \mathbf{s}_D\} \subset \mathcal{S}$ , the joint distribution  $G$  of  $\mathbf{Z} = \{Z(\mathbf{s}_1), \dots, Z(\mathbf{s}_D)\}^T$  is

$$(2) \quad G(\mathbf{z}) = \Pr(\mathbf{Z} \leq \mathbf{z}) = \exp\{-\Lambda_{\mathcal{D}}([-\infty, \mathbf{z}]^{\mathcal{C}})\}, \quad \mathbf{z} = (z_1, \dots, z_D)^T \in \mathbb{R}^D,$$

where  $[-\infty, \mathbf{z}] = [-\infty, z_1] \times \dots \times [-\infty, z_D] \subset \mathbb{R}^D$ ,  $\Lambda_{\mathcal{D}}$  is the restriction of  $\Lambda$  to the subspace  $\mathcal{D} \subset \mathcal{S}$  (i.e., taking measurable sets of  $\mathbb{R}^D$  rather than  $\mathcal{C}$  as input), and  $V_{\mathcal{D}}(\mathbf{z}) = \Lambda_{\mathcal{D}}([-\infty, \mathbf{z}]^{\mathcal{C}})$  is called the exponent function. To simplify notation, we henceforth drop the subscript  $\mathcal{D}$  in  $V_{\mathcal{D}}$  and  $\Lambda_{\mathcal{D}}$  when no confusion can arise. In the case of max-stable processes with unit Fréchet margins, i.e.,  $\Pr(Z(\mathbf{s}) \leq z) = \exp(-1/z)$ ,  $z > 0$ , (1) can be expressed more specifically through the following spectral construction:

$$(3) \quad Z(\mathbf{s}) = \sup_{i=1,2,\dots} R_i W_i(\mathbf{s}), \quad \mathbf{s} \in \mathcal{S},$$

where  $\{R_i; i = 1, 2, \dots\}$  are the points of a Poisson point process on the positive half-line  $[0, \infty]$  with mean measure  $\kappa([r, \infty)) = r^{-1}$ ,  $r > 0$ , and  $\{W_i(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$ ,  $i = 1, 2, \dots$ , are independent copies of a random process  $\{W(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  with  $\mathbb{E}[\max\{W(\mathbf{s}), 0\}] = 1$ , which are also independent of the points  $\{R_i; i = 1, 2, \dots\}$ ; see de Haan (1984) and Schlather (2002). Let  $\Phi(dw)$  be the probability distribution associated with the process  $\{W(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  (specified to be Gaussian in our model described below in §3.2). Hence, the independent random processes  $\{X_i(\mathbf{s}) = R_i W_i(\mathbf{s}); i = 1, 2, \dots\}_{\mathbf{s} \in \mathcal{S}}$  are points from a Poisson process with mean measure  $\Lambda(A) = \int_{\{r w \in A\}} r^{-2} dr \Phi(dw)$ , for measurable sets  $A \subset \mathcal{C}$ . The exponent function of the max-stable process (3) can be written as  $V(\mathbf{z}) = \mathbb{E}[\max\{W(\mathbf{s}_1)/z_1, \dots, W(\mathbf{s}_D)/z_D\}]$ ,  $\mathbf{z} = (z_1, \dots, z_D)^T \geq \mathbf{0} = (0, \dots, 0)^T$ , where  $a/0 = \infty$  for  $a > 0$ .

Using (2), the marginal distribution of  $Z$  at a given site  $\mathbf{s}_0 \in \mathcal{S}$  for general max-id processes is  $G_0(z_0) = \exp\{-\Lambda_{\mathbf{s}_0}(\{z : z > z_0\})\}$ . To focus on dependence properties, we now assume that the max-id process (1) has been standardized using the probability integral transform to have common unit Fréchet margins, such that  $\Lambda_{\mathbf{s}_0}(\{z : z > z_0\}) = 1/z_0$ ,  $z_0 > 0$ , for

all sites  $\mathbf{s}_0 \in \mathcal{S}$ . Then, for any finite collection of sites  $\mathcal{D} = \{\mathbf{s}_1, \dots, \mathbf{s}_D\} \subset \mathcal{S}$ , we define the *level-dependent extremal coefficient at (unit Fréchet) quantile level*  $z_0 > 0$  as

$$(4) \quad \theta_D(z_0) = \frac{\log\{G(z_0)\}}{\log\{G_0(z_0)\}} = \frac{\Lambda_{\mathcal{D}}([-\infty, \mathbf{z}_0]^C)}{\Lambda_{\mathbf{s}_0}(\{z : z > z_0\})} = z_0 \Lambda_{\mathcal{D}}([-\infty, \mathbf{z}_0]^C) \in [1, D],$$

$\mathbf{z}_0 = (z_0, \dots, z_0)^T \in \mathbb{R}^D$ . A similar dependence coefficient was defined by [Padoan \(2013\)](#) and [Huser, Opitz and Thibaud \(2021\)](#). It is easy to see from the definition (4) that

$$\Pr(\mathbf{Z} \leq \mathbf{z}_0) = G_0(z_0)^{\theta_D(z_0)}.$$

Therefore, the coefficient  $\theta_D(z_0)$  can be interpreted as the effective number of independent variables among  $\{Z(\mathbf{s}_1), \dots, Z(\mathbf{s}_D)\}^T$  at *quantile level*  $z_0$ . In the bivariate case,  $D = 2$ , the pair of variables  $\mathbf{Z} = \{Z(\mathbf{s}_1), Z(\mathbf{s}_2)\}^T$  turns out to be asymptotically independent if  $\lim_{z_0 \rightarrow \infty} \theta_2(z_0) = 2$ , and asymptotically dependent otherwise. It can indeed be verified that  $\Pr\{Z(\mathbf{s}_2) > z_0 \mid Z(\mathbf{s}_1) > z_0\} \sim 2 - \theta_2(z_0)$ , as  $z_0 \rightarrow \infty$ .

With max-stable distributions, the extremal coefficient  $\theta_D(z_0)$  is always constant in  $z_0$  because the exponent function  $V(\mathbf{z})$  is homogeneous of order  $-1$ , i.e.,  $V(t\mathbf{z}) = t^{-1}V(\mathbf{z})$  for all  $t > 0$ . Thus, max-stable processes cannot capture weakening dependence as events become more extreme. Moreover, they can only capture asymptotic dependence or full independence, but they cannot capture intermediate joint tail decay rates arising with asymptotic independence. The broader class of max-id processes relaxes such rigid restrictions and yields more flexible models that remain in the “neighborhood” of max-stable processes.

**3.2. A new magnitude-dependent max-id model.** For modeling temperature extremes, we build on a max-id construction proposed by [Huser, Opitz and Thibaud \(2021\)](#). It extends the spectral representation of max-stable processes in (3) and allows capturing asymptotic independence and asymptotic dependence in a single parametric model. In the max-stable case, the heavy power-law tail of the mean measure  $\kappa([r, \infty)) = r^{-1}$ ,  $r > 0$ , of the Poisson process  $\{R_i; i = 1, 2, \dots\}$ , which determines the overall magnitude of the spatial process  $\{Z(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  in (3), generates co-occurrences of very large values and leads to asymptotic dependence, while the same level of dependence persists at all quantiles (i.e.,  $\theta_D(z)$  in (4) is constant in  $z$ ). We can deploy two modifications for the dependence in  $Z(\mathbf{s})$  to weaken as the magnitude of extreme events increases. The first modification is to use a lighter-tailed intensity measure  $\kappa$  of the Poisson process  $\{R_i; i = 1, 2, \dots\}$  to attenuate the strong co-occurrence patterns at increasingly high quantiles. The second modification is to relax the independence assumption between the points  $\{R_i\}$  and the processes  $\{W_i\}$ , in such a way as to link the spatial dependence range of  $W_i$  with the magnitude of  $R_i$ , which makes the processes  $\{W_i\}$  non-identically distributed given  $\{R_i\}$ . While the first modification was already exploited by [Huser, Opitz and Thibaud \(2021\)](#), the second modification is a new idea. In this paper, we combine both modifications, in order to construct a flexible yet parsimonious and identifiable max-id model that interpolates between the (asymptotically dependent) extremal- $t$  max-stable model and asymptotic independence with a relatively fast joint probability decay and a flexible form in the bulk.

Following [Huser, Opitz and Thibaud \(2021\)](#), we use a Weibull-tailed mean measure  $\kappa$  for  $\{R_i\}$  given as

$$(5) \quad \kappa([r, \infty)) = r^{-\beta} \exp\{-\alpha(r^\beta - 1)/\beta\}, \quad r > 0, (\alpha, \beta)^T \in (0, \infty)^2.$$

We further specify  $\{W_i(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  to be standard Gaussian processes characterized by the correlation function  $\rho(\mathbf{s}_1, \mathbf{s}_2; R_i)$ , which may depend on  $R_i$ . Because  $\lim_{\beta \rightarrow 0} \kappa([r, \infty)) = r^{-\alpha}$ , this max-id process reduces to the max-stable extremal- $t$  process with  $\alpha > 0$  degrees of freedom when  $\beta \rightarrow 0$  and  $\rho(\mathbf{s}_1, \mathbf{s}_2; R_i) \equiv \rho(\mathbf{s}_1, \mathbf{s}_2)$  is independent of  $R_i$  ([Opitz, 2013](#)). Furthermore, [Huser, Opitz and Thibaud \(2021\)](#) showed that when  $\{W_i(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$ ,  $i = 1, 2, \dots$ ,



are identically distributed standard Gaussian processes with correlation function  $\rho(\mathbf{s}_1, \mathbf{s}_2)$  (independent of  $R_i$ ), the coefficient of tail dependence (Ledford and Tawn, 1996), which characterizes the joint tail decay rate for two sites  $\mathbf{s}_1, \mathbf{s}_2 \in \mathcal{S}$ , may be expressed as

$$\eta(\mathbf{s}_1, \mathbf{s}_2) = \lim_{z \uparrow \infty} \frac{\log\{1 - G_1(z)\}}{\log\{1 + G(z, z) - 2G_1(z)\}} = [\{1 + \rho(\mathbf{s}_1, \mathbf{s}_2)\}/2]^{\beta/(\beta+2)},$$

where  $G(\cdot, \cdot)$  and  $G_1(\cdot)$  represent the bivariate and univariate CDFs of  $\{Z(\mathbf{s}_1), Z(\mathbf{s}_2)\}^T$  and  $Z(\mathbf{s}_1)$  (or  $Z(\mathbf{s}_2)$ ), respectively. Hence, the parameter  $\beta$  and the correlation  $\rho$  of the Gaussian process  $W$  together strongly influence the joint tail decay rate. In particular, as  $\beta \rightarrow 0$  or  $\rho(\mathbf{s}_1, \mathbf{s}_2) \rightarrow 1$ , we get  $\eta(\mathbf{s}_1, \mathbf{s}_2) = 1$ , which yields asymptotic dependence. In all other cases, we get  $\eta(\mathbf{s}_1, \mathbf{s}_2) < 1$ , thus asymptotic independence, and we retrieve the tail decay rate of a Gaussian process as  $\beta \rightarrow \infty$ .

