

# **A model for space-time threshold exceedances with an application to extreme rainfall**

**Paola Bortot<sup>1</sup> and Carlo Gaetan<sup>2</sup>**

<sup>1</sup> Dipartimento di Scienze Statistiche, Università di Bologna, Italy

<sup>2</sup> Dipartimento di Scienze Ambientali, Informatica e Statistica, Università Ca' Foscari  
- Venezia, Italy

---

**Address for correspondence:** Carlo Gaetan, Dipartimento di Scienze Ambientali,  
Informatica e Statistica, Università Ca' Foscari di Venezia, Via Torino 155, I-30172  
Mestre (VE), Italy .

**E-mail:** [gaetan@unive.it](mailto:gaetan@unive.it).

**Phone:** (+39) 041 234 8404.

**Fax:** (+39) 041 234 8584.

---

**Abstract:** In extreme value studies models for observations exceeding a fixed high threshold have the advantage of exploiting the available extremal information, while avoiding bias from low values. In the context of space-time data, the challenge is to develop models for threshold exceedances that account for both spatial and temporal dependence. We address this issue through a modelling approach that embeds spatial dependence within a time series formulation. The model allows for different forms

of limiting dependence in the spatial and temporal domains as the threshold level increases. In particular, temporal asymptotic independence is assumed, as this is often supported by empirical evidence, especially in environmental applications, while both asymptotic dependence and asymptotic independence are considered for the spatial domain. Inference from the observed exceedances is carried out through a combination of pairwise likelihood and a censoring mechanism. For those model specifications for which direct maximization of the censored pairwise likelihood is unfeasible, we propose an indirect inference procedure which leads to satisfactory results in a simulation study. The approach is applied to a dataset of rainfall amounts recorded over a set of weather stations in the North Brabant province of the Netherlands. The application shows that the range of extremal patterns that the model can cover is wide and that it has a competitive performance with respect to an alternative existing model for space-time threshold exceedances.

---

**Key words:** Asymptotic dependence, Asymptotic independence, Gaussian spatial process, Indirect inference, Max-stable process, Student's  $t$  spatial process.

## 1 Introduction

Extreme value analyses of environmental phenomena are typically hindered by a scarcity of data, but if time series of observations are available at different sites in a region, there is the potential to alleviate this difficulty by aggregating information both spatially and temporally. Furthermore, the exchange of spatial information can be exploited to study the extremal features of the phenomenon at unobserved loca-

tions. We will focus on observations above a fixed high level as they allow the study of both the spatial pattern and temporal evolution of extremes, while avoiding potential bias from the inclusion of low values. The challenge is to develop a model for threshold exceedances that accounts for their spatial and temporal dependence. We address this issue for a data structure commonly encountered in practice, characterized by recordings at regular times over a set of sites that are close enough in space to induce a non-negligible spatial dependence.

A first step when studying dependence of threshold exceedances is to distinguish between asymptotic dependence and asymptotic independence (Coles et al., 1999). Broadly speaking, these two limiting forms are characterized, respectively, by persistence of dependence or convergence to independence among observations exceeding asymptotically increasing thresholds. More formally, let  $Y_1$  and  $Y_2$  be continuous random variables with cumulative distribution functions (CDFs)  $F_1$  and  $F_2$ , respectively, and let

$$\chi(p) := \Pr(F_2(Y_2) > p | F_1(Y_1) > p), \quad 0 \leq p < 1. \quad (1.1)$$

Then,  $Y_1$  and  $Y_2$  are said to be asymptotically independent if the limit  $\chi := \lim_{p \rightarrow 1^-} \chi(p)$  is zero and asymptotically dependent if  $\chi > 0$ , respectively. As many authors have pointed out (Ledford and Tawn, 1997; Heffernan and Tawn, 2004; Wadsworth and Tawn, 2012; Huser and Wadsworth, 2019, 2020; Simpson and Wadsworth, 2021) correct identification of the limiting dependence is a fundamental requirement, but is not enough to guarantee reliable predictions of joint upcrossings, as the rate of convergence to the asymptotic form is also important. While the value of  $\chi$  is used to discriminate between the two limiting classes, the function  $\chi(p)$  is a natural choice to assess the speed of convergence to the limit (see, for instance, Wadsworth and Tawn,

2012, Sec. 5.4).

In the context of space-time data, different forms of asymptotic dependence could arise in the spatial and temporal domains. The limiting class could itself change with the temporal and/or spatial lag. For some environmental phenomena, the assumption of asymptotic dependence in space is realistic when working within relatively small regions (Davison and Gholamrezaee, 2012; Davison et al., 2013; Bacro et al., 2016). On the other hand, for the time domain, empirical evidence in many studies supports asymptotic independence even at short time lags (Bortot and Tawn, 1998; Bortot and Gaetan, 2014; Simpson and Wadsworth, 2021). As an illustrative example, we consider a dataset of daily rainfall recorded from 01-10-1999 to 28-02-2019 at 28 meteorological stations located over the North Brabant province in the south of the Netherlands, as shown in Figure 1 (Klein Tank et al., 2002). A detailed description of this dataset will be given in Section 4. Here we focus attention on the behaviour of the empirical estimates of  $\chi(p)$  for pairs of sites at increasing spatial distances, averaging across all times, and for pairs of measurement times at increasing temporal lags, averaging across all sites. The two sets of estimates for the months from October to February are displayed in Figures 2(a) and 2(b), respectively. To study the convergence to the limiting value  $\chi$ ,  $p$  varies, taking values 0.90, 0.95, and 0.99. The shaded area around each curve represents 95% pointwise confidence intervals obtained by 200 replications of a block bootstrap of the original series with average block length of 300 days. The empirical estimates decrease with both spatial and temporal distances, i.e., a weakening of dependence occurs as one moves farther out in space and time, as expected. As  $p$  increases, corresponding to increasing threshold levels, both figures display a downward trend, but in Figure 2(b) the estimates approach the 0 lower bound quickly at all time lags, whereas in Figure 2(a) the es-

timates remain well above 0, even accounting for sample variability. These findings point strongly towards asymptotic independence in the temporal domain, while, for the spatial domain, the large empirical value of  $\chi(0.99)$  indicates that asymptotic dependence cannot be ruled out. Therefore, a model for the extremal behaviour of the rainfall process in the North Brabant province should be asymptotically independent in time and allow for asymptotic dependence in space.

Four main frameworks for modelling spatio-temporal extremes can be identified in the literature. A max-stable approach which employs max-stable processes (de Haan, 1984; Schlather, 2002; Kabluchko et al., 2009) to represent space and time interactions, with time treated as a third continuous dimension added to the two dimensions of space (Davis et al., 2013; Huser and Davison, 2014); a hierarchical approach where spatio-temporal dependence is built in by including a stochastic component in the model parameters (Sang and Gelfand, 2009; Turkman et al., 2010; Economou et al., 2014; Nieto-Barajas and Huerta, 2017; Morris et al., 2017; Bacro et al., 2020); a time series approach, where spatial dependence is embedded within a time series model (Davis and Mikosch, 2008; Meinguet, 2012; Embrechts et al., 2016); a conditional approach based on an asymptotic approximation of the conditional distribution of the space-time process given one single site and time point (Wadsworth and Tawn, 2019; Simpson et al., 2020; Simpson and Wadsworth, 2021). All models developed within the max-stable approach imply spatial and temporal asymptotic dependence or exact independence at all distances and time lags, by a fundamental property of max-stable processes (Wadsworth and Tawn, 2012; Huser and Davison, 2014). The hierarchical approach covers a wider range of limiting patterns, with asymptotic dependent models (e.g., Turkman et al., 2010; Morris et al., 2017) as well as asymptotic independent models (e.g., Sang and Gelfand, 2009; Nieto-Barajas and Huerta, 2017; Bacro et al.,

2020) being proposed in the literature. However, in the aforementioned references the models maintain the same limiting dependence in both space and time. Within the conditional approach, [Simpson et al. \(2020\)](#) and [Simpson and Wadsworth \(2021\)](#) provide examples of formulations for threshold exceedances that enable different forms of asymptotic dependence in the two domains. In this paper we follow a time series approach, as it lends itself naturally to deal with the discreteness and ordering of time. In addition, it gives a flexible framework to construct a model for threshold exceedances that are serially asymptotically independent and spatially either asymptotically independent or asymptotically dependent. One of the advantages of the proposed model is the ease of simulation which facilitates extrapolation of extremal functionals of interest. Estimation of model parameters is carried out by combining composite likelihood, a censoring scheme and an indirect inference algorithm.

The paper is structured as follows. In Section 2 the model is motivated and developed. Inferential issues are discussed in Section 3. In Section 4 the entire approach is applied to the North Brabant rainfall data and compared with an existing alternative method for space-time exceedances. Finally, Section 5 contains some concluding remarks and proposals for extensions.

## 2 Model specification

Let  $Z_t = \{Z_t(s), s \in S\}$ ,  $t = 1, 2, \dots$ , be a stationary spatio-temporal process on the geographical space  $S \subset \mathbb{R}^2$  with marginal univariate CDF  $F_Z$ . Let also

$$\chi^{(d)}(p; t_1, \dots, t_d, s_1, \dots, s_d) := \Pr(F_Z(Z_{t_2}(s_2)) > p, \dots, F_Z(Z_{t_d}(s_d)) > p | F_Z(Z_{t_1}(s_1)) > p)$$

and  $\chi^{(d)}(t_1, \dots, t_d, s_1, \dots, s_d) := \lim_{p \rightarrow 1^-} \chi^{(d)}(p; t_1, \dots, t_d, s_1, \dots, s_d)$ , for  $d = 2, 3, \dots$  and any  $t_1, \dots, t_d \in \{1, 2, \dots\}$  and  $s_1, \dots, s_d \in S$ . Extending the definitions in [Huser and Wadsworth \(2019\)](#) from the spatial to the spatio-temporal framework, we will term  $Z_t$ , respectively, asymptotically dependent or asymptotically independent in time if  $\chi^{(d)}(t_1, \dots, t_d, s, \dots, s) > 0$  or  $\chi^{(d)}(t_1, \dots, t_d, s, \dots, s) = 0$ , for all  $d = 2, 3, \dots$ , all  $s$  and all  $t_1, \dots, t_d$  (with at least one time different from the others). Similarly, the process is termed, respectively, asymptotically dependent or asymptotically independent in space if  $\chi^{(d)}(t, \dots, t, s_1, \dots, s_d) > 0$  or  $\chi^{(d)}(t, \dots, t, s_1, \dots, s_d) = 0$ , for all  $d = 2, 3, \dots$ , all  $t$  and all  $s_1, \dots, s_d$  (with at least one site different from the others). The above definitions do not cover all the possible limiting forms. Intermediate forms of asymptotic dependence could also occur, for example  $\chi^{(d)}(t_1, \dots, t_d, s, \dots, s) = 0$  for some  $d$  and some  $t_1, \dots, t_d$ , but, as [Huser and Wadsworth \(2019\)](#) point out, these are unrealistic over relatively small ranges and will not be considered here.

