

Statistical Modelling of Extremes in Space and Time Using Max-Stable Processes

Christina Katharina Steinkohl



Dissertation
TECHNISCHE UNIVERSITÄT MÜNCHEN
Lehrstuhl für Mathematische Statistik

TECHNISCHE UNIVERSITÄT MÜNCHEN
Lehrstuhl für Mathematische Statistik

Statistical Modelling of Extremes in Space and Time Using Max-Stable Processes

Christina Katharina Steinkohl

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. Folkmar Bornemann
Prüfer der Dissertation: 1. Univ.-Prof. Dr. Claudia Klüppelberg
2. Prof. Richard A. Davis, Ph.D.
Columbia University, New York City, USA
3. Prof. Dr. Thomas Mikosch
University of Copenhagen, Copenhagen, Denmark

Die Dissertation wurde am 20.11.2012 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 30.01.2013 angenommen.

Zusammenfassung

Diese Dissertation befasst sich mit der statistischen Modellierung von Extremwerten in Raum und Zeit. Dabei werden Daten durch einen stochastischen Prozess beschrieben, der auf der stetigen Raum-Zeit Menge $\mathbb{R}^d \times [0, \infty)$ definiert wird.

Für die Modellierung von räumliche Extrema haben sich max-stabile Zufallsfelder als besonders geeignet erwiesen. In dieser Arbeit werden stationäre, max-stabile Prozesse vorgestellt, die zur Modellierung von Raum-Zeit Extrema verwendet werden können. Zunächst wird ein solcher Prozess als Grenzwert von reskalierten, punktwisen Maxima unabhängiger Gauss Prozesse konstruiert. Die so entstehenden Prozesse sind mit einer Klasse von Kovarianzfunktionen der zugrundeliegenden Gauss-Felder assoziiert. Das Analogon zur Kovarianzfunktion für extreme Abhängigkeiten ist gegeben durch das so genannte Extremogramm. Der Zusammenhang zwischen der Kovarianzfunktion, die dem Gauss Prozess zugrunde liegt, und dem Extremogramm wird detailliert herausgearbeitet, und verschiedene Konstruktionsprinzipien für Raum-Zeit Kovarianzfunktionen werden analysiert. Unter anderem wird Gneiting's Klasse von Kovarianzfunktionen in diesen Kontext eingearbeitet. Des Weiteren wird Smith's Sturmprofilmodell in den Raum-Zeit Kontext übertragen und eine explizite Darstellung der bivariaten Verteilungsfunktionen bereitgestellt.

Nach Einführung von Raum-Zeit Parametern wird die paarweise Likelihood Schätzung vorgestellt, bei der die bivariate Dichte des max-stabilen Prozesses verwendet wird. Starke Konsistenz und asymptotische Normalität der Schätzer wird gezeigt unter der Annahme, dass die Beobachtungsorte auf einem regulären Grid liegen. Es werden außerdem Erweiterungen auf irregulär verteilte Beobachtungsorte diskutiert.

Des Weiteren wird eine alternative, semiparametrische Schätzmethode vorgestellt. Basierend auf dem empirischen Extremogramm in Raum und Zeit werden die Parameter mit Hilfe von gewichteter Regressionsverfahren geschätzt. Wir zeigen asymptotische Normalität der Schätzer und verwenden Bootstrap Methoden zur Konstruktion von punktwisen Konfidenzintervallen.

Eine Simulationsstudie untersucht das Verhalten auf kleinen Stichproben und vergleicht die vorgeschlagenen Schätzmethoden.

Abschließend wird das Raum-Zeit Modell und die Schätzmethoden auf Radar Regendaten angewandt, um die extremen Eigenschaften der Daten zu quantifizieren.

Abstract

This thesis deals with the statistical modelling of extreme and rare events in space and time. Max-stable processes have proved to be useful for the statistical modelling of spatial extremes. We introduce families of max-stable processes on the continuous space-time domain $\mathbb{R}^d \times [0, \infty)$.

In a first step, we construct max-stable random fields as limits of rescaled pointwise maxima of independent Gaussian processes. Specific space-time covariance models, which satisfy weak regularity assumptions are employed for the underlying Gaussian process. The analogon of the covariance function for extremal dependence is called the extremogram. We show how the spatio-temporal covariance function underlying the Gaussian process can be interpreted in terms of the extremogram. Within this context, we examine different concepts for constructing covariance functions in space and time, and analyse several specific examples, including Gneiting's class of nonseparable stationary covariance functions. In addition to the above construction, Smith's storm profile model is defined for space-time domains, and explicit expressions for the bivariate distribution functions are provided.

After introducing parameters for the max-stable space-time process, we establish pairwise likelihood estimation, where the pairwise density of the process is used to estimate the model parameters. For regular grid observations we prove strong consistency and asymptotic normality of the estimates as the joint number of spatial locations and time points tends to infinity. Furthermore, we discuss extensions to irregularly spaced locations.

As an alternative to pairwise likelihood estimation we propose a semiparametric estimation procedure based on a closed form expression of the extremogram. In particular, the extremogram is estimated nonparametrically and constrained weighted linear regression is applied to obtain the parameters of interest. We show asymptotic normality of the resulting estimates and discuss bootstrap methods to obtain pointwise confidence intervals for the parameter estimates.

A simulation study illustrates the small sample behaviour of the procedures, and compares the two different approaches for estimating the space-time parameters.

Finally, the introduced model and methods are applied to radar rainfall measurements in order to quantify the extremal properties of the space-time observations.

Acknowledgements

First of all, I would like to thank Claudia Klüppelberg and Richard Davis for being such great supervisors. Their support, confidence in me and their remarkable knowledge always pointed me in the right direction and it was such an honour to work with these excellent professors.

My special thanks go to Claudia Klüppelberg for making this thesis, including all the unforgettable experiences, possible. I am grateful for her encouragement, her helpful advice concerning all theoretical and practical matters, the excellent working conditions, and for giving me the possibility to travel to several conferences, where I was able to present our work and meet famous researchers from all over the world.

In addition, I would like to thank Richard Davis for his excellent ideas, his experience and his great support throughout this thesis. Further, I am particularly grateful to Patti and Richard for their time, the helpful advices during my stay in New York and for welcoming me to their home. I really enjoyed spending time with them in New York and of course also here in Munich.

Further, I want to express my gratefulness to Thomas Mikosch for the enlightening discussion in Denmark and for being a reviewer of my thesis.

In addition, I would like to thank Chin Man Bill Mok and Daniel Straub for their help in finding the data and their discussions regarding the results.

I owe my gratitude to the TUM-Institute of Advanced Study (TUM-IAS) in cooperation with the International Graduate School for Science and Engineering (IGSSE) for their financial support. They made it possible for me to spend several months abroad, including 4 months at Columbia University (New York, USA) and 4 weeks at the Risø National Laboratory for Sustainable Energy (Roskilde, Denmark). Special thanks go to the director of the TUM-IAS, Patrick Dewilde, for his interest in my work, the fruitful discussions, and for giving me the opportunity to be part of the TUM-IAS building opening ceremony. I will never forget this remarkable event!

Furthermore, I would like to thank my colleagues and friends at the department of statistics for their company and many helpful discussions. Especially, I want to thank Florian for sharing the office with me and helping me out several times.

Last, but not least I want to express my sincere gratitude to the most important people in my life: my parents and my sister, Stephanie, for their love and unconditional support at any time, my girls Susie and Babsi for being the best friends one can have, and my boyfriend Schorsch for his love and his patience during my stays abroad.

CONTENTS

1	Introduction	1
1.1	General introduction and motivation	1
1.2	Outline and main objective of this thesis	4
1.3	Open problems for future research	10
2	Preliminaries on extreme value theory and space-time modelling	13
2.1	Extreme value theory	13
2.1.1	Univariate extreme value theory	13
2.1.2	Point processes and regular variation	18
2.1.3	Multivariate extreme value distributions	21
2.1.4	Definition and families of max-stable processes	26
2.2	Space-time models	30
2.2.1	Fundamentals for space-time processes	30
2.2.2	Simulation of Gaussian space-time processes	31
3	Max-stable processes for extremes observed in space and time	35
3.1	Extension of extreme spatial fields to the space-time setting	36
3.1.1	Max-stable random fields based on spatio-temporal correlation functions	36
3.1.2	Extension of the storm profile model	41
3.2	Pickands dependence function and tail dependence coefficient	43
3.3	Possible correlation functions for the underlying Gaussian space-time process	44
3.3.1	Gneiting's class of correlation functions	48
3.3.2	Modelling spatial anisotropy	57
3.4	Proofs	61
3.4.1	Proof of tightness in Theorem 3.1	61

3.4.2	Derivation of the bivariate distribution function for the space-time Smith model	68
4	Composite likelihood methods for max-stable space-time processes	71
4.1	Reminder: Description of model parameters	72
4.2	Pairwise likelihood estimation	74
4.2.1	Composite likelihood estimation for the space-time setting	74
4.2.2	Pairwise likelihood estimation for regular grid observations	76
4.3	Strong consistency of the pairwise likelihood estimates for regular grid observations	78
4.3.1	Ergodic properties for max-stable processes	78
4.3.2	Consistency for large m and T	81
4.4	Asymptotic normality of the pairwise likelihood estimates for regular grid observations	85
4.4.1	Asymptotic normality and α -mixing	87
4.5	Extension to irregularly spaced locations	91
4.5.1	Deterministic irregularly spaced lattice	91
4.5.2	Random locations generated by a Poisson process	93
4.6	Proof of Lemma 4.2	99
5	A semiparametric estimation procedure	103
5.1	Derivation of the semiparametric estimation procedure	104
5.2	Estimation for the space-time Brown-Resnick process	107
5.2.1	Asymptotics of the spatial extremogram	107
5.2.2	Spatial parameter estimates and their properties	114
5.2.3	Temporal parameter estimates and their properties	118
5.3	Bootstrapping parameter estimates for the Brown-Resnick process	122
5.3.1	Spatial parameters: the unconstrained case	123
5.3.2	Spatial parameters: the constrained case	125
5.3.3	Temporal parameters	126
5.4	Some general theory for the spatial extremogram	128
6	Simulation study	133
6.1	Setup for simulation study	133
6.2	Results for pairwise likelihood estimation	134
6.3	Results for the semiparametric estimation procedure using the extremogram .	140
6.4	Comparison of pairwise likelihood and semiparametric estimation	144
7	Analysis of radar rainfall measurements in Florida	149
7.1	Description of data set	149
7.2	Daily maxima of rainfall measurements	150
7.3	Hourly rainfall measurements June 2002-September 2002	159

CHAPTER 1

INTRODUCTION

1.1 General introduction and motivation

In statistics, extreme value theory concentrates on the analysis and quantification of rare and extreme events characterized by exceptionally large (or small) magnitudes compared to the majority of the population. By nature, natural disasters like hurricanes, earthquakes, and droughts leave destruction and chaos behind while passing over certain areas. As an example, consider a tropical storm passing over some region at high wind speeds and with unusual amount of rainfall. Extreme wind or rainfall observations exhibit a spatial dependence structure, meaning that neighbouring locations within some distance show similar patterns, as well as temporal dependence, which can be seen from similar high values for two consecutive time points (e.g. within hours). This thesis aims to develop models and methods which allow for a detailed analysis of extreme events and the extremal space-time dependence structure. Before stating the main results of this thesis, we start with a motivating example concerning rainfall measurements introducing the main problems solved in this thesis.

Motivating example: Rainfall measurements

Heavy rainfall is one of the most important weather risks: Rainfall of extreme magnitude can result in flooding or landslides which in turn threatens human life, disrupt transport and communication, and damage buildings and infrastructure. Flood protection structures like

damns and sea walls are built such that they can withstand extreme hydrological events, which might occur on average only every 100 years.

Figure 1.1.1 shows spatial block maxima of radar rainfall measurements observed on a 12×12 grid of size 10 km in a region in Florida, where each square corresponds to the hourly accumulated spatial maxima of rainfall in inches. In particular, the spatial maxima is calculated for a squared region containing 25 locations and the corresponding time series for each maxima is used for further analysis. The plots show spatial maximal rainfall fields for four consecutive hours (clockwise from the top left to the bottom left). For illustration, the fields in Figure 1.1.1 were chosen such that they represent values of high magnitude with respect to the majority of the data, which can be seen from the time series in Figure 1.1.2, where the highest values are close to 2. From the figures particular patterns for rainfall processes can be detected. We see that neighbouring locations show similar magnitudes, indicating that there is extremal spatial dependence in the data. On the other hand, the rainfall fields for consecutive time points show that temporal dependence is present. Figure 1.1.2 shows the corresponding time series for one fixed grid location (7,7). We see that it is likely that a high value is followed by a value with similar magnitude. In the following, we denote by $\{Z(s, t), s \in \mathbb{R}^d, t \in [0, \infty)\}$ the rainfall process, where $Z(s, t)$ denotes the hourly rainfall at location $s \in \mathbb{R}^d$ (consisting for example of latitude and longitude) and time point $t \in [0, \infty)$. Typical questions arising for the rainfall data are the following:

- What is the probability P that the rainfall process at location s_1 and time point t_1 exceeds a high threshold z_1 , given that the process exceeds a high level z_2 at another location s_2 and time point t_2 ? In particular, how can we predict

$$P(Z(s_1, t_1) > z_1 \mid Z(s_2, t_2) > z_2)?$$

- What is the conditional return level z_c with return period $1/p_c$ of the rainfall process at location s_1 and time point t_1 , given that the process exceeds some threshold z at another space-time location (s_2, t_2) , i.e. what is the predicted level z_c for which

$$P(Z(s_1, t_1) > z_c \mid Z(s_2, t_2) > z) = p_c?$$

- How can we simulate from the extremal rainfall space-time process?

In this thesis, we develop a statistical model together with inference procedures which aim

to answer these questions.

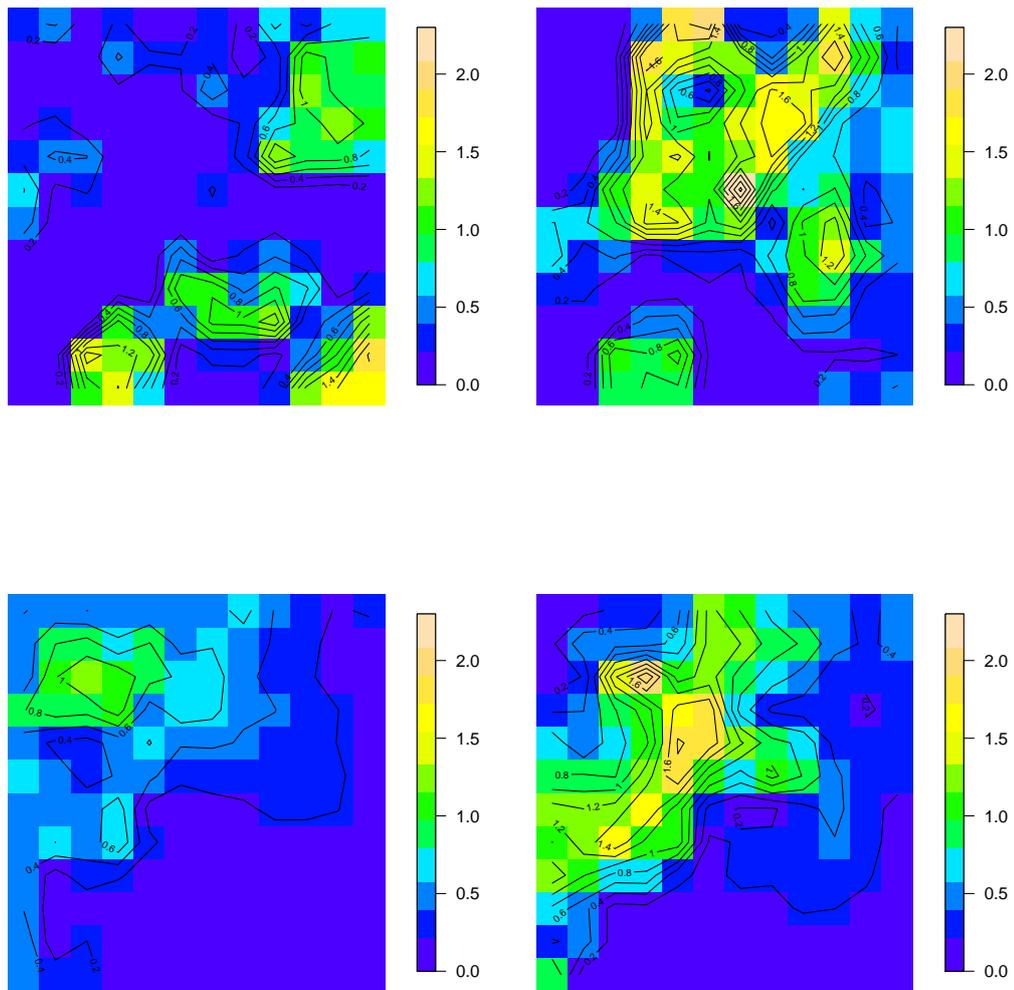


Figure 1.1.1: Spatial maxima of rainfall measurements in inches for four consecutive hours in the wet season 2002 (clockwise from the top left to the bottom left). In particular, the maximum in space is taken over a squared 5×5 region with size 2 km.

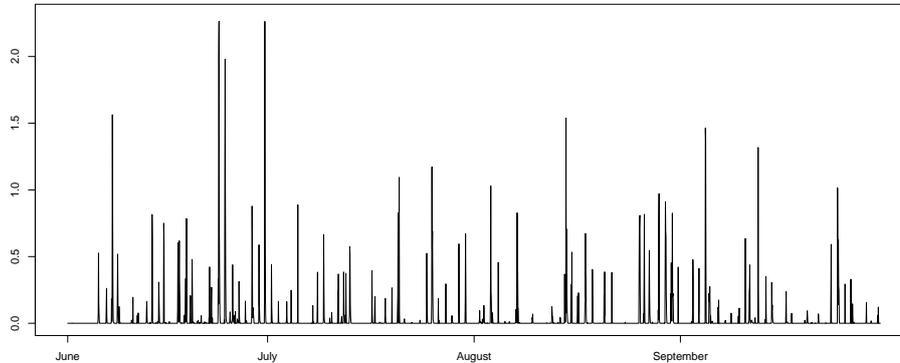


Figure 1.1.2: Hourly accumulated rainfall measurements in inches for one fixed location (grid location (7,7)) in the wet season 2002 (June to September).

1.2 Outline and main objective of this thesis

The main goal of this thesis is the development of a statistical model, which is able to capture extremal dependence in space and time. In addition, procedures for statistical inference are developed, analysed theoretically, tested for simulated data and applied to rainfall observations.

In the following, we denote by $\eta = \{\eta(s, t), s \in \mathbb{R}^d, t \in [0, \infty)\}$ the stochastic process, which is introduced to model extremes on the continuous space-time domain $\mathbb{R}^d \times [0, \infty)$. For a specific location $s \in \mathbb{R}^d$ and some time point $t \in [0, \infty)$, $\eta(s, t)$ denotes the value of interest. The technical basis for modelling extremes as continuous space-time processes is given by max-stable processes as natural extension of the generalized extreme value distributions to infinite dimensions. The general definition is as follows. A stationary continuous space-time process $\{\eta(s, t), s \in \mathbb{R}^d, t \in [0, \infty)\}$ is called max-stable if there exist sequences of constants $a_n(s, t) > 0$ and $b_n(s, t) \in \mathbb{R}$ for $n \in \mathbb{N}$, such that

$$\left\{ a_n(s, t)^{-1} \left(\bigvee_{j=1}^n \eta_j(s, t) - b_n(s, t) \right), s \in \mathbb{R}^d, t \in [0, \infty) \right\},$$

where $\bigvee_{j=1}^n x_j = \max\{x_j, j = 1, \dots, n\}$, is identical in distribution to $\{\eta(s, t), s \in \mathbb{R}^d, t \in [0, \infty)\}$, and η_1, \dots, η_n are n independent replications of η . From the general definition, we

derive special families of max-stable processes, which allow for the introduction of parameters and further properties concerning the extremal space-time dependence structure. In the following, we summarize the main results of this thesis.

Max-stable processes for extremes observed in space and time

Chapter 3 is based on the publication Davis, Klüppelberg and Steinkohl [26] and introduces the space-time process used later on for modelling extremes in space and time.

So far, max-stable processes have mostly been used for the statistical modelling of spatial data. Several examples can be found in the literature, see for example Coles [16] and Coles and Tawn [18], who model extremal rainfall fields using max-stable processes. Another application to rainfall data can be found in Padoan, Ribatet and Sisson [66], who also describe a practicable pairwise likelihood estimation procedure. An interesting application to wind gusts is shown in Coles and Walshaw [19], who use max-stable processes to model the angular dependence for wind speed directions.

In the literature first approaches concerning the analysis and quantification of the extremal behaviour of processes observed both in space and time, where a temporal dependence structure is taken into account, can be found. One idea can be found in Davis and Mikosch [22], who study the extremal properties of a moving average process, where the coefficients and the white-noise process depend on the location and the time point. Sang and Gelfand [74] propose a hierarchical modelling procedure, where on a latent stage spatio-temporal dependence is included via the parameters of the generalized extreme value distribution. Extremes of space-time Gaussian processes have been studied in Kabluchko [49]. He analyses processes of the form $\sup_{t' \in [0, t_m]} Z(s_n s, t')$ for some suitable chosen space-time Gaussian process and shows that the finite dimensional distributions of a properly scaled version converge to those of a max-stable process. An application of combined methods from univariate and bivariate extreme value theory to high frequency wind speed data measured at three masts is shown in Steinkohl, Davis and Klüppelberg [80].

The idea of constructing max-stable random fields as limits of normalized and rescaled pointwise maxima of Gaussian random fields was introduced in Kabluchko, Schlather and de Haan [50], who construct max-stable random fields associated with a class of covariance

functions. In particular, we consider the space-time process, defined by

$$\eta_n(\mathbf{s}, t) = \bigvee_{j=1}^n -\frac{1}{\log(\Phi(Z_j(s_n \mathbf{s}, t_n t)))}, \quad \mathbf{s} \in \mathbb{R}^d, t \in [0, \infty), \quad (1.1)$$

where the positive scaling sequences $(s_n)_{n \in \mathbb{N}}$ and $(t_n)_{n \in \mathbb{N}}$ tend to zero as $n \rightarrow \infty$, and $\Phi(\cdot)$ is the standard normal distribution function. In (1.1) $\{Z_j(\mathbf{s}, t), \mathbf{s} \in \mathbb{R}^d, t \in [0, \infty)\}$ denote independent replications of a Gaussian space-time process $\{Z(\mathbf{s}, t), \mathbf{s} \in \mathbb{R}^d, t \in [0, \infty)\}$ with correlation function γ satisfying

$$(\log n)(1 - \gamma(s_n \mathbf{s}, t_n t)) \rightarrow \delta(\mathbf{s}, t) > 0, \quad n \rightarrow \infty. \quad (1.2)$$

This regularity condition of the correlation function at 0 is taken from a fundamental result in Hüsler and Reiss [47]. Under Condition (1.2), the sequence η_n converges weakly to the max-stable process, defined by

$$\eta(\mathbf{s}, t) = \bigvee_{j=1}^{\infty} \xi_j \exp\{W_j(\mathbf{s}, t) - \delta(\mathbf{s}, t)\}, \quad \mathbf{s} \in \mathbb{R}^d, t \in [0, \infty), \quad (1.3)$$

where ξ_j denote points of a Poisson random measure, W_j are independent replications of a Gaussian process W with stationary increments and correlation function $\delta(\mathbf{s}_1, t_1) + \delta(\mathbf{s}_2, t_2) - \delta(\mathbf{s}_1 - \mathbf{s}_2, t_1 - t_2)$, and the function δ arises from (1.2). The max-stable process in (1.3) is called Brown-Resnick process (see Brown and Resnick [14] and Kabluchko et al. [50]). The main advantage of this approach is the fact, that we can easily simulate from the model by taking pointwise maxima of independent realizations from a Gaussian space-time process and rescaling properly.

In an earlier paper Smith [79] introduced another family of max-stable processes, which became known as the storm profile model. The process is based on points of a Poisson random measure $\{(\xi_j, \mathbf{z}_j, x_j), j \in \mathbb{N}\}$ together with a kernel function f , which in particular can be a centred Gaussian density. The max-stable process is then given by

$$\eta(\mathbf{s}, t) = \bigvee_{j=1}^{\infty} \xi_j f(\mathbf{z}_j, x_j; \mathbf{s}, t), \quad \mathbf{s} \in \mathbb{R}^d, t \in [0, \infty).$$

For the construction in (1.1) we establish an explicit connection between the limit function $\delta(\mathbf{h}, u)$, which arises from the correlation function underlying the Gaussian space-time

process $\{Z(\mathbf{s}, t), \mathbf{s} \in \mathbb{R}^d, t \in [0, \infty)\}$, and the tail dependence coefficient of the max-stable space-time process. Recently, the development of covariance models in space and time has received much attention and there is now a large literature available for the construction of a wide-range of spatio-temporal covariance functions. Examples can be found in Cressie and Huang [21], Gneiting [43], Ma [63, 61], and Schlather [76]. Within this context, we introduce a condition on correlation functions, generalized from the analysis of extremes for stationary Gaussian processes (see for instance Leadbetter et al. [57], Chapter 12). The condition is sufficient for (1.2), introduces parameters to the model and allows for an explicit expression of the δ function,

$$\delta(\mathbf{s}, t) = \theta_1 \|\mathbf{s}\|^{\alpha_1} + \theta_2 t^{\alpha_2}, \quad \mathbf{s} \in \mathbb{R}^d, t \in [0, \infty), \quad (1.4)$$

where $\theta_1, \theta_2 > 0$ and $\alpha_1, \alpha_2 \in (0, 2]$ denotes the parameters of interest, and $\|\cdot\|$ is an arbitrary norm on \mathbb{R}^d . This limit function occurs as large class of covariance functions. We show how Gneiting's class of nonseparable and isotropic covariance functions [43] fits into this framework.

In addition, we examine spatial anisotropic correlation functions, which allow for directional dependence in the spatial components. Perhaps the easiest way to introduce anisotropy in a model is to use geometric anisotropy. For a detailed introduction, we refer to Wackernagel [88], Chapter 9. Using this concept in the underlying correlation function, we can model anisotropy in the corresponding max-stable random field. Furthermore, we revisit a more elaborate way of constructing anisotropic correlation models based on Bernstein functions introduced in Porcu, Gregori and Mateu [70], called the Bernstein class.

Pairwise likelihood estimation for max-stable space-time processes

The main difficulty in deriving parameter estimates in max-stable processes is the fact that the finite-dimensional distribution functions and, thus, the densities are intractable, which precludes the use of standard maximum likelihood procedures. On the other hand, pairwise likelihood methods, where only the pairwise density is needed, can be implemented. These methods go back to Besag [8] and Godambe [44], and there is an extensive literature available dealing with applications and properties of the estimates, see for example Cox and Reid [20], Lindsay [60], Varin [86], or Varin and Vidoni [87]. Recent work concerning the application of pairwise likelihood methods to max-stable random fields can be found in Huser and Davison [46] and Padoan, Ribatet and Sisson [66].

In Chapter 4 results are presented concerning statistical inference using pairwise likelihood

methods, taken from Davis, Klüppelberg and Steinkohl [27]. In particular, we study pairwise likelihood methods for the Brown-Resnick process defined in (1.3) with δ given in (1.4). The pairwise likelihood function for a general setting with M locations and T time points is defined as a function of the parameter vector $\boldsymbol{\psi} = (\theta_1, \alpha_1, \theta_2, \alpha_2)$ by

$$PL^{(M,T)}(\boldsymbol{\psi}) = \sum_{i=1}^{M-1} \sum_{j=i+1}^M \sum_{k=1}^{T-1} \sum_{l=k+1}^T w_{i,j}^{(M)} w_{k,l}^{(T)} \log f_{\boldsymbol{\psi}}(\eta(s_i, t_k), \eta(s_j, t_l)),$$

where $w_{i,j}^{(M)} \geq 0$ and $w_{k,l}^{(T)} \geq 0$ denote spatial and temporal weights, respectively, and $f_{\boldsymbol{\psi}}$ is the bivariate density of the max-stable space-time process containing the parameter vector $\boldsymbol{\psi} = (\theta_1, \alpha_1, \theta_2, \alpha_2)$. The estimates are obtained by maximizing $PL^{(M,T)}(\boldsymbol{\psi})$ with respect to $\boldsymbol{\psi}$:

$$\hat{\boldsymbol{\psi}} = \arg \max_{\boldsymbol{\psi}} PL^{(M,T)}(\boldsymbol{\psi}).$$

We start by showing asymptotic properties of the estimates for an increasing number of space-time observations if locations lie on a regular grid, i.e. the set of locations is given by $\{(i_1, \dots, i_d), i_1, \dots, i_d \in \{1, \dots, m\}\}$, and for equidistant time points. In particular, we show strong consistency

$$\hat{\boldsymbol{\psi}} \xrightarrow{a.s.} \boldsymbol{\psi}^*, \quad M, T \rightarrow \infty,$$

where $\xrightarrow{a.s.}$ denotes almost sure convergence and $\boldsymbol{\psi}^*$ is the true parameter vector, and asymptotic normality of the estimates,

$$(MT)^{1/2}(\hat{\boldsymbol{\psi}} - \boldsymbol{\psi}^*) \xrightarrow{d} \mathcal{N}(0, F^{-1}\Sigma F^{-1\top}), \quad M, T \rightarrow \infty,$$

where \xrightarrow{d} is convergence in distribution, and $F^{-1}\Sigma F^{-1\top}$ is some covariance matrix. In contrast to previous studies we assume a spatial and temporal dependence structure and show the asymptotic properties for a jointly increasing number of spatial locations and time points. In addition, theorems in the literature addressing asymptotic properties for pairwise likelihood estimates often have restrictive assumptions, such as finite moment conditions of high order, which might not be reasonable in practical applications. For the setting considered in this thesis very weak assumptions are sufficient. In addition to the results described above, we discuss extensions to settings, where locations are irregularly spaced. For example, we consider a deterministic irregularly spaced grid, and randomly spaced locations generated by a Poisson process.

A semiparametric estimation procedure for the estimation of the parameters in a max-stable space-time process

As we will see in the simulation study, pairwise likelihood estimation is a reliable way of estimating the parameters of a max-stable process. Some disadvantages should be mentioned here. First, the computation time for evaluating the pairwise likelihood function is rather high and the maximization of the pairwise likelihood function requires optimization routines for which accurate starting values are needed. In addition, the asymptotic variance of the parameter estimates is analytically intractable which leads to the use of resampling methods like the bootstrap. These methods sample from the data, and parameters are estimated based on the resampled data. The computation time for the usage of such methods is extensive.