Here, we extend the model of Huser, Opitz and Thibaud (2021) by letting the correlation function of  $W_i$  in (3) depend on  $R_i$  such that  $\rho(\mathbf{s}_1, \mathbf{s}_2; R_i)$  decreases as  $R_i$  increases. In other words, the spatial dependence strength weakens when the overall event magnitude represented by the points  $\{R_i\}$  gets larger. In the stationary and isotropic case, one possibility is to consider the exponential correlation function

$$(6) \quad \rho(\mathbf{s}_1, \mathbf{s}_2; R_i) = \exp\{-\|\mathbf{s}_1 - \mathbf{s}_2\|(1 + R_i)^\nu/\lambda\},$$

for some baseline range parameter  $\lambda > 0$ , and ‘‘modulation’’ parameter  $\nu \in \mathbb{R}$ . When  $\nu = 0$ ,  $\rho(\mathbf{s}_1, \mathbf{s}_2; R_i) \equiv \rho(\mathbf{s}_1, \mathbf{s}_2)$  does not depend on  $R_i$  (hence retrieving the max-id models of Huser, Opitz and Thibaud (2021)), but when  $\nu > 0$ , the spatial range parameter  $\lambda(1 + R_i)^{-\nu}$  gets smaller (i.e., the dependence strength decreases) as  $R_i$  increases (and vice versa when  $\nu < 0$ ), with the value of  $\nu$  controlling the rate at which the correlation decays with larger points  $R_i$ . This modification allows us to get more flexible forms of dependence in the bulk, while keeping appealing tail dependence properties with the Huser, Opitz and Thibaud (2021) model as a special case when  $\nu = 0$ . To illustrate the flexibility of this model, Figure 3 displays the bivariate level-dependent extremal coefficient  $\theta_2(z)$  for various values of  $\beta$  and  $\nu$ . The case  $\beta = 0$  and  $\nu = 0$  yields the extremal- $t$  max-stable model, so that  $\theta_2(z)$  is constant in  $z$ . When  $\beta = 0$  but  $\nu > 0$ , we get asymptotic dependence ( $\lim_{z \rightarrow \infty} \theta_2(z) < 2$ ) with weakening dependence strength at increasing quantiles. And when  $\beta > 0$  and  $\nu \geq 0$ , we get asymptotic independence ( $\lim_{z \rightarrow \infty} \theta_2(z) = 2$ ). Moreover, the extremal coefficient grows with  $\nu$  and  $\beta$ . At any fixed value of  $\beta$ , the curvature of  $\theta_2(z)$  varies significantly for different values of  $\nu$ , which implies that introducing dependence between  $R_i$  and  $W_i$  adds considerable flexibility to the model and improves its ability to appropriately capture the dependence of moderately extreme events. In §3.3, we extend this model to the non-stationary, anisotropic case. Note that although both  $\beta$  and  $\nu$  influence the joint tail decay rate simultaneously, they have different roles:  $\beta$  controls the asymptotic dependence class (with asymptotic dependence when  $\beta = 0$  and asymptotic independence when  $\beta > 0$ ), while  $\nu$  improves the flexibility of the model to capture the dependence strength of moderate extreme events. In fact, when  $\beta = 0$ , the parameter  $\nu$  is the only one left to control how fast the spatial dependence strength weakens as we move to higher quantile levels. A deeper interpretation of how  $\beta$  and  $\nu$  act on the stochastic representation of the spatial extreme fields is as follows. While Gaussian processes are well adapted for capturing central tendencies thanks to central limit theory, they typically lack in flexibility for appropriately capturing extremal dependence since their extremal correlation decays in a relatively fast and rigid way for increasingly high quantile levels. The role of the scaling variables  $R_i$  is to lift moderate Gaussian values (where the dependence can be relatively strong) jointly towards the extreme quantiles, which then induces stronger than Gaussian dependence among extremes, and we can get arbitrarily close to max-stability arising asymptotically for extremes. The parameter  $\beta$  controls this mechanism. Moreover, in

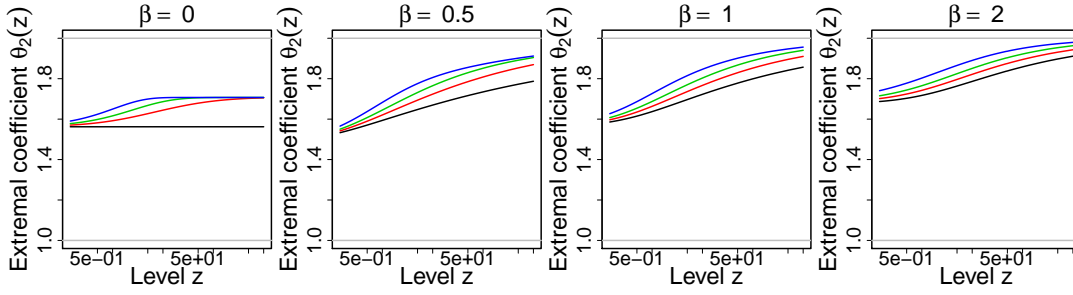


FIGURE 3. Bivariate level-dependent extremal coefficient  $\theta_2(z) = zV(z, z)$  with respect to unit Fréchet quantiles  $z$ , plotted on a logarithmic scale. Our max-id model is defined as in (3), where the mean measure  $\kappa$  of the Poisson points  $\{R_i\}$  is based on (5), here with  $\alpha = 1$  and  $\beta = 0, 0.5, 1, 2$  (left to right), and where the correlation function of the Gaussian processes  $W_i$  is here assumed to be  $\rho(\mathbf{s}_1, \mathbf{s}_2; R_i) = \exp\{-\|\mathbf{s}_1 - \mathbf{s}_2\|(1 + R_i)^\nu/\lambda\}$  with  $\lambda = 0.5$  and  $\nu = 0$  (black),  $\nu = 0.25$  (red),  $\nu = 0.5$  (green),  $\nu = 1$  (blue). The distance  $\|\mathbf{s}_1 - \mathbf{s}_2\|$  is here set to 0.5. The horizontal grey lines represent the lower and upper bounds of 1 and 2, respectively.

the distribution of maxima we can also differentiate between “bulk maxima” and “extreme maxima”, and the parameter  $\nu$  allows us to further modulate the differences in the strength of spatial correlation that may arise between bulk and extreme levels. In short, this means that the two parameters  $\beta$  and  $\nu$  are *theoretically* identifiable. In §4.3, we further demonstrate that they are also *numerically* identifiable, given that both parameters can be estimated from the data reasonably well.

By conditioning on the variables  $\{R_i\}$ , we can prove that the general form of the exponent function of  $\{Z(\mathbf{s}_1), \dots, Z(\mathbf{s}_D)\}^T$  for our proposed max-id model may be expressed as

$$(7) \quad V(\mathbf{z}) = \Lambda([-\infty, \mathbf{z}]^C) = \int_0^\infty \{1 - \Phi(\mathbf{z}/r; r)\} \kappa(dr)$$

where  $\Phi(\cdot; r)$  is the joint distribution of the Gaussian vector  $\{W_i(\mathbf{s}_1), \dots, W_i(\mathbf{s}_D)\}^T \mid \{R_i = r\}$  with correlation  $\rho(\mathbf{s}_1, \mathbf{s}_2; r)$ . Partial and full derivatives of the exponent function, which are required for likelihood-based inference, can be obtained by differentiating (7) with respect to the components of  $\mathbf{z}$  under the integral sign. Standard formulas, similar to those derived in Huser, Opitz and Thibaud (2021), can be easily obtained, although they are expressed in terms of uni-dimensional integrals that have to be numerically approximated in practice.

3.3. *Non-stationary dependence structure.* Over large study areas or long periods of time, the strength of extremal dependence, and therefore the spatial extent of clusters of extreme values, may vary. We here extend the exponential correlation model presented in §3.2 to the non-stationary context, and we show how spatio-temporal covariates may be naturally incorporated. We now index the correlation function of the process  $W_i$  in (3) by time  $t$  as  $\rho_t(\mathbf{s}_1, \mathbf{s}_2; R_i)$  to emphasize that it may vary over time. Building upon Paciorek and Schervish (2006) and Huser and Genton (2016), such a non-stationary correlation function on  $\mathbb{R}^2$  may be obtained as follows:

$$(8) \quad \rho_t(\mathbf{s}_1, \mathbf{s}_2; R_i) = |\Omega_{\mathbf{s}_1, t}(R_i)|^{1/4} |\Omega_{\mathbf{s}_2, t}(R_i)|^{1/4} \left| \frac{\Omega_{\mathbf{s}_1, t}(R_i) + \Omega_{\mathbf{s}_2, t}(R_i)}{2} \right|^{-1/2} C\{Q_{\mathbf{s}_1; \mathbf{s}_2, t}^{1/2}(R_i)\},$$

where  $\Omega_{\mathbf{s}, t}(R_i)$  is a 2-by-2 covariance matrix that may depend on spatial location  $\mathbf{s}$ , time  $t$ , and the Poisson points  $\{R_i\}$ ,  $C(h)$  is a stationary isotropic correlation function with unit range, e.g.,  $C(h) = \exp(-h)$ ,  $h \geq 0$ , and  $Q_{\mathbf{s}_1; \mathbf{s}_2, t}(R_i)$  is the quadratic form

$$Q_{\mathbf{s}_1; \mathbf{s}_2, t}(R_i) = (\mathbf{s}_1 - \mathbf{s}_2)^T \left\{ \frac{\Omega_{\mathbf{s}_1, t}(R_i) + \Omega_{\mathbf{s}_2, t}(R_i)}{2} \right\}^{-1} (\mathbf{s}_1 - \mathbf{s}_2).$$

TABLE 1  
*Interesting special cases of Model (9), categorized into stationary/non-stationary and (locally) isotropic/anisotropic models with  $R_i$  independent/dependent of  $W_i$  in (3).*

Stationarity	Anisotropy	Value of $\nu$	Model type
$\lambda_{s,t} \equiv \lambda$	$a = 1$	$\nu = 0$	Stationary, isotropic, $R_i \perp W_i$
$\lambda_{s,t} \equiv \lambda$	$a = 1$	$\nu \neq 0$	Stationary, isotropic, $R_i \not\perp W_i$
$\lambda_{s,t} \equiv \lambda$	$a \neq 1$	$\nu = 0$	Stationary, anisotropic, $R_i \perp W_i$
$\lambda_{s,t} \equiv \lambda$	$a \neq 1$	$\nu \neq 0$	Stationary, anisotropic, $R_i \not\perp W_i$
$\lambda_{s,t} \not\equiv \lambda$	$a = 1$	$\nu = 0$	Non-stationary, locally isotropic, $R_i \perp W_i$
$\lambda_{s,t} \not\equiv \lambda$	$a = 1$	$\nu \neq 0$	Non-stationary, locally isotropic, $R_i \not\perp W_i$
$\lambda_{s,t} \not\equiv \lambda$	$a \neq 1$	$\nu = 0$	Non-stationary, locally anisotropic, $R_i \perp W_i$
$\lambda_{s,t} \not\equiv \lambda$	$a \neq 1$	$\nu \neq 0$	Non-stationary, locally anisotropic, $R_i \not\perp W_i$

Covariates, such as time and elevation as used in our temperature data application in §5, can be linked to the matrix  $\Omega_{s,t}(R_i)$ . As explained in §3.2, we also allow the variables  $\{R_i\}$  to directly influence the range of spatial dependence, which is in contrast with Huser and Genton (2016). More precisely, we propose the following general model for the covariance matrix  $\Omega_{s,t}(R_i)$ , whose modeling of anisotropy through covariates is also different from the form used in Huser and Genton (2016):

(9)

$$\Omega_{s,t}(R_i) = \lambda_{s,t}^2 (1 + R_i)^{-2\nu} A(\theta), \quad A(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & a \end{bmatrix} \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}^T,$$

where  $\lambda_{s,t} > 0$  is a baseline range parameter that may vary over space and time, and  $\nu \in \mathbb{R}$  as in §3.2,  $a > 0$  is a geometric anisotropy scaling that controls the ratio of principal axes of elliptical correlation contours, and  $\theta \in [0, \pi/2]$  is a rotation angle of these elliptical contours. The value  $a = 1$  corresponds to the isotropic case, with  $A(\theta)$  reducing to the 2-by-2 identity matrix  $I_{2 \times 2}$ , such that  $\Omega_{s,t}(R_i) = \lambda_{s,t}^2 (1 + R_i)^{-2\nu} I_{2 \times 2}$  and thus (8) corresponds to the correlation (6) if  $\lambda_{s,t} \equiv \lambda > 0$ . To capture spatio-temporal variations in the dependence structure, covariates may be included in  $\lambda_{s,t}$ . For example, in our real data application in §5, we specify  $\lambda_{s,t} = \exp(\lambda_0 + \lambda_1 \times \text{elev}_s + \lambda_2 \times t)$ , where  $\lambda_0, \lambda_1, \lambda_2 \in \mathbb{R}$  are range parameters corresponding to the intercept, the effect of elevation, and the effect of time, respectively, on the spatial dependence range. More precisely, while  $\lambda_0$  measures the overall strength of spatial dependence, the parameters  $\lambda_1$  and  $\lambda_2$  determine whether the dependence structure, and thus the spatial extent of extreme events, changes according to elevation and time, respectively. Several sub-models of (9) described in Table 1 may be of interest. In §5, we specifically focus on the non-stationary, but locally isotropic case ( $a = 1$ ) with  $R_i$  and  $W_i$  potentially dependent of each other with  $\nu \in \mathbb{R}$  (i.e.,  $R_i \not\perp W_i$  when  $\nu \neq 0$ ), which already yields a rich class of models capturing complex dependence patterns.