The starting point in developing a model for threshold exceedances with the desired asymptotic properties is a parsimonious, yet flexible, specification that falls within the time series approach described in Section 1. First, we assume that the temporal dynamics on an arbitrary site  $s \in S$  are driven by a Gaussian first-order autoregressive model, namely,  $Z_t(s) = \alpha Z_{t-1}(s) + \varepsilon_t(s)$ , with  $|\alpha| < 1$  and  $\varepsilon_t(s)$ ,  $t = 1, 2, \dots$ , a sequence of independent and identically distributed random variables with  $\mathcal{N}(0, 1 - \alpha^2)$  distribution. For fixed  $s$ , all finite-dimensional distributions of the process are multivariate Gaussian with standard margins. As the multivariate Gaussian distribution is asymptotically independent ([Resnick, 2013](#), Corollary 5.8), at any location asymptotic independence in time is guaranteed. The second step is to incorporate spatial dependence by assuming that the innovations are spatially related. Formally, we

consider the stationary space-time process  $Z_t = \{Z_t(s), s \in S\}$ ,  $t = 1, 2, \dots$ , given by

$$Z_t(s) = \alpha Z_{t-1}(s) + \varepsilon_t(s), \quad s \in S, t = 1, 2, \dots, \quad (2.1)$$

where  $|\alpha| < 1$  and  $\varepsilon_t = \{\varepsilon_t(s), s \in S\}$ ,  $t = 1, 2, \dots$ , is a sequence of independent copies of a stationary random field  $\varepsilon$  on  $S$  with  $\mathcal{N}(0, 1 - \alpha^2)$  univariate margins. A valuable aspect of model (2.1) is its simple interpretation. The autoregressive component controls time dependence through  $\alpha$ , while spatial dependence is determined by  $\varepsilon_t$ . On the other hand, interpretability comes at the expense of space-time interactions, that (2.1) precludes. An extension of this formulation to accommodate interactions is discussed in Section 5.

## 2.1 Spatial dependence

While the autoregressive structure of (2.1) and the marginal Gaussianity of the innovations constrain  $Z_t$  to generate asymptotically independent exceedances in time, both asymptotic dependence and asymptotic independence are possible for the spatial domain, according to the formulation selected for  $\varepsilon$ . In particular, we will focus on two cases:  $\varepsilon$  is either a Gaussian random field or a marginally transformed asymptotically dependent random field.

If the innovations are Gaussian random fields, then  $Z_t$  is a special case of a stochastic integro-difference equation model (Wikle and Cressie, 1999; Brown et al., 2000). Under this specification, all finite-dimensional distributions of  $Z_t$  are multivariate Gaussian with standard margins and the process is asymptotically independent in both time and space. By contrast, if  $\varepsilon$  is a marginal transformation of an asymptotically dependent random field, then  $Z_t$  is itself asymptotically dependent in space. To clarify the latter point, let  $e$  be a stationary random field on  $S$ , with  $F_e(x) = \Pr(e(s) \leq x)$ ,

satisfying the asymptotic dependence requirement

$$\lim_{p \rightarrow 1^-} \Pr(F_e(e(s_2)) > p, \dots, F_e(e(s_d)) > p | F_e(e(s_1)) > p) > 0 \quad (2.2)$$

for all  $s_1, \dots, s_d$  and  $d = 2, 3, \dots$ . By letting  $\varepsilon$  in equation (2.1) be

$$\varepsilon(s) = \sqrt{1 - \alpha^2} \cdot \Phi^{-1} \{F_e(e(s))\} \quad (2.3)$$

with  $\Phi^{-1}$  denoting the quantile function of the standard Normal random variable, we derive a space-time process  $Z_t$  which is asymptotically dependent in space. It can also be shown that  $\chi^{(d)}(t_1, \dots, t_d, s_1, \dots, s_d) = 0$  if at least one of  $t_1, \dots, t_d$  is different from the others, for all  $s_1, \dots, s_d$ , and all  $d = 2, 3, \dots$ : that is, non-simultaneous exceedances are asymptotically independent. Proofs of both properties are given in the supplementary material, available at <http://www.statmod.org/smij/archive.html>. Specification (2.3) of model (2.1) is a special case of the family of processes studied in Davis and Mikosch (2008, Equation 1.1), but with innovations transformed to have Gaussian univariate margins, as opposed to regularly varying univariate margins. The transformation of the errors to the Normal scale is an essential component of the proposed modelling procedure: working with regularly varying margins as in Davis and Mikosch (2008) would yield a process that is not asymptotically independent in time (Chernick et al., 1991, Proposition 2.1), thus failing to meet one of the objectives of this work.

## 2.2 Specification of $e$

To complete the model, a formulation for the random field  $e$  in (2.3) is required. In the sequel we will restrict attention to two families of processes: Student's  $t$  random fields (Røislien and Omre, 2006) and max-stable random fields (de Haan, 1984; Smith,

1990b). Each constitutes a class of asymptotically dependent spatial processes for which flexible parametric forms are available. At the same time, they can exhibit quite different extremal behaviours so together they cover a wide range of spatial patterns of threshold exceedances (Demarta and Mcneil, 2005).

A Student's  $t$  random field, with  $\nu > 0$  degrees of freedom is defined by

$$e(s) = \sqrt{U}W(s) \quad (2.4)$$

where  $W$  is a stationary Gaussian random field on  $S$  with standard margins and  $U$  is an Inverse Gamma random variable with scale parameter  $\nu/2$  and shape parameter  $\nu/2$ , independent of  $W(s)$ . All finite-dimensional distributions of process (2.4) are multivariate  $t$  with  $\nu$  degrees of freedom, which are known to be asymptotically dependent (Chan and Li, 2008). Hence, process (2.4) satisfies condition (2.2). To specify a family of models for  $\varepsilon$ , transformation (2.3) should be applied with  $F_e$  given by the CDF of the Student's  $t$  random variable with  $\nu$  degrees of freedom. As  $\nu \rightarrow \infty$ ,  $e$  converges to a Gaussian random field. Therefore, as  $\nu$  increases,  $\varepsilon$  will increasingly resemble a Gaussian field, approaching spatial asymptotic independence.

Max-stable random fields were introduced by de Haan (1984) as an extension of max-stable distributions. Asymptotic dependence of max-stable random fields follows from Lemma 3.1 of Davis and Mikosch (2008). One of the most frequently used sub-classes of the max-stable family is the Brown-Resnick process (Brown and Resnick, 1977; Kabluchko et al., 2009) whose representation is

$$e(s) = \max_{i \geq 1} \{R_i \exp(W_i(s) - \gamma(s))\}, \quad (2.5)$$

where  $0 < R_1 < R_2 < \dots$  are the points of a Poisson process on the positive half-line with intensity  $r^{-2}dr$  and  $W_i, i = 1, 2, \dots$ , are independent copies of a Gaussian random

field  $W$  on  $S$  with stationary increments, semi-variogram  $\gamma(s) = \mathbb{E}(W(s) - W(0))^2/2$  and  $W(0) = 0$  almost surely. Random field (2.5) has unit Fréchet univariate margins, i.e.  $F_e(x) = \exp(-1/x)$ , for  $x > 0$  and  $F_e(x) = 0$ , elsewhere. Hence, transformation (2.3) becomes

$$\varepsilon(s) := \sqrt{1 - \alpha^2} \cdot \Phi^{-1} \left\{ \exp \left( -\frac{1}{e(s)} \right) \right\}. \quad (2.6)$$

An illustration of the different types of extremal patterns attainable under the proposed specifications is given in the application presented in Section 4.

### 3 Inferential aspects

To prevent non-extreme observations from introducing bias in the analysis, our proposal is to represent the extremal behaviour of the observed process through the tail of model (2.1). Let  $Y_t = \{Y_t(s), s \in S\}$ ,  $t = 1, 2, \dots$ , be the spatio-temporal process generating the data, which is assumed time-stationary and with  $\Pr(Y_t(s) \leq y) = F_s(y)$ . The process  $Y_t$  can be marginally transformed to have standard Gaussian univariate margins through

$$Z_t(s) = \Phi^{-1} \{F_s(Y_t(s))\} \quad (3.1)$$

Equation (2.1) is then assumed to hold for the transformed process when its entire space-time trajectory is above a fixed high threshold  $u_z$  on the standard Gaussian scale.

When inferring model (2.1) from the transformed threshold exceedances some issues arise, stemming mainly from the intractability of the likelihood function and the need to censor observations below the threshold. In the following sections we describe a solution based on a combination of composite likelihood and indirect inference.

### 3.1 Estimation

Let  $y_t(s_i)$  be the original observation at time  $t$ ,  $t = 1, \dots, T$ , and site  $s_i$ ,  $i = 1, \dots, n$ , and  $z_t(s_i)$  be the corresponding transformed observation on the  $\mathcal{N}(0, 1)$  scale. Transformation (3.1) from  $y_t(s_i)$  to  $z_t(s_i)$  requires knowledge of  $F_s$ . In general,  $F_s$  is unknown and should be estimated from data, but, to simplify the description of the inference for the dependence structure, throughout this section it will be treated as known. Estimation methods for  $F_s$  will be discussed in Section 4.