The structure of the Brown-Resnick process in (1.3) allows us to introduce a new semi-parametric estimation procedure for the parameters, which provides a fast method to estimate the parameters. The semiparametric estimates could be used as starting values for the optimization algorithm used to maximize the pairwise log-likelihood function. The following results are taken from Davis, Klüppelberg and Steinkohl [28]. The method is based on the extremogram, which is the natural extreme analogue of the correlation function of a stationary process. It was introduced in Davis and Mikosch [23] and extended to a spatial setting in Cho, Davis and Ghosh [15]. A special case was also considered in Fasen, Klüppelberg and Schlather [37]. The extremogram for a stationary space-time process $\{X(\mathbf{s}, t), (\mathbf{s}, t) \in \mathbb{R}^d \times [0, \infty)\}$ is defined by

$$\rho_{AB}(r, u) = \lim_{z \rightarrow \infty} \frac{P(z^{-1}X(\mathbf{s}, t) \in A, z^{-1}X(\mathbf{s} + \mathbf{h}, t + u) \in B)}{P(z^{-1}X(\mathbf{s}, t) \in A)}, \quad \min\{r = \|\mathbf{h}\|, u\} \geq 0,$$

where A and B are Borel sets bounded away from 0. The special case $A = B = (1, \infty)$ is known as the tail dependence coefficient and can be calculated for the Brown-Resnick process in (1.3) as

$$\rho_{(1, \infty)(1, \infty)}(r, u) = 2(1 - \Phi(\sqrt{\theta_1 r^{\alpha_1} + \theta_2 u^{\alpha_2}})),$$

with parameters $\theta_1, \alpha_1, \theta_2$ and α_2 , where $r = \|\mathbf{h}\|$ is the spatial lag and u is the time lag. Setting, for example, the temporal lag u equal to zero and transforming the equation leads to

$$2 \log \left(\Phi^{-1} \left(1 - \frac{1}{2} \chi(r, 0) \right) \right) = \log(\theta_1) + \alpha_1 \log(r),$$

which is a linear function in $\log(r)$. Given a nonparametric estimate for $\rho_{(1,\infty)(1,\infty)}(r, 0)$, the parameters can be estimated by the method of least squares. The same approach works for the temporal parameters θ_2 and α_2 . As for the pairwise likelihood method we show that the estimates are asymptotically normal. Since the resulting asymptotic covariance matrices for the parameter estimates are intractable we apply bootstrap procedures as done for the extremogram in Davis, Mikosch and Cribben [25] to obtain pointwise confidence intervals for the parameters.

Outline of the thesis

After stating some preliminaries on univariate and multivariate extreme value theory together with fundamentals for space-time processes in Chapter 2, we develop the max-stable space-time process, which is used throughout this thesis in Chapter 3. Statistical inference procedures are described and analysed in Chapters 4 and 5. In Chapter 4 we introduce pairwise likelihood estimation and show strong consistency and asymptotic normality of the parameter estimates. Chapter 5 introduces a semiparametric estimation procedure based on the extremogram. A simulation study in Chapter 6 illustrates the small sample behaviour of the estimation methods developed in Chapters 4 and 5. In Chapter 7 we return to the radar rainfall measurements described in the motivating example and quantify the extremal behaviour using the process and methods introduced in this thesis.

1.3 Open problems for future research

To conclude the introduction, we would like to mention some problems which are subject to future research.

Nonseparability of parameters in the extremes

In this thesis, we assume that the correlation function of the Gaussian process in the construction of the max-stable space-time process (cf. (1.1)) has an expansion around zero, such that the δ function describing the extremal space-time dependence is given by $\delta(\mathbf{h}, u) = \theta_1 \|\mathbf{h}\|^{\alpha_1} + \theta_2 |u|^{\alpha_2}$. This assumption is satisfied by a large class of correlation functions. In Chapter 3 we see that even for nonseparable correlation models like Gneiting's class, the spatial parameters θ_1 and α_1 separate from the temporal parameters θ_2 and α_2 in the described way. This property allows us to estimate the spatial and temporal parameters separately by

setting either the spatial or the temporal lag equal to zero. A possible extension of our model is to assume that some of the spatial parameters depend on the temporal lag u and that the temporal parameters are modelled as a function of the spatial lag $\|\mathbf{h}\|$. For example, we could model the δ function by

$$\delta(\mathbf{h}, u) = \theta_1(u)\|\mathbf{h}\|^{\alpha_1} + \theta_2(\|\mathbf{h}\|)u^{\alpha_2},$$

for suitable relationships for the parameters and the space-time lags.

Introducing anisotropy

In Chapter 3 we show how spatial anisotropy can be introduced to the max-stable space-time process in (1.3). In particular, using anisotropic correlation functions in the underlying Gaussian process in the construction of the max-stable process (see (1.1)) relate to an anisotropic structure in the extremogram of the limit process. For given data, it might be important to check whether the extremes in space and time have directional dependence. One possibility is as follows. First, parameters have to be introduced to the model, which describe the anisotropic behaviour. Pairwise likelihood estimation is a possible method to estimate these parameters, but one has to be careful with the identifiability of the pairwise densities, which might cast problems if more parameters are introduced. Once, the parameters are estimated, one could test whether they have a significant influence for the given data.

Nonstationary max-stable space-time processes

Throughout the thesis we assume that the considered process is stationary in space and time. This is needed as usual in all proofs concerning properties of the estimates from pairwise likelihood and the semiparametric estimation procedure. The analysis of high frequency wind speed (cf. Steinkohl et al. [80]) shows for example, that wind speed time series are nonstationary, which is clearly the case for many environmental data. A topic for future research is, therefore, the extension of the max-stable space-time process developed to the nonstationary case. One possibility is to assume a nonstationary correlation function C , which satisfies

$$(\log n)(1 - C(s_n s_1, t_n t_1; s_n s_2, t_n t_2)) \rightarrow \delta(s_1, t_1; s_2, t_2), \quad n \rightarrow \infty,$$

and δ does not only depend on the space-time lag $(s_1 - s_2, t_1 - t_2)$. One has to check whether the theoretical results still hold for this setting.

CHAPTER 2

PRELIMINARIES ON EXTREME VALUE THEORY AND SPACE-TIME MODELLING

This chapter introduces theoretical fundamentals, which will be used in the following Chapters. After reviewing well-known results in univariate and multivariate extreme value theory, max-stable processes are defined, which build the technical basis for models developed in this thesis. In addition, we state some general definitions for space-time processes, which will be frequently used. Furthermore, we describe how Gaussian space-time processes can be simulated using circular embedding.

2.1 Extreme value theory

2.1.1 Univariate extreme value theory

Classical results in extreme value theory

Univariate extreme value theory is well established and detailed introductions can be found in textbooks and lecture notes. Well-known representatives are for example Beirlant et al. [7], Coles [17], Embrechts, Klüppelberg and Mikosch [36], de Haan and Ferreira [30] and

Leadbetter, Lindgren and Rootzén [57]. We start with the fundamental theorem, which introduces the *generalized extreme value distribution* (GEV) as appropriate limit distribution for rescaled maxima of independent and identical distributed (iid) random variables.

Theorem 2.1 (Fisher Tippett [39], Gnedenko[42]). *Let X_1, \dots, X_n be iid random variables with distribution function F . Assume there exist sequences of constants $(a_n)_{n \in \mathbb{N}} > 0$ and $(b_n)_{n \in \mathbb{N}} \in \mathbb{R}$, and a non-degenerate distribution function G such that for all $x \in \mathbb{R}$, for which the limit is continuous,*

$$P \left(\frac{\bigvee_{i=1}^n X_i - b_n}{a_n} \leq x \right) \rightarrow G(x), \quad \text{as } n \rightarrow \infty, \quad (2.1)$$

where $\bigvee_{i=1}^n x_i = \max\{x_i, i = 1, \dots, n\}$. Then, G has the representation:

$$G(x) = \begin{cases} \exp \left\{ - \left(1 + \xi \frac{x-\mu}{\sigma} \right)^{-1/\xi} \right\}, & \xi \neq 0, \\ \exp \left\{ -e^{-\frac{x-\mu}{\sigma}} \right\}, & \xi = 0, \end{cases} \quad (2.2)$$

provided that $1 + \xi \frac{x-\mu}{\sigma} > 0$, $\sigma > 0$ and $\mu, \xi \in \mathbb{R}$. We say that F is in the maximum domain of attraction of G .

The variables $\sigma > 0$ and $\mu \in \mathbb{R}$ denote scale and location parameters, respectively. The shape parameter $\xi \in \mathbb{R}$ determines the tail behaviour of the distribution. Accordingly, it divides the GEV in the following three standardized families of distributions.

$$\begin{aligned} \text{Type I (Gumbel, } \xi = 0) : \quad & G(x) = \exp \left\{ -e^{-x} \right\}, \quad x \in \mathbb{R}, \\ \text{Type II (Fréchet, } \xi = \alpha^{-1} > 0) : \quad & G(x) = \begin{cases} 0, & x < 0, \\ \exp \left\{ -x^{-\alpha} \right\}, & x \geq 0, \end{cases} \\ \text{Type III (Weibull, } \xi = -\alpha^{-1} < 0) : \quad & G(x) = \begin{cases} \exp \left\{ -(-x)^\alpha \right\}, & x \leq 0, \\ 1, & x \geq 0. \end{cases} \end{aligned}$$

Figure 2.1.1 visualizes the densities of the three families.

An important characterization of extreme value distributions is max-stability.

Definition 2.1 (Max-stability). *A distribution function G is called max-stable if for all inte-*

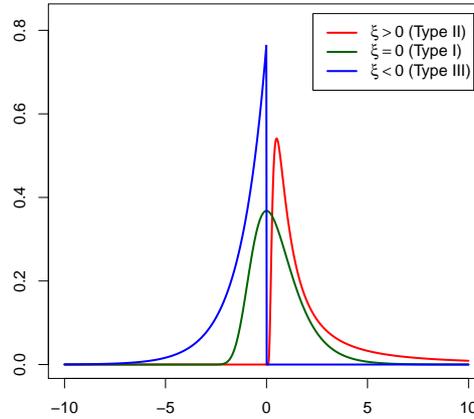


Figure 2.1.1: Densities of the GEV distributions

gers $k \in \mathbb{N}$ there exist sequences of constants $(\alpha_k)_{k \in \mathbb{N}} > 0$ and $(\beta_k)_{k \in \mathbb{N}} \in \mathbb{R}$, such that

$$G^k(\alpha_k x + \beta_k) = G(x), \quad k \in \mathbb{N}.$$

Equivalently, if X_1, \dots, X_k are iid random variables with distribution function G , then G is max-stable if there exist $(\alpha_k)_{k \in \mathbb{N}} > 0$ and $(\beta_k)_{k \in \mathbb{N}} \in \mathbb{R}$ such that $(\bigvee_{i=1}^k X_i - \beta_k) / \alpha_k$ is identical in distribution to X_1 .

The following theorem is fundamental for the generalization of univariate extreme value distributions to multivariate and infinite dimensions. A proof can be found for example in Embrechts et al. [36].

Theorem 2.2. *A distribution function G is max-stable if and only if it is a generalized extreme value distribution.*

Example 2.1. For later purposes we present results for the standard normal distribution. In particular, defining the sequences $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$ by

$$b_n = (2 \log n - \log \log n - 4 \log(4\pi))^{1/2},$$

$$a_n = 1/b_n,$$

for $n \in \mathbb{N}$, it follows that

$$\lim_{n \rightarrow \infty} n(1 - \Phi(a_n x + b_n)) = e^{-x}, \quad x \in \mathbb{R},$$

where Φ is the standard normal distribution function. This implies that the normal distribution is in the Gumbel (Type I) domain of attraction. A proof of this result can be found in de Haan and Ferreira [30].

For modelling purposes we state the following important result, which relates the asymptotic distribution of maxima to the distribution of threshold exceedances.

Theorem 2.3. *The distribution function F is in the domain of attraction of the extreme value distribution G with shape parameter ξ , if and only if there exists a positive function f such that*

$$\lim_{u \rightarrow x^*} \frac{1 - F(u + xf(u))}{1 - F(u)} = \begin{cases} (1 + \xi x)^{-1/\xi}, & \xi \neq 0, \\ e^{-x}, & \xi = 0, \end{cases}$$

where $x^* = \sup\{x : F(x) < 1\}$ is the upper right endpoint of F .

Remark 2.1. The excess distribution equals

$$\begin{aligned} P(X > x) &= P(X > u)P(X > x - u \mid X > u) = P(X > u) \frac{1 - F(u + x)}{1 - F(u)} \\ &= P(X > u) \frac{1 - F(u + xf(u)/f(u))}{1 - F(u)}. \end{aligned}$$

By using Theorem 2.3 and interpreting $1/f(u)$ as scale parameter $\tilde{\sigma} > 0$, the conditional distribution $P(X \leq x - u \mid X > u)$ can be approximated by

$$GPD(x) = \begin{cases} 1 - \left(1 + \frac{\xi x}{\tilde{\sigma}}\right)^{-1/\xi}, & \xi > 0, x \in (0, \infty) \text{ or } \xi < 0, x \in (0, -x/\xi), \\ 1 - e^{-x/\tilde{\sigma}}, & \xi = 0, x \in (0, \infty). \end{cases}$$

This distribution function is called *generalized Pareto distribution* (GPD). Note, that the parameters of the corresponding GEV distribution μ , ξ and σ are linked to the GPD parameter $\tilde{\sigma}$ through $\tilde{\sigma} = \sigma + \xi(u - \mu)$. The shape parameter ξ is the same in both distributions.

Statistical modelling of univariate extremes

We describe two methods for modelling univariate extreme values, which will be applied to the rain data in Chapter 7. First, we describe the estimation based on block maxima. Let $Z_1, \dots, Z_{k/n}$ denote sample maxima with block length k based on observations X_1, \dots, X_n , i.e. $Z_j = \bigvee_{l=k(j-1)+1}^{kj} X_l$ for $j = 1, \dots, k/n$. The parameters of the GEV can be estimated using maximum-likelihood estimation, where the log-likelihood function, given by

$$l(\mu, \sigma, \xi) = -\frac{k}{n} \log(\sigma) + (1/\xi + 1) \sum_{j=1}^{k/n} \log\left(1 + \xi \frac{Z_j - \mu}{\sigma}\right) - \sum_{j=1}^{k/n} \left(1 + \xi \frac{Z_j - \mu}{\sigma}\right)^{-1/\xi},$$

provided that $1 + \xi(Z_j - \mu)/\sigma > 0$ for $j = 1, \dots, k/n$, is maximized.

An important application of extreme value theory is the prediction of extreme *return levels*, which are defined as the level z_p for which the probability that the underlying random variable exceeds z_p is equal to a prespecified value $p \in (0, 1)$, i.e. $P(X > z_p) = p$. The value $1/p$ is called return period. Return levels with long return period, i.e. for very low values of p can be predicted by using the quantiles resulting from the fitted GEV. In particular, the predicted $1/p$ -return level is given by

$$\hat{z}_p = \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}} (1 - (-\log(1 - p)))^{-\hat{\xi}}.$$

Note, that the description above holds for $\xi \neq 0$. For shape parameter estimates close to zero one should check whether the GEV with $\xi = 0$ is more appropriate. We refer to Coles [17] for more information.

Another way to quantify the extremal behavior of data is to model threshold exceedances instead of block maxima. Let X_1, \dots, X_n be independent random variables with distribution function $F \in MDA(G)$ and define

$$N_u = \#\{i \in \{1, \dots, n\} : X_i > u\}$$

as the number of exceedances Y_1, \dots, Y_{N_u} , where $Y_j = X_j - u$, if $X_j > u$, $j = 1, \dots, N_u$. From Theorem 2.3 and Remark 2.1 the conditional excess distribution $P(X - u > x | X > u)$ can be approximated by a GPD. Using threshold exceedances, the parameters $\tilde{\sigma}$ and ξ are

estimated by maximizing the log-likelihood function

$$l(\tilde{\sigma}, \xi) = -N_u \log(\tilde{\sigma}) - (1/\xi + 1) \sum_{j=1}^{N_u} \log\left(1 + \xi \frac{Y_j}{\tilde{\sigma}}\right),$$

with respect to $\tilde{\sigma} > 0$ and $\xi \in \mathbb{R}$. An estimate for the excess distribution $P(X > x)$ is then given by

$$\hat{P}(X > x) = \frac{N_u}{n} \left(1 + \hat{\xi} \frac{x - u}{\hat{\sigma}}\right)^{-1/\hat{\xi}},$$

where N_u/n approximates the probability $P(X > u)$. Return levels are predicted by

$$\hat{z}_p = \frac{\hat{\sigma}}{\hat{\xi}} \left(\left(\frac{n}{N_u} (1 - p) \right)^{-\hat{\xi}} - 1 \right) + u.$$

2.1.2 Point processes and regular variation

For the definition of max-stable processes and further properties of extreme value distributions we introduce point processes and the relation to extreme value distributions. If not stated differently, the following definitions and results are taken from Resnick [71, 72, 73]. Suppose (Ω, \mathcal{F}, P) is a probability space and S is a subspace of the Euclidean space with Borel σ -algebra \mathcal{S} . Let further $M_p(S)$ denote the set of all measures μ on (Ω, \mathcal{F}, P) with $\mu(A) \in \mathbb{N}$ for all $A \in \mathcal{F}$, and let $\mathcal{M}_p(S)$ be the Borel σ -algebra of subsets of $M_p(S)$ generated by open sets. Formally, a *point process* N with state space S is a measurable mapping from (Ω, \mathcal{F}) to $(M_p(S), \mathcal{M}_p(S))$. A more intuitive representation of point processes is given as follows. Let $\{X_n, n \geq 0\}$ denote random points in the space S . For $X_n, n \geq 0$, we define the discrete measure $\epsilon_{X_n} : \mathcal{S} \rightarrow \{0, 1\}$ by

$$\epsilon_{X_n}(A) = \begin{cases} 1, & X_n \in A, \\ 0, & X_n \notin A, \end{cases}$$

called *Dirac measure*. A *point process* on (S, \mathcal{S}) is defined as the counting measure N with

$$N(\cdot) = \sum_{n \geq 0} \epsilon_{X_n}(\cdot),$$

and $N(A) < \infty$ for $A \in \mathcal{S}$ compact, i.e. N is a *Radon measure*. For $A \in \mathcal{F}$, $N(A)$ is the random number of points X_n falling into the set A .

Definition 2.2 (Poisson Point process). *A point process N on (S, \mathcal{S}) is called Poisson process or Poisson random measure with mean measure μ , denoted by $PRM(\mu)$, if*

1. For $A \in \mathcal{S}$

$$P(N(A) = x) = \begin{cases} \frac{1}{x!} e^{-\mu(A)} (\mu(A))^x, & \mu(A) < \infty, \\ 0, & \mu(A) = \infty. \end{cases}$$

2. The random variables $N(A_1), \dots, N(A_m)$ are independent for disjoint $A_1, \dots, A_m \in \mathcal{S}$, $m \in \mathbb{N}$.

A Poisson random measure is called *homogeneous* if the mean measure is a multiple of the Lebesgue measure.

In order to connect point processes to extreme value theory, we define point process convergence. The definition is taken from Embrechts et al. [36], Chapter 5.

Definition 2.3 (Weak convergence of point processes). *Let N, N_1, N_2, \dots denote point processes on (S, \mathcal{S}) . The sequence of point processes (N_n) converges weakly to N in $M_p(S)$, denoted by $N_n \Rightarrow N$, if*

$$P(N_n(A_1) = x_1, \dots, N_n(A_m) = x_m) \rightarrow P(N(A_1) = x_1, \dots, N(A_m) = x_m), \quad n \rightarrow \infty,$$

for all possible choices of sets $A_j \in \mathcal{S}$ for which $P(N(\partial A_j) = 0) = 1$, $j = 1, \dots, m$, $m \in \mathbb{N}$ and ∂A_j denotes the boundary of A_j , i.e. for weak convergence of point processes it is sufficient to show convergence of the finite-dimensional distributions.

We connect weak convergence of Poisson random measures with the convergence of their mean measure. To do so, we introduce vague convergence for random measures.

Definition 2.4 (Vague convergence). *Let μ_n, μ be non-negative Radon measures on (S, \mathcal{S}) , i.e. $\mu_n(K) < \infty$, $\mu(K) < \infty$ for bounded Borel sets $K \in \mathcal{S}$. μ_n converges vaguely to μ , $\mu_n \xrightarrow{v} \mu$, if*

$$\int_S f(x) \mu_n(dx) \rightarrow \int_S f(x) d\mu(dx), \quad n \rightarrow \infty,$$

for all continuous non-negative functions $f : S \rightarrow \mathbb{R}_+$ with compact support, i.e. there exists $K \subset S$ such that $f(x) = 0$ for all $x \in S \setminus K$.

Useful characterizations of vague convergence can be found in Resnick [71], Chapter 3.4. For Poisson random measures the following theorem summarizes some basic convergence facts.

Theorem 2.4. *The following statements hold*

1. *Let N, N_1, N_2, \dots denote a sequence of Poisson random measures with mean measures μ, μ_1, μ_2, \dots , i.e. $N_j = \text{PRM}(\mu_j)$. Then,*

$$N_n \Rightarrow N \quad \text{if and only if} \quad \mu_n \xrightarrow{v} \mu, \quad n \rightarrow \infty.$$

2. *Let $X_{n,j}$ be iid random variables on (S, \mathcal{S}) and $N = \text{PRM}(\mu)$ on (S, \mathcal{S}) . Define $N_n = \sum_{j=1}^n \epsilon_{X_{n,j}}$. Then,*

$$N_n \Rightarrow N \quad \text{if and only if} \quad nP(X_{n,1} \in \cdot) \xrightarrow{v} \mu(\cdot), \quad n \rightarrow \infty.$$

3. *Let $X_{n,j}$ be iid random variables on (S, \mathcal{S}) and let ξ be a PRM on $[0, \infty) \times S$ with mean measure $dt \times d\mu$, i.e. for $t > s$ and $\mu(A) < \infty$*

$$P(\xi((s, t] \times A) = x) = (t - s)e^{-\mu(A)} \mu(A)^x / x!.$$

Define further $\xi_n = \sum_{j=1}^{\infty} \epsilon_{(j/n, X_{n,j})}$. Then,

$$\xi_n \Rightarrow \xi \quad \text{if and only if} \quad nP(X_{n,1} \in \cdot) \xrightarrow{v} \mu(\cdot), \quad n \rightarrow \infty.$$

As a last step we define regularly varying functions and explain the relationship between point processes, regular variation and extreme values.

Definition 2.5 (Regular variation). *A function $f : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is called regularly varying with index $\alpha > 0$ if*

$$\frac{f(tx)}{f(x)} \rightarrow x^{-\alpha}, \quad t \rightarrow \infty.$$

The following theorem shows the connection of the Fréchet domain of attraction, regular variation and point process convergence.

Theorem 2.5. *Let X_1, X_2, \dots denote non-negative iid random variables with distribution function F . Then, the following statements are equivalent:*

1. F is in the maximum domain of attraction of a Type II (Fréchet) extreme value distribution, i.e. there exists $a_n > 0$ such that

$$F^n(a_n x) \rightarrow \exp(-x^{-\alpha}), \quad n \rightarrow \infty.$$

2. The tail distribution $1 - F$ is regularly varying with index $-\alpha$.

3. There exists a sequence $a_n > 0$ such that

$$\xi_n = \sum_{j=1}^{\infty} \epsilon_{(j/n, X_j/a_n)} \Rightarrow \xi = PRM(dt \times dv_\alpha), \quad n \rightarrow \infty,$$

where $v_\alpha(x, \infty) = x^{-\alpha}$, $x > 0$.

2.1.3 Multivariate extreme value distributions

So far, we introduced the generalized extreme value distribution as the limit distribution of maxima of univariate random variables. To extend the theory to multivariate random vectors we introduce the following notation. Let $\mathbf{X}_i = (X_{i,1}, \dots, X_{i,d})$, $i = 1, \dots, n$ denote iid d -variate random vectors with common distribution function F . The vector of componentwise maxima is defined as

$$\mathbf{M}_n = \bigvee_{i=1}^n \mathbf{X}_i = \left(\bigvee_{i=1}^n X_{i,1}, \dots, \bigvee_{i=1}^n X_{i,d} \right),$$

where $\mathbf{X} \leq \mathbf{x}$, if $X_i \leq x_i$, $i = 1, \dots, n$, componentwise. The main objective is to characterize the limit distribution G , for which

$$F^n(\mathbf{a}_n \mathbf{x} + \mathbf{b}_n) \rightarrow G(\mathbf{x}), \quad n \rightarrow \infty, \quad (2.3)$$

where $\mathbf{a}_n > 0$ and $\mathbf{b}_n \in \mathbb{R}^d$ are sequences of normalizing constants. The distribution function G in (2.3) is called *multivariate extreme value distribution*. A first observation from (2.3) is the fact that the marginal distributions of G have to be univariate extreme value distributions, i.e.

$$F_j^n(a_{j,n}x_j + b_{j,n}) \rightarrow G(\infty, \dots, \infty, x_j, \infty, \dots, \infty) = G_j(x_j), \quad n \rightarrow \infty.$$

Using the probability integral transform one can always achieve standardized marginal distributions. Assume that G is a multivariate distribution function with continuous marginal

distributions $G_j, j = 1, \dots, d$. Then, the transformations

$$\begin{aligned} T_j^{(I)}(y) &= -\log(-\log(G_j(y))), \\ T_j^{(II)}(y) &= -\frac{1}{\log G_j(y)}, \\ T_j^{(III)}(y) &= \log(G_j(y)), \end{aligned}$$

induce multivariate distributions

$$\begin{aligned} G^{(I)}(y_1, \dots, y_d) &= G(T_1^{(I)}(y_1), T_2^{(I)}(y_2), \dots, T_d^{(I)}(y_d)) \\ G^{(II)}(y_1, \dots, y_d) &= G(T_1^{(II)}(y_1), T_2^{(II)}(y_2), \dots, T_d^{(II)}(y_d)) \\ G^{(III)}(y_1, \dots, y_d) &= G(T_1^{(III)}(y_1), T_2^{(III)}(y_2), \dots, T_d^{(III)}(y_d)) \end{aligned}$$

with standard Gumbel, Fréchet and Weibull marginal distributions, respectively. In addition, $G^{(I)}$, $G^{(II)}$ and $G^{(III)}$ are multivariate extreme value distributions if and only if G is one.

As in the univariate case, *max-stable distributions* are defined as distributions for which there exist sequences of constants $\alpha_k > 0$ and $\beta_k \in \mathbb{R}^d$ such that

$$G^k(\alpha_k \mathbf{x} + \beta_k) = G(\mathbf{x})$$

for all $k \in \mathbb{N}$.

Theorem 2.6. *The class of multivariate extreme value distributions satisfying (2.3) coincides with the class of max-stable distribution functions.*

A proof of this result can be found in Resnick [71], Chapter 5. We give one characterization of multivariate extreme value distributions with standardized marginal distributions, which is due to Pickands [67] and de Haan and Resnick [32].

Theorem 2.7. *The distribution function G is a multivariate extreme value distribution with standard Fréchet marginals, i.e. $G_j(x_j) = e^{-1/x}$, $x > 0$, if and only if there exists a finite measure H on the unit sphere $\mathcal{B}_H = \{\mathbf{x} \in [\mathbf{0}, \infty) \setminus \{0\} : \|\mathbf{x}\| = 1\}$ for an arbitrary norm $\|\cdot\| \in \mathbb{R}^d$ such that*

$$G(\mathbf{x}) = \exp\{-V(\mathbf{x})\} = \exp\left\{-\int_{\mathcal{B}_H} \bigvee_{j=1}^d \frac{\omega_j}{x_j} H(d\omega)\right\}, \quad (2.4)$$

where the measure H satisfies

$$\int_{\mathcal{B}_H} \omega_j H(d\omega) = 1.$$

The measure resulting from the representation in (2.4), given by

$$v((\mathbf{0}, \mathbf{x}]^c) = V(\mathbf{x}) = \int_{\mathcal{B}_H} \bigvee_{j=1}^d \frac{\omega_j}{x_j} H(d\omega), \quad (2.5)$$

is called *exponent measure*. We introduce two well-known extremal dependence measures resulting from the representation in (2.4). For more information we refer to Beirlant et al. [7]. The *stable tail dependence function* is defined by

$$L(x_1, \dots, x_d) = V(1/x_1, \dots, 1/x_d).$$

We list some properties of the stable tail dependence function.

- L is continuous and convex.
- $\bigvee_{j=1}^d x_j \leq L(x_1, \dots, x_d) \leq \sum_{j=1}^d x_j$. The boundaries correspond to complete dependence on the left and independence on the right.
- $L(0, \dots, 0, 1, 0, \dots, 0) = 1$.
- $L(s\mathbf{x}) = sL(\mathbf{x})$, $s > 0$.
- L determines the copula of an extreme value distribution, i.e. the distribution of $(G_1(x_1), \dots, G_d(x_d))$,

$$C(u) = \exp\{-L(-\log(u_1), \dots, -\log(u_d))\}$$

Another measure of dependence is *Pickands dependence function*, which is defined on the unit simplex $\{(u_1, \dots, u_{d-1}) \in [0, 1]^{d-1} : \sum_{j=1}^{d-1} u_j \leq 1\}$ by

$$A(u_1, \dots, u_{d-1}) = \int_{\mathcal{B}_H} \max \left\{ \omega_1 u_1, \dots, \omega_{d-1} u_{d-1}, \omega_d \left(1 - \sum_{j=1}^{d-1} u_j \right) \right\} H(d\omega).$$

Using the representation of G in (2.4) Pickands dependence function determines G by the following relation.

$$G(\mathbf{x}) = \exp \left\{ - \left(\sum_{j=1}^d \frac{1}{x_j} \right) A \left(\frac{1/x_1}{\sum_{k=1}^d 1/x_k}, \dots, \frac{1/x_d}{\sum_{k=1}^d 1/x_k} \right) \right\}$$

Pickands dependence function satisfies some important properties listed below.

- A is continuous and convex.
- $1/d \leq \max\{u_1, \dots, u_{d-1}, 1 - \sum_{j=1}^{d-1} u_j\} \leq A(u_1, \dots, u_{d-1}) \leq 1$. The lower bound again corresponds to complete dependence and the upper bound to independence.
- $A(0, \dots, 0, 1, 0, \dots, 0) = 1$ and $A(\mathbf{0}) = 1$.

For the estimation of extremal dependence functions several nonparametric and parametric procedures have been developed, see Beirlant et al. [7], Chapter 8, for an overview of existing models and methods. We just mention two parametric models in the example below.