#### 4. Inference using the pairwise likelihood approach.

4.1. *Two-step modeling of marginal distributions and dependence.* We use a two-step estimation method that is known as “inference functions for margins” in the literature, and for which consistency and asymptotic normality have been established under mild conditions (Joe and Xu, 1996; Joe, 2005, 2015). Precisely, in the first step, we model marginal distributions using a parametric model, in which the model parameters are further described in terms of semi-parametric spline functions with respect to covariates, and we fit this model jointly across stations using an independence composite likelihood (Varin, Reid and Firth, 2011) built under the working assumption that the data are spatially independent (given the

covariates). In our real data application, we use the generalized extreme-value (GEV) distribution to model station-wise block maxima. We then use the fitted marginal GEV distribution functions to transform the observed data to pseudo-uniform  $\text{Unif}(0, 1)$  scores through the probability integral transform. In the second step, we fit the dependence structure (i.e., the copula) of the max-id dependence model to the transformed data using a pairwise likelihood approach, treating the margins as exactly  $\text{Unif}(0, 1)$ . This two-step inference procedure has the major advantage of making computations much simpler and significantly less intensive compared to estimating marginal and dependence parameters at once. Moreover, such an approach offers the flexibility of choosing different types of inference procedures in the two steps, e.g., using a generalized additive model with independence likelihood for margins (for which robust fitting procedures are already implemented in  $\mathbb{R}$ ) and a more specific pairwise likelihood inference procedure for the dependence structure. Furthermore, results in the Supplementary Material show that it performs very well. However, we also have to bear in mind that the two-step inference procedure assumes in the second step that the margins have been estimated with no uncertainty in the first step, and thus the overall uncertainty might be underestimated if not properly accounted for; see the Supplementary Material for further details and experiments.

*4.2. Pairwise likelihood approach.* Pairwise likelihood has become the standard inference technique for max-stable models owing to the computational intractability of full likelihood expressions in high dimensions (Padoan, Ribatet and Sisson, 2010; Padoan, 2013; Huser and Davison, 2013; Huser, Davison and Genton, 2016; Castruccio, Huser and Genton, 2016; Huser et al., 2019). The pairwise likelihood approach offers tools akin to classical likelihood inference, is much faster than a full likelihood approach, and usually retains high efficiency. We adapt this approach, here, to our max-id models. Let  $\{\mathbf{z}_k = (z_{k1}, \dots, z_{kD})^T\}_{k=1}^n$  be  $n$  independent replicates of the max-id process  $Z(\mathbf{s})$  with parameter vector  $\boldsymbol{\psi} \in \Psi \subset \mathbb{R}^p$  observed at locations  $\mathcal{D} = \{\mathbf{s}_1, \dots, \mathbf{s}_D\} \subset \mathcal{S}$ . From (2), the full likelihood is

$$(10) \quad L(\boldsymbol{\psi}; \mathbf{z}_1, \dots, \mathbf{z}_n) = \prod_{k=1}^n \left[ \exp\{-V(\mathbf{z}_k)\} \sum_{\pi \in \mathcal{P}_D} \prod_{l=1}^{|\pi|} \{-V_{\tau_l}(\mathbf{z}_k)\} \right],$$

where  $\mathcal{P}_D$  is the collection of all partitions  $\pi = \{\tau_1, \dots, \tau_{|\pi|}\}$  sets of  $\{1, \dots, D\}$ , and  $V_{\tau_l}(\mathbf{z}_k)$  denotes the partial derivatives of the exponent function  $V(\mathbf{z}_k)$  with respect to the variables  $\{z_{kj}\}_{j \in \tau_l}$ ,  $\tau_l \in \pi$ ; see, e.g., Huser et al. (2019). The number of terms in the sum in (10) grows super-exponentially with  $D$ . The pairwise likelihood approach eases the computational burden by maximizing the pairwise likelihood function  $PL(\boldsymbol{\psi}; \mathbf{z}_1, \dots, \mathbf{z}_n)$  defined as

$$(11) \quad \prod_{1 \leq j_1 < j_2 \leq D} \left[ \prod_{k=1}^n \exp\{-V(z_{kj_1}, z_{kj_2})\} \{V_1(z_{kj_1}, z_{kj_2})V_2(z_{kj_1}, z_{kj_2}) - V_{12}(z_{kj_1}, z_{kj_2})\} \right]^{\omega_{j_1, j_2}},$$

where  $\omega_{j_1, j_2} \geq 0$  are non-negative weights attributed to the pairs  $\{j_1, j_2\}$ . Here, we fit the marginal distribution first and compute the pseudo-uniform scores  $u_{kj} = \widehat{G}_{kj}(z_{kj})$ , where  $\widehat{G}_{kj}$  is the fitted marginal distribution for the  $k$ -th time point and the  $j$ -th site  $\mathbf{s}_j$ . Let  $\widehat{g}_{kj}$  be the corresponding fitted marginal density. The pairwise likelihood function  $PL(\boldsymbol{\psi}; \mathbf{u}_1, \dots, \mathbf{u}_n)$  based on pseudo-uniform scores  $\{\mathbf{u}_k = (u_{k1}, \dots, u_{kD})^T\}_{k=1}^n$  may thus be written as

$$(12) \quad PL(\boldsymbol{\psi}; \mathbf{u}_1, \dots, \mathbf{u}_n) = \prod_{1 \leq j_1 < j_2 \leq D} \left( \prod_{k=1}^n \exp \left[ -V \left\{ \widehat{G}_{kj_1}^{-1}(u_{kj_1}), \widehat{G}_{kj_2}^{-1}(u_{kj_2}) \right\} \right] \times \right.$$

$$\begin{aligned} & \times \left[ V_1 \{ \widehat{G}_{k_{j_1}}^{-1}(u_{k_{j_1}}), \widehat{G}_{k_{j_2}}^{-1}(u_{k_{j_2}}) \} V_2 \{ \widehat{G}_{k_{j_1}}^{-1}(u_{k_{j_1}}), \widehat{G}_{k_{j_2}}^{-1}(z_{k_{j_2}}) \} - V_{12} \{ \widehat{G}_{k_{j_1}}^{-1}(u_{k_{j_1}}), \widehat{G}_{k_{j_2}}^{-1}(u_{k_{j_2}}) \} \right] \times \\ & \times \left[ \widehat{g}_{k_{j_1}} \{ \widehat{G}_{k_{j_1}}^{-1}(u_{k_{j_1}}) \} \widehat{g}_{k_{j_2}} \{ \widehat{G}_{k_{j_2}}^{-1}(u_{k_{j_2}}) \} \right]^{-1} \Big)^{\omega_{j_1, j_2}}. \end{aligned}$$

Different approaches can be used to select the pairwise likelihood weights  $\omega_{j_1, j_2}$ , e.g., using binary weights  $\omega_{j_1, j_2} \in \{0, 1\}$  fixed according to the distance between sites, in order to improve both the computational and statistical efficiency (see, e.g., [Castruccio, Huser and Genton, 2016](#)). In our simulation study in §4.3, we choose  $\omega_{j_1, j_2} = I(\|\mathbf{s}_{j_1} - \mathbf{s}_{j_2}\| \leq \delta)$  for some cutoff distance  $\delta > 0$  for computational reasons, where  $I(\cdot)$  is the indicator function, whereas in our application in §5 we use the pragmatic approach of setting  $\omega_{j_1, j_2} = 1$  for all pairs  $\{j_1, j_2\}$ .

It is known that under mild regularity conditions, the pairwise likelihood estimator maximizing (11) with known margins is strongly consistent and asymptotically normal with the Godambe variance-covariance matrix, which could in principle be used to assess the variability of the estimator; see, e.g., [Varin, Reid and Firth \(2011\)](#) and [Padoan, Ribatet and Sisson \(2010\)](#) for the max-stable case. A similar asymptotic behavior holds for the estimator  $\widehat{\psi}$  based on the two-step estimator (12) (with unknown margins), though the asymptotic variance is generally slightly larger due to the uncertainty in estimating marginal distributions; see, e.g., [Genest, Ghoudi and Rivest \(1995\)](#) who treat the case where margins are estimated non-parametrically, [Joe and Xu \(1996\)](#) for the parametric case, and [Huser and Davison \(2014\)](#) and [Huser, Davison and Genton \(2016\)](#) who compare various parametric estimation schemes for extremes, including one-step and two-step pairwise likelihood estimators. However, since the computation of the asymptotic variance is intricate especially when the data contain many missing values, here, we rely on a parametric bootstrap procedure (see, e.g., [Davison and Hinkley, 1997](#), page 15) to assess the estimation uncertainty: we repeatedly sample maxima data at the data locations from the fitted max-id model (with the same sample size and with the same number of missing values inserted as in the original dataset), and we then re-estimate parameters using the same pairwise likelihood. Using 300 bootstrap samples, we can then approximate the distribution and variability of estimated parameters. Alternatively, the jackknife method originally developed by [Quenouille \(1949, 1956\)](#) and [Tukey \(1958\)](#) could also be used, treating each year as an independent (spatial) replicate, but its theoretical validity when the data are not missing-at-random can be questioned. By contrast, the parametric bootstrap has solid theoretical foundations, and is easier to apply in case of missing data (see [Davison and Hinkley, 1997](#), page 88). In the Supplementary Material, we compare our parametric bootstrap scheme with the jackknife used in our application. Although there are some differences between these two methods as expected, the estimated standard errors are reasonably consistent in general. Overall, we find that the bootstrap, which was originally inspired from the jackknife, is more suitable in our case as it better handles missing values and is more widely applicable as pointed out by [Efron \(1979\)](#). In the Supplementary Material, we further provide evidence that our proposed parametric bootstrap procedure yields satisfactory coverage probabilities in a simplified settings (namely univariate GEV-distributed data, bivariate logistic max-stable data, and spatial extremal- $t$  max-stable data similar to our real data application). This gives us confidence that it should also perform satisfactorily in our most complex spatial max-id setting, though this is difficult to verify formally due to the heavy computational burden of fitting our proposed max-id process.

4.3. *Simulation experiments.* We conducted a simulation study, in order to assess the performance of the pairwise likelihood estimator  $\widehat{\psi}$  under a non-stationary setting that resembles our real data application in §5.

We simulated data from our proposed max-id model, built from (3) using the mean measure (5) and the non-stationary correlation function (8) combined with (9), on the domain  $\mathcal{S} = (0, 1)^2$ . Here, we focused on the non-stationary, but locally isotropic case (recall Table 1), where  $\Omega_{\mathbf{s}, t} = \lambda_{\mathbf{s}, t}^2 (1 + R_i)^{-2\nu} I_{2 \times 2}$ , with  $I_{2 \times 2}$  the identity matrix, and we only considered non-stationarity in space such that  $\lambda_{\mathbf{s}, t} \equiv \lambda_{\mathbf{s}}$  only varies with spatial location  $\mathbf{s} \in \mathcal{S}$ . To mimic the effect of a “mountain range”, we used the covariate defined as  $x_{\mathbf{s}} = 2\phi(s_x; 0.5, 0.25) - 1$ , where  $\mathbf{s} = (s_x, s_y)^T$  and  $\phi(\cdot; \mu, \sigma)$  denotes the Gaussian  $\mathcal{N}(\mu, \sigma^2)$  density, and then defined  $\lambda_{\mathbf{s}, t} \equiv \lambda_{\mathbf{s}} = \exp(\lambda_0 + \lambda_1 x_{\mathbf{s}})$ , with  $\lambda_0 = -0.5$  and  $\lambda_1 \in \{-0.5, -0.25, 0\}$ . Negative values of  $\lambda_1$  correspond to weaker dependence at higher “elevation”, here represented by the covariate  $x_{\mathbf{s}}$ , while  $\lambda_1 = 0$  corresponds to a stationary model. For the mean measure of the Poisson points  $\{R_i\}$  defined in (5), we chose  $\alpha = 1$  and  $\beta \in \{0, 0.5, 1\}$  (from asymptotic dependence with  $\beta = 0$  to asymptotic independence with  $\beta > 0$ ), while we selected  $\nu \in \{0.25, 0.5, 1\}$  to control the interaction between the points  $\{R_i\}$  and the processes  $\{W_i\}$  in (3). Overall, this yields 27 simulation scenarios (3 values of  $\lambda_1$  times 3 values of  $\beta$  times 3 values of  $\nu$ ), and we then jointly estimated the 5 dependence parameters, namely  $\psi = (\alpha, \beta, \lambda_0, \lambda_1, \nu)^T \in (0, \infty)^2 \times \mathbb{R}^3$ , treating margins as known here for simplicity. For each case, we generated a dataset comprising 50 independent replicates of the process at 49 sites on  $\mathcal{S}$ , roughly located on a  $7 \times 7$  grid with some small additional random perturbations. The dependence parameters were jointly estimated using the pairwise likelihood approach described in §4.2 with the pairwise likelihood weights  $\omega_{j_1, j_2}$  set to be 0 when the distance between the sites  $s_{j_1}$  and  $s_{j_2}$  exceeds the cutoff distance  $\delta = 0.375$  and  $\omega_{j_1, j_2} = 1$  otherwise, in order to ease the computation burden. We repeated the above steps 200 times to assess the variability and bias of estimated parameters. Figure 4 reports the results and shows boxplots of estimated parameters in the case  $\nu = 0.25$ , with one display for each  $\{\lambda_1, \beta\}$ -scenario. Results for  $\nu = 0.5$  and  $\nu = 1$  are similar and reported in the Supplementary Material for completeness. The parameters are well estimated overall without any strong biases, overly large variances, or notable outliers, which suggests that they are well identifiable. The true values (red dots) are close to the medians and always within the interquartile range in all scenarios. While the estimated parameters for  $\alpha$  and  $\beta$  seem comparatively more variable, especially when  $\beta$  is large, the covariate effect  $\lambda_1$  always has a fairly small variability. Finally, we can notice that the estimated values of  $\nu$  are always positive even if the domain of definition for this parameter is fixed to the whole real line in our implementation. This result shows that it is easy to identify that the dependence strength is weakening (rather than strengthening) as the severity of extreme events increases. Moreover, bivariate scatterplots of estimated parameters displayed in the Supplementary Material do not reveal any significant correlation between  $\beta$  and  $\nu$ . In our simulations, correlations turn out to be stronger between the dependence range  $\lambda_0$  and  $\alpha$ , and between  $\lambda_0$  and  $\nu$ , but in the application such correlation values are quite weak. Overall, by considering all simulation results summarized in Figure 4 and the Supplementary Material, it is clear that all parameters—and in particular  $\beta$  and  $\nu$ —can be reasonably well identified from the data in a wide range of parameter settings. This confirms that  $\beta$  and  $\nu$  indeed play a different role and that both of them help improve the flexibility of the model’s bulk and tail dependence structures in different ways.