We assume that a parametric model  $\varepsilon^\theta$  has been selected for  $\varepsilon$  in (2.1), where the parameter vector  $\theta = (\alpha, \psi)$  comprises the autoregressive parameter  $\alpha$  and the vector  $\psi$  of spatial parameters. Evaluation of the likelihood function for  $\theta$  is impractical, even with a moderate number of observed locations, as the overall dimensionality of the problem is determined by both space and time. We resort to a composite likelihood approach (Lindsay, 1988; Varin et al., 2011) based on the bivariate marginal distributions of the process. In addition, to account for model (2.1) being assumed valid only above a fixed threshold  $u_z$ , a censoring scheme similar to that originally proposed by Smith et al. (1997), and later employed by Ledford and Tawn (1997), Huser and Davison (2014) and Bacro et al. (2020) among others, is embedded within the composite likelihood. The resulting estimating function for  $\theta$  is given by the following pairwise weighted log-likelihood (PL)

$$\text{PL}(\theta; z) = \sum_{t=1}^T \sum_{k=0}^{C_T} \sum_{i=1}^n \sum_{j=i+1-k}^n \log h_{s_i, s_j, k}(z_t(s_i), z_{t+k}(s_j); \theta) w_{(s_i, s_j)} \quad (3.2)$$

where  $z$  is the set of available observations on the standard Gaussian scale,  $w_{s_i, s_j}$  is a weight such that  $w_{s_i, s_j} = 1$  if  $\|s_i - s_j\| \leq C_S$  and  $w_{s_i, s_j} = 0$  otherwise (Davis et al., 2013; Huser and Davison, 2014), and  $C_S > 0$  and  $C_T \in \{1, 2, \dots\}$  are, respectively, the

spatial and temporal cut-off points. The PL contributions are obtained as logarithmic transformations of

$$h_{s_i, s_j, k}(z_1, z_2; \theta) = \begin{cases} f_{s_i, s_j, k}(z_1, z_2; \theta) & \text{if } \min\{z_1, z_2\} > u_z \\ \frac{\partial}{\partial z_1} F_{s_i, s_j, k}(z_1, u_z; \theta) & \text{if } z_1 > u_z, z_2 \leq u_z \\ \frac{\partial}{\partial z_2} F_{s_i, s_j, k}(u_z, z_2; \theta) & \text{if } z_1 \leq u_z, z_2 > u_z \\ F_{s_i, s_j, k}(u_z, u_z; \theta) & \text{if } \max\{z_1, z_2\} \leq u_z \end{cases} \quad (3.3)$$

with  $f_{s_i, s_j, k}$  and  $F_{s_i, s_j, k}$  being, respectively, the density function and CDF of  $(Z_t(s_i), Z_{t+k}(s_j))$ . When  $\varepsilon^\theta$  is a Gaussian random field, the associated PL, which will be denoted by  $\text{PL}^G(\theta; z)$ , can be easily computed and  $\theta$  estimated by maximization of (3.2). If either a marginally transformed Student's  $t$  or max-stable random field is chosen for  $\varepsilon^\theta$ , no analytical expression is available for either the marginal distribution of  $Z_t$  or the joint distribution of  $Z_t$  and  $Z_{t+k}$  for any  $k = 1, 2, \dots$ . Evaluation of the terms in (3.3) therefore entails complicated numerical integration. To avoid this computational burden, we propose the indirect inference procedure described below.

### 3.2 Indirect inference

Indirect inference, introduced by [Smith \(1990a\)](#) and later extended by [Gourieroux et al. \(1993\)](#) and [Gallant and Tauchen \(1996\)](#), is a simulation-based approach to classical estimation pre-dating the main developments of Approximate Bayesian Computation (ABC), with which it shares similar goals. For details on the connections between ABC and indirect inference see, for example, [Drovandi \(2019\)](#). Indirect inference proves to be particularly effective in situations where the likelihood function, or any other criterion function that forms the basis of estimation, is too difficult to evaluate, but simulation from the model is feasible. Model (2.1) with innovations that are either a marginally transformed Student's  $t$  process or max-stable process

satisfies both conditions, as it can be sampled but its PL is intractable.

Equation (6) of [Heggland and Frigessi \(2004\)](#) summarizes the indirect inference approach advocated here. An auxiliary model depending on a set of parameters  $\beta$  is identified and used to construct an auxiliary estimating function  $Q(\beta; z)$  whose analytical expression is available. Let  $\theta$  be the set of parameters of the model we wish to infer, hereafter termed the target model, and  $z^{obs}$  be the original sample. For fixed  $\theta$ ,  $M$  independent samples,  $z^{(m,\theta)}$ ,  $m = 1, \dots, M$ , of the same size as  $z^{obs}$  are simulated from the target model. The auxiliary estimating function is maximized with respect to  $\beta$  over the set of simulated samples, yielding the estimate  $\hat{\beta}(\theta) = \operatorname{argmax}_{\beta} \sum_{m=1}^M Q(\beta; z^{(m,\theta)})$ . Finally,  $\theta$  is estimated through

$$\hat{\theta}_I = \operatorname{argmax}_{\theta} \left[ Q(\hat{\beta}(\theta); z^{obs}) \right]. \quad (3.4)$$

By results in [Gourieroux and Monfort \(2018\)](#), if the target model is stationary and satisfies suitable mixing conditions, as the size of the observed sample diverges  $\hat{\beta}(\theta)$  converges almost surely and uniformly to a deterministic limit  $b(\theta)$ , often referred to as the binding function. If the function  $b(\theta)$  is one-to-one,  $\hat{\theta}_I$  is a consistent and asymptotically Gaussian estimator of  $\theta$ .

In applying the above procedure, we selected process (2.1) with Gaussian innovations as the auxiliary model, and let  $\varepsilon = \varepsilon^{\beta}$  and  $Q(\beta; z) = \operatorname{PL}^G(\beta; z)$ . This choice is based on two arguments. First,  $\operatorname{PL}^G$  is readily computed and the auxiliary parameter  $\beta$  can be estimated directly. Second, the tail of a Gaussian random field provides a reasonable approximation of a wide range of exceedance patterns. The auxiliary model is therefore likely to be close enough to the target model for the algorithm to perform well. Equation (3.4) involves the solution of two nested maximization problems, an outer maximization with respect to  $\theta$  and an inner maximization with

respect to  $\beta$ , whose steps are outlined below.

Indirect Inference Algorithm (IIA)

Step 1.1: For given  $\theta$ , simulate  $M$  independent samples

$$z^{(m,\theta)} = \{z_t^{(m,\theta)}(s_i), i = 1, \dots, n, t = 1, \dots, T\}, \quad m = 1, \dots, M,$$

from model (2.1) with  $\varepsilon = \varepsilon^\theta$ .

Step 1.2: Maximize  $\sum_{m=1}^M \text{PL}^G(\beta; z^{(m,\theta)})$  with respect to  $\beta$  to obtain  $\hat{\beta}(\theta)$ .

Step 1.3: Evaluate  $\text{PL}^G(\hat{\beta}(\theta); z^{obs})$ .

Repeat Steps 1.1 to 1.3 till maximization of  $\text{PL}^G(\hat{\beta}(\theta); z^{obs})$  to obtain  $\hat{\theta}_1$ .

For  $|\alpha| < 1$ , process (2.1) is time-stationary and  $\alpha$ -mixing, which guarantees almost sure convergence of  $\hat{\beta}(\theta)$  to the binding function  $b(\theta)$  as  $T \rightarrow \infty$ . The limit  $b(\theta)$  is unknown; however, it is reasonable to assume that it satisfies the one-to-one property required for consistency of  $\hat{\theta}_1$ , as  $\theta$  and  $\beta$  have the same number of components with similar interpretations.

In the implementation of IIA we used the Nelder-Mead routine for the inner maximization and a simulated annealing routine for the outer maximization. Simulated annealing typically requires more iterations than the Nelder-Mead routine, but is more robust and suited especially for irregular surfaces like  $\text{PL}^G(\hat{\beta}(\theta); z^{obs})$ . The performance of IIA has been assessed through a simulation study whose details are available as supplementary material. We have found that in all the scenarios considered it leads to entirely satisfactory results.

### 3.3 Approximation of the estimator distribution

In this section, the issue of quantifying the variability of the proposed estimators is addressed.

When  $\varepsilon^\theta$  is a Gaussian random field,  $\theta$  is estimated directly from  $\text{PL}^G$  without resorting to indirect inference and a parametric bootstrap is a viable way to approximate the estimator distribution. This consists of repeatedly simulating a sample of the same size as  $z^{obs}$  from the fitted Gaussian model and re-estimating  $\theta$  on each simulated sample. The resulting estimates are treated as realizations of the estimator.

When  $\varepsilon^\theta$  is either a spatial  $t$  or a max-stable process marginally transformed to the Gaussian scale, i.e. when estimation is carried out via the indirect inference approach, we approximate the estimator distribution by exploiting the output from the iterations of IIA. We assume that the binding function  $b(\theta)$  is bijective and denote its inverse by  $t(\beta)$ . Let  $\hat{\beta}$  be the estimate of  $\beta$  obtained by maximizing  $Q$  on the original sample:  $\hat{\beta} = \text{argmax}_\beta Q(\beta; z^{obs})$ . The rationale behind the proposed procedure is that for large  $T$  the distribution of  $\hat{\theta}_T$  can be approximated by the distribution of  $t(\hat{\beta})$ . The distribution of  $\hat{\beta}$  can be estimated via a parametric bootstrap similar to the one described above: samples are simulated from the target model with  $\theta = \hat{\theta}_T$  and  $\beta$  is estimated on each sample by direct maximization of  $\text{PL}^G$ . A simulation-based estimate of  $t$  can be obtained as follows. Suppose that  $N$  iterations of IIA have been run before convergence. The  $N$  repetitions of Steps 1.1 and 1.2 output the set of pairs  $\{(\theta_j, \hat{\beta}(\theta_j)), j = 1, \dots, N\}$  from which a semiparametric estimate  $\hat{t}$  of  $t$  can be derived. Note that the choice of a simulated annealing routine for the outer maximization of IIA yields a relatively high number of iterations which allows

an accurate reconstruction of  $t$ .

#### Standard Error Algorithm

Step 2.1: Compute a semiparametric estimate  $\hat{t}(\beta)$  of  $t(\beta)$  from the intermediate output  $\{(\theta_j, \hat{\beta}(\theta_j)), j = 1, \dots, N\}$  of IIA.

Step 2.1: Simulate  $L$  samples of the same size as  $z^{obs}$  from model (2.1) with  $\varepsilon = \varepsilon^{\hat{\theta}_1}$  and obtain the bootstrap estimates  $\hat{\beta}_1, \dots, \hat{\beta}_L$  by maximizing  $Q$  on each sample.

Step 2.3: Calculate  $\hat{t}(\hat{\beta}_1), \dots, \hat{t}(\hat{\beta}_L)$  as approximate realizations from the distribution of  $\hat{\theta}_1$ .