Example 2.2. The first example is the multivariate symmetric logistic model, defined through the stable tail dependence function by

$$L(x_1, \dots, x_d) = (x_1^{1/\alpha} + \dots + x_d^{1/\alpha})^\alpha$$

with parameter $0 < \alpha \leq 1$. Originally, the logistic model was defined in Gumbel [45]. An extension of the logistic model, which is able to allow for the exchangeability of the parameters in two dimensions is given in Tawn [84]. The so-called *asymmetric logistic model* is defined by

$$L(x_1, x_2) = (1 - \theta_1)x_1 + (1 - \theta_2)x_2 + ((\theta_1 x_1)^{1/\alpha} + (\theta_2 x_2)^{1/\alpha})^\alpha$$

with parameters $0 < \theta_1, \theta_2, \alpha \leq 1$. A multivariate extension of the asymmetric logistic model can be found in Tawn [85]. Figure 2.1.2 shows Pickands dependence function for the logistic (left) and the asymmetric logistic (right) model in the two dimensional case.

To complete the theory of multivariate extreme values we introduce the notion of multivariate regular variation and make the connection to point processes. In Resnick [73] multivariate regular variation is defined for distributions on cones. We restrict ourselves to the cone $[\mathbf{0}, \infty) \setminus \{\mathbf{0}\}$, since this is the corresponding domain for extreme value distributions.

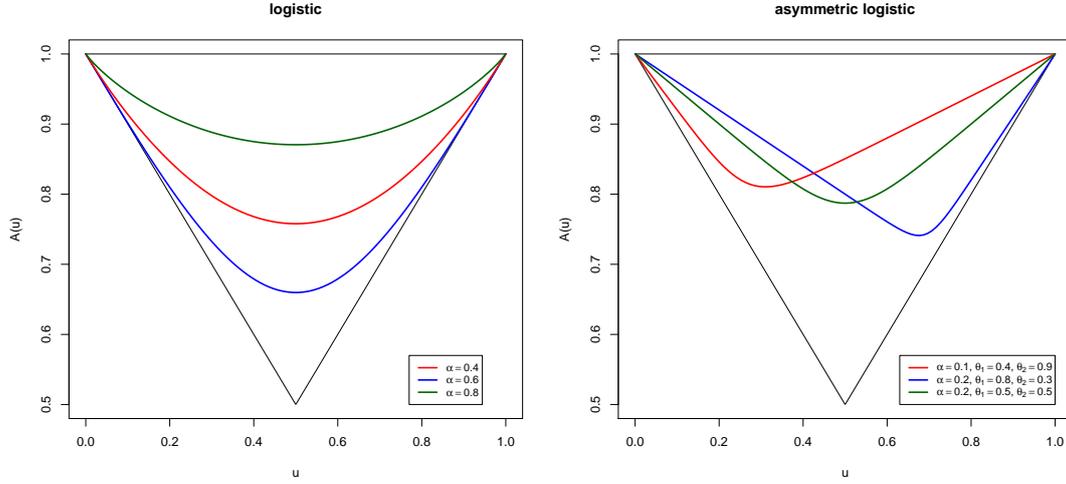


Figure 2.1.2: Pickands dependence function for the logistic (left) and the asymmetric logistic model (right) and different sets of parameters.

Definition 2.6 (Multivariate regular variation). *The random vector $\mathbf{X} = (X_1, \dots, X_d)$ is called multivariate regularly varying if there exists a sequence $a_n \rightarrow \infty$, $a_n > 0$ and a nonzero Radon measure ν on $\mathcal{B}([\mathbf{0}, \infty) \setminus \{\mathbf{0}\})$, called the limit measure, such that*

$$nP(a_n^{-1}\mathbf{X} \in \cdot) \xrightarrow{\nu} \nu(\cdot), \quad n \rightarrow \infty$$

on $\mathcal{B}([\mathbf{0}, \infty) \setminus \{\mathbf{0}\})$. Equivalently, \mathbf{X} is regularly varying if there exists $a(t) \rightarrow \infty$ and a nonzero Radon measure ν such that

$$tP(a(t)^{-1}\mathbf{X} \in \cdot) \xrightarrow{\nu} \nu(\cdot), \quad t \rightarrow \infty.$$

The sequence $(a_n)_{n \in \mathbb{N}}$ can be defined by

$$a_n = \inf\{z \geq 0 : P(\|\mathbf{X}\| > a_n) \sim n^{-1}\},$$

where \sim denotes asymptotic equivalence. Assume, the marginal distributions are standard Fréchet, that the maximum norm is used, and that $c > 0$ is some constant. Then,

$$nP(\|\mathbf{X}\| > a_n) = nP\left(\max_{j=1, \dots, d} X_j > a_n\right) = nP(\exists j : X_j > a_n)$$

$$\leq n \sum_{j=1}^d P(X_j > a_n) = nd(1 - e^{-1/a_n}) \sim \frac{nd}{a_n} \rightarrow c, n \rightarrow \infty,$$

if and only if $a_n \sim n$.

In order to make statements about the domain of attraction of some distribution function F , we need to standardize the marginal distributions of F as well. For the Fréchet domain of attraction consider the transformation $U_j = 1/(1 - F_j)$, $j = 1, \dots, d$. Then,

$$F_*(\mathbf{x}) = F(U_1^{\leftarrow}(x_1), \dots, U_d^{\leftarrow}(x_d)) \quad (2.6)$$

is in the domain of attraction of some extreme value distribution G with Fréchet marginals, if and only if F_* is in the domain of attraction of G . The following proposition describes the relations between multivariate regular variation and the multivariate extreme value distribution. It is taken from Resnick [71], Proposition 5.17.

Proposition 2.1. *Assume that G is a multivariate extreme value distribution with standard Fréchet marginal distributions. Further let \mathbf{X}_* denote a random vector with distribution function F_* as defined in (2.6). The following statements are equivalent.*

1. F_* is in the maximum domain of attraction of G .
2. \mathbf{X}_* is multivariate regularly varying on $[\mathbf{0}, \infty) \setminus \{\mathbf{0}\}$ with limit distribution ν , i.e. $nP(n^{-1}\mathbf{X}_* \in \cdot) \xrightarrow{\nu} \nu(\cdot)$, where ν is the exponent measure in (2.5).
3. $nP((n^{-1}\|\mathbf{X}_*\|, \|\mathbf{X}_*\|^{-1}\mathbf{X}_*) \in \cdot) \xrightarrow{\nu} r^{-2}dr \times H$ on $(0, \infty] \times \mathcal{B}_H$, where H is the measure in (2.4).

2.1.4 Definition and families of max-stable processes

Max-stable processes form the natural extension of the multivariate extreme value distribution to infinite dimensions. Detailed introductions and different families of stationary max-stable processes have been developed for example in Brown and Resnick [14], Deheuvels [33], de Haan [29], de Haan and Pickands [31], Kabluchko, Schlather and de Haan [50] and Schlather [75]. We start with the definition of max-stable processes.

Definition 2.7 (Max-stable process). *In the following, let $\{X_t, t \in T\}$ denote a continuous stationary stochastic process, where T is an arbitrary index set. The process is called max-stable if there exist sequences of constants $a_{n,t} > 0$ and $b_{n,t} \in \mathbb{R}$ for $n \geq 1$ and $t \in T$, such*

that

$$\left\{ \frac{\bigvee_{j=1}^n X_{j,t} - b_{n,t}}{a_{n,t}}, t \in T \right\},$$

is identical in distribution to $\{X_t, t \in T\}$, where $X_{1,t}, \dots, X_{n,t}$ are n independent replications of the process $\{X_t, t \in T\}$.

Remark 2.2. From the definition it immediately follows, that all finite-dimensional distributions must be multivariate extreme value distributions. In particular, the marginal distributions are GEVs. In the following we will assume that the marginal distributions are standard Fréchet, i.e.

$$P(X_t \leq x) = \exp\{-1/x\}, \quad x > 0, t \in T.$$

The definition of max-stable processes simplifies to the following condition:

The process $\{X_t, t \in T\}$ with standard Fréchet marginals is max-stable if $\{\bigvee_{j=1}^n X_{j,t}/n, t \in T\}$ is identical in distribution to $\{X_t, t \in T\}$. To show that a process with standard Fréchet marginals is max-stable, it suffices to verify that

$$(P(X_{t_1} \leq kx_1, \dots, X_{t_n} \leq kx_n))^k = P(X_{t_1} \leq x_1, \dots, X_{t_n} \leq x_n). \quad (2.7)$$

for all $0 \leq t_1 < \dots < t_n, k, n \in \mathbb{N}$ and $x_1, \dots, x_n > 0$.

We describe three families of max-stable processes which have been suggested in the literature.

Smith's storm profile model

The first model was originally introduced in de Haan [29] and further analysed in Smith [79]. Let $\{(\xi_j, x_j), j \geq 1\}$ denote points of a Poisson random measure on $(0, \infty) \times X$ with intensity measure $\xi^{-2} d\xi \times \lambda(dx)$, where X is a measurable set and λ denotes Lebesgue measure on X . Further assume that $f_t, t \in T$ is a non-negative function on X for which

$$\int_X f_t(x) dx = 1, \quad t \in T.$$

Then, the process defined by

$$\eta_t = \bigvee_{j=1}^{\infty} \xi_j f_t(x_j), \quad t \in T, \quad (2.8)$$

is max-stable. To see this, one shows (2.7) by using properties of the Poisson process (see Section 2.1.2). For $t_1, \dots, t_n \in T$ and $y_1, \dots, y_n > 0$ it holds

$$\begin{aligned}
 & (P(\eta_{t_1} \leq ky_1, \dots, \eta_{t_n} \leq ky_n))^k \\
 &= \left(P\left(\xi_j \leq \bigwedge_{l=1}^n \frac{ky_l}{f_{t_l}(x_j)}, j = 1, 2, \dots \right) \right)^k \\
 &= \left(P\left(\text{no points of the Poisson process above the function } g(x) = \bigwedge_{l=1}^n ky_l / f_{t_l}(x) \right) \right)^k \\
 &= \left(\exp \left\{ - \int_X \int_0^\infty \mathbb{1}_{\{\xi > \bigwedge_{l=1}^n ky_l / f_t(x)\}} \xi^{-2} d\xi dx \right\} \right) = \exp \left\{ - \int_X \bigvee_{l=1}^n \frac{f_{t_l}(x)}{y_l} dx \right\} \\
 &= P(\eta_{t_1} \leq y_1, \dots, \eta_{t_n} \leq y_n).
 \end{aligned}$$

The above derivation also shows how the finite-dimensional distribution function can be calculated. The most popular choice for the function f_t is the density of a normal distribution, i.e.

$$f_t(x) = f_0(x - t) = \frac{1}{\sqrt{2\pi\sigma}} \exp \left\{ -\frac{(x - t)^2}{2\sigma^2} \right\},$$

leading to a *Gaussian extreme value process*. Other choices are possible, for examples we refer to Smith [79]. The model is often called Smith's storm profile model resulting from the interpretation of the components in terms of rainfall storms. In particular, the points x_j can be seen as center of storm j , ξ_j is the corresponding intensity, the function f_t describes the shape of the storm, and the product $\xi_j f_t(x_j)$ is the rainfall or wind speed at time t from storm j . Smith's storm profile model will be put into the space-time context in Section 3.1.2.

Brown-Resnick process

Another family of max-stable process was introduced in Brown and Resnick [14] and later generalized in Kabluchko et al. [50]. Let $\{\xi_j, j \geq 1\}$ denote points of a Poisson random measure on $[0, \infty)$ with intensity measure $\xi^{-2} d\xi$. Further let W_j be independent replicates of a Gaussian process with stationary increments and covariance function γ . The process defined by

$$\eta_t = \bigvee_{j=1}^{\infty} \xi_j \exp\{W_j(t) - \delta(t)\}, \quad t \in T, \quad (2.9)$$

where $\delta(t) = \gamma(0) - \gamma(t)$ is the variogram of the Gaussian process W , is a max-stable process with Fréchet marginals known as the Brown-Resnick process. In contrast to Smith's storm profile model, there is no storm center (in Smith's model denoted by x_j). But since the deterministic function f_t is replaced by a stochastic process, it allows for random shapes in the storm.

Schlather model

A generalization of the Brown-Resnick process is described in Schlather [75], where the process $\exp\{W_j(t) - \delta(t)\}$ in (2.9) is replaced by some stationary process. Let $\{\xi_j, j \geq 1\}$ be an enumeration of points of a Poisson process on $(0, \infty)$ with intensity measure $\xi^{-2}d\xi$. Further assume that $\{Y_t, t \in T\}$ is a stationary process such that $\mathbb{E}[\max\{0, Y_t\}] = 1$, $t \in T$. Let $Y_{j,t}, j \geq 1$ be independent replications of $\{Y_t, t \in T\}$ and independent of $\{\xi_j, j \geq 1\}$. The process, defined by

$$\eta_t = \bigvee_{j=1}^{\infty} \xi_j \max\{0, Y_{j,t}\}, \quad t \in T,$$

is max-stable with unit Fréchet marginal distributions.

2.2 Space-time models

In this section, we introduce some basic concepts for Gaussian space-time processes, which will be used throughout this thesis. In addition, we explain how Gaussian space-time processes can be simulated.

2.2.1 Fundamentals for space-time processes

We denote by $\{Z(\mathbf{s}, t), \mathbf{s} \in \mathbb{R}^d, t \in [0, \infty)\}$ a space-time process, where $\mathbf{s} \in \mathbb{R}^d$ denote d -dimensional locations and $t \in [0, \infty)$ is the time. The correlation function of a space-time process is defined by

$$C(\mathbf{s}_1, t_1; \mathbf{s}_2, t_2) = \frac{\text{Cov}(Z(\mathbf{s}_1, t_1), Z(\mathbf{s}_2, t_2))}{\sqrt{\text{Var}(Z(\mathbf{s}_1, t_1))\text{Var}(Z(\mathbf{s}_2, t_2))}}.$$

For simplicity, we assume that the variance equals one, i.e. $\text{Var}(Z(\mathbf{s}, t)) = 1$ for all $\mathbf{s} \in \mathbb{R}^d$ and $t \in [0, \infty)$. In subsequent chapters we will make use of the following basic concepts for space-time processes.

Definition 2.8 (Basic concepts for space-time correlation functions). *We call the correlation function C*

- stationary, if C only depends on the spatial and the temporal lag. In particular, for all $\mathbf{h} \in \mathbb{R}^d$ and $u \in [0, \infty)$,

$$C(\mathbf{s}_1, t_1) = C(\mathbf{s}_1 + \mathbf{h}, t_1 + u) =: \gamma(\mathbf{h}, u).$$

- isotropic, if the stationary correlation function only depends on the absolute spatial and temporal lag, i.e. there exists a correlation function $\tilde{\gamma}$ such that

$$\gamma(\mathbf{h}, u) =: \tilde{\gamma}(\|\mathbf{h}\|, |u|).$$

- separable, if C can be separated into a spatial correlation function C_1 and a temporal correlation function C_2 ,

$$C(\mathbf{s}_1, t_1; \mathbf{s}_2, t_2) = C_1(\mathbf{s}_1, \mathbf{s}_2)C_2(t_1, t_2) \quad \text{or}$$

$$C(\mathbf{s}_1, t_1; \mathbf{s}_2, t_2) = C_1(\mathbf{s}_1, \mathbf{s}_2) + C_2(t_1, t_2).$$

A space-time process $\{Z(\mathbf{s}, t), \mathbf{s} \in \mathbb{R}^d, t \in [0, \infty)\}$ is Gaussian, if all finite-dimensional distributions are multivariate Gaussian, i.e. for all $k \in \mathbb{N}$, $\mathbf{s}_1, \dots, \mathbf{s}_k \in \mathbb{R}^d$ and $t_1, \dots, t_k \in [0, \infty)$, the vector

$$(Z(\mathbf{s}_1, t_1), \dots, Z(\mathbf{s}_k, t_k))$$

is multivariate normally distributed.

2.2.2 Simulation of Gaussian space-time processes

Assume, we want to obtain realizations from a stationary Gaussian space-time process $\{Z(\mathbf{s}, t), \mathbf{s} \in \mathbb{R}^d, t \in [0, \infty)\}$ with mean $\boldsymbol{\mu}$ and stationary correlation function γ (assume that the variance equals 1). To simulate from Gaussian processes with M locations $\mathbf{s}_1, \dots, \mathbf{s}_M$ and T time points t_1, \dots, t_T we need to draw values from a multivariate normal distribution with mean vector $\boldsymbol{\mu} \in \mathbb{R}^{MT}$ and positive-definite covariance matrix $\Sigma \in \mathbb{R}^{MT} \times \mathbb{R}^{MT}$, given by

$$\Sigma[(i-1)T + (1+k), (j-1)T + (1+l)] = \gamma(\mathbf{s}_i - \mathbf{s}_j, t_k - t_l)$$

for $i, j = 1, \dots, M, k, l = 1, \dots, T$. In general, one starts by simulating a vector of standard normal distributed random variables $\mathbf{Z}^{(0)} = (Z_1^{(0)}, \dots, Z_{MT}^{(0)})$ with independent components. This can be done by using the Box-Muller transform (see Box and Muller [12]). In a second step the covariance matrix is decomposed such that $\Sigma = LL^\top$, where L is a lower triangular matrix, by using for instance the Cholesky decomposition. By transforming $\mathbf{Z}^{(0)}$ to $\mathbf{Z} = \boldsymbol{\mu} + L\mathbf{Z}^{(0)}$ we obtain the desired simulated random vector.

Even for a relatively small number of locations and time points, the correlation matrix can be huge and the Cholesky decomposition invisible. Therefore, approximation methods incorporating the structure of correlation matrices are used to simulate such processes. If space-time locations lie on a regular grid, one can use *circulant embedding*, introduced in Wood and Chan [92]. We shortly describe the procedure in the simplest case for the simulation of a stationary Gaussian random field $Z(\mathbf{s})$ with one-dimensional locations $\mathbf{s} = 1, \dots, m$. The

correlation matrix in this case is given by

$$\begin{pmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(m-1) \\ \vdots & \vdots & \cdots & \vdots \\ \gamma(m-1) & \gamma(m-2) & \cdots & \gamma(0) \end{pmatrix}.$$

The idea is to embed the correlation matrix into a larger circulant matrix, i.e. there exists a function g such that the entries in the circulant matrix G satisfy $G[i, j] = g(j - i)$. The new matrix is of size $\tilde{m} \times \tilde{m}$, where $\tilde{m} = 2^f > 2(m - 1)$ for some integer f , and is defined by

$$G = \begin{pmatrix} g(0) & g(1) & \cdots & g(\tilde{m} - 1) \\ g(\tilde{m} - 1) & g(0) & \cdots & g(\tilde{m} - 2) \\ \vdots & \vdots & \cdots & \vdots \\ g(1) & g(2) & \cdots & g(0) \end{pmatrix},$$

where

$$g(j) = \begin{cases} \gamma(k) & 0 \leq k \leq \tilde{m}/2, \\ \gamma(\tilde{m} - k) & \tilde{m}/2 < k \leq \tilde{m} - 1. \end{cases}$$

As pointed out in Brockwell and Davis [13], Section 4.5, the circulant matrix G can be diagonalized using the relation

$$QGQ^\top = \text{diag}\{\lambda_0, \dots, \lambda_{\tilde{m}-1}\},$$

where Q is the matrix of eigenvectors and $\lambda_0, \dots, \lambda_{\tilde{m}-1}$ are the eigenvalues of M . Using this result, it follows that

$$\mathbf{Z} = Q(\text{diag}\{\lambda_0^{1/2}, \dots, \lambda_{\tilde{m}-1}^{1/2}\})Q^\top \mathbf{Z}^{(0)} \sim \mathcal{N}_{\tilde{m}}(\mathbf{0}, G), \quad (2.10)$$

where $\mathbf{Z}^{(0)}$ is a standard normally distributed random vector. The first m entries of the resulting vector have the desired distribution. Special properties of circulant matrices, see Brockwell and Davis [13], allow for the use of the discrete Fourier transform in the algorithm to calculate the eigenvalues of G and to evaluate the first equation in (2.10). The procedure can be extended to random fields and space-time processes. For details see Kozintsev [52]. Figure 2.2.1 shows realizations of Gaussian space-time processes for four consecutive time points (clockwise from the top left to the bottom left). The fields were simulated using the R-package `RandomFields`, where the circulant embedding method is implemented.

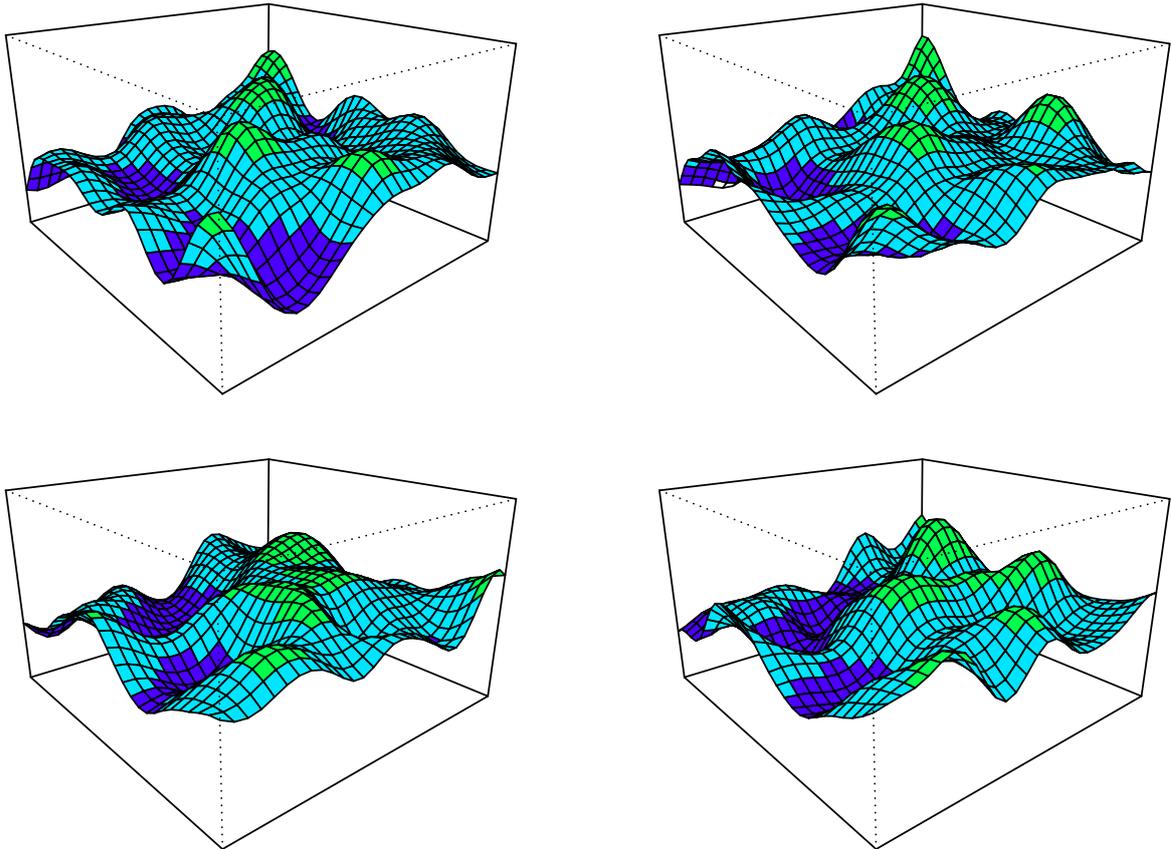


Figure 2.2.1: Simulated Gaussian space-time processes for four consecutive time points (clockwise from the top left to the bottom left) using the circulant embedding method.

CHAPTER 3

MAX-STABLE PROCESSES FOR EXTREMES OBSERVED IN SPACE AND TIME

This chapter is based on Davis, Klüppelberg and Steinkohl [26]. In Section 2.1.4, max-stable processes were introduced as natural extension of the generalized extreme value distribution to infinite dimensions. This chapter deals with the construction of max-stable processes defined on the continuous space-time domain $\mathbb{R}^d \times [0, \infty)$. We follow the approach introduced in Kabluchko et al. [50], who construct max-stable processes as infinite maximum of rescaled and transformed replications of Gaussian processes. We extend this concept to a space-time domain and relate the process to an extended version of Smith's storm profile model (see Section 2.1.4). Basic properties regarding extremal dependence and possible choices for the underlying correlation function of the Gaussian process are discussed.

The chapter is organized as follows. Max-stable space-time processes are developed in Section 3.1. In Section 3.2, we show how Pickands dependence function and the tail dependence coefficient relate to the correlation model used in the underlying Gaussian process. Further correlation models are discussed in Section 3.3 and simulations based on a set of different parameters are visualized. Section 3.3.2 analyses anisotropic correlation functions, where one can see directional movements in the storm profile model, which are not possible in the isotropic case.

3.1 Extension of extreme spatial fields to the space-time setting

Max-stable processes form the natural extension of multivariate extreme value distributions to infinite dimensions. In the literature, different families of max-stable processes have been considered. In Section 3.1.1 we describe the construction introduced in Kabluchko, Schlather and de Haan [50], which is based on the limit of pointwise maxima from an array of independent Gaussian random fields. Furthermore, we extend the approach introduced in de Haan [29] and interpreted by Smith [79] as the storm profile model to a space-time setting in Section 3.1.2.

3.1.1 Max-stable random fields based on spatio-temporal correlation functions

Before presenting the construction of a max-stable Gaussian random field in space and time, we recall the definition of the Brown-Resnick space-time process with Fréchet marginals (see Section 2.1.4). Let $\{\xi_j, j \geq 1\}$ denote points of a Poisson random measure on $[0, \infty)$ with intensity measure $\xi^{-2}d\xi$ and let $Y_j(s, t), j = 1, 2, \dots$, be independent replications of some space-time process $\{Y(s, t), (s, t) \in \mathbb{R}^d \times [0, \infty)\}$ with $\mathbb{E}(Y(s, t)) < \infty$, and $Y(s, t) \geq 0$ a.s., which are also independent of ξ_j . The space-time process, defined by

$$\eta(s, t) = \bigvee_{j=1}^{\infty} \xi_j Y_j(s, t), \quad (s, t) \in \mathbb{R}^d \times [0, \infty), \quad (3.1)$$

is a max-stable process with Fréchet marginals and often referred to as the Brown-Resnick process (see Kabluchko et al. [50]). The finite-dimensional distributions can be calculated using a point process argument as done in de Haan [29]. For example, if $(s_1, t_1), \dots, (s_K, t_K)$ are distinct space-time locations (duplicates in the space or the time components are allowed), then

$$\begin{aligned} P(\eta(s_1, t_1) \leq y_1, \dots, \eta(s_K, t_K) \leq y_K) &= P\left(\xi_j \bigvee_{k=1}^K \frac{Y_j(s_k, t_k)}{y_k} \leq 1, \forall j = 1, 2, \dots\right) \\ &= P(N(A) = 0) = \exp\left\{-\mathbb{E}\left(\bigvee_{k=1}^K \frac{Y(s_k, t_k)}{y_k}\right)\right\}, \end{aligned} \quad (3.2)$$

where $A = \{(u, v); uv > 1\}$ and N is the Poisson random measure with points at

$$\left\{ \left(\xi_j, \bigvee_{k=1}^K \frac{Y_j(\mathbf{s}_k, t_k)}{y_k} \right) \right\}.$$

As we will see below, the Brown-Resnick process appears as limit of a sequence of pointwise maxima of independent Gaussian space-time processes. In the following, let $\{Z(\mathbf{s}, t), \mathbf{s} \in \mathbb{R}^d, t \in [0, \infty)\}$ denote a space-time Gaussian process with covariance function given by

$$\tilde{C}(\mathbf{s}_1, t_1; \mathbf{s}_2, t_2) = \text{Cov}(Z(\mathbf{s}_1, t_1), Z(\mathbf{s}_2, t_2)),$$

for two locations $\mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d$ and time points $t_1, t_2 \in [0, \infty)$. We assume stationarity in space and time, so that we can write

$$\tilde{C}(\mathbf{s}_1, t_1; \mathbf{s}_2, t_2) = C(\mathbf{s}_1 - \mathbf{s}_2, t_1 - t_2) = C(\mathbf{h}, u),$$

where $\mathbf{h} = \mathbf{s}_1 - \mathbf{s}_2$ and $u = t_1 - t_2$. Furthermore, let $\gamma(\mathbf{h}, u) = C(\mathbf{h}, u)/C(\mathbf{0}, 0)$ denote the corresponding correlation function. We will assume smoothness conditions on $\gamma(\cdot, \cdot)$ near $(\mathbf{0}, 0)$. This assumption is natural in the context of spatio-temporal processes, since it basically relates to the smoothness of the underlying process.

Assumption 3.1. *There exist two nonnegative sequences of constants $s_n \rightarrow 0$, $t_n \rightarrow 0$ as $n \rightarrow \infty$ and a nonnegative function δ such that*

$$(\log n)(1 - \gamma(s_n(\mathbf{s}_1 - \mathbf{s}_2), t_n(t_1 - t_2))) \rightarrow \delta(\mathbf{s}_1 - \mathbf{s}_2, t_1 - t_2) \in (0, \infty), \quad n \rightarrow \infty,$$

for all $(\mathbf{s}_1, t_1) \neq (\mathbf{s}_2, t_2)$, $\mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d, t_1, t_2 \in [0, \infty)$.

Examples of such correlation functions are given in Section 3.3. If $(\mathbf{s}_1, t_1) = (\mathbf{s}_2, t_2)$, it follows that the correlation function equals one, $\gamma(\mathbf{0}, 0) = 1$, which implies $\delta(\mathbf{0}, 0) = 0$. The following theorem regarding limits of finite-dimensional distributions stems from Theorem 1 in Hüsler and Reiss [47] and Theorem 17 in Kabluchko et al. [50]. In the following let $C(\mathbb{R}^d \times [0, \infty))$ denote the space of continuous functions on $\mathbb{R}^d \times [0, \infty)$, where convergence is defined as uniform convergence on compact subsets K of $\mathbb{R}^d \times [0, \infty)$.