## 5. Application to European temperature extremes.

5.1. *Spatial and temporal trends in marginal distributions.* The first step of our statistical analysis is to adequately model the non-stationary marginal distributions of maxima. We assume the annual temperature maxima follow a generalized extreme-value (GEV) distribution, which arises as the only possible asymptotic distribution for block maxima of univariate random variables (Fisher and Tippett, 1928); it is defined as

$$(13) \quad G(z) = \exp \left[ - \{1 + \xi(z - \mu)/\sigma\}_+^{-1/\xi} \right], \quad \xi \neq 0,$$

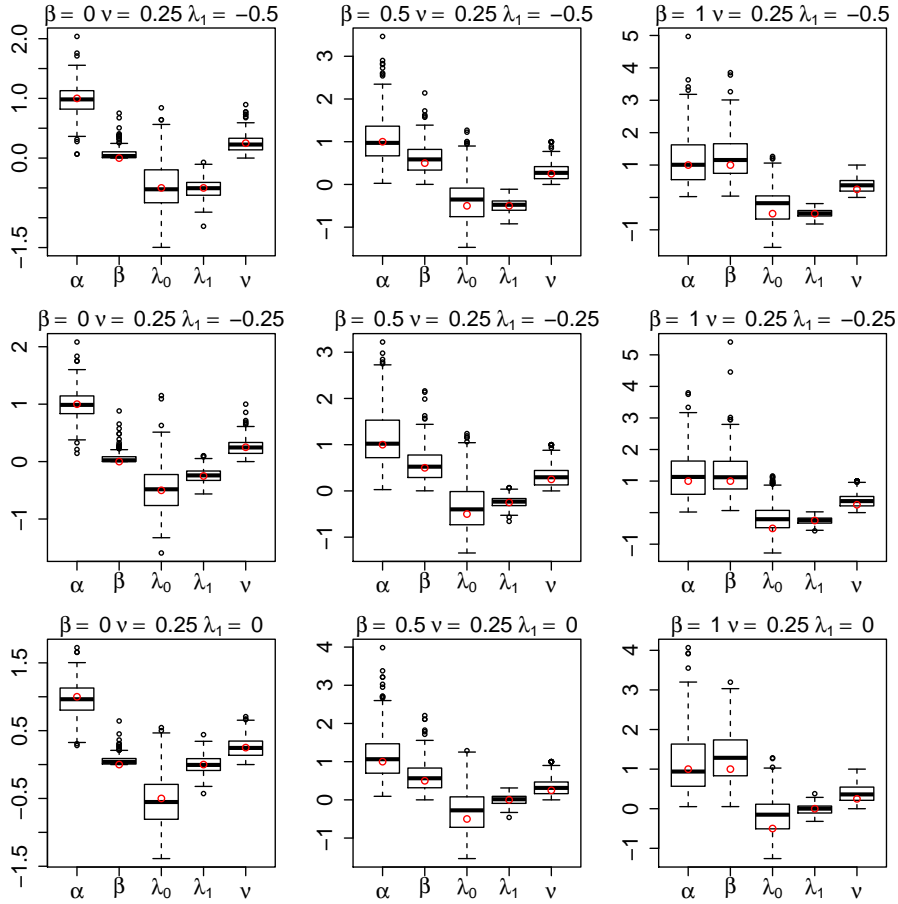


FIGURE 4. Boxplots of estimated parameters for the simulation study described in §4.3. Each panel corresponds to a different simulation scenario with  $\lambda_1 = -0.5, -0.25, 0$  (top to bottom) and  $\beta = 0, 0.5, 1$  (left to right) (see details in the text), and shows boxplots for each of the 5 parameters based on 200 experiments. Red dots indicate the true values.

with support  $\{z : 1 + \xi(z - \mu)/\sigma > 0\}$ , where  $a_+ = \max(0, a)$ , and the Gumbel distribution  $\exp[-\exp\{-(z - \mu)/\sigma\}]$ ,  $z \in \mathbb{R}$ , is obtained as  $\xi \rightarrow 0$ . Here,  $\mu, \sigma > 0$ , and  $\xi$  are location, scale and shape parameters, respectively. We distinguish three types of GEV distributions depending on the value of  $\xi$ : Fréchet, Gumbel and reversed Weibull, corresponding to  $\xi > 0$  (heavy-tailed),  $\xi \rightarrow 0$  (light-tailed) and  $\xi < 0$  (bounded tail), respectively. All three parameters could potentially depend on longitude, latitude, elevation, as well as time. To reduce the uncertainty in estimated marginal parameters, we pool the data together in a single generalized additive model with penalized cubic regression splines to accurately describe the time trend and spatial variation of GEV parameters, and then estimate parameter surfaces by maximizing an independence composite likelihood (Varin, Reid and Firth, 2011). As explained in §4.1, this approach provides valid inference for marginal parameters. Specifically, let  $\{Z(s, t)\}_{s \in \mathcal{S}, t \in [0, 1]}$  denote the spatio-temporal process of annual maxima (defined over the space-time domain  $\mathcal{S} \times [0, 1]$ ), and let  $Z_{kj} = Z\{s_j, k/(n + 1)\}$  be the annual maximum for the  $k$ -th year at the  $j$ -th station ( $k = 1, \dots, n$ ,  $j = 1, \dots, D$ ). After some experimentation, our final model assumes that  $Z(s, t)$  has a marginal GEV distribution with location parameter  $\mu_{s, t}$ , and constant scale and shape parameters,  $\sigma > 0$  and  $\xi$ , respectively. Since the location parameter  $\mu_{s, t}$  determines the overall magnitude of extreme values, we link it with the covariates  $\text{lon}_s, \text{lat}_s, \text{elev}_s$  representing longitude, latitude and elevation, respectively,

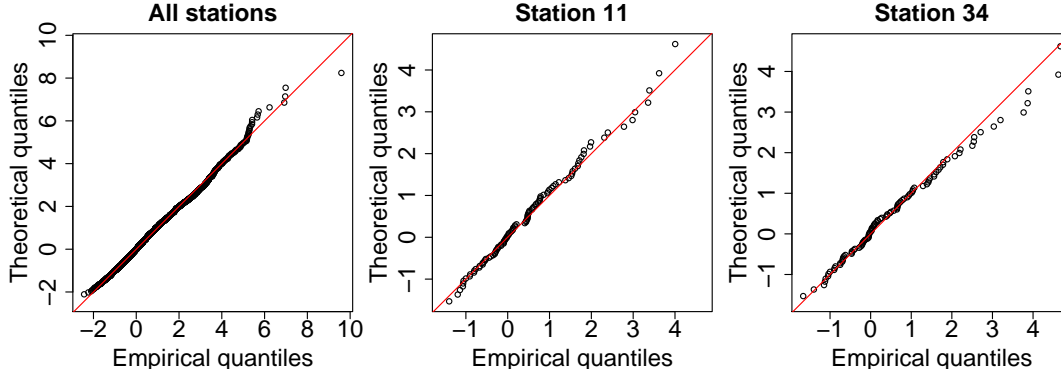


FIGURE 5. *QQ-plots of maxima transformed to the standard Gumbel scale based on the fitted marginal model. Left: all stations pooled together. Middle and right: stations  $s_{11}$  and  $s_{34}$ .*

and (rescaled) time  $t \in [0, 1]$ . We also tried to let  $\sigma$  vary over space and time, but this did not improve the model significantly. We can thus formulate the marginal model as

$$(14) \quad Z(s, t) \sim \text{GEV}(\mu_{s,t}, \sigma, \xi), \quad \mu_{s,t} = \text{ti}(\text{lon}_s, \text{lat}_s, \text{elev}_s) + \text{ti}(t), \quad s \in \mathcal{S}, t \in [0, 1],$$

where “ti” refers to the tensor product of penalized cubic regression splines. This marginal model was chosen to provide a good balance between flexibility (for a good model fit) and parsimony (for robustness and to avoid overfitting), and diagnostics described below suggest that the margins are appropriately modeled over the study region. For each observation, (14) yields  $Z_{kj} \sim \text{GEV}(\mu_{kj}, \sigma, \xi)$ , where  $\mu_{kj} = \mu_{s_j, k/(n+1)}$ , and we fit the marginal model jointly combining all observations by pretending that the  $Z_{kj}$ ’s are independent. Because the geographical location is jointly determined by  $\text{lon}_s$ ,  $\text{lat}_s$  and  $\text{elev}_s$ , they are put together in (14) to account for interaction effects, while we keep the time  $t$  separate to avoid an overly complex model with too many spline coefficients to be estimated. Here we take 4 spline knots for each dimension, which is rich enough to provide good marginal fits as demonstrated below. Therefore,  $\text{ti}(\text{lon}_s, \text{lat}_s, \text{elev}_s)$  has  $4^3 = 64$  spline knots in total. The estimated scale and shape parameters are  $\hat{\sigma} = 17.7$  with 95% confidence interval (17.2, 18.1) and  $\hat{\xi} = -0.20$  with 95% confidence interval  $(-0.22, -0.18)$ , respectively, where the uncertainty in this marginal setting is here estimated using the delta method as done in the function `gam` of the R package `mgcv`. Because  $\hat{\xi}$  is negative, the temperature distribution is estimated to have a finite upper endpoint, which is meaningful in view of the results obtained in similar studies about extreme temperatures (see, e.g., Davison and Gholamrezaee, 2012; Huser and Genton, 2016). The estimated endpoint  $\hat{\mu}_{s,t} - \hat{\sigma}/\hat{\xi}$  varies with the covariates (longitude, latitude, elevation, time) according to the estimated location surface  $\hat{\mu}_{s,t}$ . To check the marginal goodness-of-fit, we then transform the maxima  $\hat{Z}_{kj}$  to the standard Gumbel scale as  $\hat{\xi}^{-1} \log\{1 + \hat{\xi}(Z_{kj} - \hat{\mu}_{kj})/\hat{\sigma}\}$  by plugging in estimated parameters  $\hat{\mu}_{s,t}, \hat{\sigma}, \hat{\xi}$ , and we produce marginal quantile-quantile (QQ)-plots based on theoretical and empirical standard Gumbel quantiles for each station, and by pooling all stations together. Figure 5 displays QQ-plots for the pooled dataset and two randomly selected stations ( $s_{11}$  and  $s_{34}$ ). Overall, the marginal goodness-of-fit looks satisfactory, with the dots well aligned along the main diagonal for the vast majority of stations. To further examine the quality of the marginal fit, we perform a two-sided Kolmogorov-Smirnov test for the data from each station and for the pooled dataset, by comparing the empirical distributions and the fitted distributions. All the Kolmogorov-Smirnov tests fail to reject the null hypothesis (i.e., equality of distributions) with large p-values. Most of the p-values are greater than 0.9, and the smallest is 0.28, which does not provide evidence against