## 4 Application

The approach described in Sections 2 and 3 was applied to recordings of daily rainfalls from 01-10-1999 to 28-02-2019 over the North Brabant province in the south of the Netherlands. The dataset was downloaded from the European Climate Assessment (ECA) & Dataset website (<https://www.ecad.eu>), which provides a zipped file containing a separate ASCII file for each station. In the reference region 40 stations are installed, but 10 of them were discarded for having more than 10% missing values. Of the 30 remaining stations, 28 are manual rainfall stations, and two, corresponding to Eindhoven and Gilze Rijen weather stations, are automatic. We noticed non-negligible differences in the magnitude of the recordings between the automatic stations and the nearby manual ones. This is a known phenomenon (Brandsma, 2014) which led us to exclude from the analysis the two automatic stations. Figure 1 shows a map of the 28 manual stations analyzed.

Flood events are mainly determined by aggregations in space and/or time of extreme rainfalls (Richards et al., 2021). A spatio-temporal study of threshold exceedances of the rainfall process therefore has important practical implications in evaluating flood risk. Figure 2 suggests that the exceedances of the rainfall process are spatially and temporally related, though in ways that change between the two domains. Dependence weakens rapidly with threshold level in the temporal domain and slowly in the spatial domain where it retains a significant strength even at the most extreme quantiles. These patterns are compatible with asymptotic independence in time but leave ambiguity about the type of asymptotic dependence in space. We investigate these features by fitting model (2.1) with Gaussian innovations and asymptotically dependent innovations.

An exploratory analysis of the dataset was carried out in order to evaluate possible departures from the model assumptions. No evidence of an annual trend was found over the 21-year period, but a seasonal variation was detected. Deseasonalizing each series would only partially address this issue as it would not account for seasonal changes in the spatial dependence. A separate-season modelling is therefore a safer approach. In the following we will focus on the period from October to February within which approximate time-stationarity appears to hold. Anisotropy at extreme levels was also assessed. The assessment, whose details are reported in the supplementary material, revealed no significant differences across directions for the summary statistics considered. To investigate presence of interactions between space and time, Figure 3 shows site-specific estimates and 95% confidence intervals for the autoregressive parameter  $\alpha$ , computed from observations above the 0.90 quantile. Also included in Figure 3 is a vertical line representing the estimate of  $\alpha$  obtained by combining all the series into a single one and assuming a common autoregressive coefficient over the region.

The horizontal line intersects all of the sitewise confidence intervals, supporting the hypothesis of a constant  $\alpha$  for the North Brabant data.

## 4.1 Marginal and dependence modelling

Fitting model (2.1) requires the original data to be transformed to the Gaussian scale through equation (3.1) which, in turn, requires specification and estimation of  $F_s(y) = \Pr(Y_t(s) \leq y)$ . As the inferential procedure of Section 3 involves only threshold exceedances, the problem can be simplified to estimating the tail of  $F_s$ . It is common practice in studies of spatial exceedances (Eastoe and Tawn, 2009; Northrop and Jonathan, 2011) to model the marginal tails via the Generalized Pareto (GP) distribution whose parameters vary spatially, namely

$$F_s(y) = \begin{cases} \zeta(s) + (1 - \zeta(s)) \left(1 + \xi(s) \frac{y - u(s)}{\sigma(s)}\right)_+^{-1/\xi(s)} & y > u(s), \\ \zeta(s) & y \leq u(s) \end{cases} \quad (4.1)$$

where  $(a)_+ = \max(0, a)$  and, for all  $s \in S$ ,  $\zeta(s) = \Pr(Y_t(s) \leq u(s))$ ,  $u(s)$  is the threshold level,  $\sigma(s)$  a positive scale parameter and  $\xi(s)$  a real shape parameter, respectively. In principle, estimation of marginal and dependence parameters could be carried out simultaneously, by explicitly including the transformation from the GP scale to the Gaussian scale in expression (3.3). However, simultaneous estimation increases substantially the computational burden, especially when implementing IIA. For this reason, we resort to a two-step procedure, whereby model (4.1) is first estimated and the dependence model is then fitted to the transformed exceedances.

For the first step, a preliminary analysis showed that the GP model provides a good fit above the 0.90 quantile at all sites. We therefore set  $\zeta(s) = 0.90$  for all  $s \in S$ . The threshold level  $u(s)$  was then estimated by quantile regression, as proposed by

Northrop and Jonathan (2011). The shape and scale parameters were modelled via the semi-parametric approach advocated by Youngman (2019) and implemented in the companion R package `evgam`. In particular, we considered

$$\log \sigma(s) = \lambda_0^\sigma + \lambda_1^\sigma(\text{lon}(s), \text{lat}(s)) \quad \text{and} \quad \xi(s) = \lambda_0^\xi + \lambda_1^\xi(\text{lon}(s), \text{lat}(s)) \quad (4.2)$$

where  $\lambda_0^\sigma$  and  $\lambda_0^\xi$  are real constants, and each of  $\lambda_1^\sigma$  and  $\lambda_1^\xi$  is a thin plate regression spline (Wood, 2003) with  $\text{lon}(s)$  and  $\text{lat}(s)$  denoting longitude and latitude, respectively. By fitting model (4.1)-(4.2) through the `evgam` package, we infer that the following simplifications hold for the North Brabant data:  $\log \sigma(s) = \lambda_0^\sigma$  and  $\xi(s) = \lambda_0^\xi$ . Working with this reduced form, the observed threshold exceedances,  $y_t(s_i)$  with  $y_t(s_i) > u(s_i)$ ,  $i = 1, \dots, 28$ , were transformed to the standard Gaussian scale through  $z_t(s_i) = \Phi^{-1}(\hat{F}_{s_i}(y_t(s_i)))$ , where  $\hat{F}_{s_i}$  denotes the estimated GP distribution.

In the second step of the procedure, process (2.1) was fitted to the transformed exceedances. For the innovation process  $\varepsilon$ , in light of the findings from the exploratory analysis, we considered four isotropic parametric specifications  $\varepsilon^\theta$ , each characterized by a three-dimensional vector of parameters  $\theta = (\alpha, \psi)$ , with  $\psi = (\psi_1, \psi_2)$ . These are:

- Model 1: a spatial Gaussian model with  $\mathcal{N}(0, 1 - \alpha^2)$  univariate margins and powered exponential correlation function

$$\rho(h; \psi) = \exp\left(-(h/\psi_1)^{\psi_2}\right), \quad h \geq 0, \psi_1 > 0, 0 < \psi_2 < 2, \quad (4.3)$$

- Model 2: a spatial  $t$  process as in (2.4) with  $\nu = 5$  degrees of freedom and  $W$  having correlation function (4.3), marginally transformed to the  $\mathcal{N}(0, 1 - \alpha^2)$  scale through equation (2.3).

- Model 3: a spatial  $t$  process as in (2.4) with  $\nu = 2$  degrees of freedom and  $W$  having correlation function (4.3), marginally transformed to the  $\mathcal{N}(0, 1 - \alpha^2)$  scale through equation (2.3).
- Model 4: a spatial Brown-Resnick process as in (2.5) with  $W$  having a power-law semi-variogram (Chilès and Delfiner, 2002, pag. 266)  $\gamma(h; \psi) = (h/\psi_1)^{\psi_2}$ ,  $h \geq 0, \psi_1 > 0, 0 < \psi_2 < 2$ , marginally transformed to the  $\mathcal{N}(0, 1 - \alpha^2)$  scale through equation (2.6).

The degrees of freedom of the spatial  $t$  process cannot be estimated by the proposed indirect inference algorithm. If  $\nu$  itself were treated as an unknown dependence parameter, the parametric space of the auxiliary model would have smaller dimension than that of the target model, failing to meet the identifiability condition of [Gourieroux and Monfort \(2018\)](#). We therefore profiled over a grid of reasonable integer values for  $\nu$  which comprises  $\nu = 2, 5$  and 10. Results for  $\nu = 10$  are omitted in the sequel as we found only slight differences with respect to the Gaussian specification.

The interpretation of  $\theta$  remains the same across all four formulations:  $\psi_1$  is a spatial range parameter,  $\psi_2$  a spatial smoothness parameter and  $\alpha$  the temporal autoregressive parameter. The numbering of the models reflects the strength of spatial asymptotic dependence, with Model 1 being asymptotically independent, and Models 2-4 having increasing degree of asymptotic dependence for a common value of  $\theta$ .

Model 1 was estimated by direct maximization of  $PL^G$ , while Models 2-4 were estimated via IIA, using Model 1 as the auxiliary model and selecting  $M = 10$ , as in the simulation study described in the supplementary material. For all models, the temporal cut-off point  $C_T$  in (3.2) was set at 1 on account of the first-order Markov

nature of the generating process and efficiency results obtained by [Davis and Yau \(2011\)](#). The spatial cut-off point  $C_S$  was set at the maximum distance between observed sites, so that all pairs of sites were included in the evaluation of PL. Estimates of  $\theta$  for all four models are reported in [Table 1](#), with standard errors in parentheses. For Model 1, standard errors were computed by the parametric bootstrap procedure described in [Section 3.3](#) with 200 bootstrapped samples. For Models 2-4 standard errors were computed by the standard error algorithm, with  $L = 200$  in [Step 2.2](#) and estimating  $t$  in [Step 2.1](#) by a second-order polynomial regression with trivariate response. Estimates of  $\alpha$  are essentially constant across all four models, as would be expected given their common autoregressive construction. Estimates of  $\psi_2$  are also stable, denoting a similar smoothness in the spatial dependence decay with distance. The greatest variation is shown by the estimates of  $\psi_1$ , which decrease with model number, thus counterbalancing the increasing strength of asymptotic spatial dependence. An example of code for estimating the previous models is given in the supplementary material.

An alternative model for space-time exceedances proposed by [Simpson and Wadsworth \(2021\)](#), hereafter referred to as the SW model, was also fitted to the data. [Simpson and Wadsworth \(2021\)](#) extend to the space-time setting the conditional approach developed by [Wadsworth and Tawn \(2019\)](#) for spatial extremes. As the SW model enables different forms of limiting dependence in space and time, it provides a benchmark to assess the performance of Models 1-4. The parametric specification selected for the SW model is given by equations (2)–(4) and (7) of [Simpson and Wadsworth \(2021\)](#). It involves, among others, the parameters  $\Delta_S$  and  $\Delta_T$ , determining the asymptotic dependence class of the spatial and temporal domains, respectively. On the basis of the same diagnostic tools as those used for Models 1-4 and described in the next

section, we found that the best fitting formulation within the SW class has  $\Delta_S$  equal to the maximum spatial distance between observed sites, corresponding to asymptotic dependence in space, and  $\Delta_T = 0$ , corresponding to asymptotic independence in time, respectively. Model parameters were estimated by composite likelihood (see [Simpson and Wadsworth, 2021](#), for details), with the threshold fixed at the 0.90 marginal quantile as for Models 1-4.