Theorem 3.1. *Let $Z_j(\mathbf{s}, t)$, $j = 1, 2, \dots$ be independent replications from a stationary Gaussian space-time process with mean 0, variance 1 and correlation model γ satisfying Assump-*

tion 3.1 with limit function δ . Assume there exists a metric D on $\mathbb{R}^d \times [0, \infty)$ such that

$$\delta(\mathbf{s}_1 - \mathbf{s}_2, t_1 - t_2) \leq (D((\mathbf{s}_1, t_1), (\mathbf{s}_2, t_2)))^2, \quad \mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d, t_1, t_2 \in [0, \infty), \quad (3.3)$$

and set

$$\eta_n(\mathbf{s}, t) = \frac{1}{n} \bigvee_{j=1}^n -\frac{1}{\log(\Phi(Z_j(\mathbf{s}_n \mathbf{s}, t_n t)))}, \quad \mathbf{s} \in \mathbb{R}^d, t \in [0, \infty). \quad (3.4)$$

Then,

$$\eta_n(\mathbf{s}, t) \Rightarrow \eta(\mathbf{s}, t), \quad n \rightarrow \infty, \quad (3.5)$$

where \Rightarrow denotes weak convergence in $C(\mathbb{R}^d \times [0, \infty))$ and $\{\eta(\mathbf{s}, t), (\mathbf{s}, t) \in \mathbb{R}^d \times [0, \infty)\}$ is a max-stable space-time process. The bivariate distribution functions for $\eta(\mathbf{s}, t)$ have an explicit form given by

$$F(y_1, y_2) = \exp \left\{ -\frac{1}{y_1} \Phi \left(\frac{\log \frac{y_2}{y_1}}{2\sqrt{\delta(\mathbf{h}, u)}} + \sqrt{\delta(\mathbf{h}, u)} \right) - \frac{1}{y_2} \Phi \left(\frac{\log \frac{y_1}{y_2}}{2\sqrt{\delta(\mathbf{h}, u)}} + \sqrt{\delta(\mathbf{h}, u)} \right) \right\}. \quad (3.6)$$

Remark 3.1. Condition (3.1) is sufficient to prove tightness of the sequence $(\eta_n)_{n \in \mathbb{N}}$ in $C(\mathbb{R}^d \times [0, \infty))$. As shown in the proof of Theorem 17 in Kabluchko et al. [50] the limit process η turns out to be a Brown-Resnick process with Y in (3.1) given by

$$\exp\{W(\mathbf{s}, t) - \delta(\mathbf{s}, t)\}, \quad \mathbf{s} \in \mathbb{R}^d, t \in [0, \infty),$$

where $\{W(\mathbf{s}, t), \mathbf{s} \in \mathbb{R}^d, t \in [0, \infty)\}$ is a Gaussian process with mean 0 and covariance function

$$\text{Cov}(W(\mathbf{s}_1, t_1), W(\mathbf{s}_2, t_2)) = \delta(\mathbf{s}_1, t_1) + \delta(\mathbf{s}_2, t_2) - \delta(\mathbf{s}_1 - \mathbf{s}_2, t_1 - t_2). \quad (3.7)$$

In particular, δ is a variogram leading to a valid covariance function in (3.7).

Proof. Although this proof is similar to the one given in Kabluchko et al. [50], we provide a sketch of the arguments for completeness. We start with the bivariate distributions. From classical extreme value theory (see Example 2.1 in Chapter 2), we have for

$$b_n = \sqrt{2 \log n} - \frac{\log \log n + \log(4\pi)}{2\sqrt{2 \log n}}, \quad n \in \mathbb{N}, \quad (3.8)$$

that

$$\lim_{n \rightarrow \infty} \Phi^n \left(b_n + \frac{\log(y)}{b_n} \right) = e^{-1/y}, \quad y > 0.$$

By using the standard arguments as in Embrechts et al. [36], it follows that

$$\Phi^{-1} \left(e^{-1/ny} \right) \sim \frac{\log y}{b_n} + b_n,$$

where \sim denotes asymptotic equivalence. By applying this relation and Theorem 1 in Hüsler and Reiss [47] to the random variables $\eta_n(s_1, t_1)$ and $\eta_n(s_2, t_2)$ for fixed $s_1, s_2 \in \mathbb{R}^d, t_1, t_2 \in [0, \infty)$, we obtain for $y_1, y_2 > 0$

$$\begin{aligned} & P(\eta_n(s_1, t_1) \leq y_1, \eta_n(s_2, t_2) \leq y_2) \\ &= P \left(\bigvee_{j=1}^n \frac{1}{\log(\Phi(Z_j(s_n s_1, t_n t_1)))} \leq ny_1, \bigvee_{j=1}^n \frac{1}{\log(\Phi(Z_j(s_n s_2, t_n t_2)))} \leq ny_2 \right) \\ &= P \left(\bigvee_{j=1}^n Z_j(s_n s_1, t_n t_1) \leq \Phi^{-1} \left(e^{-1/(ny_1)} \right), \bigvee_{j=1}^n Z_j(s_n s_2, t_n t_2) \leq \Phi^{-1} \left(e^{-1/(ny_2)} \right) \right) \\ &\sim P^n \left(Z_1(s_n s_1, t_n t_1) \leq \frac{\log(y_1)}{b_n} + b_n, Z_1(s_n s_2, t_n t_2) \leq \frac{\log(y_2)}{b_n} + b_n \right) \\ &\sim \exp \left\{ -\frac{1}{y_1} - \frac{1}{y_2} + nP \left(Z_1(s_n s_1, t_n t_1) > \frac{\log(y_1)}{b_n} + b_n, Z_1(s_n s_2, t_n t_2) > \frac{\log(y_2)}{b_n} + b_n \right) \right\} \end{aligned}$$

for $n \rightarrow \infty$. The vector $(Z_1(s_n s_1, t_n t_1), Z_2(s_n s_2, t_n t_2))$ is bivariate normally distributed with mean $\mathbf{0}$ and covariance matrix given by $\gamma(s_n(s_1 - s_2), t_n(t_1 - t_2))$. Using the properties of the conditional normal distribution and Assumption 3.1, it can be shown that the last expression converges to (3.6). Similarly to the procedure above, the finite-dimensional limit distributions of beyond second order can be calculated by using Theorem 2 in Hüsler and Reiss [47].

It remains to show that the sequence (η_n) is tight in $C(\mathbb{R}^d \times [0, \infty))$. Following Kabluchko et al. [50], the main step of the proof is to show that the conditional family of processes $\{Y_n^\omega(s, t), (s, t) \in \mathbb{R}^d \times [0, \infty)\}$, defined by

$$Y_n^\omega(s, t) = (b_n(Z(s_n s, t_n t) - b_n) - \omega) \mid (b_n(Z(\mathbf{0}, 0) - b_n) = \omega), \quad \omega \in [-c, c], \quad n \in \mathbb{N},$$

is tight in $C(K)$, where K is any compact subset of $\mathbb{R}^d \times [0, \infty)$ and b_n is defined in (3.8).

This is achieved by calculating an upper bound for the variance of the distance between the process at two spatio-temporal locations, which in our case is given by Assumption (3.3). That is, for large n ,

$$\begin{aligned} \mathbb{V}ar(Y_n^\omega(\mathbf{s}_1, t_1) - Y_n^\omega(\mathbf{s}_2, t_2)) &\leq 2b_n^2(1 - \gamma(s_n(\mathbf{s}_1 - \mathbf{s}_2), t_n(t_1 - t_2))) \\ &\leq 2B\delta(\mathbf{s}_1 - \mathbf{s}_2, t_1 - t_2) \leq 2B(D((\mathbf{s}_1, t_1), (\mathbf{s}_1, t_2)))^2, \end{aligned}$$

where $B > 0$ is some constant. The rest of the proof follows analogously to the proof in [50] and is shown in full detail in Section 3.4.1. \square

Remark 3.2. Kabluchko [49] studies the limit behaviour of rescaled space-time processes of the form

$$\sup_{t' \in [0, m]} Z(s_n \mathbf{s}, t'),$$

and shows that a rescaled version converges in the sense of finite-dimensional distributions to a space-time Brown-Resnick process. The assumptions on the covariance function in the underlying Gaussian space-time random field are similar to those we use in Section 3.3. The approach differs from ours in the sense that we analyse the pointwise maxima of independent replications of space-time random fields, rather than the supremum over time of a single random field.

Remark 3.3. In applications, the marginal distributions are often fitted by a generalized extreme value distribution and are then transformed to standard Fréchet. Sometimes it may be useful to think about other marginal distributions, such as the Gumbel or Weibull. In order to use Gumbel marginals, we need

$$\eta_n(\mathbf{s}, t) = \bigvee_{j=1}^n -\log(-\log(\Phi(Z_j(s_n \mathbf{s}, t_n t)))) - \log(n), \quad (3.9)$$

and obtain the bivariate distribution function in (3.6) with $1/y_1$ and $1/y_2$ replaced by e^{-y_1} and e^{-y_2} . If Weibull marginals should be used, we define

$$\eta_n(\mathbf{s}, t) = n \bigvee_{j=1}^n \log(\Phi(Z_j(s_n \mathbf{s}, t_n t))), \quad (3.10)$$

leading to the same bivariate distribution function as in (3.6), but with $1/y_1$ and $1/y_2$ replaced by y_1 and y_2 , respectively.

3.1.2 Extension of the storm profile model

In this section, we extend the max-stable process introduced in Section 2.1.4, see Equation (2.8), to the space-time setting. The process was interpreted by Smith [79] as a model for storms, where each component can be interpreted as elements of a storm, like intensity or centre. In later papers, including for instance Schlather and Tawn [77] this process is called the storm profile model. We extend the concept to a space-time setting, where extremes are observed at certain locations through time. For simplicity of presentation we assume without loss of generality that \mathbb{R}^2 is the space domain. Assume, that we have a domain for point processes of storm centres $Z \subset \mathbb{R}^2$ and a time domain $X \subset [0, \infty)$, for which the storm is strongest at its centres. Further, let $\{(\xi_j, \mathbf{z}_j, x_j), j \geq 1\}$ denote points of a Poisson random measure on $(0, \infty) \times Z \times X$ with intensity measure $\xi^{-2} d\xi \times \lambda_2(d\mathbf{z}) \times \lambda_1(dx)$, where λ_d denotes Lebesgue measure on \mathbb{R}^d for $d = 1, 2$. Each ξ_j represents the intensity of storm j . Moreover, let $f(\mathbf{z}, x; \mathbf{s}, t)$ for $(\mathbf{z}, x) \in Z \times X$ and $(\mathbf{s}, t) \in \mathbb{R}^2 \times [0, \infty)$ be a non-negative function with

$$\int_{Z \times X} f(\mathbf{z}, x; \mathbf{s}, t) \lambda_2(d\mathbf{z}) \lambda_1(dx) = 1, \quad (\mathbf{s}, t) \in \mathbb{R}^2 \times [0, \infty).$$

The function f represents the shape of the storm. Define

$$\eta(\mathbf{s}, t) = \bigvee_{j=1}^{\infty} \{\xi_j f(\mathbf{z}_j, x_j; \mathbf{s}, t)\}, \quad (\mathbf{s}, t) \in \mathbb{R}^2 \times [0, \infty). \quad (3.11)$$

The product $\xi_j f(\mathbf{z}_j, x_j; \mathbf{s}, t)$ can be interpreted as the wind speed at location \mathbf{s} and time point t from storm j with intensity ξ_j , spatial location of the center \mathbf{z}_j and maximum wind speed at time x_j at the centre. The distribution function of $(\eta(\mathbf{s}_1, t_1), \dots, \eta(\mathbf{s}_K, t_K))$ for fixed $\mathbf{s}_1, \dots, \mathbf{s}_K \in \mathbb{R}^d, t_1, \dots, t_K \in [0, \infty)$ and $y_1, \dots, y_K > 0$, is given through the spectral representation calculated in de Haan [29] by

$$F(y_1, \dots, y_K) = \exp \left\{ - \int_{Z \times X} \bigvee_{k=1}^K \frac{f(\mathbf{z}, x; \mathbf{s}_k, t_k)}{y_k} \lambda_2(d\mathbf{z}) \lambda_1(dx) \right\}. \quad (3.12)$$

To connect the storm model with the Brown-Resnick process that arises in Theorem 3.1, we assume a trivariate Gaussian density for the function f with mean (\mathbf{z}, x) and covariance

matrix $\tilde{\Sigma}$, i.e.

$$f(\mathbf{z}, x; \mathbf{s}, t) = f_0(\mathbf{z} - \mathbf{s}, x - t),$$

where f_0 is a Gaussian density with mean $\mathbf{0}$ and covariance matrix $\tilde{\Sigma}$. We assume that the spatial dependence is modelled through the matrix Σ and the temporal dependence is given through σ_3^2 , leading to the covariance matrix

$$\tilde{\Sigma} = \begin{pmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \sigma_3^2 \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & 0 \\ \sigma_{12} & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{pmatrix}. \quad (3.13)$$

In the following theorem, we calculate a closed form of the bivariate distribution function resulting from the setting defined above. The derivation of the bivariate distribution function in a purely spatial setting can be found in Padoan, Ribatet and Sisson [66] and we stick closely to their notation. The idea of the proof is widely known and for completeness, details are given in Section 3.4.

Theorem 3.2. *With the setting defined above, the max-stable space-time process*

$$\eta(\mathbf{s}, t) = \bigvee_{j=1}^{\infty} \left\{ \xi_j f_0(\mathbf{z}_j - \mathbf{s}; x_j - t) \right\}, \quad \mathbf{s} \in \mathbb{R}^2, t \in [0, \infty), \quad (3.14)$$

has the bivariate distribution function given by

$$\begin{aligned} F(y_1, y_2) &= P(\eta(\mathbf{s}_1, t_1) \leq y_1, \eta(\mathbf{s}_2, t_2) \leq y_2) \\ &= \exp \left\{ -\frac{1}{y_1} \Phi \left(\frac{2\sigma_3^2 \log\left(\frac{y_2}{y_1}\right) + \sigma_3^2 a(\mathbf{h})^2 + u^2}{2\sigma_3 \sqrt{\sigma_3^2 a(\mathbf{h})^2 + u^2}} \right) - \frac{1}{y_2} \Phi \left(\frac{2\sigma_3^2 \log\left(\frac{y_1}{y_2}\right) + \sigma_3^2 a(\mathbf{h})^2 + u^2}{2\sigma_3 \sqrt{\sigma_3^2 a(\mathbf{h})^2 + u^2}} \right) \right\}, \end{aligned} \quad (3.15)$$

where $\mathbf{h} = \mathbf{s}_1 - \mathbf{s}_2$ is the space lag, $u = t_1 - t_2$ is the time lag and $a(\mathbf{h}) = (\mathbf{h}^T \Sigma^{-1} \mathbf{h})^{1/2}$.

Note, that if the time lag u equals zero, the formula reduces to

$$F(y_1, y_2) = \exp \left\{ -\frac{1}{y_1} \Phi \left(\frac{a(\mathbf{h})}{2} + \frac{\log\left(\frac{y_2}{y_1}\right)}{a(\mathbf{h})} \right) - \frac{1}{y_2} \Phi \left(\frac{a(\mathbf{h})}{2} + \frac{\log\left(\frac{y_1}{y_2}\right)}{a(\mathbf{h})} \right) \right\},$$

which is the bivariate distribution of a Gaussian max-stable random field as calculated in Padoan, Ribatet and Sisson [66]. If the space lag \mathbf{h} is zero, the bivariate distribution function is given by

$$F(y_1, y_2) = \exp \left\{ -\frac{1}{y_1} \Phi \left(\frac{1}{u} \left(\log \left(\frac{y_2}{y_1} \right) + \frac{u^2}{2\sigma_3^2} \right) \right) - \frac{1}{y_2} \Phi \left(\frac{1}{u} \left(\log \left(\frac{y_1}{y_2} \right) + \frac{u^2}{2\sigma_3^2} \right) \right) \right\}.$$

By comparing the bivariate distributions from the Smith model with those of the approach discussed in (3.6) in Section 3.1.1, we recognize that the functions are the same, if

$$\delta(\mathbf{h}, u) = \frac{1}{4}a(\mathbf{h})^2 + \frac{1}{\sigma_3^2}u^2. \quad (3.16)$$

In Section 3.3, where we study a more detailed representation of the function δ , we return to this point.

3.2 Pickands dependence function and tail dependence coefficient

The Pickands dependence function (Pickands [67]) was introduced in Section 2.1.3 as a measure of extremal dependence and is related to the exponent measure. In particular, the joint distribution of the max-stable space-time process can be expressed by the exponent measure V ,

$$P(\eta(s_1, t_1) \leq y_1, \eta(s_2, t_2) \leq y_2) = \exp \{-V(y_1, y_2; \delta(s_1 - s_2, t_1 - t_2))\},$$

where V is given through the bivariate distribution function (3.5) by

$$V(y_1, y_2; \delta(\mathbf{h}, u)) = \frac{1}{y_1} \Phi \left(\frac{\log \frac{y_2}{y_1}}{2\sqrt{\delta(\mathbf{h}, u)}} + \sqrt{\delta(\mathbf{h}, u)} \right) + \frac{1}{y_2} \Phi \left(\frac{\log \frac{y_1}{y_2}}{2\sqrt{\delta(\mathbf{h}, u)}} + \sqrt{\delta(\mathbf{h}, u)} \right)$$

and depends on the space and time lags \mathbf{h} and u . In the bivariate case, the Pickands dependence function is defined through

$$\exp \{-V(y_1, y_2, \delta(\mathbf{h}, u))\} = \exp \left\{ -\left(\frac{1}{y_1} + \frac{1}{y_2} \right) A \left(\frac{y_1}{y_1 + y_2} \right) \right\}.$$

Setting $\lambda = y_1/(y_1 + y_2)$, hence $1 - \lambda = y_2/(y_1 + y_2)$, we obtain

$$\begin{aligned} A(\lambda; \delta(\mathbf{h}, u)) &= \lambda(1 - \lambda)V(\lambda, 1 - \lambda; \delta(\mathbf{h}, u)) \\ &= \lambda\Phi\left(\frac{\log \frac{\lambda}{1-\lambda}}{2\sqrt{\delta(\mathbf{h}, u)}} + \sqrt{\delta(\mathbf{h}, u)}\right) + (1 - \lambda)\Phi\left(\frac{\log \frac{1-\lambda}{\lambda}}{2\sqrt{\delta(\mathbf{h}, u)}} + \sqrt{\delta(\mathbf{h}, u)}\right). \end{aligned}$$

A useful summary measure for extremal dependence is the tail-dependence coefficient, which goes back to Geffroy [40, 41] and Sibuya [78]. It is defined by

$$\chi = \lim_{x \rightarrow \infty} P\left(\eta(\mathbf{s}_1, t_1) > F_{\eta(\mathbf{s}_1, t_1)}^{\leftarrow}(x) \mid \eta(\mathbf{s}_2, t_2) > F_{\eta(\mathbf{s}_2, t_2)}^{\leftarrow}(x)\right),$$

where $F_{\eta(\mathbf{s}, t)}^{\leftarrow}$ is the generalized inverse of the marginal distribution for fixed location $\mathbf{s} \in \mathbb{R}^d$ and time point $t \in [0, \infty)$. For the stationary isotropic limit process in Theorem 3.1 χ is given as a function of the spatial lag $r = \|\mathbf{h}\|$ and the (positive) time lag u by

$$\chi(r, u) = 2\left(1 - \Phi\left(\sqrt{\delta(\mathbf{h}, u)}\right)\right). \quad (3.17)$$

The tail dependence coefficient is a special case of the extremogram introduced in Davis and Mikosch [23] (Section 1.4), with A and B defined by $(1, \infty)$. The two cases $\chi(r, u) = 0$ and $\chi(r, u) = 1$ correspond to the boundary cases of asymptotic independence and complete dependence. Thus, if $\delta(\mathbf{h}, u) \rightarrow 0$, the marginal components in the bivariate case are completely dependent and if $\delta(\mathbf{h}, u) \rightarrow \infty$, the components become independent. In the following section, we examine the relationship between the underlying correlation function and the tail dependence coefficient.

3.3 Possible correlation functions for the underlying Gaussian space-time process

Provided the correlation function of the underlying Gaussian process is sufficiently smooth near $(\mathbf{0}, 0)$, then Assumption 3.1 holds for some sequences s_n and t_n . One such condition is given below. Throughout this section let $\mathbf{h} = s_1 - s_2$ denote the space lag and $u = t_1 - t_2$ the time lag.

Assumption 3.2. Assume that the correlation function allows for the following expansion

$$\gamma(\mathbf{h}, u) = 1 - \theta_1 \|\mathbf{h}\|^{\alpha_1} - \theta_2 |u|^{\alpha_2} + O(\|\mathbf{h}\|^{\alpha_1} + |u|^{\alpha_2})$$

around $(\mathbf{0}, 0)$, where $\alpha_1, \alpha_2 \in (0, 2]$ and $\theta_1, \theta_2 \geq 0$ are constants independent of \mathbf{h} and u .

Remark 3.4. The parameters α_1 and α_2 relate to the smoothness of the sample paths in the underlying Gaussian space-time process. The special case $\alpha_1 = \alpha_2 = 2$ corresponds to a mean-square differentiable process. (For further reference see for example Adler [1], Chapter 2). The characterizing condition is given in terms of the correlation function, i.e. a process is mean-square differentiable if and only if all second-order partial derivatives of the correlation function exist in $\mathbf{0}$, i.e. for a spatial dimension $d = 2$,

$$\left. \frac{\partial^2 \gamma(h_1, h_2, u)}{\partial h_1^{m_1} \partial h_2^{m_2} \partial u^{m_3}} \right|_{(0,0,0)},$$

where $\mathbf{m} = (m_1, m_2, m_3) \in \mathbb{N}_0^3$ and $m_1 + m_2 + m_3 = 2$, exist and are finite. It is obvious, that the special case $\alpha_1 = \alpha_2 = 2$ in Assumption 3.2 corresponds to a mean-square differentiable process. For $\alpha_1 < 2$ or $\alpha_2 < 2$, one of the partial derivatives of the correlation function does not exist in $\mathbf{0}$, and thus the process is not mean-square differentiable. The condition for a.s. differentiability in the Gaussian case is given through the second derivative of the correlation function. In the univariate case, a Gaussian process is a.s. differentiable, if for some constants $K_1, K_2 > 0$

$$-\gamma''(u) = K_1 - K_2 |u|^\beta + o(|u|^\beta), \quad \beta \in (0, 2].$$

With the expansion in Assumption 3.2 this is not possible, since $\beta = \alpha_2 - 2 \in (-2, 0]$ (set the spatial lag equal to zero). Furthermore, a Gaussian process is a.s. continuous, if $\alpha_1, \alpha_2 \in (0, 2]$, which is the case in Assumption 3.2.

Under Assumption 3.2, the scaling sequences in Assumption 3.1 can be chosen as $s_n = (\log n)^{-1/\alpha_1}$ and $t_n = (\log n)^{-1/\alpha_2}$. It follows that

$$(\log n)(1 - \gamma(s_n \mathbf{h}, t_n u)) \rightarrow \theta_1 \|\mathbf{h}\|^{\alpha_1} + \theta_2 |u|^{\alpha_2} = \delta(\mathbf{h}, u), \quad \text{as } n \rightarrow \infty. \quad (3.18)$$

The condition for the tightness in (3.3) can be obtained by setting

$$D((s_1, t_1), (s_2, t_2)) = \max \left\{ \|s_1 - s_2\|^{\alpha_1/2}, |t_1 - t_2|^{\alpha_2/2} \right\},$$

which is a metric in $\mathbb{R}^d \times [0, \infty)$, since $\alpha_1, \alpha_2 \in (0, 2]$.

Smith's storm profile model can recover a subset of the class of correlation functions specified in Assumption 3.2. Choosing $\alpha_1 = \alpha_2 = 2$ and

$$\sigma_{12} = 0, \quad \sigma_1^2 = \sigma_2^2 = \frac{1}{4\theta_1}, \quad \text{and} \quad \sigma_3^2 = \frac{1}{4\theta_2}$$

in the Smith model, we find that $\delta(\mathbf{h}, u)$ is of the form given in Assumption 3.2,

$$\delta(\mathbf{h}, u) = \frac{1}{4(\sigma_1^2\sigma_2^2 - \sigma_{12}^2)} (\sigma_2^2 h_1^2 - 2\sigma_{12} h_1 h_2 + \sigma_1^2 h_2^2) + \frac{1}{4\sigma_3^2} u^2,$$

and, hence, has the same finite-dimensional distributions as the limit process in Theorem 3.1.

In the following, we analyse several correlation models used in the literature for modelling Gaussian space-time processes. In recent years, the interest in spatio-temporal correlation models has been growing significantly; especially in the construction of valid covariance functions in space and time. A simple way to construct such a model is to take the product of a spatial correlation function $\gamma_1(\mathbf{h})$ and a temporal correlation function $\gamma_2(u)$, i.e. $\gamma(\mathbf{h}, u) = \gamma_1(\mathbf{h})\gamma_2(u)$ (see for example Cressie and Huang [21]). Such a model is called separable and Assumption 3.2 is satisfied if the spatial and the temporal correlation functions have expansions around zero of the form

$$\gamma_1(\mathbf{h}) = 1 - \theta_1 \|\mathbf{h}\|^{\alpha_1} + O(\|\mathbf{h}\|^{\alpha_1}), \quad \gamma_2(u) = 1 - \theta_2 |u|^{\alpha_2} + O(|u|^{\alpha_2}),$$

respectively.

Example 3.1. A more sophisticated method to obtain covariance models is given on a process-based level. An interesting example in this context is presented in Baxevani, Podgórski and Rychlik [6] and Baxevani, Caires and Rychlik [5], who construct Gaussian space-time processes in a continuous setup using moving averages of spatial random fields over time, given by

$$X(\mathbf{s}, t) = \int_{-\infty}^{\infty} f(t-u) \Phi(\mathbf{s}, du),$$

where $\Phi(\cdot, du)$ is a Gaussian random field valued measure (see Appendix 5.1 in [5] for a definition) which is uniquely determined by the stationary correlation function $\gamma_1(\mathbf{h})$ and f is a deterministic kernel function. Note, that the process stated here is the special case of the

process introduced in [6].

Consider the kernel function $f(t) = e^{-\lambda t} \mathbb{1}_{\{t \geq 0\}}$ and the stationary spatial covariance model $\gamma_1(\mathbf{h}) = \exp\{-\|\mathbf{h}\|^2/C\}$. Using equation (5) in [6], the covariance function for the spatial lag \mathbf{h} and the temporal lag $u > 0$ can be calculated as

$$\begin{aligned} \tilde{\gamma}(\mathbf{h}, u) &= \gamma_1(\mathbf{h}) \int_{-\infty}^{\infty} e^{-\lambda(u-y)} \mathbb{1}_{\{u-y \geq 0\}} e^{\lambda y} \mathbb{1}_{\{y \leq 0\}} dy = \gamma_1(\mathbf{h}) \int_{-\infty}^0 e^{-\lambda u + 2\lambda y} dy \\ &= \gamma_1(\mathbf{h}) \frac{1}{2\lambda} e^{-\lambda u} = \frac{1}{2\lambda} \exp\left\{-\frac{\|\mathbf{h}\|^2}{C} - \lambda u\right\}, \end{aligned}$$

where the temporal dependence is of Ornstein-Uhlenbeck type (see Example 2 in [6]). The corresponding correlation function satisfies

$$\gamma(\mathbf{h}, u) = 1 - \frac{1}{C} \|\mathbf{h}\|^2 - \lambda|u| + O(\|\mathbf{h}\|^2 + |u|).$$

Separable space-time models do not allow for any interaction between space and time. Disadvantages of this assumption are pointed out for example in Cressie and Huang [21]. Therefore, nonseparable model constructions have been developed. One approach for combining purely spatial and temporal covariance functions leading to nonseparable covariance models is introduced in Ma [61, 62], given in terms of correlation functions by

$$\gamma(\mathbf{h}, u) = \int_0^{\infty} \int_0^{\infty} \gamma_1(\mathbf{h}v_1) \gamma_2(uv_2) dG(v_1, v_2),$$

where G is a bivariate distribution function on $[0, \infty) \times [0, \infty)$. Using the expansions above, it follows that

$$\gamma(\mathbf{h}, u) = 1 - \theta_1 \int_0^{\infty} v_1^{\alpha_1} dG_1(v_1) \|\mathbf{h}\|^{\alpha_1} - \theta_2 \int_0^{\infty} v_2^{\alpha_2} dG_2(v_2) |u|^{\alpha_2} + O(\|\mathbf{h}\|^{\alpha_1} + |u|^{\alpha_2}),$$

where G_1 and G_2 denote the marginal distributions of G , respectively. From this representation, the components in Assumption 3.2 can be defined.

3.3.1 Gneiting's class of correlation functions

A more elaborate class of nonseparable, stationary correlation functions is given by Gneiting's class [43]. This family of covariance functions is based on completely monotone functions, which are defined as functions φ on $(0, \infty)$ with existing derivatives of all orders $\varphi^{(n)}$, $n = 0, 1, \dots$ and

$$(-1)^n \varphi^{(n)}(t) \geq 0, \quad t > 0, \quad n = 0, 1, \dots$$

For our purpose we use a slightly different definition of Gneiting's class.

Definition 3.1 (Gneiting's class of correlation functions [43]). *Let $\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}$ be completely monotone and let $\psi : \mathbb{R}_+ \rightarrow \mathbb{R}$ be a positive function with completely monotone derivative. Further assume that $\psi(0)^{-d/2} \varphi(0) = 1$, where d is the spatial dimension, and $\beta_1, \beta_2 \in (0, 1]$. The function*

$$\gamma(\mathbf{h}, u) = \frac{1}{\psi(|u|^{2\beta_2})^{d/2}} \varphi\left(\frac{\|\mathbf{h}\|^{2\beta_1}}{\psi(|u|^{2\beta_2})}\right), \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}_+,$$

defines a non-separable, isotropic space-time correlation function with $\gamma(\mathbf{0}, 0) = 1$.

Compared to the original definition in [43], we included the parameters β_1 and β_2 , which is not a restriction since we can simply change the norms by defining $\|\cdot\|_*$ and $|\cdot|_*$ in terms of the old ones through

$$\|\mathbf{h}\|_* = \|\mathbf{h}\|^{\beta_1}, \quad \text{and} \quad |u|_* = |u|^{\beta_2}.$$

These new quantities are still norms since $\beta_1, \beta_2 \in (0, 1]$. In the next step, we provide an expansion of the correlation function around zero to obtain Assumption 3.2. The following proposition generalizes a result by Xue and Xiao [93] (Proposition 6.1).

Proposition 3.1. *Assume that $\psi'(0) \neq 0$. The correlation function taken from the Gneiting class satisfies Assumption 3.2 with $\alpha_1 = 2\beta_1$, $\alpha_2 = 2\beta_2$ and*

$$\theta_1 = \psi(0)^{-1} \left(\int_0^\infty z dF_\varphi(z) \middle/ \int_0^\infty dF_\varphi(z) \right), \quad \theta_2 = \frac{d}{2} \psi(0)^{-1} \psi'(0), \quad (3.19)$$

where F_φ is a non-decreasing bounded function with $F_\varphi(0) \neq 0$ and $\int_0^\infty z dF_\varphi(z) < \infty$.