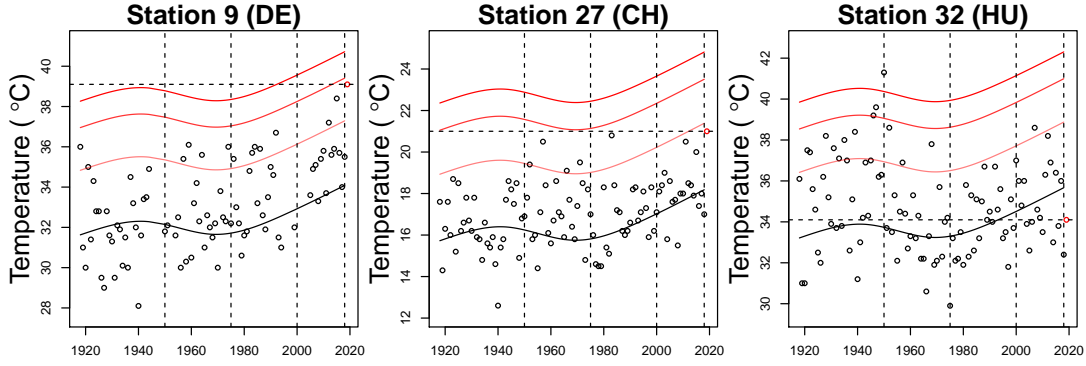


FIGURE 6. Plot of observed annual maxima (dots), estimated time trend (solid black), and estimated 10, 100 and 1000 year-return level curves (from light to dark red) for 3 selected stations located in Germany, Switzerland and Hungary (left to right). The red dot in each panel represents the observed annual maximum for 2019 (not used to fit the model). The four vertical dashed lines correspond to the reference years of 1950, 1975, 2000, and 2018.

good marginal fits at all monitoring stations. A marginal homogeneity analysis (see the Supplementary Material) also concur that there is not sufficient evidence to justify making the model more complex by including covariates or splines into the scale and shape parameters.

We then examine the fitted time trend, as well as the estimated time-varying  $M$ -year return level  $z_{s,t}^M$ , defined for each site  $s \in \mathcal{S}$  as the  $(1 - 1/M)$ -quantile from the fitted GEV distribution, i.e.,  $z_{s,t}^M = \widehat{\mu}_{s,t} - \widehat{\sigma}[\{-\log(1 - 1/M)\}^{-\widehat{\xi}} - 1]/\widehat{\xi}$ . Under temporal stationary conditions, the  $M$ -year return level is expected to be exceeded once every  $M$  years (at each site). With global warming, return levels from the past may be exceeded much more frequently in the present and future. In other words, observations that were extreme in the past may no longer be as rare under the current conditions. The effect of climate change can thus be assessed based on return levels. Figure 6 exhibits the estimated time trend and the corresponding 10, 100 and 1000 year-return level curves for 3 selected stations located in Germany, Switzerland and Hungary. The estimated time trend is at its lowest around 1975 and its highest in 2018, which corroborates other studies about climate change. The red dot in each plot represents the observed annual maximum for 2019 (not used to fit the model). For station 9, the 2019 annual maximum exceeds the 1000 year-return levels corresponding to 1950 and 1975. However, it barely reaches the 100 year-return level for 2018. For station 27, the 2019 annual maximum approximately corresponds to the 100-year event when taking 1975 as the reference year, but it becomes a 10-year event when taking 2018 as the reference. For these two stations in Germany and Switzerland, the 2019 heatwave was therefore very extreme compared to mid-20th century conditions, but only moderately extreme with respect to current climate. As for the station 32 in Hungary, our model suggests that the 2019 annual maximum was not very extreme overall (both with respect to past and current conditions).

*5.2. Spatial dependence structure and model comparison.* We now use the estimated marginal distributions and transform the data to the standard uniform  $\text{Unif}(0, 1)$  scale. We next estimate the dependence structure (i.e., the copula) through maximum pairwise likelihood inference using the max-id model introduced in §3.2–3.3. The most complex model that we fit is the non-stationary, but locally isotropic dependence structure, specified in §3.3 and Table 1, and we also fit several sub-models for comparison. Specifically, our most general model assumes that  $\Omega_{s,t}(R_i)$  in (9) has the form  $\Omega_{s,t}(R_i) = \lambda_{s,t}^2 (1 + R_i)^{-2\nu} I_{2 \times 2}$ ,  $\lambda_{s,t} = \exp(\lambda_0 + \lambda_1 \times \text{elev}_s + \lambda_2 \times t)$ . Using (5) for the mean measure of the Poisson points  $\{R_i\}$  arising in the spectral representation (3), the parameter vector to be estimated is thus

TABLE 2

Parameter estimates for the six max-id models fitted to annual European temperature maxima, with 95% confidence intervals (indicated as subscripts) based on the parametric bootstrap procedure described in §4.2 using 300 replications. Here,  $\hat{\lambda}_1$  and  $\hat{\lambda}_2$  represent the increase in  $\log \hat{\lambda}_{s,t}$  (log-range) per km in elevation, and per century in time, respectively.

	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}_0$	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\nu}$
Model 1	5.0 <sub>(3.5,10.0)</sub>	0	0.04 <sub>(-0.31,0.71)</sub>	0	0	0
Model 2	5.1 <sub>(3.7,10.0)</sub>	0	0.09 <sub>(-0.29,0.92)</sub>	-0.31 <sub>(-0.44,-0.13)</sub>	0.31 <sub>(-0.40,0.89)</sub>	0
Model 3	2.5 <sub>(0.5,6.5)</sub>	1.5 <sub>(0.4,3.9)</sub>	-0.35 <sub>(-0.60,0.19)</sub>	0	0	0
Model 4	2.5 <sub>(0.6,6.4)</sub>	1.5 <sub>(0.3,4.2)</sub>	-0.28 <sub>(-0.55,0.45)</sub>	-0.40 <sub>(-0.56,-0.17)</sub>	0.30 <sub>(-0.43,0.76)</sub>	0
Model 5	5.0 <sub>(0.5,9.9)</sub>	2.3 <sub>(1.1,9.8)</sub>	1.85 <sub>(0.60,3.88)</sub>	0	0	2.9 <sub>(1.2,6.0)</sub>
Model 6	5.5 <sub>(2.7,8.3)</sub>	2.4 <sub>(1.0,7.3)</sub>	2.12 <sub>(1.71,2.92)</sub>	-0.31 <sub>(-0.43,-0.12)</sub>	0.23 <sub>(-0.55,0.83)</sub>	3.2 <sub>(2.6,4.3)</sub>

$\psi = (\alpha, \beta, \lambda_0, \lambda_1, \lambda_2, \nu)^T \in (0, \infty)^2 \times \mathbb{R}^4$ . We compare this model to the fits of five simpler models, contained as special cases (or limiting cases) of our model, leading to features such as stationarity or max-stability. Specifically, Model 1 corresponds to the stationary extremal- $t$  max-stable process, and Model 3 to a stationary max-id model proposed by [Huser, Opitz and Thibaud \(2021\)](#). Models 2 and 4 are their non-stationary counterparts. Finally, Models 5 and 6 are our new stationary and non-stationary max-id models, with an explicit magnitude-dependent range of dependence. These six different models are specified with the following parameter configurations:

$$\left\{ \begin{array}{l} \text{Model 1 : } \{\alpha > 0, \beta \downarrow 0, \lambda_0 \in \mathbb{R}, \lambda_1 = 0, \lambda_2 = 0, \nu = 0\}, \text{ stationary max-stable} \\ \text{Model 2 : } \{\alpha > 0, \beta \downarrow 0, \lambda_0 \in \mathbb{R}, \lambda_1 \in \mathbb{R}, \lambda_2 \in \mathbb{R}, \nu = 0\}, \text{ non-stationary max-stable} \\ \text{Model 3 : } \{\alpha > 0, \beta > 0, \lambda_0 \in \mathbb{R}, \lambda_1 = 0, \lambda_2 = 0, \nu = 0\}, \text{ stationary simple max-id} \\ \text{Model 4 : } \{\alpha > 0, \beta > 0, \lambda_0 \in \mathbb{R}, \lambda_1 \in \mathbb{R}, \lambda_2 \in \mathbb{R}, \nu = 0\}, \text{ non-stationary simple max-id} \\ \text{Model 5 : } \{\alpha > 0, \beta > 0, \lambda_0 \in \mathbb{R}, \lambda_1 = 0, \lambda_2 = 0, \nu \in \mathbb{R}\}, \text{ stationary general max-id} \\ \text{Model 6 : } \{\alpha > 0, \beta > 0, \lambda_0 \in \mathbb{R}, \lambda_1 \in \mathbb{R}, \lambda_2 \in \mathbb{R}, \nu \in \mathbb{R}\}, \text{ non-stationary general max-id} \end{array} \right.$$

To assess the uncertainty of estimated parameters in the spatial dependence fit (without accounting for marginal uncertainty), we used the parametric bootstrap procedure with 300 bootstrap samples for each model as described in §4.2. The estimates and the 95% bootstrap confidence intervals are reported in Table 2. For comparison, the Supplementary Material also reports the standard errors of parameter estimates when approximated using the jackknife, whereby each replicate (i.e., one year) is left out at a time and the model refitted sequentially to the remaining  $n - 1$  years and all  $D = 44$  stations. The jackknife variance estimate is then obtained as a rescaled version of the sample variance of these  $n - 1$  parameter estimates. There are some differences when comparing the jackknife and bootstrap estimates, but generally speaking the estimated uncertainty is roughly of the same order in both cases, and our general conclusions about the asymptotic dependence class remain the same. Since the parametric bootstrap has been shown to be more widely applicable than the jackknife, and better deals with missing values, we believe that the confidence intervals shown in Table 2 are reliable. The estimates for  $\alpha$  and  $\beta$  are relatively large with lower confidence bounds clearly above 0, indicating that the data are asymptotically independent. Moreover, in Models 5 and 6, we obtain relatively large estimates  $\hat{\nu}$  with lower confidence bounds above 1 and 2, respectively, which suggests that the range of spatial dependence is substantially smaller for more severe extreme events. In all non-stationary models, the estimates for the elevation coefficient  $\lambda_1$  are significantly negative, such that the range of dependence diminishes in subregions with higher elevations. From our new Model 6,  $\hat{\lambda}_1 = -0.31$ , so the spatial extent of heatwaves is estimated to be about  $\exp(0.31) \approx 1.36$  smaller 1km higher in elevation. The estimates of  $\lambda_2$  are positive in all three non-stationary models, hinting that the spatial extent of heatwaves has increased in recent years, and Model 6 suggests that it increases by a factor

TABLE 3

Ranking of the 10 models using the cross-validated logarithmic score (15) for pairwise predictions. Lower rank means better predictive performance.

Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
9	8	4	3	2	1	10	7	6	5

about  $\exp(0.23) \approx 1.26$  per century. Qualitatively, this confirms the general findings of previous studies such as [Osborn and Briffa \(2006\)](#) and [Lyon et al. \(2019\)](#), though here this effect is not significant based on the available data.