Model	$\alpha$	$\psi_1$	$\psi_2$
Model 1	0.33 (0.035)	887.1 (109.9)	0.74 (0.020)
Model 2	0.33 (0.039)	719.2 (75.4)	0.73 (0.019)
Model 3	0.33 (0.039)	458.7 (73.42)	0.78 (0.028)
Model 4	0.32 (0.038)	148.0 (15.60)	0.72 (0.022)

Table 1: Estimates and standard errors (in parentheses) of model parameters.

## 4.2 Diagnostics and results

As a tool to assess and compare the quality of fit of the estimated models, we analyzed the behaviour of

$$\begin{aligned}
 \chi(p; l, \|h\|) &= \Pr(F_s(Y_{t+l}(s)) > p | F_{s+h}(Y_t(s+h)) > p) \\
 &= \Pr(Z_{t+l}(s) > \Phi^{-1}(p) | Z_t(s+h) > \Phi^{-1}(p))
 \end{aligned}
 \tag{4.4}$$

as  $\|h\|$ ,  $l$  and  $p$  vary. The function  $\chi(p; l, \|h\|)$ , which gives the conditional probability of observing an exceedance at time  $t+l$  at one site given that an exceedance has occurred at time  $t$  at an inter-location distance  $\|h\|$ , is used as a summary measure of the bivariate dependence of the process. [Figures 4 and 5](#) compare estimates of  $\chi(p; l, \|h\|)$  from Models 1, 2, 4 and SW with empirical estimates. Results for Model

3 are not shown, since they are very close to those of Model 4. For Models 2-4,  $\chi(p; l, \|h\|)$  was evaluated by simulation, as no closed-form expression is available; analytical calculations are possible for Model 1 and the SW model. To investigate space-time patterns,  $\chi(p; l, \|h\|)$  is plotted as a function of  $\|h\|$ , for different choices of  $l$ , with  $l = 0$  in Figure 4 and  $l = 1$  in Figure 5, respectively. To verify the goodness-of-fit at the selected threshold, as well as the quality of extrapolations to higher levels, within each plot different values of  $p$  are considered, with  $p \in \{0.90, 0.95, 0.99, 0.995\}$ . For  $l = 0$ , Model 4 and the SW model fail to capture the decay of dependence at increasing distances for fixed  $p$  and for increasing  $p$  at fixed distances. Under both formulations, extrapolations to higher values of  $p$  exhibit a stability that causes departures from the empirical estimates. This is not surprising, given that, for  $\Delta_S$  equal to the maximum observed distance, the spatial dependence of the SW model is driven by a Brown-Resnick process (Wadsworth and Tawn, 2019) as for Model 4. Models 1 and 2 provide a better overall fit, but for  $p = 0.99$  and  $p = 0.995$  Model 2 outperforms Model 1, which shows an increasing tendency to underestimate dependence between simultaneous exceedances. For  $l = 1$ , the empirical estimates decay rapidly with  $p$ . Models 1, 2 and 4 follow this behaviour closely; SW model estimates also decrease with  $p$  for fixed  $\|h\|$ , as a consequence of the temporal asymptotic independence, but exhibit positive bias for  $p = 0.90$  and  $p = 0.95$ . The case  $l = 2$  was also considered. All four formulations perform satisfactorily in this case, as shown by the results summarized in the supplementary material. Figure 6 focuses on the temporal dynamics by displaying empirical estimates and model-based estimates from Models 1, 2, 4 and the SW model of  $\chi(p; l, \|h\|)$  as a function of  $l$ , for  $p \in \{0.90, 0.95, 0.99, 0.995\}$  and  $\|h\| = 0$ , i.e. on a single site. Unlike process (2.1), the SW model is non-Markovian in time. The lag-1 dependence is slightly overestimated by all models, but other-

wise Models 1, 2 and 4 approximate the empirical pattern well; by contrast, the SW model yields too fast a decline for both  $p = 0.90$  and  $p = 0.95$ . The first-order Markov assumption therefore seems adequate for the North Brabant data.

All five models were also fitted at a higher threshold level, corresponding to the 0.95 marginal quantile. Qualitatively, no changes occurred in the fitting and extrapolations for Models 1 and 2. A net improvement was found in the performance of Models 3 and 4 and SW model for  $p = 0.95$ , but for  $p > 0.95$  we observed the same limitations as above. Diagnostics therefore point to Models 1 and 2 as providing the best formulations for this application, with a preference for Model 2 when inferring dependence between pairs of simultaneous exceedances of extremely high thresholds. These findings give additional support to asymptotic dependence in the spatial domain.

Model 2 is now used to further investigate extremal features of the rainfall phenomenon in the North Brabant region. Under this formulation, analytical computations are generally unfeasible, but simulation is simple and fast. As an example, Figure 7 displays the temporal evolution of a simulated realization of Model 2 over the entire North Brabant area. Dots represent the observed stations. At time  $t - 1$  (top-left panel) no upcrossings of the 0.90 marginal quantiles occur on the whole region: as the model governs only the exceedances, the map shows the thresholds  $u(s)$  obtained by quantile regression, as described in Section 4.1. At time  $t$  (top-right panel) an extreme event takes place, affecting a sub-region, and subsequently evolves to produce the map at time  $t + 1$  (bottom-left panel). At time  $t + 2$  (bottom-right panel) the extreme event dies out and again only the thresholds  $u(s)$  are displayed in the map. Stochastic properties of functionals of space-time trajectories lying above 0.90 marginal quantiles, e.g. rainfall aggregates over areas involved in an extreme

event, can be studied by repeated simulations. This approach was applied to analyze the number of observed stations exceeding a high threshold in a single day  $N_t(p, 0) = \sum_{i=1}^n \mathbb{I}(\hat{F}_{s_i}(Y_t(s_i)) > p)$ , where  $\mathbb{I}(A)$  is the indicator function of the event  $A$ , and in two consecutive days  $N_t(p, 1) = \sum_{i=1}^n \prod_{l=0}^1 \mathbb{I}(\hat{F}_{s_i}(Y_{t+l}(s_i)) > p)$ . Conditioning upon  $N_t(p, 0) > 0$ , i.e. on having at least one exceedance in a day,  $N_t(p, 0)$  and  $N_t(p, 1)$  can be seen as measures of spatial clustering of extreme rainfalls on a single day and over a two-day period, respectively. In Figure 8, Model 2 estimates of  $E(N_t(p, 0) | N_t(p, 0) > 0)$  and  $E(N_t(p, 1) | N_t(p, 0) > 0)$  are plotted against  $p$ . For the empirical estimates,  $p$  ranges from 0.90 to 0.975, whereas model-based estimates extrapolate up to  $p = 0.99$ . Also shown on each figure are 95% confidence bands for the empirical estimates obtained by block bootstrapping the original data. As  $p \rightarrow 1$ , estimates in Figure 8(b) converge to 0, by the asymptotic independence in time, and to a strictly positive value in Figure 8(a), by the asymptotic dependence in space.

## 5 Conclusions

Motivated by extremal features that can be observed in some environmental studies, we developed a model for space-time exceedances which comprises serial asymptotic independence and either asymptotic dependence or asymptotic independence in space. The proposed model has the advantages of exploiting all the information above a fixed high threshold, possessing an easy-to-interpret structure and respecting the ordered and discrete nature of the measurement times. In the application to a dataset of rainfall amounts over the North Brabant province, it was found to cover a wide range of extremal patterns. In particular, among the four parametric specifications

considered, two appear to provide the best fit to the data: the one with Gaussian errors and the one with Student  $t_5$  errors, with a superiority of the latter when focussing on extrapolations to higher threshold levels. These findings support the conjecture that rainfall amounts in the North Brabant province are asymptotically independent in time, but asymptotically dependent in space.

The model can be extended to include cases that fall outside the studied framework. For example, an extension that accommodates more complex forms of space-time interactions consists in letting the autoregressive parameter  $\alpha$  depend on space. Model (2.1) would then become  $Z_t(s) = \alpha(s)Z_{t-1}(s) + \varepsilon_t(s)$ , where  $\alpha(s)$  is formulated according to the pattern of spatial variations observed in the temporal dynamics. As an example, we could specify a logistic-type function  $\alpha(s) = 1 - 2 \exp\{\alpha_0 + \alpha_1 x + \alpha_2 y\} / \{1 + \exp\{\alpha_0 + \alpha_1 x + \alpha_2 y\}\}$ , with  $\alpha_0, \alpha_1, \alpha_2 \in \mathbb{R}$  and  $s = (x, y) \in \mathbb{R}^2$ , which ensures that the condition  $|\alpha(s)| < 1$  for time-stationarity at each location is satisfied. An extension of the model that would enlarge its range of applicability would be to avoid an a priori choice of the limiting dependence class for the temporal dimension. One way is to replace the Gaussian autoregressive component in (2.1) with a Markov chain whose transitions encompass both asymptotic dependence and asymptotic independence. In this setting, the data itself would drive the selection between the two dependence classes. The [Ledford and Tawn \(1997\)](#) procedure or the conditional extreme value approach of [Heffernan and Tawn \(2004\)](#) could be used for the specification of the chain transitions. We are currently exploring these possibilities.

## Acknowledgment

Carlo Gaetan's research was partially funded by CORILA-research program VENEZIA 2021. The authors would like to thank Emma Simpson for sharing the R code used in [Simpson and Wadsworth \(2021\)](#), Gianmarco Callegher for programming assistance, Stuart Coles for helpful discussions and the referees for their constructive comments.

## References

- Bacro, J. N., Gaetan, C., and Toulemonde, G. (2016). A flexible dependence model for spatial extremes. *Journal of Statistical Planning and Inference*, **172**, 36–52.
- Bacro, J.-N., Gaetan, C., Opitz, T., and Toulemonde, G. (2020). Hierarchical space-time modeling of exceedances with an application to rainfall data. *Journal of the American Statistical Association*, **115**, 555–569.
- Bortot, P. and Tawn, J. (1998). Models for the extremes of Markov chains. *Biometrika*, **85**, 851–867.
- Bortot, P. and Gaetan, C. (2014). A latent process model for temporal extremes. *Scandinavian Journal of Statistics*, **41**, 606–621.
- Brandsma, T. (2014). Comparison of automatic and manual precipitation networks in the Netherlands. TR-347, KNMI. De Bilt.
- Brown, B. and Resnick, S. (1977). Extremes values of independent stochastic processes. *Journal of Applied Probability*, **14**, 732–739.