Proof. Since the function φ is completely monotone, Bernstein's theorem (see for example Feller [38], Chapter 13) gives

$$\varphi(x) = \int_0^{\infty} e^{-xz} dF_{\varphi}(z), \quad x \geq 0.$$

From the properties of the correlation function $\psi(0)^{-d/2}\varphi(0) = 1$, it follows that $\psi(0) \neq 0$ and $\varphi(0) \neq 0$. We apply a Taylor expansion to the functions $\psi(\cdot)^{-d/2}$ and the exponential in the representation of φ :

$$\begin{aligned} \psi(u)^{-d/2} &= \psi(0)^{-d/2} - \frac{d}{2}\psi(0)^{-d/2-1}\psi'(0)u + o(u), \quad u \rightarrow 0, \\ \varphi(x) &= \int_0^{\infty} (1 - xz + o(x))dF_{\varphi}(z) = \int_0^{\infty} dF_{\varphi}(z) - x \int_0^{\infty} zdF_{\varphi}(z) + o(x), \quad x \rightarrow 0. \end{aligned}$$

Using the expansions in the correlation function and replacing u by $|u|^{2\beta_2}$ and x by $\|\mathbf{h}\|^{2\beta_1}/\psi(|u|^{2\beta_2})$, we obtain

$$\begin{aligned} \gamma(\mathbf{h}, u) &= \left(\psi(0)^{-d/2} - \frac{d}{2}\psi(0)^{-d/2-1}\psi'(0)|u|^{2\beta_2} + o(|u|^{2\beta_2}) \right) \\ &\quad \times \left(\int_0^{\infty} dF_{\varphi}(z) - \int_0^{\infty} zdF_{\varphi}(z)\|\mathbf{h}\|^{2\beta_1} \left[\psi(0)^{-1} - \psi(0)^{-2}\psi'(0)|u|^{2\beta_2} \right. \right. \\ &\quad \left. \left. + o(|u|^{2\beta_2}) \right] + o(\|\mathbf{h}\|^{2\beta_1}) \right) \\ &= \psi(0)^{-d/2}\varphi(0) - \frac{d}{2}\psi(0)^{-d/2-1}\varphi(0)\psi'(0)|u|^{2\beta_2} - \psi(0)^{-d/2-1} \int_0^{\infty} zdF_{\varphi}(z)\|\mathbf{h}\|^{2\beta_1} \\ &\quad + O(\|\mathbf{h}\|^{2\beta_1} + |u|^{2\beta_2}) \\ &= 1 - \frac{d}{2}\psi(0)^{-1}\psi'(0)|u|^{2\beta_2} - \psi(0)^{-1} \left(\int_0^{\infty} zdF_{\varphi}(z) \middle/ \int_0^{\infty} dF_{\varphi}(z) \right) \|\mathbf{h}\|^{2\beta_1} \\ &\quad + O(\|\mathbf{h}\|^{2\beta_1} + |u|^{2\beta_2}) \\ &= 1 - \theta_1\|\mathbf{h}\|^{2\beta_1} - \theta_2|u|^{2\beta_2} + O(\|\mathbf{h}\|^{2\beta_1} + |u|^{2\beta_2}), \end{aligned}$$

where θ_1 and θ_2 are defined as in (3.19). Note that we used $O(\|\mathbf{h}\|^{2\beta_1}|u|^{2\beta_2}) + o(\|\mathbf{h}\|^{2\beta_1}) +$

$o(|u|^{2\beta_2}) = O(\|\mathbf{h}\|^{2\beta_1} + |u|^{2\beta_2})$, which follows immediately from

$$\frac{\|\mathbf{h}\|^{2\beta_1}|u|^{2\beta_2} + \|\mathbf{h}\|^{2\beta_1} + |u|^{2\beta_2}}{\|\mathbf{h}\|^{2\beta_1} + |u|^{2\beta_2}} = \frac{1}{\frac{1}{\|\mathbf{h}\|^{2\beta_1}} + \frac{1}{|u|^{2\beta_2}}} + 1 < C, \quad \|\mathbf{h}\|, |u| \rightarrow 0.$$

□

Remark 3.5. For various choices of $\beta_1, \beta_2 \in (0, 1]$, the correlation functions in the Gneiting class have the flexibility to model different levels of smoothness of the underlying Gaussian process.

Example 3.2. We illustrate with a specific example, where the functions ϕ and ψ are taken from [43]; namely

$$\varphi(x) = (1 + bx)^{-\nu_1}, \quad \psi(x) = (1 + ax)^{\nu_2},$$

where $a, b, \nu_1 > 0$ and $0 < \nu_2 \leq 1$. The function φ is the Laplace transform of a gamma probability density function with shape $\nu_1 > 0$ and scale $b > 0$, which has mean $b\nu_1$. The first-order derivative of ψ at zero is given by $\psi'(0) = a\nu_2$. We choose $\beta_1 = \beta_2 = 1$ leading to $\alpha_1 = \alpha_2 = 2$ and, thus, a mean-square differentiable Gaussian random field. The constants θ_1 and θ_2 are given by

$$\theta_1 = b\nu_1, \quad \text{and } \theta_2 = \frac{d}{2}a\nu_2.$$

Figure 3.3.1 shows contour plots of the correlation function and the resulting tail dependence coefficient as in (3.17) based on different values for a and b with $\nu_1 = 3/2$ and $\nu_2 = 1$ fixed as a function of the space-lag $\|\mathbf{h}\|$ and time lag $|u|$. We see that the tail dependence function exhibits virtually the identical pattern of the underlying correlation function under a compression of the space-time scale. In particular, the extremal dependence dies out more quickly for large space and time lags than for the correlation function.

In a second step, we simulate processes in space and time using the above defined correlation model with $a = b = 0.03$, $\nu_1 = 3/2$ and $\nu_2 = 1$. We start the simulation procedure with $n = 100$ replications of a Gaussian random field with correlation function $\gamma(s_n s, t_n u)$ using the simulation routine in the R-package `RandomFields` by Schlather [75], where the circulant embedding method introduced in Section 2.2.2 is implemented. The fields are then transformed to standard Fréchet and the pointwise maximum is taken over the 100 replications. Figure 3.3.3 shows image and perspective plots (using the R-package `fields` for

visualization) of the simulated random fields for four consecutive time points. Figures 3.3.4 and 3.3.5 show the resulting random fields, if the margins are transformed to standard Gumbel and Weibull instead of Fréchet distributions. Clearly, in both cases the peaks are not as high as in the Fréchet case, leading to a smoother appearance of the resulting random field. One still sees the isolation of the peaks in the Fréchet case, which are well-known from the storm model of Smith using a centered Gaussian density for the function f . However, in the other two cases, they are not as pronounced.

3 Max-stable processes for extremes observed in space and time

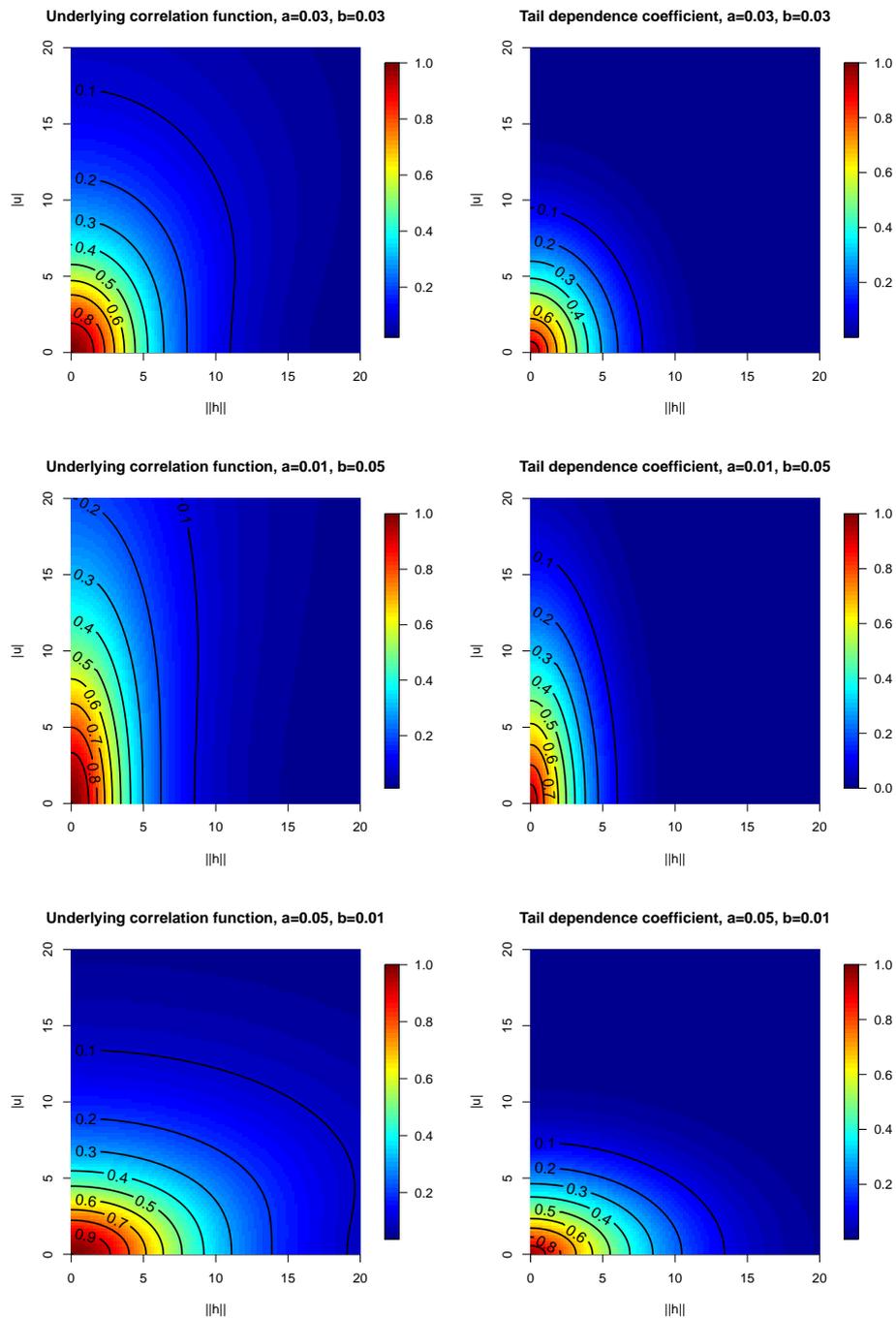


Figure 3.3.1: Contour plots for the underlying correlation function (left) and the resulting tail dependence coefficient (right) depending on the absolute space lag $\|h\|$ and time lag $|u|$ for different values of the scaling parameters a (time) and b (space), $\nu_1 = 3/2$ and $\nu_2 = 1$.

3.3 Possible correlation functions for the underlying Gaussian space-time process

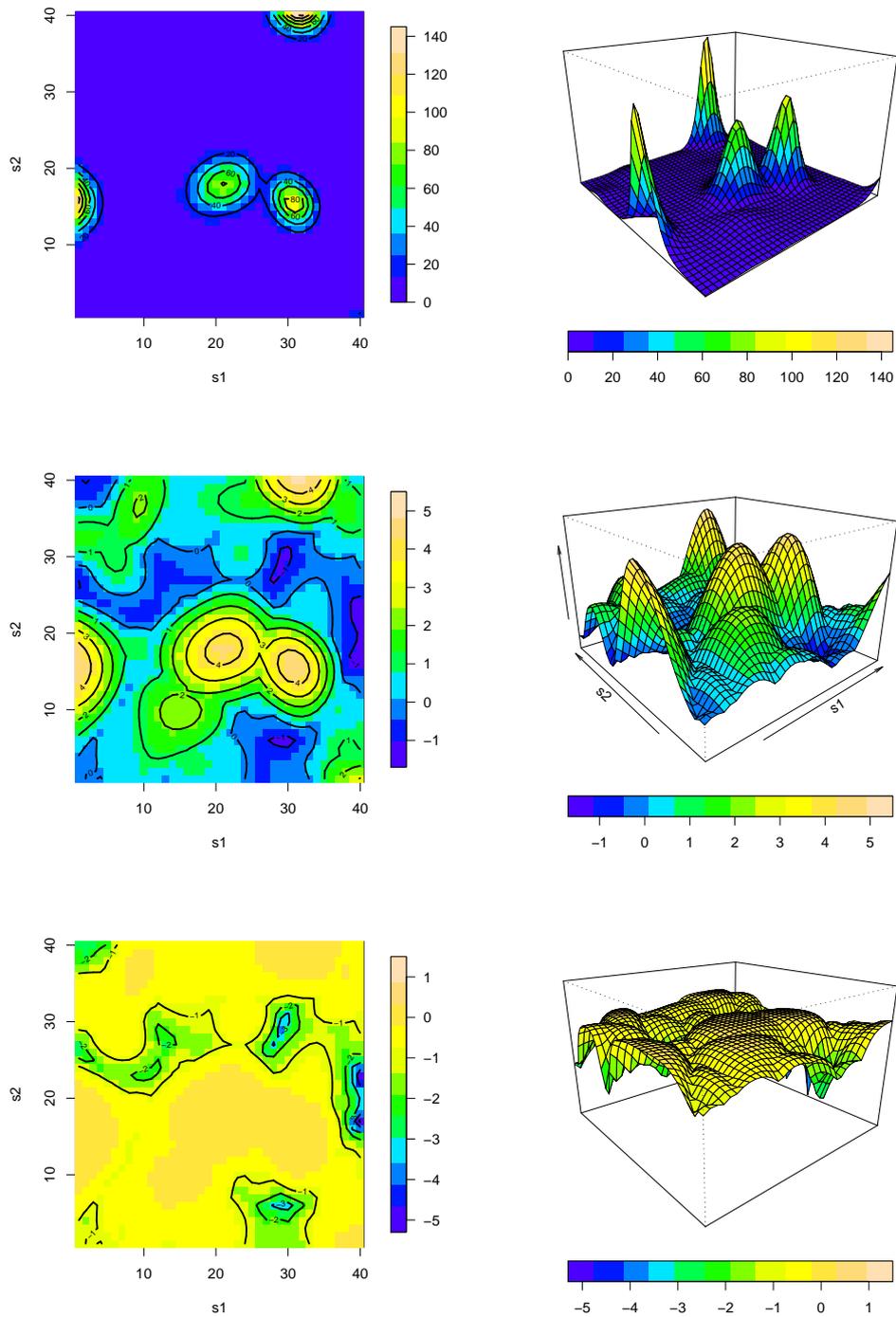


Figure 3.3.2: Simulated max-stable random fields (for fixed time point) with Fréchet (top), Gumbel (middle) and Weibull (bottom) marginal distributions.

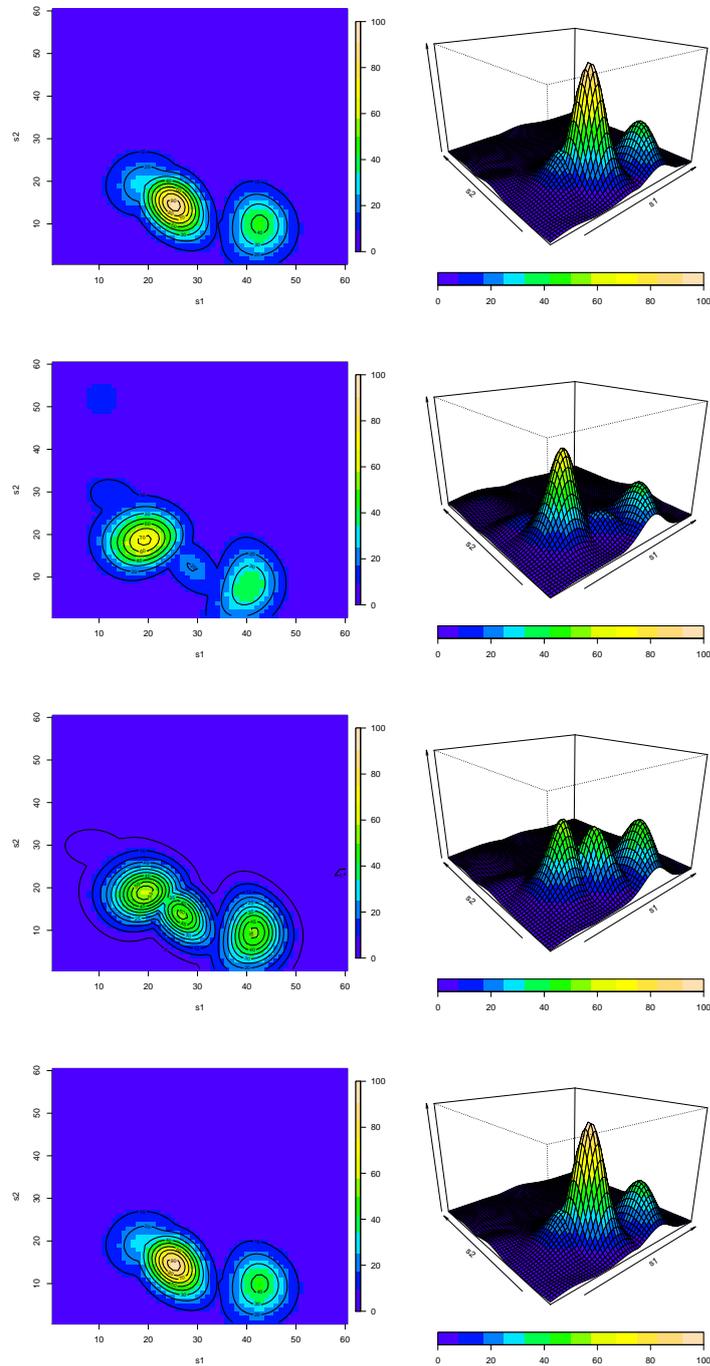


Figure 3.3.3: Simulated max-stable random fields with Type II (Fréchet) marginals for four consecutive time points (from the top to the bottom).

3.3 Possible correlation functions for the underlying Gaussian space-time process

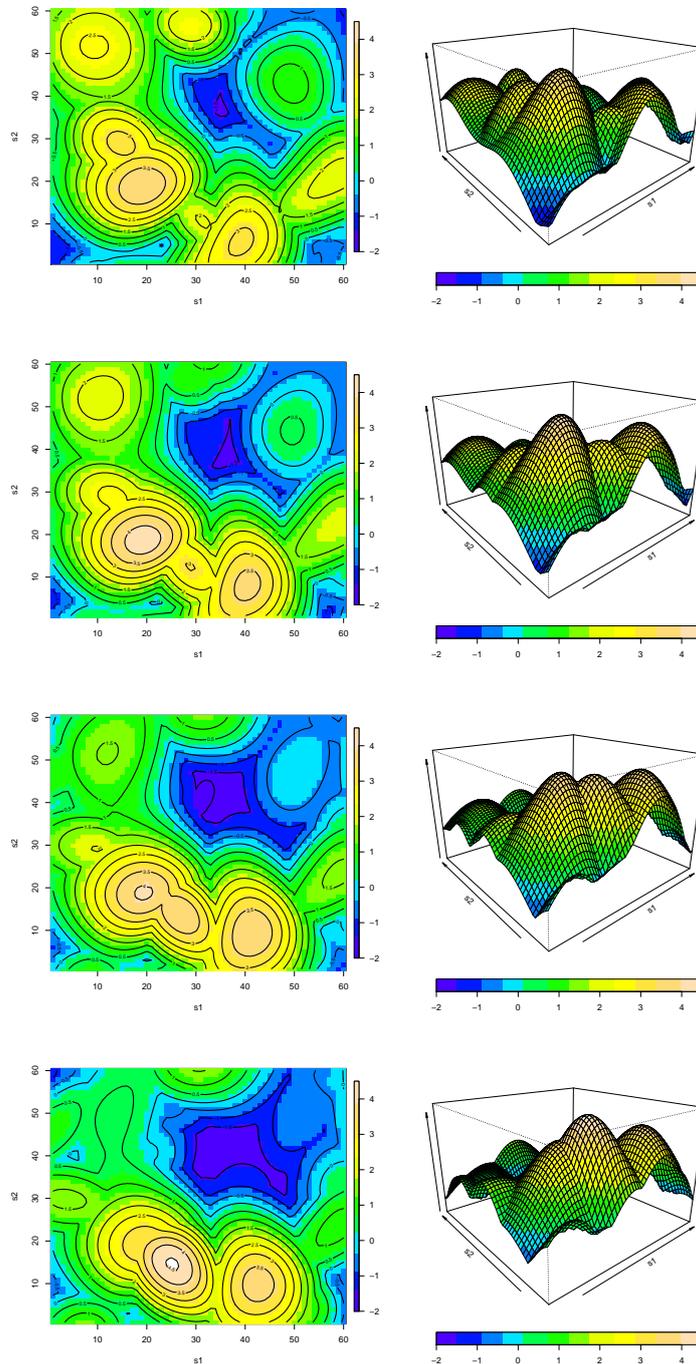


Figure 3.3.4: Simulated max-stable random fields with Type I (Gumbel) marginals for four consecutive time points (from the top to the bottom).

3 Max-stable processes for extremes observed in space and time

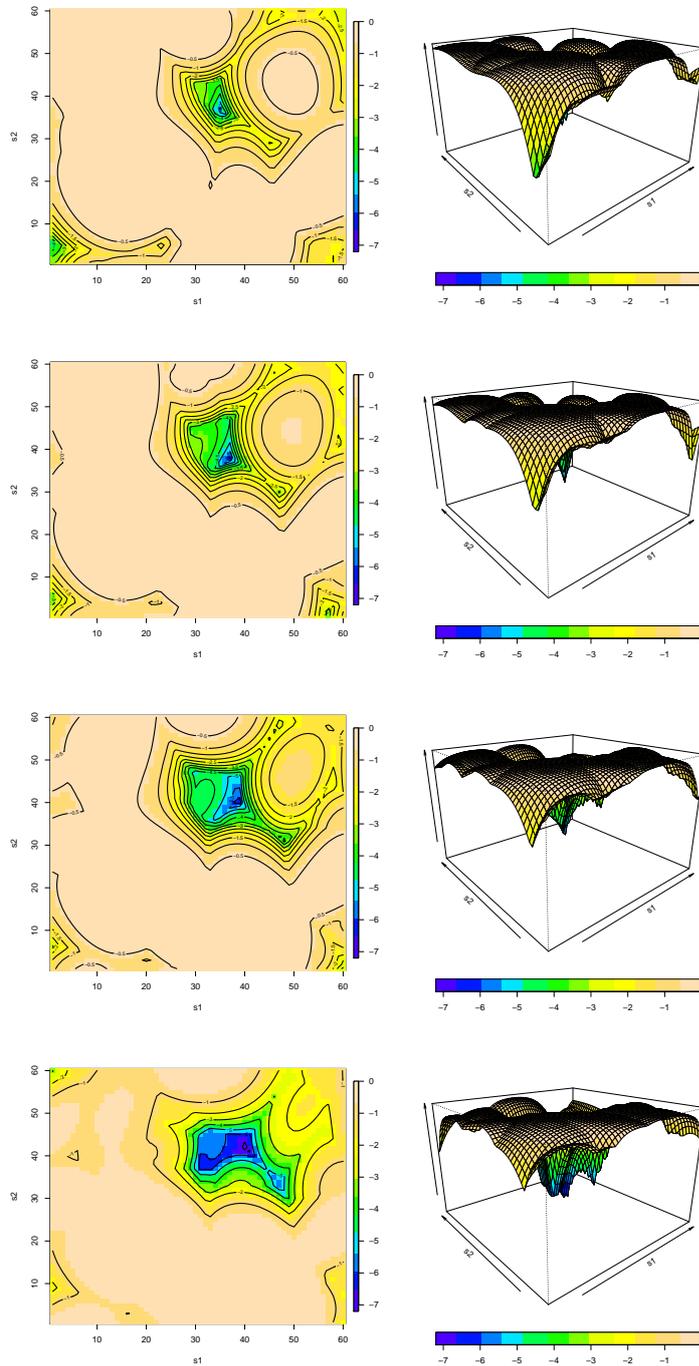


Figure 3.3.5: Simulated max-stable random fields with Type III (Weibull) marginals for four consecutive time points (from the top to the bottom).

3.3.2 Modelling spatial anisotropy

The correlation functions of the underlying Gaussian random fields in the previous sections were assumed to be spatially isotropic, meaning that the correlation function only depends on the absolute space and time lags $\|\mathbf{h}\|$ and $|u|$. An easy way to introduce spatial anisotropy to a model is given by geometric anisotropy, i.e.,

$$\tilde{\gamma}(\mathbf{h}, u) = \gamma(\|A\mathbf{h}\|, |u|),$$

where A is a transformation matrix.

In the two dimensional case geometric anisotropy in space can be modelled by the transformation matrix $A = TR$, with rotation and distance matrix where

$$R = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}, \quad T = \begin{pmatrix} 1/a_{\max} & 0 \\ 0 & 1/a_{\min} \end{pmatrix}.$$

Geometric anisotropy directly relates to the tail dependence coefficient

$$\chi(\mathbf{h}, u) = 2(1 - \Phi(\sqrt{\delta(A\mathbf{h}, u)})).$$

Figure 3.3.6 compares isotropic and anisotropic correlation functions and the corresponding tail dependence coefficients as function of the space lag components $\mathbf{h} = (h_1, h_2)^\top$, where the isotropic correlation is the same as in Example 3.2 with $\nu_1 = 3/2$, $\nu_2 = 1$ and $a = b = 0.03$. For the anisotropic parameters we choose $a_{\min} = 1$, $a_{\max} = 3$ and $\alpha = 45^\circ$. It can be seen, that the structure in the correlation function translates to the tail dependence coefficient. Corresponding max-stable random fields with Fréchet margins are shown for four consecutive time points in Figure 3.3.7. From the image plots, one clearly sees that the dependence is stronger in one direction. The perspective plots show that the isolated peaks are now stretched in one direction. In reality, this could correspond to wind speed peaks coming for example from a storm shaped particular in this wind direction.

A more complex way of introducing anisotropy in space is given by the Bernstein class, which is introduced in Porcu et al. [70] and revisited in Mateu et al. [64]. The covariance model is defined by

$$C(\mathbf{h}, u) = \int_0^\infty \int_0^\infty \exp \left\{ - \sum_{i=1}^d \psi_i(|h_i|) v_1 - \psi_t(|u|) v_2 \right\} dF(v_1, v_2),$$

where F is a bivariate distribution function and $\psi_i, i = 1, \dots, d$ and ψ_t are positive functions on $[0, \infty)$ with completely monotone derivatives, also called Bernstein functions. We assume that $\psi_i, i = 1, \dots, d$ and ψ_t are standardized, such that $\psi_i(0) = \psi_t(0) = 1, i = 1, \dots, d$. Assumption 3.1 can directly be derived for the corresponding correlation function.

$$\begin{aligned} \gamma(\mathbf{h}, u) &= C(\mathbf{h}, u) / C(\mathbf{0}, 0) \\ &= \left(1 - \sum_{i=1}^d \psi_i(|h_i|) \int_0^\infty \int_0^\infty v_1 dF(v_1, v_2) - \psi_t(|u|) \int_0^\infty \int_0^\infty v_2 dF(v_1, v_2) \right) \Bigg/ \\ &\quad \left(1 - d \int_0^\infty \int_0^\infty v_1 dF(v_1, v_2) - \int_0^\infty \int_0^\infty v_2 dF(v_1, v_2) \right) \\ &= \left(1 - \sum_{i=1}^d (1 - \theta_1 |h_i|^{\alpha_1} + o(|h_i|^{\alpha_1})) \int_0^\infty v_1 dF_{v_1}(v_1) - (1 - \theta_2 |u|^{\alpha_2} + O(|u|^{\alpha_2})) \right. \\ &\quad \left. \times \int_0^\infty v_2 dF_{v_2}(v_2) \right) \Bigg/ \left(1 - d \int_0^\infty v_1 dF_{v_1}(v_1) - \int_0^\infty v_2 dF_{v_2}(v_2) \right) \\ &= 1 - \theta_1 \int_0^\infty v_1 dF_{v_1}(v_1) \sum_{i=1}^d |h_i|^{\alpha_1} - \theta_2 \int_0^\infty v_2 dF_{v_2}(v_2) |u|^{\alpha_2} + O\left(\sum_{i=1}^d |h_i|^{\alpha_1}\right) + O(|u|^{\alpha_2}). \end{aligned}$$

3.3 Possible correlation functions for the underlying Gaussian space-time process

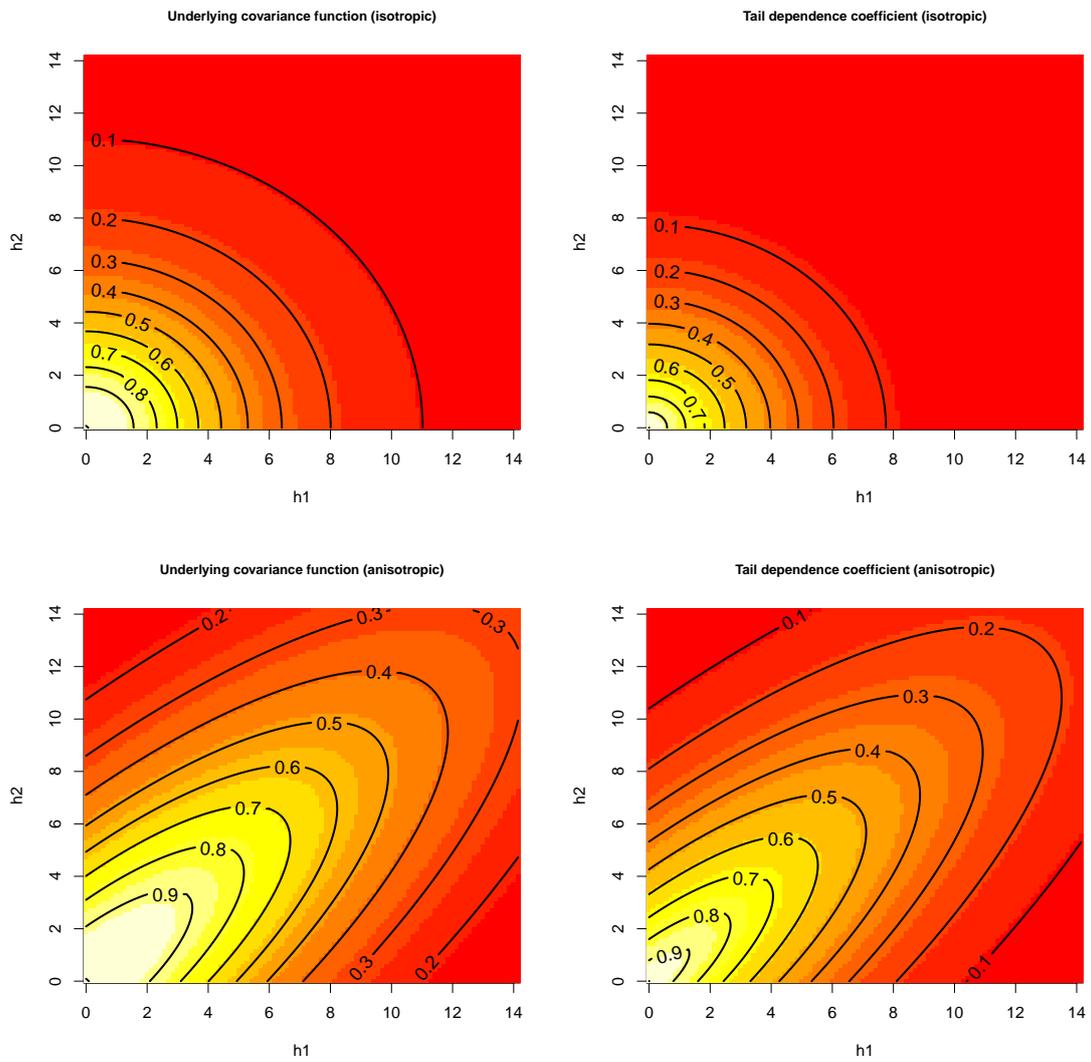


Figure 3.3.6: Contour plots for covariance functions and tail dependence coefficients depending on the spatial lag components h_1 and h_2 in the isotropic case (top) and for included geometric anisotropy (bottom) for $a = b = 0.03$, $a_{\min} = 1$, $a_{\max} = 3$ and $\alpha = 45^\circ$.