To assess the relative goodness-of-fit and test the predictive performance of the six models, we use a cross-validation scheme, whereby each station  $s_{j_0}$ ,  $j_0 = 1, \dots, D$ , is left out at a time and the six models refitted. We then compare the models using the (negatively oriented) logarithmic score,

$$(15) \quad \text{LogS}_{j_0} = \sum_{j \neq j_0} \left[ \sum_{k=1}^n V(z_{kj}, z_{kj_0}) - \log\{V_1(z_{kj}, z_{kj_0})V_2(z_{kj}, z_{kj_0}) - V_{12}(z_{kj}, z_{kj_0})\} \right],$$

which is the sum of the negated log pairwise-densities by considering only the pairs composed of the left-out station  $s_{j_0}$  and one of the other stations  $s_j$ ,  $j \neq j_0$ . Logarithmic scores are strictly proper in the sense of [Gneiting and Raftery \(2007\)](#), such that they enable us to appropriately compare the predictive power of different models. The final score of a model is obtained by summing scores for all stations, i.e., using  $\text{LogS} = \sum_{j_0=1}^D \text{LogS}_{j_0}$ . The lower LogS, the better the model at predicting pairwise interactions. In our model comparison, we also include traditional geostatistical models from the spatial statistics literature, which do not have the strong theoretical motivation from Extreme-Value Theory. Precisely, we also fit the Gaussian copula and the Student- $t$  copula models with  $\alpha > 0$  degrees of freedom, using the same stationary or non-stationary correlation function as before. For consistency, we use the same pairwise likelihood inference approach. We label these models as follows: Model 7 is the stationary Gaussian copula model; Model 8 is its non-stationary counterpart; Model 9 is the stationary Student- $t$  copula model; Model 10 is its non-stationary counterpart.

The final ranking of all models based on the logarithmic score is reported in [Table 3](#). Interestingly, the ‘‘traditional’’ models from spatial statistics and Extreme-Value Theory, namely the Gaussian copula (Models 7, 8) and max-stable (Models 1, 2) models, perform worst. Furthermore, the non-stationary Gaussian copula (Model 8) outperforms its max-stable counterpart (Model 2) despite the additional parameters of the latter, which casts strong doubts about the max-stability assumption and suggests that the dependence strength of maxima weakens at higher quantiles. The four estimated max-id (but not max-stable) models have the best results, and the most complex model that we propose (Model 6), which includes covariate effects of elevation and time, as well as the magnitude-dependent probabilistic structure, performs the best overall. Finally, the non-stationary Student- $t$  copula (Model 10) ranks 5th, right behind the max-id (non-max-stable) models. Its flexible structure—being at the same time in the domain of attraction of the max-stable extremal- $t$  limit, and also very close to the Gaussian copula for large degrees of freedom—seems to compensate for some of the weaknesses of max-stable and Gaussian copula models.

We then conduct a parametric bootstrap simulation experiment, in order to confirm our conclusions from this model comparison, assess the uncertainty of the ranking, and remove any model selection bias. Precisely, we simulate 50 datasets according to the best model (Model 6), where we use the same sample size and structure of missing values as in the real dataset. For each of the 50 simulated datasets, we then refit the 10 different models and recompute the ranking based on the logarithmic score, LogS. This gives 50 rankings for the

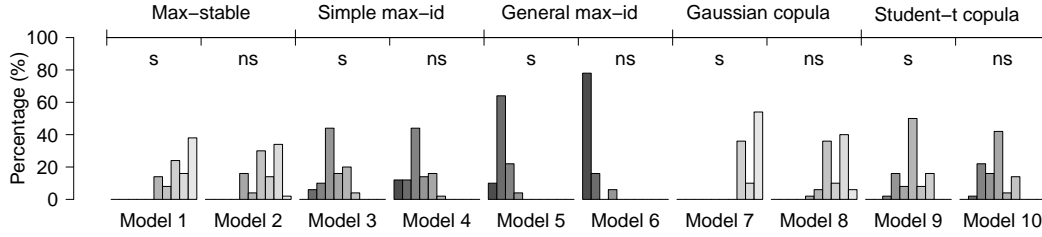


FIGURE 7. Logarithmic score (15) ranks computed for 50 bootstrap simulations from the fitted Model 6, for each of the 10 models. The bars show the percentage of ranks ranging from 1 (best score) to 10 (worst score) for each of the Models 1–10, with darker grey (towards the left side of each histogram) indicating better rank. The letter “s” means “stationary”, whereas “ns” means “non-stationary”.

models 1–10. Figure 7 shows the percentage of times that a given model was ranked 1st to 10th, with high bars towards the left being generally better. We note that although the most complex model (Model 6) was used to resample the datasets in this bootstrap experiment, the cross-validation criterion that we use is designed to identify the model that gives the best prediction of bivariate interactions, which could well be a simpler, more parsimonious and robust model. It is indeed well known that the best-fitting models are not necessarily the best ones for prediction. In other words, our parametric bootstrap is fair as our criterion does not only take the model fit into consideration but it also implicitly penalizes model complexity. From Figure 7, we clearly see that our most complex max-id models (Models 5 and 6) have the best performances, and Model 6 is ranked 1st overall in about 40 out of the 50 cases. This Monte Carlo experiment therefore confirms our initial findings and the advantage of the very flexible dependence structure of our proposed Model 6 with respect to the other models.

If a model appropriately captures the dependence structure of the data, it is expected that the fitted extremal coefficients  $\hat{\theta}_D(z)$  from the model are close to the empirical extremal coefficients  $\hat{\theta}_D^{\text{emp}}(z) = -z \log[\widehat{\text{Pr}}\{Z(s_1) \leq z, \dots, Z(s_D) \leq z\}]$  at level  $z$  (assuming here unit Fréchet marginals), where  $\widehat{\text{Pr}}$  is the empirical probability measure. Since Models 2, 4 and 6 are non-stationary, empirical extremal coefficients are more tricky to estimate accurately in these cases. Therefore, for simplicity, we here only compare the fitted extremal coefficients of Models 1, 3 and 5, which are the stationary versions of the max-stable model, the simple max-id model of [Huser, Opitz and Thibaud \(2021\)](#) and our proposed general max-id model, respectively, with their empirical counterparts in dimensions  $D = 2-20$ . In dimensions  $D = 2$  and 3, we computed extremal coefficients for all pairs and triplets of the 44 stations, whereas in higher dimensions, we only computed coefficients for a maximum of 1000 randomly sampled combinations of stations among the 22 stations without missing values. Figure 8 shows the average absolute difference between the empirical and fitted extremal coefficients  $\theta_D(z)$  in dimensions  $D = 2-20$  at unit Fréchet quantile levels  $z = -1/\log(q)$  with  $q = 0.25, 0.5, 0.75$  and 0.95 for Models 1, 3 and 5. Notice that under stationarity, these levels are on average marginally exceeded 3 times in 4 years, once in 2 years, once in 4 years, and once in 20 years, respectively, so they correspond to moderately extreme events. All three models are comparable for moderate quantiles  $q = 0.5$  and  $q = 0.75$  representing the behavior in the bulk of the max-id distribution. The relatively complex max-id Model 5 (green curve) performs sensibly better than the max-stable model (black curve) and the simple max-id model (red curve) at quantile levels  $q = 0.25$  and 0.95, especially in higher dimensions. Model 5 thus better captures the dependence structure of spatial extreme events of relatively small and large magnitudes. Throughout, the observed absolute differences are not excessively large compared to the theoretical range  $[1, D]$  of extremal coefficients.

To further assess the goodness-of-fit and verify the fidelity of our fitted max-id Model 5 to the data, Figure 9 compares boxplots of empirical bivariate coefficients  $\theta_2(z)$  for all pairs of

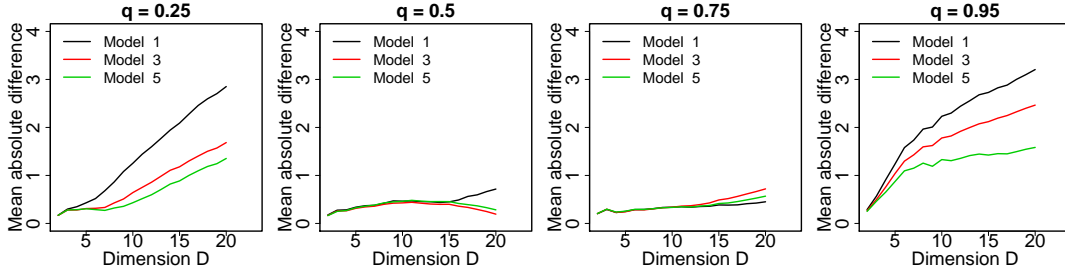


FIGURE 8. Mean absolute difference between empirical and fitted extremal coefficients  $\theta_D(z)$ , plotted with respect to dimension  $D = 2-20$ , for Models 1 (black), 3 (red), and 5 (green), at unit Fréchet quantile levels  $z = -1/\log(q)$  with  $q = 0.25, 0.5, 0.75$  and  $0.95$  (left to right).

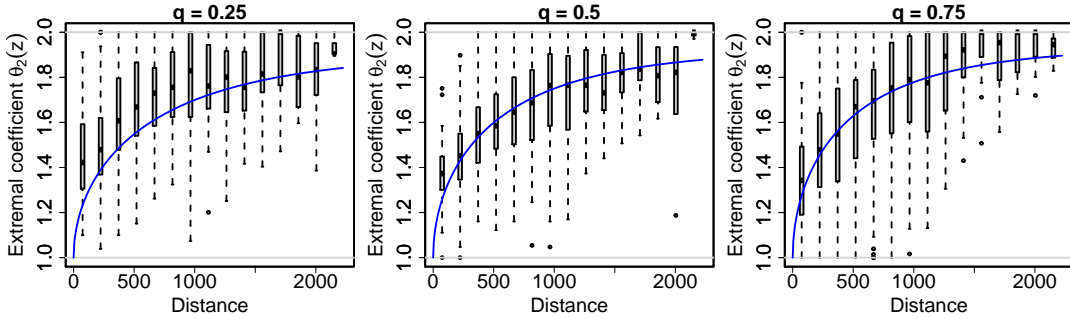


FIGURE 9. Boxplots of empirical bivariate extremal coefficients  $\theta_2(z)$  computed for all pairs of sites within 15 equally spaced distance classes, and theoretical curve (blue) based on the fitted Model 5 plotted with respect to spatial distance, for unit Fréchet quantile levels  $z = -1/\log(q)$  with  $q = 0.25, 0.5$  and  $0.75$  (left to right).

sites within 15 equally spaced distance classes to their model-based counterparts, for three different quantile levels  $z$ . Although the variability of estimated bivariate extremal coefficients is quite high, the fitted curves are quite close to the boxplot medians and seem to adequately capture the decay of spatial dependence with distance. Our fitted model suggests that extremal dependence persists at very large distances, which is consistent with heatwaves being large-scale phenomena with the potential of simultaneously affecting large parts of Europe. We also verify the goodness-of-fit in higher dimension, which is important here since the spatial extent of extreme episodes (e.g., the size of hotspot areas) may be driven by highly multivariate (and not only bivariate) extremal behavior. Figure 10 shows scatterplots of empirical versus fitted extremal coefficients  $\theta_D(z)$  for Model 5 in dimensions  $D = 2, 5, 10$  for unit Fréchet quantile levels  $z = -1/\log(q)$  with  $q = 0.25, 0.5$  and  $0.75$ . The dots tend to concentrate around the main diagonal, especially in high dimensions, which confirms a satisfactory model fit. Nevertheless, the fitted model tends to be slightly smoother in general than empirical data in terms of the range of values of empirical coefficients, but such behavior can be expected since our model cannot perfectly capture all the non-stationarities of extremal dependence arising over this very large and geographically heterogeneous study region. While the stationary max-id Model 5 already produces a very decent fit, our non-stationary Model 6 is expected to perform even better.