- Brown, P. E., Roberts, G. O., Kåresen, K. F., and Tonellato, S. (2000). Blur-generated non-separable space–time models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **62**, 847–860.
- Chan, Y. and Li, H. (2008). Tail dependence for multivariate t-copulas and its monotonicity. *Insurance: Mathematics and Economics*, **42**, 763–770.
- Chernick, M. R., Hsing, T., and McCormick, W. P. (1991). Calculating the extremal index for a class of stationary sequences. *Advances in Applied Probability*, **23**, 835–850.
- Chilès, J.-P. and Delfiner, P. (2002). *Geostatistics: Modeling Spatial Uncertainty*. Wiley, New York, second edition.
- Coles, S., Heffernan, J., and Tawn, J. A. (1999). Dependence measures for extreme value analyses. *Extremes*, **2**, 339–365.
- Davis, R. A. and Mikosch, T. (2008). Extreme value theory for space-time processes with heavy-tailed distributions. *Stochastic Processes and their Applications*, **118**, 560–584.
- Davis, R. A. and Yau, C. Y. (2011). Comments on pairwise likelihood in time series models. *Statistica Sinica*, **21**, 255–277.
- Davis, R. A., Klüppelberg, C., and Steinkohl, C. (2013). Statistical inference for max-stable processes in space and time. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **75**, 791–819.
- Davison, A. C. and Gholamrezaee, M. M. (2012). Geostatistics of extremes. *Proceedings of the Royal Society London: Series A*, **468**, 581–608.

Davison, A. C., Huser, R., and Thibaud, A. (2013). Geostatistics of dependent and asymptotically independent extremes. *Mathematical Geosciences*, **45**, 511–529.

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

Demarta, S. and Mcneil, A. J. (2005). Modelling non-stationary extremes with application to surface level ozone. *International Statistical Review*, **73**, 111–129.

Drovandi, C. C. (2019). ABC and indirect inference. In *Handbook of Approximate Bayesian Computation*. Chapman & Hall/CRC, Boca Raton.

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

Economou, T., Stephenson, D. B., and Ferro, C. A. T. (2014). Spatio-temporal modelling of extreme storms. *Annals of Applied Statistics*, **8**, 2223–2246.

Embrechts, P., Koch, E., and Robert, C. (2016). Space-time max-stable models with spectral separability. *Advances in Applied Probability*, **48**, 77–97.

Gallant, A. R. and Tauchen, G. (1996). Which moments to match ? *Econometric Theory*, **4**, 657–681.

Gourieroux, C. and Monfort, A. (2018). Composite indirect inference with application to corporate risk. *Econometrics and Statistics*, **7**, 30–45.

Gourieroux, C., Monfort, A., and Renault, E. (1993). Indirect inference. *Journal of Applied Econometrics*, **8**, S85–S118.

- Heffernan, J. E. and Tawn, J. A. (2004). A conditional approach for multivariate extreme values (with discussion). *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **66**, 497–546.
- Heggland, K. and Frigessi, A. (2004). Estimating functions in indirect inference. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **66**, 447–462.
- 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. 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: Computational Statistics*, page e1537.
- Kabluchko, Z., Schlather, M., and de Haan, L. (2009). Stationary max-stable fields associated to negative definite functions. *The Annals of Probability*, **37**, 2042–2065.
- Klein Tank, A., Wijngaard, J., Können, G., Böhm, R., Demarée, G., Gocheva, A., Mileta, M., Pashiardis, S., Hejkrlik, L., Kern-Hansen, C., et al. (2002). Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. *International Journal of Climatology: a Journal of the Royal Meteorological Society*, **22**, 1441–1453.
- Ledford, A. W. and Tawn, J. A. (1997). Modelling dependence within joint tail regions. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **59**, 475–499.

- Lindsay, B. (1988). Composite likelihood methods. *Contemporary Mathematics*, **80**, 221–239.
- Meinguet, T. (2012). Maxima of moving maxima of continuous functions. *Extremes*, **15**, 267–297.
- Morris, S. A., Reich, B. J., Thibaud, E., and Cooley, D. (2017). A space-time skew-t model for threshold exceedances. *Biometrics*, **73**, 749–758.
- Nieto-Barajas, L. E. and Huerta, G. (2017). Spatio-temporal Pareto modelling of heavy-tail data. *Spatial Statistics*, **20**, 92–109.
- Northrop, P. J. and Jonathan, P. (2011). Threshold modelling of spatially dependent non-stationary extremes with application to hurricane-induced wave heights. *Environmetrics*, **22**, 799–809.
- Resnick, S. I. (2013). *Extreme Values, Regular Variation and Point Processes*. Springer, New York.
- Richards, J., Tawn, J. A., and Brown, S. (2021). Modelling extremes of spatial aggregates of precipitation using conditional methods. *arXiv*, **2102.10906**.
- Røislien, J. and Omre, H. (2006). T-distributed random fields: a parametric model for heavy-tailed well-log data. *Mathematical Geology*, **38**, 821–849.
- Sang, H. and Gelfand, A. (2009). Hierarchical modeling for extreme values observed over space and time. *Environmental and Ecological Statistics*, **16**, 407–426. ISSN 1352-8505.
- Schlather, M. (2002). Models for stationary max-stable random fields. *Extremes*, **5**, 33–44.

- Simpson, E. S. and Wadsworth, J. L. (2021). Conditional modelling of spatio-temporal extremes for Red Sea surface temperatures. *Spatial Statistics*, **41**, 100482.
- Simpson, E. S., Optiz, T., and Wadsworth, J. L. (2020). High-dimensional modeling of spatial and spatio-temporal conditional extremes using INLA and the SPDE approach. *arXiv*, **2011.04486**.
- Smith, A. A. (1990a). *Three essays on the solution and estimation of dynamic macroeconomic models*. PhD thesis, Duke University.
- Smith, R. L. (1990b). Max-stable processes and spatial extremes. Preprint. University of Surrey.
- Smith, R., Tawn, J. A., and Coles, S. (1997). Markov chain models for threshold exceedances. *Biometrika*, **84**, 249–268.
- Turkman, K., Amaral Turkman, M., and Pereira, J. (2010). Asymptotic models and inference for extremes of spatio-temporal data. *Extremes*, **13**, 375–397.
- Varin, C., Reid, N., and Firth, D. (2011). An overview of composite likelihood methods. *Statistica Sinica*, **21**, 5–42.
- Wadsworth, J. and Tawn, J. A. (2012). Dependence modelling for spatial extremes. *Biometrika*, **99**, 253–272.
- Wadsworth, J. and Tawn, J. A. (2019). Higher-dimensional spatial extreme via single-site conditioning. *arXiv*, **1912.06560**.
- Wikle, C. K. and Cressie, N. (1999). A dimension-reduced approach to space-time Kalman filtering. *Biometrika*, **86**, 815–829.

Wood, S. N. (2003). Thin plate regression splines. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **65**, 95–114.

Youngman, B. D. (2019). Generalized additive models for exceedances of high thresholds with an application to return level estimation for US wind gusts. *Journal of the American Statistical Association*, **114**, 1865–1879.

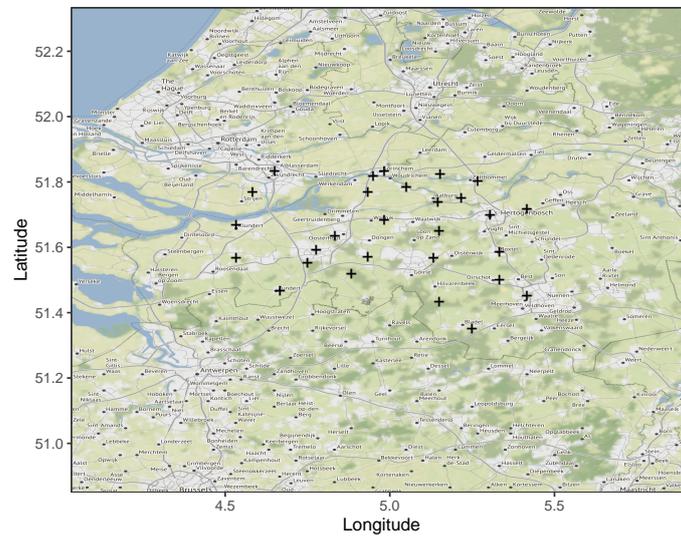


Figure 1: Map of the 28 manual weather stations analyzed in the North Brabant province. Longitude and latitude are in decimal degrees.

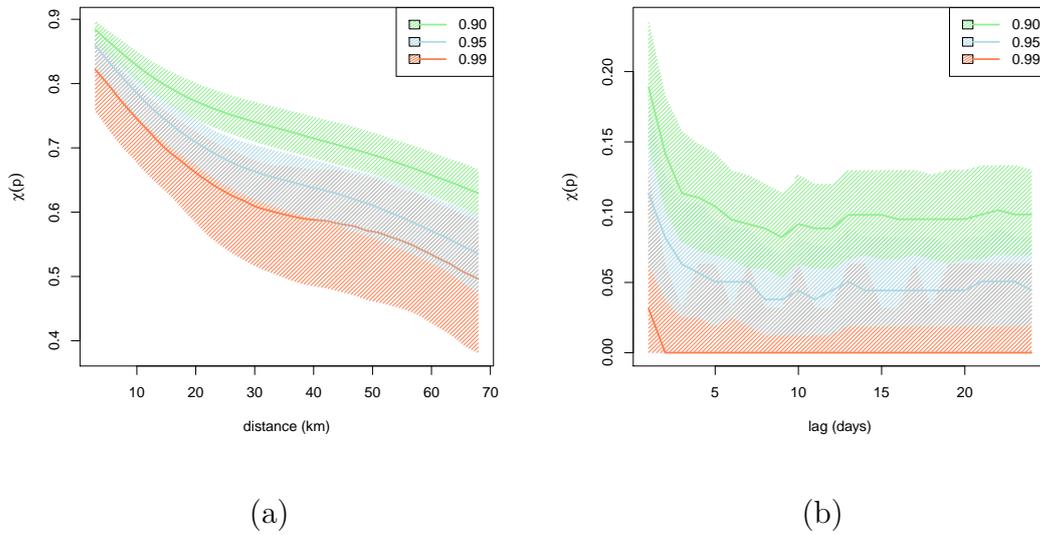


Figure 2: For the North Brabant rainfall data, empirical estimates of  $\chi(p)$ ,  $p = 0.90, 0.95, 0.99$ , for pairs of observations at increasing spatial distances (in km) in (a) and temporal lags (in days) in (b), respectively. The shaded areas represent approximate 95% confidence regions based on a stationary bootstrap procedure. In (a), estimates and bounds are smoothed.