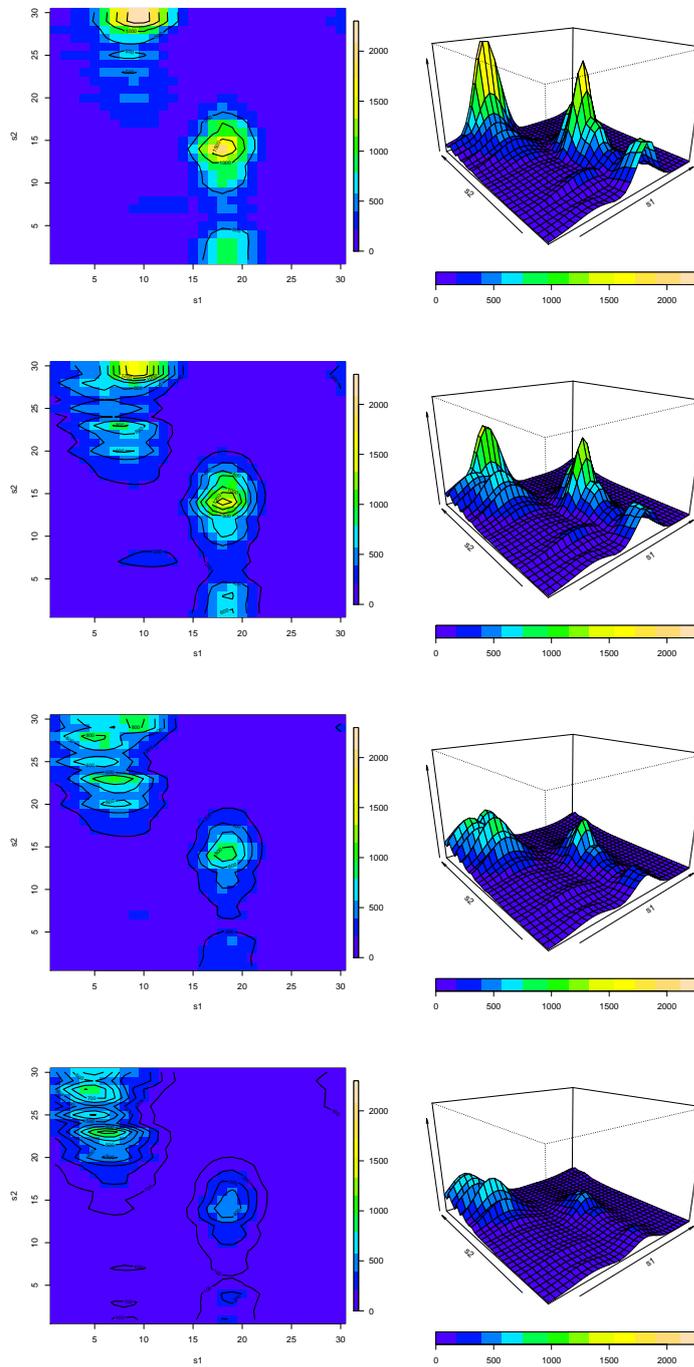


Figure 3.3.7: Simulated anisotropic max-stable random fields with Fréchet margins from Example 3.2 ($a = 0.03$, $b = 0.03$, $\nu_1 = 3/2$, $\nu_2 = 1$) with anisotropic parameters $a_{\min} = 1$, $a_{\max} = 3$ and $\alpha = 45^\circ$.

3.4 Proofs

3.4.1 Proof of tightness in Theorem 3.1 (see Kabluchko et al. [50])

The objective is to show that the sequence η_n , $n \in \mathbb{N}$, defined by

$$\eta_n(s, t) = \bigvee_{j=1}^n -\frac{1}{\log \Phi(Z_j(s_n s, t_n t))}, \quad s \in \mathbb{R}^d, t \in [0, \infty),$$

is tight in $C(K)$, where $K \subset \mathbb{R}^d \times [0, \infty)$ is a compact set and $C(K)$ is the space of continuous functions on K . The proof is taken from Kabluchko et al. [50] and extended to the space-time setting here. First, note that

$$\begin{aligned} P(\eta_n(s, t) \leq x) &= P^n(Z(s_n s, t_n t) \leq \Phi^{-1}(e^{-1/nx})) \\ &\sim P^n\left(Z(s_n s, t_n t) \leq \frac{\log(x)}{b_n} + b_n\right) = P\left(\bigvee_{j=1}^n b_n(Z_j(s_n s, t_n t) - b_n) \leq \log(x)\right), \end{aligned}$$

where

$$b_n = \sqrt{2 \log n} - \frac{\log \log n + \log(4\pi)}{2 \sqrt{\log n}}.$$

Define for $n \in \mathbb{N}$

$$Y_n(s, t) = b_n(Z(s_n s, t_n t) - b_n), \quad s \in \mathbb{R}^d, t \in [0, \infty).$$

The key point in the proof is the following lemma:

Lemma 3.1. *For $\omega \in \mathbb{R}$, let Y_n^ω denote the conditional process $Y_n - \omega$ given $Y_n(\mathbf{0}, 0) = \omega$. Then,*

1. *The family of processes $Y_n^\omega - \mu_n^\omega$, $\omega \in \mathbb{R}$, $n \in \mathbb{N}$, is tight in $C(K)$, where*

$$\mu_n^\omega(s, t) = \mathbb{E}[Y_n^\omega(s, t)] = \mathbb{E}[Y_n(s, t) - \omega \mid Y_n(\mathbf{0}, 0) = \omega].$$

2. *For every $c > 0$, the family of processes Y_n^ω , $\omega \in [-c, c]$, $n \in \mathbb{N}$, is tight in $C(K)$.*

Proof of Lemma 3.1. Since

$$\begin{aligned} Y_n(\mathbf{s}, t) - \omega &= b_n Z(s_n \mathbf{s}, t_n t) - b_n^2 - \omega \stackrel{d}{\sim} \mathcal{N}(-b_n^2 - \omega, b_n^2), \\ Y_n(\mathbf{0}, 0) &= b_n Z(\mathbf{0}, 0) - b_n^2 \stackrel{d}{\sim} \mathcal{N}(-b_n^2, b_n^2), \end{aligned}$$

where $\stackrel{d}{\sim}$ denotes asymptotic equivalence in distribution, it follows for all $(\mathbf{s}, t) \in K$ that

$$\begin{aligned} &\mu_n^\omega(\mathbf{s}, t) \\ &= \mathbb{E}[Y_n(\mathbf{s}, t) - \omega] + \text{Cor}(Y_n(\mathbf{s}, t) - \omega, Y_n(\mathbf{0}, 0)) \sqrt{\text{Var}(Y_n(\mathbf{s}, t) - \omega)} \frac{\omega - \mathbb{E}[Y_n(\mathbf{0}, 0)]}{\sqrt{\text{Var}(Y_n(\mathbf{0}, 0))}} \\ &= -b_n^2 - \omega + b_n^2 \gamma(s_n \mathbf{s}, t_n t) \frac{\omega + b_n^2}{b_n^2} \\ &= -(b_n^2 + \omega)(1 - \gamma(s_n \mathbf{s}, t_n t)) \\ &\sim -\log n \frac{\delta(\mathbf{s}, t)}{\log n} - \frac{\omega \delta(\mathbf{s}, t)}{\log n} \rightarrow -\delta(\mathbf{s}, t), \quad n \rightarrow \infty, \end{aligned}$$

where we used $(\log n)(1 - \gamma(s_n \mathbf{s}, t_n t)) \rightarrow \delta(\mathbf{s}, t), n \rightarrow \infty$. Furthermore, for $(\mathbf{s}_1, t_1), (\mathbf{s}_2, t_2) \in K$,

$$\begin{aligned} \text{Cov}(Y_n^\omega(\mathbf{s}_1, t_1), Y_n^\omega(\mathbf{s}_2, t_2)) &= b_n^2 (\gamma(s_n(\mathbf{s}_1 - \mathbf{s}_2), t_n(t_1 - t_2)) - \gamma(s_n \mathbf{s}_1, t_n t_1) \gamma(s_n \mathbf{s}_2, t_n t_2)) \\ &\sim (\log n) (-(1 - \gamma(s_n(\mathbf{s}_1 - \mathbf{s}_2))) + 1 - \gamma(s_n \mathbf{s}_1, t_n t_1) \gamma(s_n \mathbf{s}_2, t_n t_2)) \\ &\sim -\delta(\mathbf{s}_1 - \mathbf{s}_2, t_1 - t_2) + \log n \left(\frac{\delta(\mathbf{s}_1, t_1)}{\log n} + \frac{\delta(\mathbf{s}_2, t_2)}{\log n} - \frac{\delta(\mathbf{s}_1, t_1) \delta(\mathbf{s}_2, t_2)}{(\log n)^2} \right) \\ &\rightarrow -\delta(\mathbf{s}_1 - \mathbf{s}_2, t_1 - t_2) + \delta(\mathbf{s}_1, t_1) + \delta(\mathbf{s}_2, t_2), \quad n \rightarrow \infty. \end{aligned}$$

We derive an upper bound for the variance of the difference of $Y_n^\omega(\mathbf{s}_1, t_1)$ and $Y_n^\omega(\mathbf{s}_2, t_2)$ for $(\mathbf{s}_1, t_1), (\mathbf{s}_2, t_2) \in K$,

$$\begin{aligned} \text{Var}(Y_n^\omega(\mathbf{s}_1, t_1) - Y_n^\omega(\mathbf{s}_2, t_2)) &= b_n^2 (1 - \gamma(s_n \mathbf{s}_1, t_n t_1))^2 + b_n^2 (1 - \gamma(s_n \mathbf{s}_2, t_n t_2))^2 \\ &\quad - 2b_n^2 (\gamma(s_n(\mathbf{s}_1 - \mathbf{s}_2), t_n(t_1 - t_2)) - \gamma(s_n \mathbf{s}_1, t_n t_1) \gamma(s_n \mathbf{s}_2, t_n t_2)) \\ &= b_n^2 (2 - 2\gamma(s_n(\mathbf{s}_1 - \mathbf{s}_2), t_n(t_1 - t_2)) - (\gamma(s_n \mathbf{s}_1, t_n t_1) - \gamma(s_n \mathbf{s}_2, t_n t_2))^2) \\ &\leq 2b_n^2 (1 - \gamma(s_n(\mathbf{s}_1 - \mathbf{s}_2), t_n(t_1 - t_2))) \leq K_1 \delta(\mathbf{s}_1 - \mathbf{s}_2, t_1 - t_2) \\ &\leq K_1 (D((\mathbf{s}_1, t_1), (\mathbf{s}_2, t_2)))^2, \end{aligned} \tag{3.20}$$

where $K_1 > 0$ is some constant and the last inequality follows from (3.3). The following corollary is taken from Ledoux and Talagrand [58], Corollary 11.7, and will be used to show Lemma 3.1.

Corollary 3.1 (Ledoux and Talagrand [58]). *Let (T, d) be compact and let ψ be a Young function (convex, increasing, $\lim_{t \rightarrow \infty} \psi(t) = \infty$, $\psi(0) = 0$). Assume, that d is a pseudo-metric ($d(x, y) \geq 0$, $d(x, y) = d(y, x)$, $d(x, z) \leq d(x, y) + d(y, z)$ and $d(x, x) = 0$, not necessarily: $d(x, y) = 0 \Rightarrow x = y$). Furthermore,*

$$\int_0^D \psi^{-1}(N(T, d, \epsilon)) d\epsilon < \infty,$$

where $D = \sup \{d(x, y), x, y \in T\}$ and $N(T, d; \epsilon)$ denotes the smallest number of open balls of radius ϵ in the metric d which form a covering of T . Let \mathcal{X} be a family of separable random processes $X = (X(t))_{t \in T}$ in $L_1(\Omega, \mathcal{A}, P)$, such that for all $t_1, t_2 \in T$ and $A \in \mathcal{A}$ measurable

$$\int_A |X(t_1) - X(t_2)| dP \leq d(t_1, t_2) P(A) \psi^{-1}(1/P(A)).$$

Then, each element of \mathcal{X} defines a tight probability distribution on $C(T)$ and \mathcal{X} is weakly relatively compact if and only if, for some $t \in T$, $\{X(t); X \in \mathcal{X}\}$ is weakly relatively compact.

Now it follows with (3.20) that

$$\begin{aligned} & \mathbb{E} \left[\left| Y_n^\omega(s_1, t_1) - \mu_n^\omega(s_1, t_1) - Y_n^\omega(s_2, t_2) + \mu_n^\omega(s_2, t_2) \right|^2 \right] \\ & \leq \mathbb{E} \left[\left| Y_n^\omega(s_1, t_1) - \mu_n^\omega(s_1, t_1) - Y_n^\omega(s_2, t_2) + \mu_n^\omega(s_2, t_2) \right|^2 \right] \\ & = \text{Var}(Y_n^\omega(s_1, t_1) - Y_n(s_2, t_2)) \leq K_1 (D((s_1, t_1), (s_2, t_2)))^2. \end{aligned}$$

Set $\psi(x) = x^2$ in Corollary 3.1. It follows that the family of processes $Y_n^\omega - \mu_n^\omega$, $\omega \in \mathbb{R}$, $n \in \mathbb{N}$ is tight in $C(K)$. Since $\mu_n^\omega(s, t) \rightarrow -\delta(s, t)$, $n \rightarrow \infty$ uniformly in $(s, t) \in K$ and $\omega \in [-c, c]$ it follows, that the family μ_n^ω is tight in $C(K)$. Together with part (a), the second statement of the lemma follows. \square

Together with Lemma 3.1 the tightness of the sequence η_n is shown by using Theorem 8.3 in Billingsley [10]. In particular, the following two statements are verified.

(I) The sequence $\eta_n(\mathbf{0}, 0)$, $n \in \mathbb{N}$ is tight in \mathbb{R} .

(II) For all $\epsilon > 0$, $a > 0$, there exist $\delta > 0$ and $N \in \mathbb{N}$, such that

$$P \left(\sup_{\substack{(s_1, t_1), (s_2, t_2) \in K \\ \|(s_1, t_1), (s_2, t_2)\| \leq \delta}} \|\eta_n(s_1, t_1) - \eta_n(s_2, t_2)\| > a \right) < \epsilon, \text{ for all } n > N.$$

The first assertion (I) is obvious, since $\eta_n(\mathbf{0}, 0)$ converges weakly to a Fréchet distribution for $n \rightarrow \infty$, and is therefore tight in \mathbb{R} . Now, define for $n \in \mathbb{N}$

$$\omega_\delta(\eta_n) = \sup_{\substack{(s_1, t_1), (s_2, t_2) \in K \\ \|(s_1, t_1), (s_2, t_2)\| \leq \delta}} \|\eta_n(s_1, t_1) - \eta_n(s_2, t_2)\| \quad \text{and} \quad K_n = \{\omega_\delta(\eta_n) > a\}.$$

To prove (II) the following statements are shown.

1. a) $K_n \subset G_n \cup H_n$, where

$$G_n := \left\{ \exists (s, t) \in K : \bigvee_{j=1}^n Y_{j,n}(s, t) \neq \sup_{\substack{j=1, \dots, n \\ |Y_{j,n}(\mathbf{0}, 0)| < c_2}} Y_{j,n}(s, t) \right\},$$

$$H_n := \bigcup_{j=1}^n C_{j,n}$$

$$:= \bigcup_{j=1}^n \left\{ Y_{j,n}(\mathbf{0}, 0) \in [-c_2, c_2], \quad \sup_{\substack{(s_1, t_1), (s_2, t_2) \in K \\ \|(s_1, t_1), (s_2, t_2)\| \leq \delta}} |Y_{j,n}(s_1, t_1) - Y_{j,n}(s_2, t_2)| > a \right\}.$$

b) $G_n \subset E_n \cup F_n \cup \bigcup_{j=1}^n B_{j,n}$ with

$$E_n := \left\{ \inf_{(s, t) \in K} \bigvee_{j=1}^n Y_{j,n}(s, t) < -c_1 \right\},$$

$$F_n := \bigcup_{j=1}^n \{Y_{j,n}(\mathbf{0}, 0) > c_2\} \Rightarrow P(F_n) = P(\eta_n(\mathbf{0}, 0) > c_2) < \epsilon, \quad n > N,$$

$$B_{j,n} := \left\{ Y_{j,n}(\mathbf{0}, 0) < -c_2, \quad \sup_{(s, t) \in K} Y_{j,n}(s, t) > -c_1 \right\}$$

c) $E_n \subset \bigcap_{j=1}^n A_{j,n}^c$, where

$$A_{j,n}^c := \bigcap_{j=1}^n \left\{ Y_{j,n}(\mathbf{0}, 0) \in [-c_0, c_0], \inf_{(s,t) \in K} Y_{j,n}(s, t) - Y_{j,n}(\mathbf{0}, 0) \geq c_0 - c_1 \right\}.$$

2. In a second step the positive constants c_0 , c_1 and c_2 are chosen such that the probabilities $P(A_{j,n})$, $P(F_n)$, $P(B_{j,n})$ and $P(C_{j,n})$ are bounded by ϵ .

We show 1.a) (i.e. $K_n \subset G_n \cup H_n$). Assume there exists $\omega \in \Omega$ such that $\omega \notin G_n \cup H_n$. Then,

$$\bigvee_{j=1}^n Y_{j,n}(s, t)(\omega) = \sup_{\substack{j=1, \dots, n \\ |Y_{j,n}(\mathbf{0}, 0)(\omega)| < c_2}} Y_{j,n}(s, t)(\omega), \quad \text{for all } (s, t) \in K,$$

and for all $j = 1, \dots, n$

$$\begin{aligned} Y_{j,n}(\mathbf{0}, 0)(\omega) &\notin [-c_2, c_2] \\ \sup_{\substack{(s_1, t_1), (s_2, t_2) \in K \\ \|(s_1, t_1), (s_2, t_2)\| \leq \delta}} |Y_{j,n}(s_1, t_1)(\omega) - Y_{j,n}(s_2, t_2)(\omega)| &\leq a. \end{aligned}$$

It follows that

$$\begin{aligned} &\sup_{\substack{(s_1, t_1), (s_2, t_2) \in K \\ \|(s_1, t_1), (s_2, t_2)\| \leq \delta}} \|\eta_n(s_1, t_1)(\omega) - \eta_n(s_2, t_2)(\omega)\| \\ &= \sup_{\substack{(s_1, t_1), (s_2, t_2) \in K \\ \|(s_1, t_1), (s_2, t_2)\| \leq \delta}} \left| \sup_{\substack{j=1, \dots, n \\ |Y_{j,n}(\mathbf{0}, 0)(\omega)| < c_2}} Y_{j,n}(s_1, t_1)(\omega) - \sup_{\substack{j=1, \dots, n \\ |Y_{j,n}(\mathbf{0}, 0)(\omega)| < c_2}} Y_{j,n}(s_2, t_2)(\omega) \right| \\ &\leq \sup_{\substack{(s_1, t_1), (s_2, t_2) \in K \\ \|(s_1, t_1), (s_2, t_2)\| \leq \delta}} \sup_{\substack{j=1, \dots, n \\ |Y_{j,n}(\mathbf{0}, 0)(\omega)| < c_2}} |Y_{j,n}(s_1, t_1)(\omega) - Y_{j,n}(s_2, t_2)(\omega)| \leq a, \end{aligned}$$

which implies $\omega \notin K_n$.

We show 1.b) (i.e. $G_n \subset E_n \cup F_n \cup \bigcup_{j=1}^n B_{j,n}$). Since $G_n \subset E_n \cup F_n \cup G_n \setminus (E_n \cup F_n)$, it is sufficient to verify that

$$G_n \setminus (E_n \cup F_n) \subset \bigcup_{j=1}^n B_{j,n}.$$

Let $\omega \in G_n \setminus (E_n \cup F_n)$. Then, it follows that there exists $(s, t) \in K$ such that

$$\bigvee_{j=1}^n Y_{j,n}(s, t)(\omega) \neq \sup_{\substack{j=1, \dots, n \\ |Y_{j,n}(\mathbf{0}, 0)(\omega)| < c_2}} Y_{j,n}(s, t)(\omega).$$

In addition,

(a) $\forall j = 1, \dots, n : Y_{j,n}(\mathbf{0}, 0)(\omega) \leq c_2$ and

(b) $\inf_{(s,t) \in K} \bigvee_{j=1}^n Y_{j,n}(s, t)(\omega) \geq -c_1$.

From (a) it follows that $Y_{j,n}(\mathbf{0}, 0)(\omega) < -c_2$, since otherwise

$$\bigvee_{j=1}^n Y_{j,n}(s, t)(\omega) = \sup_{\substack{j=1, \dots, n \\ |Y_{j,n}(\mathbf{0}, 0)(\omega)| < c_2}} Y_{j,n}(s, t)(\omega) \text{ for all } (s, t) \in K. \text{ From (b) there exists } j =$$

$1, \dots, n$ such that

$$\sup_{(s,t) \in K} Y_{j,n}(s, t)(\omega) > -c_1.$$

Therefore, $\omega \in \bigcup_{j=1}^n B_{j,n}$.

To show 1.c) assume that $\omega \notin \bigcap_{j=1}^n A_{j,n}^c$. Then, there exists $\tilde{j} \in \{1, \dots, n\}$ such that

$$Y_{\tilde{j},n}(\mathbf{0}, 0)(\omega) \in [-c_0, c_0] \quad \text{and} \quad \inf_{(s,t) \in K} Y_{\tilde{j},n}(s, t)(\omega) - Y_{\tilde{j},n}(\mathbf{0}, 0)(\omega) \geq c_0 - c_1.$$

Therefore,

$$\inf_{(s,t) \in K} Y_{\tilde{j},n}(s, t)(\omega) \geq -c_1,$$

which implies

$$\inf_{(s,t) \in K} \bigvee_{j=1}^n Y_{j,n}(s, t)(\omega) \geq -c_1,$$

and, thus, $\omega \notin E_n$.

By using the tightness of the conditional family of processes $Y_n^\omega, \omega \in [-c_0, c_0], n \in \mathbb{N}$, the constants c_0, c_1 and c_2 can be chosen, such that the probabilities $P(A_{j,n}), P(F_n), P(B_{j,n})$ and $P(C_{j,n})$ are bounded by ϵ . First,

$$P(A_{j,n}) = P(Y_{j,n}(\mathbf{0}, 0) \in [-c_0, c_0], \inf_{(s,t) \in K} Y_{j,n}(s, t) - Y_{j,n}(\mathbf{0}, 0) \geq c_0 - c_1)$$

$$\begin{aligned}
 &= \int_{-c_0}^{c_0} P\left(\inf_{(s,t) \in K} Y_{j,n}(s,t) - \omega \geq c_0 - c_1 \mid Y_{j,n}(\mathbf{0},0) = \omega\right) f_{Y_{j,n}(\mathbf{0},0)}(\omega) d\omega \\
 &= \int_{-c_0}^{c_0} \frac{1}{\sqrt{2\pi}b_n} e^{-\frac{(\omega-b_n)^2}{(2b_n)^2}} P\left(\inf_{(s,t) \in K} Y_{j,n}^\omega(s,t) \geq c_0 - c_1\right) d\omega \\
 &= \frac{1}{\sqrt{2\pi}b_n} e^{-b_n^2/2} \int_{-c_0}^{c_0} e^{-\omega - \omega^2/(2b_n^2)} P\left(\inf_{(s,t) \in K} Y_{j,n}^\omega(s,t) \geq c_0 - c_1\right) d\omega
 \end{aligned}$$

The constant c_0 is chosen large enough such that $2e^{-c_0} < \epsilon$. Since $Y_n^\omega, \omega \in [-c_0, c_0], n \in \mathbb{N}$, is tight, it follows that $c_1 > 0$ can be chosen, such that

$$P\left(\inf_{(s,t) \in K} Y_n^\omega(s,t) < c_0 - c_1\right) < 1/2.$$

Therefore,

$$P(A_{j,n}) \geq \frac{1}{4n} \int_{-c_0}^{c_0} e^{-\omega - \omega^2/(2b_n^2)} d\omega \geq \frac{c_0}{n}, \quad n > N,$$

which implies

$$P(E_n) \leq P\left(\bigcap_{j=1}^n A_{j,n}^c\right) \leq \left(1 - \frac{c_0}{n}\right)^n \leq 2e^{-c_0} < \epsilon.$$

We analyse the probability $P(B_{j,n})$. Note that for large n there exist constants $c_4, \kappa > 0$ such that

$$\mu_n(s,t) \leq c_4 - \frac{\omega}{2}, \quad \text{and} \quad \text{Var}(Y_n^\omega(s,t)) \leq \kappa^2, \quad (s,t) \in K.$$

Further, there exists $c_3 > 0$ such that

$$P\left(\sup_{(s,t) \in K} (Y_n^\omega(s,t) - \mu_n^\omega(s,t)) > c_3\right) < \frac{1}{2}, \quad \omega \in \mathbb{R}, \quad n \in \mathbb{N}.$$

Using Borell's inequality (Theorem D.1 in [58]), it follows

$$P\left(\sup_{(s,t) \in K} Y_n^\omega(s,t) > -c_1 - \omega\right) < 2\bar{\Phi}\left(-\frac{-c_1 - \omega/2 - c_3 - c_4}{\kappa}\right).$$

Now, with $\omega < -4(c_1 + c_3 + c_4)$ and $\psi(x) \leq e^{-x^2/2}$,

$P\left(\sup_{(s,t) \in K} Y_n^\omega(s,t) > -c_1 - \omega\right) < 2e^{-\omega^2/(32\kappa^2)}$, $n > N$, and, thus,

$$\begin{aligned} P(B_{j,n}) &= P(Y_{j,n}(\mathbf{0}, 0) < -c_2, \sup_{(s,t) \in K} Y_{j,n}(s,t) > -c_1) \\ &= \frac{1}{\sqrt{2\pi b_n}} e^{-b_n^2/2} \int_{-\infty}^{-c_2} e^{-\omega - \omega^2/(2b_n^2)} P\left(\sup_{(s,t) \in K} Y_n^\omega(s,t) > -c_1 - \omega\right) d\omega \\ &\leq \frac{4}{n} \int_{-\infty}^{-c_2} e^{-\omega} e^{-\omega^2/(32\kappa^2)} d\omega. \end{aligned}$$

The constant c_2 is chosen large enough, such that $nP(B_{j,n}) < \epsilon$, $n > N$. For the probability $P(C_{j,n})$ it finally follows

$$\begin{aligned} P(C_{j,n}) &= P(Y_n(\mathbf{0}, 0) \in [-c_2, c_2], \omega_\delta(Y_{j,n}) > a) \\ &= \frac{1}{\sqrt{2\pi b_n}} e^{-b_n^2/2} \int_{-c_2}^{c_2} e^{-\omega - \omega^2/(2b_n^2)} P(\omega_\delta(Y_n^\omega) > a) d\omega, \end{aligned}$$

and δ can be chosen sufficiently small such that

$$P(C_{j,n}) < \frac{\epsilon}{n}.$$

□

3.4.2 Derivation of the bivariate distribution function for the space-time Smith model

Proof of Theorem 3.2. Since space and time are independent, we can write

$$f_0(\mathbf{z}, x) = f_1(\mathbf{z})f_2(x), \quad \mathbf{z} \in \mathbb{R}^2, x \in \mathbb{R}$$

where f_1 is the density of a bivariate normal distribution with mean $\mathbf{0}$ and covariance matrix Σ , and f_2 is the density of a normal distribution with mean 0 and variance σ_3^2 . Starting from

equation (3.12) with $K = 2$, we obtain

$$\begin{aligned}
 F(y_1, y_2) &= \exp \left\{ - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{f_0(\mathbf{z}, x)}{y_1} \right) \vee \left(\frac{f_0(\mathbf{z} - \mathbf{h}, x - u)}{y_2} \right) dz dx \right\} \\
 &= \exp \left\{ - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{f_0(\mathbf{z}, x)}{y_1} \mathbb{1} \left\{ \frac{f_0(\mathbf{z}, x)}{y_1} \geq \frac{f_0(\mathbf{z} - \mathbf{h}, x - u)}{y_2} \right\} dz dx \right. \\
 &\quad \left. - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{f_0(\mathbf{z} - \mathbf{h}, x - u)}{y_2} \mathbb{1} \left\{ \frac{f_0(\mathbf{z} - \mathbf{h}, x - u)}{y_2} \geq \frac{f_0(\mathbf{z}, x)}{y_1} \right\} dz dx \right\} \\
 &= \exp \{-(\text{I}) - (\text{II})\}
 \end{aligned}$$

$$\begin{aligned}
 (\text{I}) &= \int_{-\infty}^{\infty} f_2(x) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{f_1(\mathbf{z})}{y_1} \mathbb{1} \left\{ \frac{f_1(\mathbf{z})f_2(x)}{y_1} \geq \frac{f_1(\mathbf{z} - \mathbf{h})f_2(x - u)}{y_2} \right\} dz dx \\
 &= \int_{-\infty}^{\infty} f_2(x) \frac{1}{y_1} \mathbb{E} \left[\mathbb{1} \left\{ \frac{f_1(\mathbf{Z})f_2(x)}{y_1} \geq \frac{f_1(\mathbf{Z} - \mathbf{h})f_2(x - u)}{y_2} \right\} \right] dx,
 \end{aligned}$$

where \mathbf{Z} has a normal density with mean $\mathbf{0}$ and variance Σ . Now note that

$$\begin{aligned}
 \frac{f_1(\mathbf{Z})f_2(x)}{y_1} \geq \frac{f_1(\mathbf{Z} - \mathbf{h})f_2(x - u)}{y_2} &\Leftrightarrow f_1(\mathbf{Z}) \geq f_1(\mathbf{Z} - \mathbf{h}) \frac{y_1}{y_2} \frac{f_2(x - u)}{f_2(x)} \\
 &\Leftrightarrow (2\pi)^{-d/2} |\Sigma|^{-1} \exp \left\{ -\frac{1}{2} \mathbf{Z}^T \Sigma^{-1} \mathbf{Z} \right\} \\
 &\geq (2\pi)^{-d/2} |\Sigma|^{-1} \exp \left\{ -\frac{1}{2} (\mathbf{Z} - \mathbf{h})^T \Sigma^{-1} (\mathbf{Z} - \mathbf{h}) \right\} \frac{y_1}{y_2} \frac{f_2(x - u)}{f_2(x)} \\
 &\Leftrightarrow \mathbf{Z}^T \Sigma^{-1} \mathbf{Z} \\
 &< \mathbf{Z}^T \Sigma^{-1} \mathbf{Z} - 2 \mathbf{Z}^T \Sigma^{-1} \mathbf{h} + \mathbf{h}^T \Sigma^{-1} \mathbf{h} - 2 \log \left(\frac{y_1}{y_2} \right) - 2 \log \left(\frac{f_2(x - u)}{f_2(x)} \right) \\
 &\Leftrightarrow \mathbf{Z}^T \Sigma^{-1} \mathbf{h} \leq \frac{1}{2} \mathbf{h}^T \Sigma^{-1} \mathbf{h} - \log \left(\frac{y_1}{y_2} \right) - \log \left(\frac{f_2(x - u)}{f_2(x)} \right).
 \end{aligned}$$