In order to visualize the spatio-temporal variation in the estimated extremal dependence structure, and to assess whether the spatial extent of heatwaves has changed a lot over time due to climate change, we then compute the effective extremal dependence range for 1918 and 2018, based on the fitted non-stationary Model 6. We define the effective extremal dependence range (at a given point in space and time) as the minimum spatial distance (from that

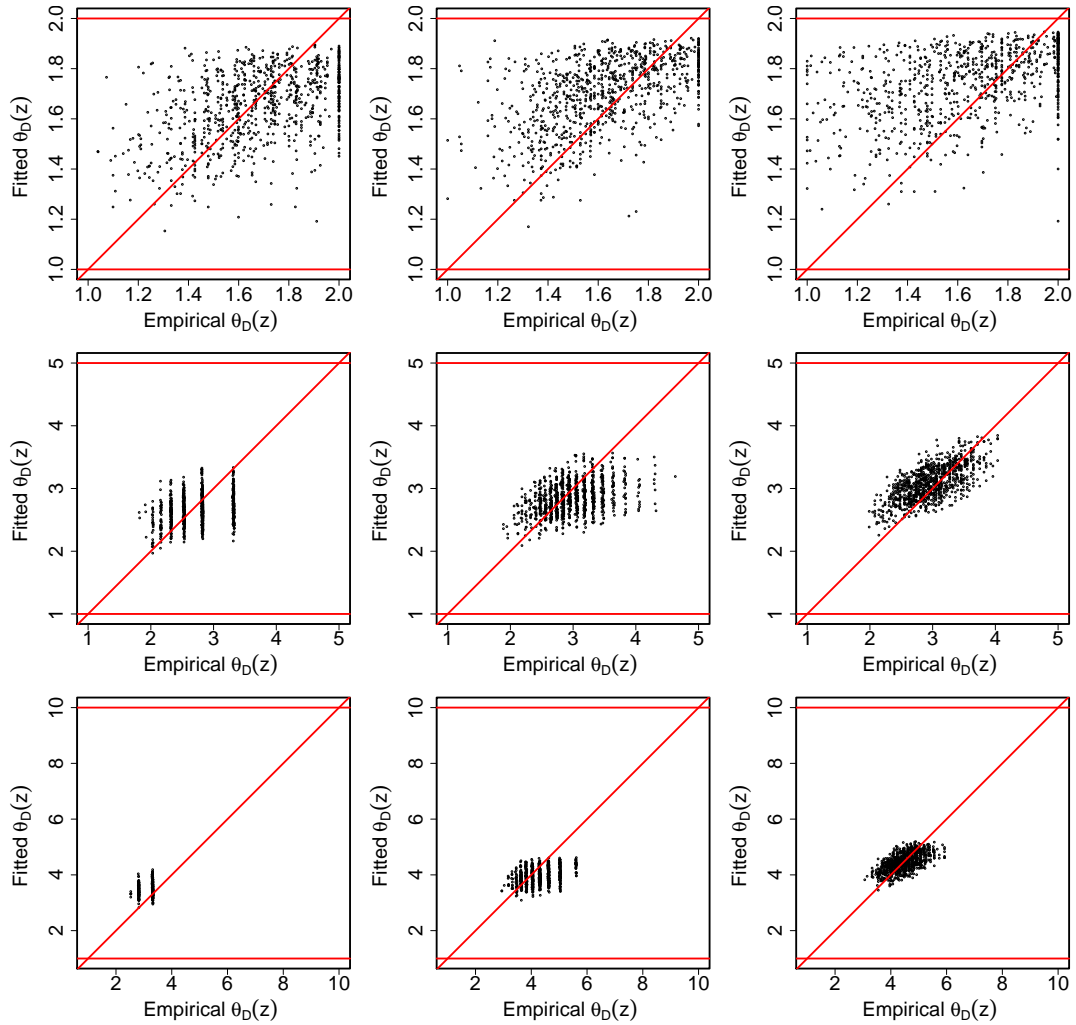


FIGURE 10. Scatterplots of empirical versus fitted extremal coefficients  $\theta_D(z)$  for Model 5 in dimensions  $D = 2, 5, 10$  (top to bottom) for unit Fréchet quantile levels  $z = -1/\log(q)$  with  $q = 0.25, 0.5, 0.75$  (left to right). The main diagonal indicates a perfect fit.

point) such that  $\theta_2(z) = 1.95$  for a given level  $z$ , under constant covariate values. Figure 11 displays a map of the results for 1918 taking  $z$  as the level  $z = -1/\log(0.9)$ , as well as the difference between the results for 2018 and 1918. From the top panel, we can see that the effective extremal dependence range varies from about 400km at high elevations to 1500km at low elevations. Elevation is thus a major (significant) covariate, though we should also be careful not to over-interpret the results as the model may not perfectly extrapolate far outside the range of covariate values used for fitting. From the bottom panel, we see that our Model 6 estimates the change in extremal dependence range over the last century to be between about 150km at high elevations and 400km at low elevations. Heatwaves might therefore have become slightly larger in extent, especially at low elevations. In the Supplementary Material, we compare the effective extremal dependence range estimated for Models 2, 4, 5, and 6, in order to show how our assessment of risk based on Model 6 differs because (i) it is asymptotically independent; (ii) it has a magnitude-dependent correlation structure, and (iii) it is non-stationary. The results show major differences across models, suggesting that all the assumptions (i), (ii) and (iii) are key in practice. This shows once again the importance of

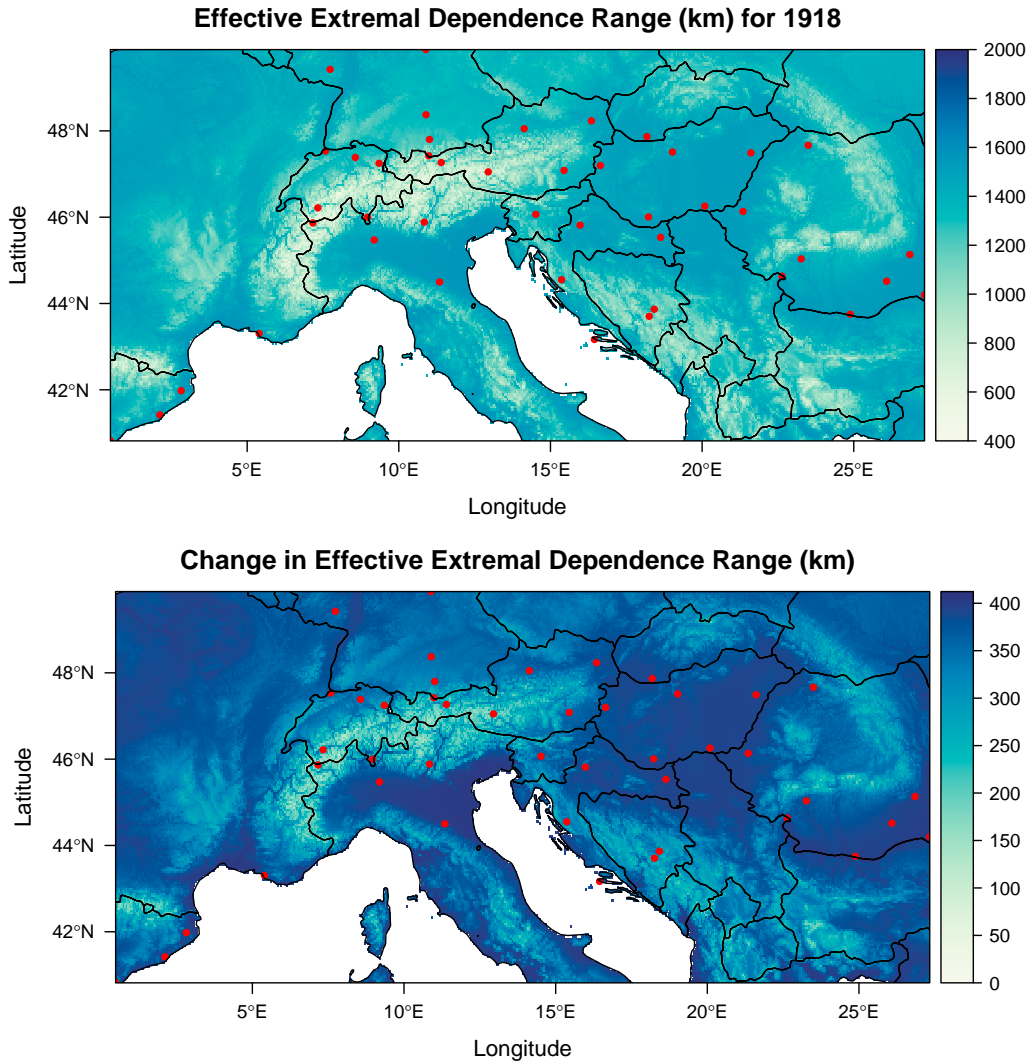


FIGURE 11. Map of the effective extremal dependence range (km) for 1918 (top), and of the difference between 2018 and 1918 (bottom), based on Model 6. The extremal dependence range is here defined as the minimum spatial distance such that  $\theta_2(z) = 1.95$  at the level  $z = -1/\log(0.9)$ . Monitoring stations are shown as red dots.

correctly specifying the dependence structure for an accurate risk assessment of spatial extreme events, and we believe that Model 6 is the most suitable in our case based on various model diagnostics.

*5.3. Probabilistic assessment of the 2019 European heatwaves.* We conclude our real data analysis with a probabilistic assessment of the extremes observed during the 2019 European heatwaves, which affected large parts of Europe. Over the summer 2019, many monitoring stations across Europe indeed recorded the highest temperature in almost a century. A natural question is whether this could have been anticipated from historical data. To assess the severity of the 2019 European heatwaves, we here complement the marginal analysis of §5.1, by simulating  $10^5$  replicates from our best fitted non-stationary Model 6 at 31 stations for which the 2019 annual maxima are available, and transforming these simulated data to their estimated marginal GEV scales. From these  $10^5$  replicates, we then compute empirical return periods for the spatial maximum, spatial minimum and spatial average of the observed 2019

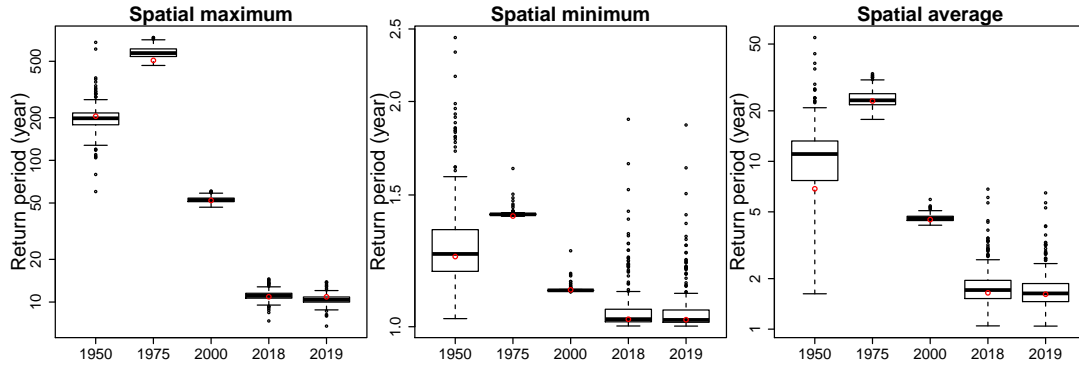


FIGURE 12. Boxplots of bootstrapped return periods (on logarithmic scale) of the 2019 annual maximum, computed for the spatial maximum, minimum and average (left to right) over 31 stations, based on  $10^5$  random fields simulated from our best non-stationary spatial Model 6 (fitted to the annual maxima for the period 1918–2018), with respect to 1950, 1975, 2000, 2018 and 2019 as reference years. Red dots are pointwise estimates of these return periods.

maxima, with respect to the reference years 1950, 1975, 2000, 2018 and 2019. To estimate the variability of our return period estimates, we use the 300 bootstrap fits and recompute these return periods. Figure 12 shows boxplots of the bootstrapped return periods, as well as the point estimates (red dots). Due to the estimated time trend (both in margins and dependence), return periods are always highest when compared to 1975 and lowest when compared to 2018–2019. When considering return periods for the spatial maximum (left panel), which is large when at least one site experiences an extreme event, we get a return period of about 500 years when compared to the climatic conditions of 1975, but only about 10 years when compared to current climatic conditions. When considering the spatial average, we get a return period of about 20–30 years when compared to 1975, but only 2 years for 2018–2019. Finally, when considering the spatial minimum, which is large only when all sites experience simultaneous extreme events, and which is usually observed at one of the locations in the Alps, the 2019 heatwaves were not especially extreme, corresponding only to a 1–1.5 year event for all reference years.

**6. Conclusion.** We have proposed a non-stationary max-id spatial model for block maxima, which embeds spatio-temporal covariates in its dependence structure, while having a very flexible form of weakening dependence strength at increasingly high quantiles, in order to model extreme temperatures over Southern Europe. A novelty of our proposed approach is that, unlike contributions from the existing climate literature, it exploits extreme-value theory and real measurements from a century-spanning dataset to estimate the spatial extent of heatwaves through the estimated effective extremal dependence range. Our fitted models reveal that the dependence structure of temperature annual maxima is significantly weaker at higher elevations, and similarly for more severe heatwaves. The estimated parameters of our models with temporal non-stationarity also suggest that the spatial extent of heatwaves has become wider in recent years, confirming the general findings of previous studies, though this effect was not significant based on our parametric bootstrap procedure.