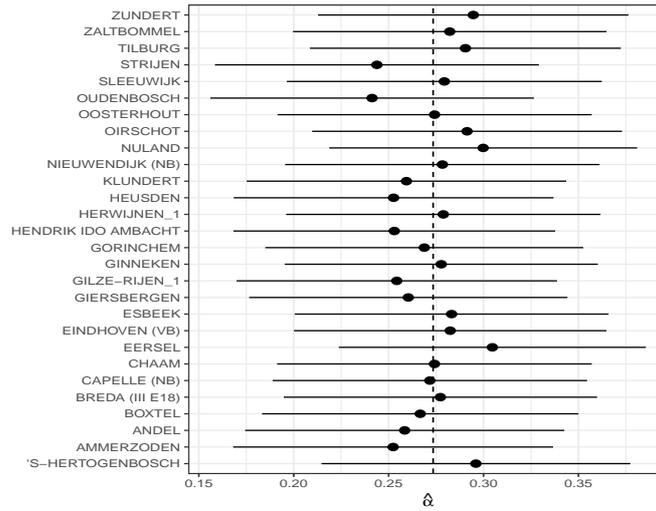


Figure 3: Sitewise estimates and 95% confidence intervals for  $\alpha$  computed from observations above the 0.90 quantile. The vertical line corresponds to the estimate of  $\alpha$  obtained from all the exceedances of the 0.90 marginal quantile without distinction of location.

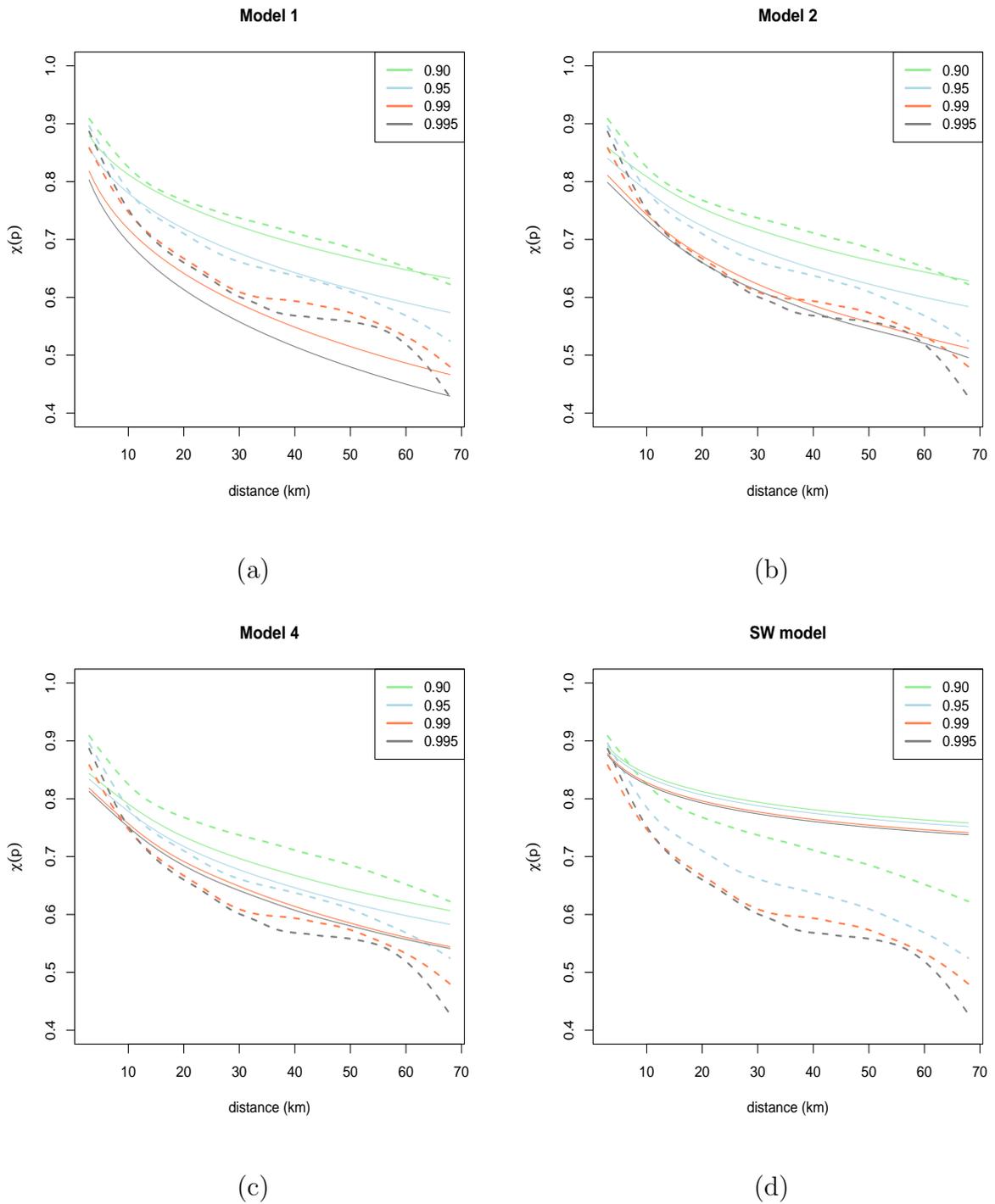


Figure 4: Estimates of  $\chi(p; 0, \|h\|)$  as a function of  $\|h\|$ , for  $p \in \{0.90, 0.95, 0.99, 0.995\}$ . In each plot, the dashed line corresponds to smoothed empirical estimates. In (a) the continuous line corresponds to Model 1 estimates, in (b) to Model 2 estimates, in (c) to Model 4 estimates and in (d) to SW model estimates, respectively.

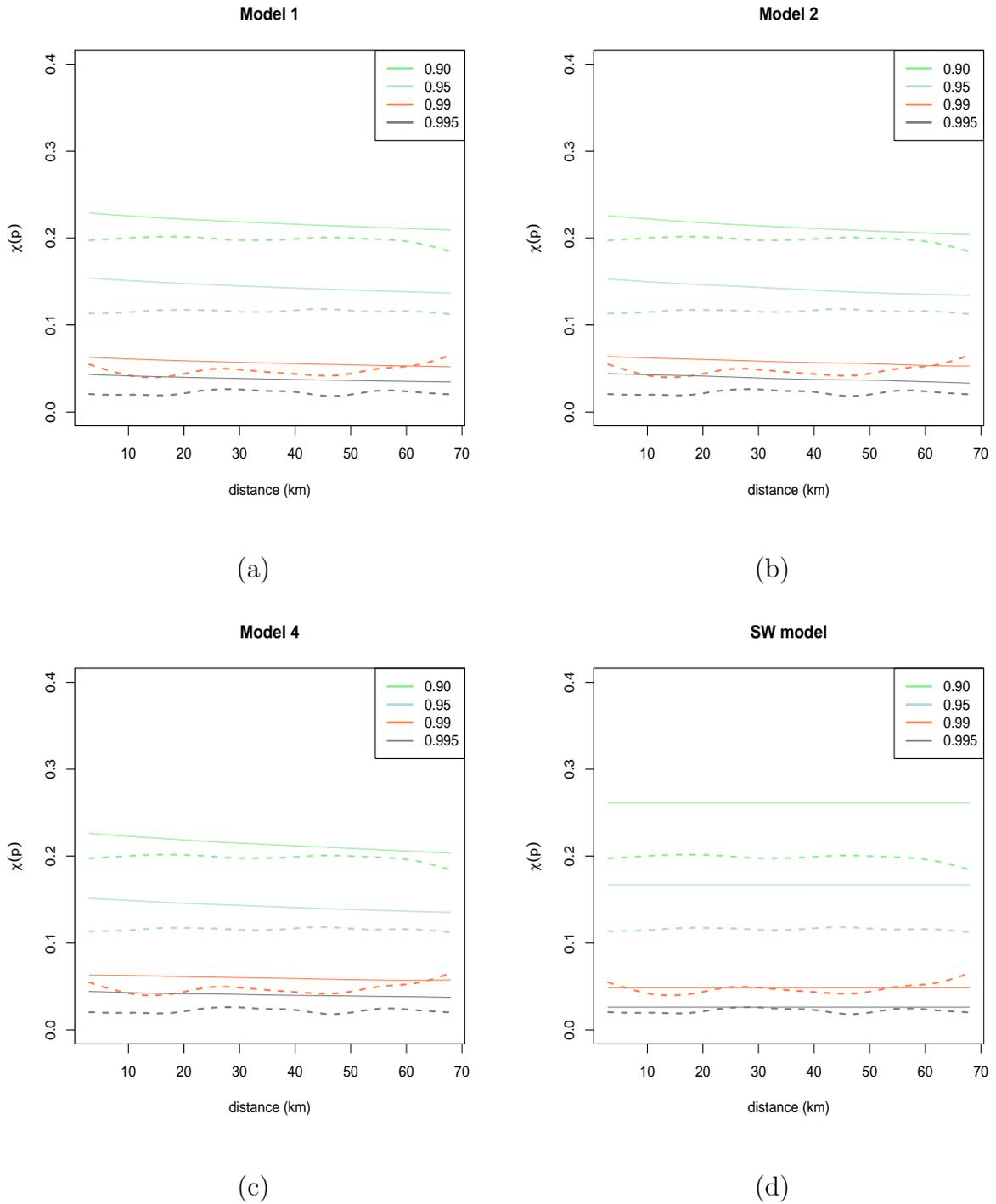


Figure 5: Estimates of  $\chi(p; 1, \|h\|)$  as a function of  $\|h\|$ , for  $p \in \{0.90, 0.95, 0.99, 0.995\}$ . In each plot, the dashed line corresponds to smoothed empirical estimates. In (a) the continuous line corresponds to Model 1 estimates, in (b) to Model 2 estimates, in (c) to Model 4 estimates and in (d) to SW model estimates, respectively.