The random variable $\mathbf{Z}^T \Sigma^{-1} \mathbf{h} =: \mathbf{Z}^T \mathbf{B}$ is normally distributed with mean $\mathbf{0}$ and variance

$$\mathbf{B}^T \Sigma \mathbf{B} = \mathbf{h}^T \Sigma^{-1} \Sigma \Sigma^{-1} \mathbf{h} = \mathbf{h}^T \Sigma^{-1} \mathbf{h}.$$

Since f_2 is the density of a zero mean normal distribution with variance σ_3^2 , we obtain

$$\log\left(\frac{f_2(x-u)}{f_2(x)}\right) = -\frac{1}{2\sigma_3^2}(u^2 - 2ux).$$

With $a(\mathbf{h}) = (\mathbf{h}^T \Sigma^{-1} \mathbf{h})^{1/2}$ and $\mathbf{Z}^T \Sigma^{-1} \mathbf{h} / a(\mathbf{h}) \sim \mathcal{N}(0, 1)$ it follows

$$\begin{aligned} & P\left(\mathbf{Z}^T \Sigma^{-1} \mathbf{h} \leq \frac{1}{2} \mathbf{h}^T \Sigma^{-1} \mathbf{h}^T - \log\left(\frac{y_1}{y_2}\right) - \log\left(\frac{f_2(x-u)}{f_2(x)}\right)\right) \\ &= P\left(\frac{\mathbf{Z}^T \Sigma^{-1} \mathbf{h}}{a(\mathbf{h})} \leq \frac{a(\mathbf{h})}{2} - \frac{\log(y_1/y_2)}{a(\mathbf{h})} + \frac{1}{2\sigma_3^2 a(\mathbf{h})}(u^2 - 2ux)\right) \\ &= \Phi\left(\frac{a(\mathbf{h})}{2} + \frac{\log(y_2/y_1)}{a(\mathbf{h})} + \frac{u^2 - 2ux}{2\sigma_3^2 a(\mathbf{h})}\right). \end{aligned}$$

Altogether, for independent random variables N and X with N standard normally distributed and X normally distributed with mean 0 and variance σ_3^2 , it holds

$$\begin{aligned} \text{(I)} &= \frac{1}{y_1} \int_{-\infty}^{\infty} f_2(x) \Phi\left(\frac{a(\mathbf{h})}{2} + \frac{\log(y_2/y_1)}{a(\mathbf{h})} + \frac{u^2}{2\sigma_3^2 a(\mathbf{h})} - \frac{u}{\sigma_3^2 a(\mathbf{h})} x\right) dx \\ &= \frac{1}{y_1} P\left(N + \frac{u}{\sigma_3 a(\mathbf{h})} \frac{X}{\sigma_3} \leq \frac{a(\mathbf{h})}{2} + \frac{\log(y_2/y_1)}{a(\mathbf{h})} + \frac{u^2}{2\sigma_3^2 a(\mathbf{h})}\right) \\ &= \frac{1}{y_1} \Phi\left(\frac{\frac{a(\mathbf{h})}{2} + \frac{\log(y_2/y_1)}{a(\mathbf{h})} + \frac{u^2}{2\sigma_3^2 a(\mathbf{h})}}{\sqrt{1 + \frac{u^2}{\sigma_3^2 a(\mathbf{h})^2}}}\right) = \frac{1}{y_1} \Phi\left(\frac{2\sigma_3^2 \log(y_2/y_1) + \sigma_3^2 a(\mathbf{h})^2 + u^2}{2\sigma_3 \sqrt{\sigma_3^2 a(\mathbf{h})^2 + u^2}}\right), \end{aligned}$$

since $N + u/(\sigma_3 a(\mathbf{h}))(X/\sigma_3)$ is normally distributed with mean 0 and variance $1 + u^2/(\sigma_3^2 a(\mathbf{h})^2)$.

Analogously to (I), using the substitution $\mathbf{Z} \rightarrow \mathbf{Z} + \mathbf{h}$, we obtain the second term (II) in (3.15). \square

CHAPTER 4

COMPOSITE LIKELIHOOD METHODS FOR MAX-STABLE SPACE-TIME PROCESSES

The results in this chapter are based on the findings in Davis, Klüppelberg and Steinkohl [27]. We follow the approach described in Chapter 3, where the max-stable process introduced in Kabluchko et al. [50] is extended to a space-time setting.

As it is well-known for max-stable processes, the full likelihood function is computationally intractable and other methods have to be used to derive parameter estimates. Standard procedures for such cases are composite likelihood including pairwise likelihood estimation.

Since the observations in a space-time setting are correlated, we use special properties of max-stable processes to show strong consistency and asymptotic normality of the estimates. First, it is assumed that the locations lie on a regular lattice and that the time points are equidistant. The spatial and the temporal dimension, i.e., the number of spatial locations and time points, increases to infinity. The main step in the proof of strong consistency is the derivation of a strong law of large numbers for the pairwise likelihood function. Stoev [81] analysed ergodic properties for max-stable processes in time resulting from extremal integral representations for max-stable processes that were introduced in Stoev and Taqqu [82]. The extension to a spatial setting and the resulting strong law of large numbers was shown by Wang, Roy and Stoev [91]. By combining these two results we obtain a strong law of large

numbers for a jointly increasing space-time domain.

In addition to strong consistency, we prove asymptotic normality for the pairwise likelihood estimates. A first result concerning asymptotic normality of pairwise likelihood estimates for max-stable space-time processes can be found Huser and Davison [46], who fix the number of locations and let the number of time points tend to infinity. We formulate asymptotic normality for the space-time setting and use Bolthausen's theorem [11] together with strong mixing properties, shown by Dombry and Eyi-Minko [34], to prove asymptotic normality for a jointly increasing number of space-time locations.

The chapter is organized as follows. In Section 4.1 we recall the definition of the max-stable space-time process for which inference properties will be considered in subsequent sections. Section 4.2 describes pairwise likelihood estimation and the particular setting for our model. In Sections 4.3 and 4.4 we prove strong consistency and asymptotic normality, when locations lie on a regular grid and for equidistant time points. In Section 4.5 we discuss two possible ways of redefining the set of spatial locations, which can be irregularly spaced, for which consistency and asymptotic normality of the pairwise likelihood estimates still hold.

4.1 Reminder: Description of model parameters

We shortly recall the definition of the model which was fully developed in Section 3.1. We calculate the bivariate density needed for the definition of the pairwise likelihood function. Let $\{Z_j(s, t), s \in \mathbb{R}^d, t \in [0, \infty)\}$, $j = 1, \dots, n$, be independent replications of a space-time Gaussian process with correlation function γ satisfying the following condition.

Condition 4.1. *The correlation function γ satisfies*

$$(\log n)(1 - \gamma(s_n \mathbf{h}, t_n u)) \rightarrow \delta(\mathbf{h}, u) = \theta_1 \|\mathbf{h}\|^{\alpha_1} + \theta_2 |u|^{\alpha_2} > 0, \text{ as } n \rightarrow \infty,$$

where $s_n = (\log n)^{-1/\alpha_1}$, $t_n = (\log n)^{-1/\alpha_2}$, $\alpha_1, \alpha_2 \in (0, 2]$ and $\theta_1, \theta_2 > 0$.

Let further $\{\xi_j, j \in \mathbb{N}\}$ denote points of a Poisson random measure on $[0, \infty)$ with intensity measure $\xi^{-2} d\xi$. By Theorem 3.1 the random fields $\{\eta_n(s, t), s \in \mathbb{R}^d, t \in [0, \infty)\}$, defined for $n \in \mathbb{N}$ by

$$\eta_n(s, t) = \bigvee_{j=1}^n -\frac{1}{\log(\Phi(Z_j(s_n s, t_n t)))}, \quad s \in \mathbb{R}^d, t \in [0, \infty), \quad (4.1)$$

converge weakly on the space of continuous functions on $\mathbb{R}^d \times [0, \infty)$ to the stationary Brown-Resnick process

$$\eta(\mathbf{s}, t) = \bigvee_{j=1}^{\infty} \xi_j \exp\{W_j(\mathbf{s}, t) - \delta(\mathbf{s}, t)\}, \quad (4.2)$$

where the deterministic function δ is defined in Condition 4.1 and $\{W_j(\mathbf{s}, t), \mathbf{s} \in \mathbb{R}^d, t \in [0, \infty)\}$, $j \in \mathbb{N}$ are independent replications of a Gaussian process with stationary increments, $W(\mathbf{0}, 0) = 0$, $\mathbb{E}(W(\mathbf{s}, t)) = 0$ and covariance function for $\mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d, t_1, t_2 \in [0, \infty)$

$$\text{Cov}(W(\mathbf{s}_1, t_1), W(\mathbf{s}_2, t_2)) = \delta(\mathbf{s}_1, t_1) + \delta(\mathbf{s}_2, t_2) - \delta(\mathbf{s}_1 - \mathbf{s}_2, t_1 - t_2).$$

The bivariate distribution function of η can be expressed in closed form and is based on a fundamental result by Hüsler and Reiss [47];

$$F(x_1, x_2) = \exp\left\{-\frac{1}{x_1} \Phi\left(\frac{\log \frac{x_2}{x_1}}{2\sqrt{\delta(\mathbf{h}, u)}} + \sqrt{\delta(\mathbf{h}, u)}\right) - \frac{1}{x_2} \Phi\left(\frac{\log \frac{x_1}{x_2}}{2\sqrt{\delta(\mathbf{h}, u)}} + \sqrt{\delta(\mathbf{h}, u)}\right)\right\},$$

where Φ denotes the standard normal distribution function.

From the bivariate distribution function we calculate the bivariate density. For later purposes we state the closed form expression in the following lemma. For simplicity we suppress the argument (x_1, x_2) .

Lemma 4.1. *Set $\delta := \delta(\mathbf{h}, u)$ and define for $x_1, x_2 > 0$*

$$q_1 := \frac{\log(x_2/x_1)}{2\sqrt{\delta}} + \sqrt{\delta} \quad q_2 := \frac{\log(x_1/x_2)}{2\sqrt{\delta}} + \sqrt{\delta}, \quad (4.3)$$

$$V := \frac{1}{x_1} \Phi(q_\psi^{(1)}) + \frac{1}{x_2} \Phi(q_\psi^{(2)}). \quad (4.4)$$

The partial derivatives of q_1 and q_2 are given by

$$\frac{\partial q_1}{\partial x_1} = -\frac{1}{2\sqrt{\delta}x_1}, \quad \frac{\partial q_1}{\partial x_2} = \frac{1}{2\sqrt{\delta}x_2}, \quad \frac{\partial q_2}{\partial x_2} = -\frac{1}{2\sqrt{\delta}x_2}, \quad \frac{\partial q_2}{\partial x_1} = \frac{1}{2\sqrt{\delta}x_1}.$$

The first and second order partial derivatives of V are given by

$$\begin{aligned} V_1 &= \frac{\partial V}{\partial x_1} = -\frac{1}{x_1^2} \Phi(q_1) - \frac{1}{2\sqrt{\delta}x_1^2} \varphi(q_1) + \frac{1}{2\sqrt{\delta}x_1x_2} \varphi(q_2), \\ V_2 &= \frac{\partial V}{\partial x_2} = -\frac{1}{x_2^2} \Phi(q_2) - \frac{1}{2\sqrt{\delta}x_2^2} \varphi(q_2) + \frac{1}{2\sqrt{\delta}x_1x_2} \varphi(q_\psi^{(1)}), \\ V_{12} &= \frac{\partial^2 V}{\partial x_1 \partial x_2} = -\frac{2\sqrt{\delta} - q_1}{4\delta x_1^2 x_2} \varphi(q_1) - \frac{2\sqrt{\delta} - q_2}{4\delta x_1 x_2^2} \varphi(q_2). \end{aligned}$$

Finally, the bivariate log-density is

$$\log f(x_1, x_2) = -V + \log(V_1 V_2 - V_{12}). \quad (4.5)$$

4.2 Pairwise likelihood estimation

In this section, we describe the pairwise likelihood procedure for estimating the parameters of the Brown-Resnick process (4.2), when the underlying correlation function satisfies Condition 4.1. Composite likelihood methods have been used, whenever the full likelihood is not available or intractable. We present the general definition of composite and pairwise likelihood functions for a space-time setting in Section 4.2.1. Afterwards, we rewrite the pairwise likelihood for regular grid observations.

4.2.1 Composite likelihood estimation for the space-time setting

Composite likelihood methods go back to Besag [8] and Lindsay [60] and there is vast literature available, from a theoretical and an applied point of view. For more information we refer to Varin [86], who presents an overview of existing models and inference including an extensive number of references. In the most general setting the composite log-likelihood function is given by

$$l_c(\boldsymbol{\psi}, \mathbf{x}) = \sum_{i=1}^q w_i \log f_{\boldsymbol{\psi}}(\mathbf{x} \in A_i),$$

where for $i = 1, \dots, p$ the sets A_i describe measurable events and the w_i are non-negative weights associated to the events. From this general form special composite likelihood func-

tions can be derived. We define the (*weighted*) *pairwise log-likelihood function* by

$$PL(\boldsymbol{\psi}; \mathbf{x}) = \sum_{i=1}^n \sum_{j=1}^n w_{i,j} \log f_{\boldsymbol{\psi}}(x_i, x_j), \quad (4.6)$$

where $\mathbf{x} = (x_1, \dots, x_n)$ is the data vector, $f_{\boldsymbol{\psi}}(x_i, x_j)$ is the density of the bivariate observations (x_i, x_j) , and the $w_{i,j}$ are weights which can be used for example to reduce the number of pairs included in the estimation. The parameter estimates are obtained by maximizing (4.6).

As noted in Cox and Reid [20], for dependent observations, estimates based on the composite likelihood need not be consistent or asymptotically normal. This is important for space-time applications, since all components may be highly dependent across space and time. We describe the pairwise likelihood estimation for observations from the Brown-Resnick process (4.2), where the underlying correlation function satisfies Condition 4.1. The resulting parameter vector is given by $\boldsymbol{\psi} = (\theta_1, \alpha_1, \theta_2, \alpha_2)$. The pairwise likelihood for a general setting with M locations s_1, \dots, s_M and T time points $0 \leq t_1 < \dots < t_T < \infty$ is given by

$$PL^{(M,T)}(\boldsymbol{\psi}) = \sum_{i=1}^{M-1} \sum_{j=i+1}^M \sum_{k=1}^{T-1} \sum_{l=k+1}^T w_{i,j}^{(M)} w_{k,l}^{(T)} \log f_{\boldsymbol{\psi}}(\eta(s_i, t_k), \eta(s_j, t_l)), \quad (4.7)$$

where $w_{i,j}^{(M)} \geq 0$ and $w_{k,l}^{(T)} \geq 0$ denote spatial and temporal weights, respectively, and $f_{\boldsymbol{\psi}}$ is the bivariate density in (4.5). Since it is expected that space-time pairs, which are far apart in space or in time, have only little influence on the dependence parameters to be estimated, we define the weights, such that in the estimation only pairs with a maximal spatio-temporal distance of (r, p) are included, i.e.,

$$w_{i,j}^{(M)} = \mathbb{1}_{\{\|s_i - s_j\| \leq r\}}, \quad w_{k,l}^{(T)} = \mathbb{1}_{\{|t_k - t_l| \leq p\}}, \quad (4.8)$$

where $\|\cdot\|$ denotes any arbitrary norm on \mathbb{R}^d . The pairwise likelihood estimates are given by

$$(\hat{\theta}_1, \hat{\alpha}_1, \hat{\theta}_2, \hat{\alpha}_2) = \arg \max_{(\theta_1, \alpha_1, \theta_2, \alpha_2)} PL^{(M,T)}(\theta_1, \alpha_1, \theta_2, \alpha_2). \quad (4.9)$$

Using the definition of the weights in (4.8), the log-likelihood function in (4.7) can be rewrit-

ten as

$$PL^{(M,T)}(\boldsymbol{\psi}) = \sum_{i=1}^{M-1} \sum_{\substack{j=i+1 \\ \|s_i - s_j\| \leq r}}^M \sum_{k=1}^{T-p} \sum_{l=k+1}^{\min\{k+p, T\}} \log f_{\boldsymbol{\psi}}(\eta(s_i, t_k), \eta(s_j, t_l)). \quad (4.10)$$

4.2.2 Pairwise likelihood estimation for regular grid observations

The proof of strong consistency and asymptotic normality in Sections 4.3 and 4.4 is based on the assumption that locations lie on a regular grid and that time points are equidistant. The following condition summarizes the sampling scheme.

Condition 4.2. *We assume that the locations lie on a regular d -dimensional lattice,*

$$S_m = \left\{ (i_1, \dots, i_d), i_1, \dots, i_d \in \{1, \dots, m\} \right\}.$$

Further assume that the time points are equidistant and given by the set $\{1, \dots, T\}$.

For later purposes, we rewrite the pairwise log-likelihood function under Condition 4.2 in the following way. Define \mathcal{H}_r as the set of all vectors with non-negative integer-valued components \mathbf{h} without the $\mathbf{0}$ -vector, which point to other sites in the set of locations within distance r , i.e.,

$$\mathcal{H}_r = \mathbb{N}^d \cap B(\mathbf{0}, r) \setminus \{\mathbf{0}\},$$

where $B(\mathbf{0}, r) = \{\mathbf{s} : \|\mathbf{s}\| < r\}$. Nott and Rydén [65] call this the *design mask*. We denote by $|\mathcal{H}_r|$ the cardinality of the set \mathcal{H}_r . In our application, we will use design masks according to the Euclidean distance; for example with $d = 2$ (cf. Figure 4.2.1),

$$\mathcal{H}_3 = \{(1, 0), (0, 1), (1, 1), (0, 2), (2, 0), (1, 2), (2, 1), (2, 2), (0, 3), (3, 0)\}.$$

Using Condition 4.2 and the design mask, the *pairwise log-likelihood function* in (4.9) can be rewritten as

$$\begin{aligned} PL^{(m,T)}(\boldsymbol{\psi}) &= \sum_{s \in S_m} \sum_{t=1}^T \sum_{\substack{\mathbf{h} \in \mathcal{H}_r \\ s + \mathbf{h} \in S_m}} \sum_{\substack{u=1 \\ t+u \leq T}}^p \log f_{\boldsymbol{\psi}}(\eta(s, t), \eta(s + \mathbf{h}, t + u)) \\ &= \sum_{s \in S_m} \sum_{t=1}^T g_{\boldsymbol{\psi}}(s, t; r, p) - \mathcal{R}^{(m,T)}(\boldsymbol{\psi}), \end{aligned} \quad (4.11)$$

where

$$g_{\psi}(s, t; r, p) = \sum_{\mathbf{h} \in \mathcal{H}_r} \sum_{u=1}^p \log f_{\psi}(\eta(s, t), \eta(s + \mathbf{h}, t + u)), \quad (4.12)$$

and $\mathcal{R}^{(m,T)}(\psi)$ is a boundary term, given by

$$\mathcal{R}^{(m,T)}(\psi) = \sum_{s \in S_m} \sum_{t=1}^T \sum_{\substack{\mathbf{h} \in \mathcal{H}_r \\ s + \mathbf{h} \notin S_m}} \sum_{u=1}^p \log f_{\psi}(\eta(s, t), \eta(s + \mathbf{h}, t + u)). \quad (4.13)$$

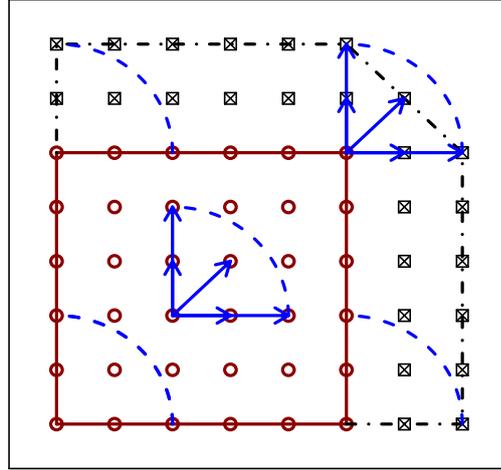


Figure 4.2.1: Visualization of the boundary term $\mathcal{R}^{(m,T)}$ for $d = 2$, $m = 6$, $r = 2$, and any fixed time point; the set S_m of locations is the inner square and the outer polygon represents the endpoints of pairs in the boundary.

Figure 4.2.1 depicts a spatial grid with side length $m = 6$, where the inner square is the set of observed locations S_m and the points in the outer polygon are endpoints of pairs which are in the boundary term $\mathcal{R}^{(m,T)}$. The figure visualizes the case \mathcal{H}_2 , which is represented by the quarter circles.

4.3 Strong consistency of the pairwise likelihood estimates for regular grid observations

In this section we establish strong consistency for the pairwise likelihood estimates based on regular grid observations introduced in Section 4.2.2. For univariate time series models Davis and Yau [24] proved strong consistency of the composite likelihood estimates in full detail. For max-stable random fields with replicates, which are independent in time, Padoan et al. [66] showed consistency and asymptotic normality of the pairwise likelihood estimates. In contrast to previous studies, where either the spatial or the time domain increases, we show strong consistency as the space-time domain increases jointly.

4.3.1 Ergodic properties for max-stable processes

Stoev and Taqqu [82] introduced extremal integrals as an analogy to sum-stable integrals. We briefly explain the notion of an extremal integral. The basis for the definition are α -Fréchet sup-measures. Given a measure space (E, \mathcal{E}, μ) with σ -finite, positive measure μ , the set-indexed random process $\{M_\alpha(A), A \in \mathcal{E}\}$ is called an *independently scattered α -Fréchet sup-measure with control measure μ* , if

1. for disjoint $A_1, \dots, A_n \in \mathcal{E}$, the random variables $M_\alpha(A_1), \dots, M_\alpha(A_n)$ are independent,

2. for $A \in \mathcal{E}$

$$P(M_\alpha(A) \leq x) = \exp\{-\mu(A)x^{-\alpha}\} \mathbb{1}_{\{x>0\}},$$

i.e. $M_\alpha(A)$ is α -Fréchet distributed with scale parameter $\mu(A)^{1/\alpha}$,

3. for disjoint $A_j \in \mathcal{E}, j \in \mathbb{N}$, with $\bigcup_{j \in \mathbb{N}} A_j \in \mathcal{E}$,

$$M_\alpha\left(\bigcup_{j \in \mathbb{N}} A_j\right) = \bigvee_{j \in \mathbb{N}} M_\alpha(A_j).$$

4.3 Strong consistency of the pairwise likelihood estimates for regular grid observations

For a non-negative simple function $f : E \rightarrow \mathbb{R}$, $f(x) = \sum_{j=1}^n a_j \mathbb{1}_{A_j}(x)$, where $A_1, \dots, A_n \in \mathcal{E}$ are disjoint, the *extremal integral* \int_E^e is defined by

$$\int_E^e f(x) M_\alpha(dx) := \bigvee_{j=1}^n a_j M_\alpha(A_j),$$

and the integral is independent of the representation of f . This definition can be extended stepwise from simple functions to nondecreasing sequences of simple functions and finally to any non-negative function $f : E \rightarrow \mathbb{R}$ satisfying $\int_E (f(x))^\alpha \mu(dx) < \infty$. Based on the extremal integral representation of max-stable processes Stoev [81] establishes conditions under which a max-stable process is ergodic. Wang et al. [91] extend these results to a spatial setting. In the following, let $\tau_{(h_1, \dots, h_d, u)}$ denote the multiparameter shift-operator. In accordance with the definitions and results in Wang et al. [91], we define ergodic and mixing space-time processes.

Definition 4.1. Let $\{\eta(s, t), s \in \mathbb{R}^d, t \in [0, \infty)\}$ be a strictly stationary space-time process. The process is called

1. ergodic, if for all $A, B \in \sigma\{\eta(s, t), s \in \mathbb{R}^d, t \in [0, \infty)\}$

$$\lim_{m_1, \dots, m_d, T \rightarrow \infty} \frac{1}{m_1 \cdots m_d T} \sum_{h_1=1}^{m_1} \cdots \sum_{h_d=1}^{m_d} \sum_{u=1}^T P(A \cap \tau_{(h_1, \dots, h_d, u)}(B)) = P(A)P(B), \quad (4.14)$$

where $m_1, \dots, m_d, T \rightarrow \infty$ means that each individual component of (m_1, \dots, m_k, T) tends to infinity.

2. mixing, if

$$\lim_{n \rightarrow \infty} P(A \cap \tau_{(s_{1,n}, \dots, s_{d,n}, t_n)}(B)) = P(A)P(B), \quad (4.15)$$

for all sequences $\{(s_{1,n}, \dots, s_{d,n}, t_n), n \in \mathbb{N}\}$ with $\max\{|s_{1,n}|, \dots, |s_{d,n}|, |t_n|\} \rightarrow \infty$.

Note in (4.14) that in contrast to the ergodic theorem in Wang et al. [91], the number of terms in each sum is not equal, since we have an additional sum for the time component. Using Theorem 6.1.2 in Krengel [53], we can relate the conventional definition of ergodicity to

the one given above. We focus on max-stable processes with extremal integral representation

$$\eta(s_1, \dots, s_d, t) = \int_E^e U_{(s_1, \dots, s_d, t)}(f) dM_1, \quad (4.16)$$

where $U_{(s_1, \dots, s_d, t)} : L^1(\mu) \rightarrow L^1(\mu)$ given by $U_{(s_1, \dots, s_d, t)}(f) = f \circ \tau_{(s_1, \dots, s_d, t)}$ is a group of max-linear automorphisms with $U_{(0, \dots, 0, 0)}(f) = f$, M_1 is an independently scattered 1–Fréchet random sup-measure with control measure μ , where (E, μ) can be chosen as the standard Lebesgue space (\mathbb{R}, λ) . The following result is a direct extension of the uniparameter theorem established in Stoev [81], Theorem 3.4, and its multiparameter counterpart:

Proposition 4.1 (Wang et al. [91], Theorem 5.3). *The max-stable process defined in (4.16) is mixing, if and only if*

$$\int_E U_{(s_{1,n}, \dots, s_{d,n}, t_n)}(f) \wedge U_{(0, \dots, 0, 0)}(f) d\mu = \int_E U_{(s_{1,n}, \dots, s_{d,n}, t_n)}(f) \wedge f d\mu \rightarrow 0, \quad (4.17)$$

for all sequences $\{(s_{1,n}, \dots, s_{d,n}, t_n)\}$ with $\max\{|s_{1,n}|, \dots, |s_{d,n}|, |t_n|\} \rightarrow \infty$ as $n \rightarrow \infty$.

Wang et al. [91] showed that the ergodic theorem stated above holds for mixing max-stable processes with extremal integral representation (4.16) in the case of $T = m$. The extension to the multiparameter case where $T \neq m$ is a simple generalization using Theorem 6.1.2 in Krenzel [53], which is a multiparameter extension of the Ackoglu’s ergodic theorem. Ergodic properties of Brown-Resnick processes have been studied for the uniparameter case in Stoev and Taqqu [82] and Wang and Stoev [90]. The Brown-Resnick process (4.2) has a stochastic representation

$$\left\{ \int_E^e \exp\{W(s, t) - \delta(s, t)\} dM_1, \quad s \in \mathbb{R}^d, t \in [0, \infty) \right\}, \quad (4.18)$$

where M_1 is a random 1-Fréchet sup-measure on the probability space (Ω, \mathcal{E}, P) on which the Gaussian process W is defined. The intensity is P , the probability measure which defines the Gaussian process W . We summarize the results in the following proposition.

Proposition 4.2. *If δ satisfies Condition 4.1, the Brown-Resnick process given above in (4.18) is mixing in space and time. The strong law of large numbers holds: for every measurable*

function $g : \mathbb{R} \rightarrow \mathbb{R}$ such that $\mathbb{E} [|g(\eta(s_1, 1))|] < \infty$;

$$\begin{aligned} \frac{1}{m^d T} \sum_{i_1=1}^m \cdots \sum_{i_d=1}^m \sum_{t=1}^T g(\eta((i_1, \dots, i_d), t)) &= \sum_{s \in S_m} \sum_{t=1}^T g(\eta(s, t)) \\ &\xrightarrow{a.s.} \mathbb{E} [g(\eta((1, \dots, 1), 1))] = \mathbb{E} [\eta(s_1, 1)] \quad m, T \rightarrow \infty. \end{aligned}$$

4.3.2 Consistency for large m and T

In the following we show that the pairwise likelihood estimate resulting from maximizing (4.11) for the Brown-Resnick process (4.2) is strongly consistent.