Modeling approaches in classical Gaussian-based geostatistics and spatial extreme-value analysis often use a setting where the dependence structure is stationary over both space and time. This assumption is problematic when spatial and temporal scales are large and lead to heterogeneous regional and temporal characteristics in co-occurrence patterns of extreme values, and even more so when we aim to detect and analyze such patterns. The max-id models developed in this paper are a step forward towards more accurate inference while keeping



parsimonious specifications. Trends in dependence are notoriously difficult to estimate when data are not abundant, and one has to carefully avoid confusion with marginal trends. Indeed, the accurate modeling of marginal trends in extremes remains of paramount importance, and it is a prerequisite to avoid estimating spurious trends in dependence models. In our real data application, we implemented semi-parametric spline functions for capturing marginal trends in the GEV parameters, and we opted for a flexible tensor product specification to allow interaction of trends arising in latitude, longitude and elevation. We also assessed spatial return periods associated with the 2019 Europe heatwaves over Southern Europe, and concluded that the summer 2019 was very extreme when considering the spatial maximum over the monitoring stations (especially compared to mid-20th century conditions), moderately extreme when considering the spatial average, and not especially extreme when considering the spatial minimum. Furthermore, our analysis provided clear evidence for climate change and its impact on spatial extreme temperature events.

Finally, we underline the main methodological novelty of building magnitude-dependent max-id models, where the spatial dependence range becomes shorter as events become more extreme. Our construction *explicitly* accounts for this behavior, and allows us to capture in a single parsimonious parametric model: (i) max-stable asymptotic dependence; (ii) weakening asymptotic dependence; (iii) weakening asymptotic independence. By keeping a flexible max-stable process on the boundary of the parameter space, our proposed model achieves the subtle trade-off of combining the strength of theoretically-motivated max-stable models together with the pragmatism of flexible max-id extensions with weakening dependence strength. Our sophisticated extreme-value model, combined with covariates and geometric anisotropy, thus provides a very rich class of models for spatially-indexed block maxima, and opens the door to more realistic risk assessment of extreme environmental events.

**Acknowledgements.** The authors would like to thank the Editor, Associate Editor and two referees for valuable suggestions that have improved the manuscript. This publication is based upon work supported by the King Abdullah University of Science and Technology (KAUST) Office of Sponsored Research (OSR) under Award No. OSR-CRG2017-3434.

## SUPPLEMENTARY MATERIAL

### **Supplementary Material for ‘Modeling Non-Stationary Temperature Maxima Based on Extremal Dependence Changing with Event Magnitude’.**

We provide further results in our simulation study for alternative parameter settings, and also provide further details and simulation results on the performance of our proposed parametric bootstrap inference scheme in the context of our real data application. We also conduct a detailed homogeneity analysis to assess the validity of the marginal model structure assumed in our data application, and provide further results on the effect of model assumptions on the estimated spatial extremal dependence range.

## REFERENCES

- BALKEMA, A. A., DE HAAN, L. and KARANDIKAR, R. L. (1993). Asymptotic distribution of the maximum of  $n$  independent stochastic processes. *Journal of Applied Probability* **30** 66–81.
- BALLESTER, F., CORELLA, D., PÉREZ-HOYOS, S., SÁEZ, M. and HERVÁS, A. (1997). Mortality as a function of temperature. A study in Valencia, Spain, 1991–1993. *International Journal of Epidemiology* **26** 551–561.
- BOPP, G., SHABY, B. A. and HUSER, R. (2020). A hierarchical max-infinitely divisible spatial model for extreme precipitation. *Journal of American Statistical Association* **116** 93–106.
- BROWN, A. (2016). Heatwave mortality. *Nature Climate Change* **6** 821.
- BROWN, B. M. and RESNICK, S. I. (1977). Extreme values of independent stochastic processes. *Journal of Applied Probability* **14** 732–739.

- 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** 1212–129.
- DAVISON, A. C. and GHOLAMREZAEI, M. M. (2012). Geostatistics of extremes. *Proceedings of the Royal Society A: Mathematical, Physical & Engineering Sciences* **468** 581–608.
- DAVISON, A. C. and HINKLEY, D. V. (1997). *Bootstrap Methods and their Application*. Cambridge University Press.
- DAVISON, A. C. and HUSER, R. (2015). Statistics of extremes. *Annual Review of Statistics and its Application* **2** 203–235.
- DAVISON, A. C., HUSER, R. and THIBAUD, E. (2019). Spatial extremes. In *Handbook of Environmental and Ecological Statistics* (A. E. Gelfand, M. Fuentes, J. A. Hoeting and R. L. Smith, eds.) 711–744. CRC Press.
- DAVISON, A. C., PADOAN, S. and RIBATET, M. (2012). Statistical modelling of spatial extremes (with Discussion). *Statistical Science* **27** 161–186.
- DAVISON, A. C. and SMITH, R. L. (1990). Models for exceedances over high thresholds (with discussion). *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **52** 393–442.
- DE HAAN, L. (1984). A spectral representation for max-stable processes. *The Annals of Probability* **12** 1194–1204.
- DOMBRY, C. and EYI-MINKO, F. (2013). Regular conditional distributions of continuous max-infinitely divisible random fields. *Electronic Journal of Probability* **18** 1–21.
- EFRON, B. (1979). Bootstrap methods: another look at the jackknife. *Annals of Statistics* **7** 1–26.
- FIELD, C. B., BARROS, V., STOCKER, T. F., QIN, D., DOKKEN, D. J., EBI, K. L., MAstrandrea, M. D., MACH, K. J., PLATTNER, G. K., ALLEN, S. K., TIGNOR, M. and MIDGLEY, P. M. (2012). *IPCC, 2012: Managing the risks of extreme events and disasters to advance climate change adaptation. A special report of Working Groups I and II of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK.
- FISHER, R. A. and TIPPETT, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society* **24** 180.
- GENEST, C., GHOUDI, K. and RIVEST, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* **82** 543–552.
- GINÉ, E., HAHN, M. G. and VATAN, P. (1990). Max-infinitely divisible and max-stable sample continuous processes. *Probability Theory and Related Fields* **87** 139–165.
- GNEITING, T. and RAFTERY, A. E. (2007). Strictly Proper Scoring Rules, Prediction, and Estimation. *Journal of the American Statistical Association* **102** 359–378.
- GORDO, O. and SANZ, J. J. (2010). Impact of climate change on plant phenology in Mediterranean ecosystems. *Global Change Biology* **16** 1082–1106.
- GRABHERR, G., GOTTFRIED, M. and PAULI, H. (2010). Climate change impacts in alpine environments. *Geography Compass* **4** 1133–1153.
- HUSER, R. and DAVISON, A. C. (2013). Composite likelihood estimation for the Brown–Resnick process. *Biometrika* **100** 511–518.
- HUSER, R. and DAVISON, A. C. (2014). Space-time modelling of extreme events. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **76** 439–461.
- HUSER, R., DAVISON, A. C. and GENTON, M. G. (2016). Likelihood estimators for multivariate extremes. *Extremes* **19** 79–103.
- 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., OPITZ, T. and THIBAUD, E. (2021). Max-infinitely divisible models and inference for spatial extremes. *Scandinavian Journal of Statistics* **48** 321–348.
- HUSER, R. and WADSWORTH, J. L. (2019). Modeling spatial processes with unknown extremal dependence class. *Journal of the American Statistical Association* **114** 434–444.
- HUSER, R. and WADSWORTH, J. L. (2020). Advances in statistical modeling of spatial extremes. *Wiley Interdisciplinary Reviews (WIREs): Computational Statistics* **e1537**. To appear.
- HUSER, R., DOMBRY, C., RIBATET, M. and GENTON, M. G. (2019). Full likelihood inference for max-stable data. *Stat* **8** e218.
- JOE, H. (2005). Asymptotic efficiency of the two-stage estimation method for copula-based models. *Journal of Multivariate Analysis* **94** 401–419.
- JOE, H. (2015). *Dependence Modeling with Copulas*. Chapman and Hall/CRC, New York.
- JOE, H. and XU, J. J. (1996). The estimation method of inference functions for margins for multivariate models Technical Report #166, The University of British Columbia, Vancouver, Canada.

- JUN, M., KNUTTI, R. and NYCHKA, D. W. (2008). Spatial analysis to quantify numerical model bias and dependence: How many climate models are there? *Journal of the American Statistical Association* **103** 934–947.
- KABLUCHKO, Z., SCHLATHER, M. and DE HAAN, L. (2009). Stationary max-stable fields associated to negative definite functions. *Annals of Probability* **37** 2042–2065.
- KLAUSMEYER, K. R. and SHAW, M. R. (2009). Climate change, habitat loss, protected areas and the climate adaptation potential of species in Mediterranean ecosystems worldwide. *PloS one* **4** e6392.
- KLEIN TANK, A. M. G., WIJNGAARD, J. B., KÖNNEN, G. P., BÖHM, R., DEMARÉE, G., GOCHEVA, A., MILETA, M., PASHIARDIS, S., HEJKRLIK, L., KERN-HANSEN, C., HEINO, R., BESSEMOULIN, P., MÜLLER-WESTERMEIER, G., TZANAKOU, M., SZALAI, S., PÁLSDÓTTIR, T., FITZGERALD, D., RUBIN, S., CAPALDO, M., MAUGERI, M., LEITASS, A., BUKANTIS, A., ABERFELD, R., VAN ENGELEN, A. F. V., FORLAND, E., MIETUS, M., COELHO, F., MARES, C., RAZUVAEV, V., NIEPLOVA, E., CEGNAR, T., LÓPEZ, J. A., DAHLSTRÖM, B., MOBERG, A., KIRCHHOFER, W., CEYLAN, A., PACHALIUK, O., ALEXANDER, L. V. and PETROVIC, P. (2002). Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. *International Journal of Climatology* **22** 1441–1453.
- LEDFORD, A. W. and TAWN, J. A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika* **83** 169–187.
- LYON, B., BARNSTON, G. ANTHONY, COFFEL, E. and HORTON, R. M. (2019). Projected increase in the spatial extent of contiguous US summer heat waves and associated attributes. *Environmental Research Letters* **14** 114029.
- MITCHELL, D., HEAVISIDE, C., VARDOLAKIS, S., HUNTINGFORD, C., MASATO, G. and GUILLOD, B. P. ET AL. (2016). Attributing human mortality during extreme heat waves to anthropogenic climate change. *Environmental Research Letters* **11** 074006.
- 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.
- OSBORN, T. J. and BRIFFA, K. R. (2006). The Spatial Extent of 20th-Century Warmth in the Context of the Past 1200 Years. *Science* **311** 841–844.
- PACIOREK, C. J. and SCHERVISH, M. (2006). Spatial modelling using a new class of nonstationary covariance functions. *Environmetrics* **17** 483–506.
- PADOAN, S. A. (2013). Extreme dependence models based on event magnitude. *Journal of Multivariate Analysis* **122** 1 - 19.
- PADOAN, S. A., RIBATET, M. and SISSON, S. A. (2010). Likelihood-based inference for max-stable processes. *Journal of the American Statistical Association* **105** 263–277.
- QUENOUILLE, M. H. (1949). Approximate tests of correlation in time-series. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **11** 68–84.
- QUENOUILLE, M. H. (1956). Notes on bias in estimation. *Biometrika* **43** 353–360.
- REICH, B. J. and SHABY, B. A. (2012). A hierarchical max-stable spatial model for extreme precipitation. *Annals of Applied Statistics* **6** 1430–1451.
- RESNICK, S. I. (1987). *Extreme Values, Regular Variation and Point Processes*. Springer-Verlag New York.
- SCARROTT, C. and MACDONALD, A. (2012). A Review of Extreme Value Threshold Estimation And Uncertainty Quantification. *REVSTAT* **10** 33–60.
- SCHLATHER, M. (2002). Models for stationary max-stable random fields. *Extremes* **5** 33–44.
- SHOOTER, R., ROSS, E., TAWN, J. A. and JONATHAN, P. (2019). On spatial conditional extremes for ocean storm severity. *Environmetrics* **e2562**.
- THEURILLAT, J.-P. and GUISAN, A. (2001). Potential impact of climate change on vegetation in the European Alps: a review. *Climatic change* **50** 77–109.
- TUKEY, J. W. (1958). Bias and confidence in not quite large samples (Abstract). *Annals of Mathematical Statistics* **29** 614.
- VARIN, C., REID, N. and FIRTH, D. (2011). An overview of composite likelihood methods. *Statistica Sinica* **21** 5–42.
- WADSWORTH, J. L. and TAWN, J. A. (2012). Dependence modelling for spatial extremes. *Biometrika* **99** 253–272.
- WADSWORTH, J. L. and TAWN, J. A. (2019). Higher-dimensional spatial extremes via single-site conditioning. arXiv preprint 1912.06560.