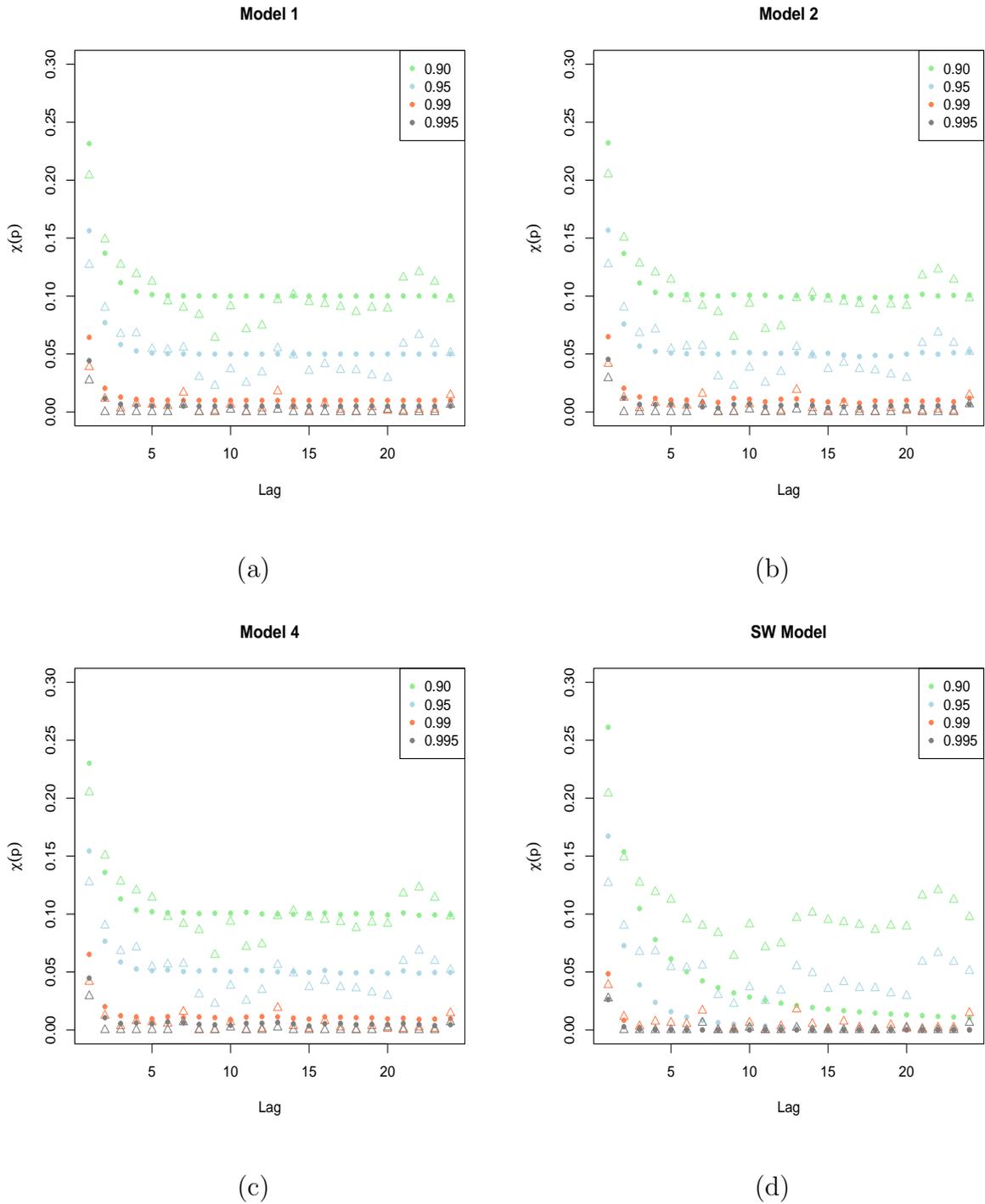


Figure 6: Estimates of  $\chi(p; l, 0)$  as a function of  $l$ , for  $p \in \{0.90, 0.95, 0.99, 0.995\}$ . In each plot, triangles correspond to empirical estimates. In (a) circles correspond to Model 1 estimates, in (b) to Model 2 estimates, in (c) to Model 4 estimates and in (d) to SW model estimates, respectively.

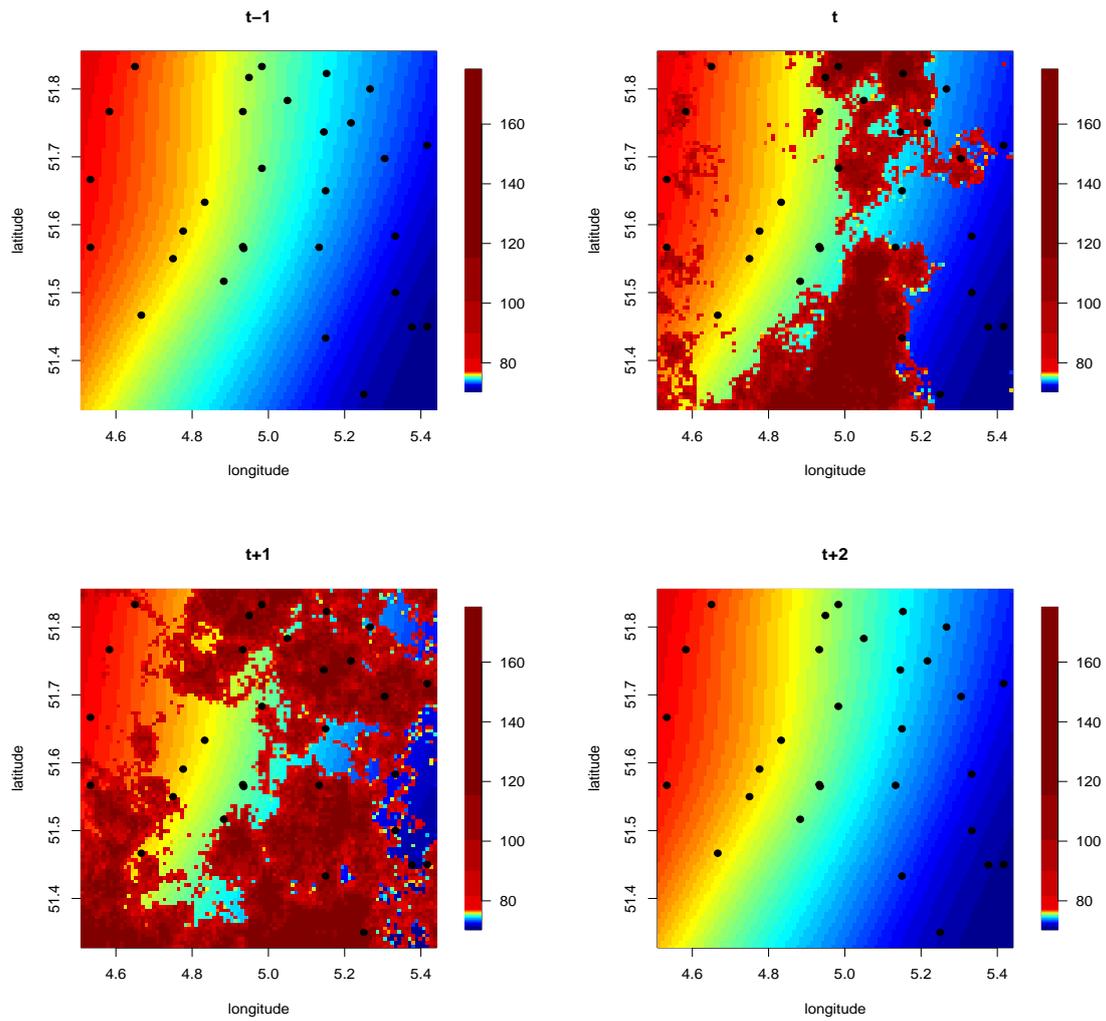


Figure 7: Simulations from the estimated Model 2. Highlighted are the exceedances over the 0.90 marginal quantile. Dots denote the observed stations.

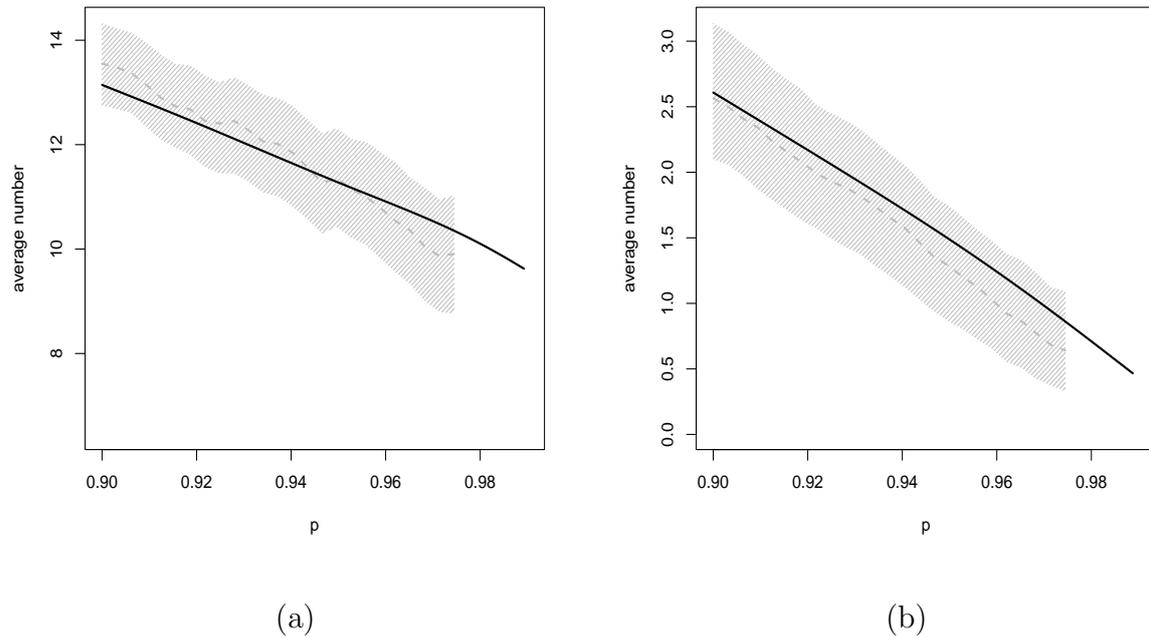


Figure 8: For the North Brabant data, in (a), the continuous line corresponds to Model 2 estimates of  $E(N_t(p, 0) | N_t(p, 0) > 0)$ , and, in (b), of  $E(N_t(p, 1) | N_t(p, 0) > 0)$ , plotted against  $p$ , for  $p \in [0.90, 0.99]$ . The dashed line and the filled region represent smoothed empirical estimates and approximate 95% confidence bands on the empirical estimates, respectively.