Theorem 4.1. *Assume that the correlation function γ satisfies Condition 4.1 with parameter vector $\boldsymbol{\psi} = (\theta_1, \alpha_1, \theta_2, \alpha_2)$. Suppose further that the true parameter vector $\boldsymbol{\psi}^* = (\theta_1^*, \alpha_1^*, \theta_2^*, \alpha_2^*)$ lies in a compact set Ψ , which does not contain $\mathbf{0}$ and which satisfies for some $c > 0$*

$$\Psi \subseteq \{\min\{\theta_1, \theta_2\} > c, \alpha_1, \alpha_2 \in (0, 2]\}. \quad (4.19)$$

Assume also that the identifiability condition

$$\boldsymbol{\psi} = \tilde{\boldsymbol{\psi}} \Leftrightarrow f_{\boldsymbol{\psi}}(\eta(s_1, t_1), \eta(s_2, t_2)) = f_{\tilde{\boldsymbol{\psi}}}(\eta(s_1, t_1), \eta(s_2, t_2)), \quad (4.20)$$

is satisfied for all $(s_1, t_1), (s_2, t_2)$. It then follows that the pairwise likelihood estimate

$$\hat{\boldsymbol{\psi}} = \arg \max_{\boldsymbol{\psi} \in \Psi} PL^{(m, T)}(\boldsymbol{\psi}) \quad (4.21)$$

for observations from the Brown-Resnick process (4.2) is strongly consistent, i.e. $\hat{\boldsymbol{\psi}} \xrightarrow{a.s.} \boldsymbol{\psi}^*$ as $m, T \rightarrow \infty$.

Remark 4.1. For the identifiability condition (4.20) we consider different cases according to the maximal space-time lag (r, p) included in the composite likelihood. The pairwise density depends on the spatial distance \mathbf{h} and the time lag u only through the function $\delta(\mathbf{h}, u) = \theta_1 \|\mathbf{h}\|^{\alpha_1} + \theta_2 |u|^{\alpha_2}$. For specific combinations of (r, p) not all parameters are identifiable. Strong consistency still holds for the remaining parameters. Table 4.1 lists the various scenarios.

Proof of Theorem 4.1. To show strong consistency of the estimates (4.21) we follow the

Maximal spatial lag r	Maximal temporal lag p	Identifiable parameters
0	1	θ_2
0	> 1	θ_2, α_2
1	0	θ_1
> 1	0	θ_1, α_1
1	1	θ_1, θ_2
1	> 1	$\theta_1, \theta_2, \alpha_2$
> 1	1	$\theta_1, \alpha_1, \theta_2$
> 1	> 1	$\theta_1, \alpha_1, \theta_2, \alpha_2$

Table 4.1: Identifiable parameters corresponding to different maximal space-time lags (r, p) included in the pairwise likelihood function.

method of Wald [89]. From (4.11) we show

$$\frac{1}{m^d T} PL^{(m,T)}(\boldsymbol{\psi}) = \frac{1}{m^d T} \left(\sum_{s \in \mathcal{S}_m} \sum_{t=1}^T g_{\boldsymbol{\psi}}(s, t; r, p) - \mathcal{R}^{(m,T)}(\boldsymbol{\psi}) \right) \xrightarrow{a.s.} PL(\boldsymbol{\psi}),$$

as $m, T \rightarrow \infty$, where $PL(\boldsymbol{\psi}) := \mathbb{E} [g_{\boldsymbol{\psi}}(s_1, 1; r, p)]$ and $g_{\boldsymbol{\psi}}$ and $\mathcal{R}^{(m,T)}(\boldsymbol{\psi})$ are defined in (4.12) and (4.13), respectively.

We use the following three steps.

(C1) Strong law of large numbers: Uniformly on the compact parameter space Ψ ,

$$\frac{1}{m^d T} \sum_{s \in \mathcal{S}_m} \sum_{t=1}^T g_{\boldsymbol{\psi}}(s, t; r, p) \xrightarrow{a.s.} PL(\boldsymbol{\psi}) = \mathbb{E} [g_{\boldsymbol{\psi}}(s_1, 1; r, p)], \quad m, T \rightarrow \infty,$$

(C2) $\frac{1}{m^d T} \mathcal{R}^{(m,T)}(\boldsymbol{\psi}) \xrightarrow{a.s.} 0, \quad m, T \rightarrow \infty.$

(C3) and the limit function $PL(\boldsymbol{\psi})$ in (C1) is uniquely maximized at the true parameter vector $\boldsymbol{\psi}^* \in \Psi$.

We first prove (C1). For fixed $\boldsymbol{\psi} \in \Psi$ the convergence in (C1) follows immediately from Proposition 4.2 together with the fact that $g_{\boldsymbol{\psi}}$ in (4.12) is a measurable function of lagged versions of $\eta(s, t)$. To prove uniform convergence we have from (3.6) for $x_1, x_2 > 0$

$$\begin{aligned} \log f_{\boldsymbol{\psi}}(x_1, x_2) &= -V(x_1, x_2) + \log(V_1(x_1, x_2)V_2(x_1, x_2) - V_{12}(x_1, x_2)), \\ V(x_1, x_2) &= \Phi(q_1)/x_1 + \Phi(q_2)/x_2, \end{aligned}$$

4.3 Strong consistency of the pairwise likelihood estimates for regular grid observations

$$V_1(x_1, x_2) = \frac{\partial V(x_1, x_2)}{\partial x_1}, \quad V_2(x_1, x_2) = \frac{\partial V(x_1, x_2)}{\partial x_2}, \quad V_{12}(x_1, x_2) = \frac{\partial^2 V(x_1, x_2)}{\partial x_1 \partial x_2},$$

and

$$q_1 = q_1(x_1, x_2) = \frac{\log(x_2/x_1)}{2\sqrt{\delta(\mathbf{h}, u)}} + \sqrt{\delta(\mathbf{h}, u)} \quad \text{and}$$

$$q_2 = q_2(x_1, x_2) = \frac{\log(x_1/x_2)}{2\sqrt{\delta(\mathbf{h}, u)}} + \sqrt{\delta(\mathbf{h}, u)}.$$

For $x_1, x_2 > 0$ the log-density $\log f_\psi(x_1, x_2)$ can be bounded as follows.

$$\begin{aligned} |\log f_\psi(x_1, x_2)| &= |-V(x_1, x_2) + \log(V_1(x_1, x_2)V_2(x_1, x_2) - V_{12}(x_1, x_2))| \\ &\leq |\Phi(q_1)/x_1| + |\Phi(q_2)/x_2| + |V_1(x_1, x_2)V_2(x_1, x_2) - V_{12}(x_1, x_2)| \\ &\leq \frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_1^2 x_2^2} + \frac{1}{2\sqrt{\delta(\mathbf{h}, u)}} \left(\frac{1}{x_1^2 x_2^2} + \frac{1}{x_1^3 x_2} + \frac{1}{x_1^2 x_2^2} + \frac{1}{x_1 x_2^3} + \frac{1}{x_1^2 x_2} + \frac{1}{x_1 x_2^2} \right) \\ &\quad + \frac{1}{4\delta(\mathbf{h}, u)} \left(\frac{1}{x_1^2 x_2^2} + \frac{1}{x_1^3 x_2} + \frac{1}{x_1 x_2^3} + \frac{1}{x_1^2 x_2} + \left| \frac{q_1}{x_1^2 x_2} + \frac{q_2}{x_1 x_2^2} \right| \right), \end{aligned}$$

where $\Phi(\cdot) \leq 1$ was used. Finally note that

$$\frac{q_1}{4\delta(\mathbf{h}, u)x_1^2 x_2} = \frac{\log(x_2/x_1) + 2\delta(\mathbf{h}, u)}{8(\delta(\mathbf{h}, u))^{3/2} x_1^2 x_2} \leq \frac{1}{8(\delta(\mathbf{h}, u))^{3/2} x_1^3} + \frac{1}{4\sqrt{\delta(\mathbf{h}, u)} x_1 x_2^2}.$$

Since the marginal distributions of the Brown-Resnick process (4.2) are assumed to be standard Fréchet, it follows that for every fixed location $s \in S_m$ and fixed time point $t \in \{1, \dots, T\}$ the random variable $1/\eta(s, t)$ is standard exponentially distributed with all moments finite. Using Hölder's inequality, it follows that

$$\mathbb{E} \left[\left| \log f_\psi(\eta(s_1, t_1), \eta(s_2, t_2)) \right| \right] \leq K_1 + \frac{K_2}{2\sqrt{\delta(\mathbf{h}, u)}} + \frac{K_3}{4\delta(\mathbf{h}, u)} + \frac{K_4}{8(\delta(\mathbf{h}, u))^{3/2}},$$

where $K_1, K_2, K_3, K_4 > 0$ are finite constants. Since the parameter space Ψ is assumed to be compact and together with assumption (4.19), δ can be bounded away from zero, i.e.

$$\delta(\mathbf{h}, u) \geq \min\{\theta_1, \theta_2\} (\|\mathbf{h}\|^{\alpha_1} + |u|^{\alpha_2}) > c(\|\mathbf{h}\|^{\alpha_1} + |u|^{\alpha_2}) > \tilde{c} > 0, \quad (4.22)$$

where $\tilde{c} > 0$ is independent of the parameters. Therefore,

$$\mathbb{E} \left[\left| \log f_{\psi}(\eta(s_1, t_1), \eta(s_2, t_2)) \right| \right] < K_1 + \frac{K_2}{2\sqrt{\tilde{c}}} + \frac{K_3}{4\tilde{c}} + \frac{K_4}{8\tilde{c}^{3/2}} =: K_5 < \infty, \quad (4.23)$$

where $K_5 > 0$. From (4.22) and (4.23) it follows that

$$\mathbb{E} \left[\sup_{\psi \in \Psi} \left| \log f_{\psi}(\eta(s_1, 1), \eta(s_1 + \mathbf{h}, 1 + u)) \right| \right] < \infty,$$

which implies $\mathbb{E} \left[\sup_{\psi \in \Psi} |g_{\psi}(s_1, 1, r, p)| \right] < \infty$. By Theorem 2.7 in Straumann [83] uniform convergence in (C1) follows.

Turning to (C2), note from (4.13) that by similar arguments as above

$$\begin{aligned} & \mathbb{E} \left[\left| \frac{1}{m^d T} \mathcal{R}^{(m, T)}(\psi) \right| \right] \\ & \leq \frac{1}{m^d T} \sum_{s \in S_m} \sum_{\substack{\mathbf{h} \in \mathcal{H}_r \\ s + \mathbf{h} \notin S_m}} \sum_{t=1}^T \sum_{\substack{u=1 \\ t+u > T}}^p \mathbb{E} \left[\left| \log f_{\psi}(\eta(s, t), \eta(s + \mathbf{h}, t + u)) \right| \right] \\ & \leq \frac{1}{m^d T} \sum_{s \in S_m} \sum_{\substack{\mathbf{h} \in \mathcal{H}_r \\ s + \mathbf{h} \notin S_m}} \sum_{t=1}^T \sum_{\substack{u=1 \\ t+u > T}}^p K_5 \leq \frac{K_5 K_6}{mT} \rightarrow 0, \quad m, T \rightarrow \infty, \end{aligned}$$

where we used the bound derived in (4.23) and the fact that the number of space-time points in the boundary is of order m^{d-1} (independent of T) and, therefore, can be bounded by $K_6 m^{d-1}$ with $K_6 > 0$ a constant independent of m and T .

We denote by $\mathcal{B}_{m, T}$ the set of ‘‘boundary’’ points, i.e.

$$\mathcal{B}_{m, T} = \{s \in S_m : s + \mathbf{h} \notin \mathcal{H}_r\} \times \{t \in \{1, \dots, T\} : t + u > T\}.$$

Then,

$$\mathcal{R}^{(m, T)}(\psi) = \sum_{\mathbf{h} \in \mathcal{H}_r} \sum_{u=1}^p \sum_{(s, t) \in \mathcal{B}_{m, T}} \log f_{\psi}(\eta(s, t), \eta(s + \mathbf{h}, t + u)).$$

By Proposition 4.2 and (4.23) it follows uniformly on Ψ , that

$$\sum_{\mathbf{h} \in \mathcal{H}_r} \sum_{u=1}^p \frac{1}{|\mathcal{B}_{m, T}|} \sum_{(s, t) \in \mathcal{B}_{m, T}} \log f_{\psi}(\eta(s, t), \eta(s + \mathbf{h}, t + u))$$

$$\rightarrow \mathbb{E} \left[\sum_{\mathbf{h} \in \mathcal{H}_r} \sum_{u=1}^p \log f_{\boldsymbol{\psi}}(\eta(\mathbf{s}_1, 1), \eta(\mathbf{s}_1 + \mathbf{h}, 1 + u)) \right], \quad m, T \rightarrow \infty.$$

Therefore,

$$\frac{1}{m^d T} \mathcal{R}^{(m, T)}(\boldsymbol{\psi}) \leq \frac{K_6}{mT} \sum_{\mathbf{h} \in \mathcal{H}_r} \sum_{u=1}^p \frac{1}{|\mathcal{B}_{m, T}|} \sum_{(s, t) \in \mathcal{B}_{m, T}} \log f_{\boldsymbol{\psi}}(\eta(\mathbf{s}, t), \eta(\mathbf{s} + \mathbf{h}, t + u)) \xrightarrow{a.s.} 0,$$

since $\mathbb{E} [|\log f_{\boldsymbol{\psi}}(\eta(\mathbf{s}, t), \eta(\mathbf{s} + \mathbf{h}, t + u))|] < \infty$. This proves (C2).

To prove (C3), note that by Jensen's inequality

$$\mathbb{E} \left[\log \left(\frac{f_{\boldsymbol{\psi}}(x_1, x_2)}{f_{\boldsymbol{\psi}^*}(x_1, x_2)} \right) \right] \leq \log \left(\mathbb{E} \left[\frac{f_{\boldsymbol{\psi}}(x_1, x_2)}{f_{\boldsymbol{\psi}^*}(x_1, x_2)} \right] \right) = 0$$

and, hence,

$$PL(\boldsymbol{\psi}) \leq PL(\boldsymbol{\psi}^*)$$

for all $\boldsymbol{\psi} \in \Psi$. So, $\boldsymbol{\psi}^*$ maximizes $PL(\boldsymbol{\psi})$ and is the unique optimum if and only if there is equality in Jensen's inequality. However, this is precluded by (4.20). \square

4.4 Asymptotic normality of the pairwise likelihood estimates for regular grid observations

In order to prove asymptotic normality of the pairwise likelihood estimates resulting from maximizing (4.11) we need the following results for the pairwise log-density. The proofs can be found in Section 4.6.

Lemma 4.2. *Consider the Brown-Resnick process in (4.2), where the underlying correlation function satisfies Condition 4.1. Further assume that all conditions from Theorem 4.1 hold.*

(1) *The gradient of the bivariate log-density satisfies*

$$\mathbb{E} \left[\left| \nabla_{\boldsymbol{\psi}} \log f_{\boldsymbol{\psi}}(\eta(\mathbf{s}_1, t_1), \eta(\mathbf{s}_2, t_2)) \right|^3 \right] < \infty$$

(2) The Hessian of the pairwise log-density satisfies

$$\mathbb{E} \left[\sup_{\psi \in \Psi} \left| \nabla_{\psi}^2 \log f_{\psi}(\eta(s_1, t_1), \eta(s_2, t_2)) \right| \right] < \infty.$$

The absolute values of the vector in (1) and the matrix in (2) are perceived componentwise.

Assuming asymptotic normality of the pairwise score function $\nabla_{\psi} PL^{(m,T)}(\psi)$ it is relatively routine to show that the pairwise likelihood estimates are asymptotically normal. We formulate the first result.

Theorem 4.2. *Assume that the conditions of Theorem 4.1 hold. In addition, assume that a central limit theorem holds for the gradient of g_{ψ} defined in (4.12) in the following sense*

$$\frac{1}{m^{d/2} \sqrt{T}} \sum_{s \in S_m} \sum_{t=1}^T \nabla_{\psi} g_{\psi^*}(s, t; r, p) \xrightarrow{d} \mathcal{N}(0, \Sigma), \quad m, T \rightarrow \infty, \quad (4.24)$$

where ψ^* is the true parameter vector and Σ is some covariance matrix. Then it follows that the pairwise likelihood estimate in (4.21) satisfies

$$m^{d/2} \sqrt{T}(\hat{\psi} - \psi^*) \xrightarrow{d} \mathcal{N}(0, F^{-1} \Sigma (F^{-1})^T), \quad m, T \rightarrow \infty,$$

where

$$F = \mathbb{E} \left[-\nabla_{\psi}^2 g_{\psi^*}(s_1, 1; r, p) \right].$$

Proof. We use a standard Taylor expansion of the pairwise score function around the true parameter vector and obtain

$$\begin{aligned} m^{d/2} \sqrt{T}(\hat{\psi} - \psi^*) &= - \left(\frac{1}{m^{dT}} \nabla_{\psi}^2 PL^{(m,T)}(\tilde{\psi}) \right)^{-1} \left(\frac{1}{m^{d/2} \sqrt{T}} \nabla_{\psi} PL^{(m,T)}(\psi^*) \right) \\ &= - \left(\frac{1}{m^{dT}} \sum_{s \in S_m} \sum_{t=1}^T \nabla_{\psi}^2 g_{\tilde{\psi}}(s, t; r, p) - \frac{1}{m^{dT}} \nabla_{\psi}^2 \mathcal{R}^{(m,T)}(\tilde{\psi}) \right)^{-1} \\ &\quad \times \left(\frac{1}{m^{d/2} \sqrt{T}} \sum_{s \in S_m} \sum_{t=1}^T \nabla_{\psi} g_{\psi^*}(s, t; r, p) - \frac{1}{m^{d/2} \sqrt{T}} \nabla_{\psi} \mathcal{R}^{(m,T)}(\psi^*) \right) \\ &= -(I_1 - I_2)^{-1} (J_1 - J_2), \end{aligned}$$

where $\tilde{\psi} \in [\hat{\psi}, \psi^*]$. By (4.24) J_1 converges weakly to a normal distribution with mean 0 and covariance matrix Σ . By using the same arguments as in the proof of Theorem 4.1 together with (4.24) we have that $J_2 \xrightarrow{P} 0$. Since the underlying space-time process in the likelihood function is mixing, it follows that the process $\{\nabla_{\psi}^2 g_{\psi}(s, t; r, p), s \in \mathbb{Z}^d, t \in \mathbb{N}\}$ is mixing as a measurable function of mixing and lagged processes. To prove the uniform convergence we verify that

$$\mathbb{E} \left[\sup_{\psi \in \Psi} \left| \nabla_{\psi}^2 g_{\psi}(s_1, 1; r, p) \right| \right] < \infty.$$

This follows immediately from Lemma 4.2. Putting this together with the fact that $\tilde{\psi} \in [\hat{\psi}, \psi^*]$, and because of the strong consistency of $\hat{\psi}$, it follows that

$$I_1 \xrightarrow{a.s.} \mathbb{E} \left[\nabla_{\psi}^2 g_{\psi^*}(s_1, 1; r, p) \right] =: -F.$$

Using the strong law of large numbers for $\{\nabla_{\psi}^2 \log f_{\psi}(\eta(s, t), \eta(s + \mathbf{h}, t + u))\}$ it follows in the same way as in the proof of Theorem 4.1 that $I_2 \xrightarrow{a.s.} 0$ as $m, T \rightarrow \infty$. Combining these results, we obtain by Slutsky's lemma

$$m^{d/2} \sqrt{T}(\hat{\psi} - \psi^*) \xrightarrow{d} \mathcal{N}(0, F^{-1}\Sigma(F^{-1})^T), \quad m, T \rightarrow \infty.$$

□

In the next section we provide a sufficient condition for (4.24).

4.4.1 Asymptotic normality and α -mixing

In this section we consider asymptotic normality of the parameters estimates for the Brown-Resnick process in (4.2). Under the assumption of α -mixing of the random field the key is to show asymptotic normality for the pairwise score function. For an increasing time domain and fixed number of locations asymptotic normality of the pairwise likelihood estimates was shown in Huser and Davison [46]. The main difference between a temporal setting and a space-time setting is the definition of the α -mixing coefficients and the resulting assumptions to obtain a central limit theorem for the score function.

We apply the central limit theorem for random fields established in Bolthausen [11] to the pairwise score function of the pairwise likelihood in our model. In a second step we verify the α -mixing conditions for the Brown-Resnick process (4.2), where the underlying correlation

function satisfies Condition 4.1. First, we define the α -mixing coefficients in a space-time setting as follows. Define the distances

$$d((s_1, t_1), (s_2, t_2)) = \max \left\{ \max_{1 \leq i \leq d} |s_1(i) - s_2(i)|, |t_1 - t_2| \right\}, \quad s_1, s_2 \in \mathbb{Z}^d, t_1, t_2 \in \mathbb{N},$$

$$d(\Lambda_1, \Lambda_2) = \inf \{d((s_1, t_1), (s_2, t_2)), (s_1, t_1) \in \Lambda_1, (s_2, t_2) \in \Lambda_2\},$$

where $\Lambda_1, \Lambda_2 \subset \mathbb{Z}^d \times \mathbb{N}$, and $s_k = (s_k(1), \dots, s_k(d))$, $k = 1, 2$. Let further $\mathcal{F}_{\Lambda_i} = \sigma\{\eta(s, t), (s, t) \in \Lambda_i\}$ for $i = 1, 2$. The mixing coefficients are defined for $k, l \in \mathbb{N}$ and $n \in \mathbb{N} \cup \{\infty\}$ by

$$\alpha_{k,l}(n) = \sup \left\{ |P(A_1 \cap A_2) - P(A_1)P(A_2)| : A_i \in \mathcal{F}_{\Lambda_i}, |\Lambda_1| \leq k, |\Lambda_2| \leq l, d(\Lambda_1, \Lambda_2) \geq n \right\} \quad (4.25)$$

and depend on the sizes and the distance of the sets Λ_1 and Λ_2 .

A space-time process is called α -mixing, if $\alpha_{k,l}(n) \rightarrow 0$ as $n \rightarrow \infty$ for all $k, l > 0$. We assume that the process $\{\eta(s, t), s \in \mathbb{Z}^d, t \in \mathbb{N}\}$ is α -mixing with mixing coefficients defined in (4.25), from which it follows that the space-time process

$$\{\nabla_{\psi} g_{\psi}(s, t; r, p), s \in \mathbb{Z}^d, t \in \mathbb{N}\} \quad (4.26)$$

is α -mixing for all $\psi \in \Psi$. We apply Bolthausen's central limit theorem this process. By adjusting the assumptions on the α -mixing coefficients we obtain the following proposition.

Proposition 4.3. *We consider the Brown-Resnick process (4.2) with $\delta(\mathbf{h}, u) = \theta_1 \|\mathbf{h}\|^{\alpha_1} + \theta_2 |u|^{\alpha_2}$. Assume, that the following conditions hold:*

(1) *The process $\{(\eta(s, t), s \in \mathbb{Z}^d, t \in \mathbb{N})\}$ is α -mixing.*

(2) *The α -mixing coefficients in (4.25) satisfy*

$$\sum_{n=1}^{\infty} n^d \alpha_{k,l}(n) < \infty \text{ for } k+l \leq 4(|\mathcal{H}_r|+1)(p+1) \text{ and } \alpha_{(|\mathcal{H}_r|+1)(p+1), \infty}(n) = o(n^{-(d+1)}).$$

(3) *There exists some $\beta > 0$ such that*

$$\mathbb{E} \left[\left| \nabla_{\psi} g_{\psi^*}(s, t; r, p) \right|^{2+\beta} \right] < \infty \text{ and}$$

$$\sum_{n=1}^{\infty} n^d \alpha_{(|\mathcal{H}_r|+1)(p+1), (|\mathcal{H}_r|+1)(p+1)}(n)^{\beta/(2+\beta)} < \infty.$$

4.4 Asymptotic normality of the pairwise likelihood estimates for regular grid observations

Then,

$$\frac{1}{m^{d/2} \sqrt{T}} \sum_{s \in \mathcal{S}_m} \sum_{t=1}^T \nabla_{\psi} g_{\psi^*}(s, t; r, p) \xrightarrow{d} \mathcal{N}(0, \Sigma), \quad m, T \rightarrow \infty,$$

where $\Sigma = \sum_{s \in \mathbb{Z}^d} \sum_{t \in \mathbb{N}} \text{Cov}(\nabla_{\psi} g_{\psi^*}(s_1, 1; r, p), \nabla_{\psi} g_{\psi^*}(s, t; r, p))$.

Recent work by Dombry and Eyi-Minko [34] deals with strong mixing properties for max-stable random fields. By using a point process representation of max-stable processes together with coupling techniques, they show that the α -mixing coefficients can be bounded by a function of the tail dependence coefficient. A direct extension to the space-time setting gives the following lemma.

Lemma 4.3 (Dombry and Eyi-Minko [34], Corollary 2.2). *Consider a stationary max-stable space-time process*

$\{\eta(s, t), s \in \mathbb{Z}^d, t \in \mathbb{N}\}$ *with arbitrary tail dependence coefficient* $\chi(\mathbf{h}, u)$. *The α -mixing coefficients (4.25) satisfy*

$$\alpha_{k,l}(n) \leq kl \sup_{\max\{\|\mathbf{h}\|, |u|\} \geq n} \chi(\mathbf{h}, u) \quad \text{and} \quad \alpha_{k,\infty}(n) \leq k \sum_{\max\{\|\mathbf{h}\|, |u|\} \geq n} \chi(\mathbf{h}, u).$$

In the following we show that Proposition 4.3 applies for the Brown-Resnick process (4.2) with tail dependence coefficient χ given by

$$\chi(\mathbf{h}, u) = 2(1 - \Phi(\sqrt{\theta_1 \|\mathbf{h}\|^{\alpha_1} + \theta_2 |u|^{\alpha_2}})).$$

By using the inequality for the normal tail probability $1 - \Phi(x) = \bar{\Phi}(x) \leq e^{-x^2/2}$ for $x > 0$ it follows that

$$\begin{aligned} \alpha_{k,l}(n) &\leq 4kl \sup_{\max\{\|\mathbf{h}\|, |u|\} \geq n} (1 - \Phi(\sqrt{\delta(\mathbf{h}, u)})) \leq 4kl \sup_{\max\{\|\mathbf{h}\|, |u|\} \geq n} \exp\left\{-\frac{\delta(\mathbf{h}, u)}{2}\right\} \\ &= 4kl \sup_{\max\{\|\mathbf{h}\|, |u|\} \geq n} \exp\left\{-\frac{1}{2}(\theta_1 \|\mathbf{h}\|^{\alpha_1} + \theta_2 |u|^{\alpha_2})\right\} \\ &\leq 4kl \sup_{\max\{\|\mathbf{h}\|, |u|\} \geq n} \exp\left\{-\frac{1}{2} \min\{\theta_1, \theta_2\} (\max\{\|\mathbf{h}\|, |u|\})^{\min\{\alpha_1, \alpha_2\}}\right\}. \end{aligned}$$

For $n \rightarrow \infty$, the right hand side tends to zero for all $k, l \geq 0$. Thus, $\{\eta(s, t), s \in \mathbb{Z}^d, t \in \mathbb{N}\}$ is

α -mixing. Furthermore, for $k + l \leq 4(|\mathcal{H}_r| + 1)(p + 1)$ the coefficients satisfy

$$\begin{aligned} \sum_{n=1}^{\infty} n^d \alpha_{k,l}(n) &\leq 4kl \sum_{n=1}^{\infty} n^d \sup_{\max\{\|\mathbf{h}\|, |u|\} \geq n} \exp \left\{ -\frac{1}{2} (\theta_1 \|\mathbf{h}\|^{\alpha_1} + \theta_2 |u|^{\alpha_2}) \right\} \\ &\leq 4kl \sum_{n=1}^{\infty} n^d \exp \left\{ -\frac{1}{2} \min\{\theta_1, \theta_2\} n^{\min\{\alpha_1, \alpha_2\}} \right\} < \infty. \end{aligned}$$

In addition,

$$n^{d+1} \alpha_{(|\mathcal{H}_r|+1)(p+1), \infty}(n) \leq n^{d+1} (|\mathcal{H}_r| + 1)(p + 1) \sum_{x \geq n} \exp \left\{ -\frac{1}{2} \min\{\theta_1, \theta_2\} x^{\min\{\alpha_1, \alpha_2\}} \right\},$$

where the right hand side converges to zero as $n \rightarrow \infty$, which finally proves (2). As for (3), from Lemma 4.2 and using $\beta = 1$ we know that

$$\mathbb{E} \left[\left| \nabla_{\psi} g_{\psi^*}(s, t; r, p) \right|^{(2+\beta)} \right] < \infty.$$

By the same arguments as in the proof of (2) above the second condition in (3) holds.

By combining the above results with Theorem 4.2 we obtain asymptotic normality for the parameter estimates $\hat{\psi}$ for an increasing number of space-time locations. We summarize this result as follows.

Theorem 4.3. *Assume that the conditions of Theorem 4.1 hold. Then,*

$$(m^d T)^{1/2} (\hat{\psi} - \psi^*) \xrightarrow{d} \mathcal{N}(0, F^{-1} \Sigma (F^{-1})^\top), \quad m, T \rightarrow \infty,$$

where

$$F = \mathbb{E} \left[-\nabla_{\psi}^2 g_{\psi^*}(s_1, 1; r, p) \right],$$

and

$$\Sigma = \sum_{s \in \mathbb{Z}^d} \sum_{t \in \mathbb{N}} \text{Cov} \left(\nabla_{\psi} g_{\psi^*}(s_1, 1; r, p), \nabla_{\psi} g_{\psi^*}(s, t; r, p) \right).$$

Remark 4.2. Unfortunately, we cannot provide a closed form expression for the asymptotic covariance matrix. The matrix F is the expected Hessian matrix of the pairwise log-likelihood

function and an estimate is given by its empirical version

$$\hat{F} = - \sum_{s \in \mathcal{S}_m} \sum_{t=1}^T \sum_{\substack{h \in \mathcal{H}_r \\ s+h \in \mathcal{S}_m}} \sum_{\substack{u=1 \\ t+u \leq T}}^p \nabla_{\psi}^2 \log f_{\hat{\psi}}(\eta(s, t), \eta(s + \mathbf{h}, t + u)),$$

which can be obtained numerically from the optimization routine used to maximize the pairwise likelihood function. The calculation of Σ or estimates for Σ seems to be a difficult task. We therefore rely on resampling methods like bootstrap or jackknife for obtaining estimates of the variance and confidence regions. For example a block bootstrap procedure could be applied which approximates the distribution of $\hat{\psi} - \psi$. The situation here is similar to the estimation of the extremogram, where bootstrap methods have been suggested to construct asymptotically correct confidence bands (see Davis and Mikosch [23] and Davis et al. [25]). The justification of resampling methods is the subject of another study.

4.5 Extension to irregularly spaced locations

So far we have assumed that the spatial sampling is a regular grid. In the following we discuss two settings, where the locations are irregularly spaced.

4.5.1 Deterministic irregularly spaced lattice

One way to extend our results to irregularly spaced locations is to invoke the ideas in Bai et al. [4] as adapted from Jenish and Prucha [48]. Let

$$D \subset \mathbb{R}^d \times [0, \infty) \times \mathbb{R}^d \times [0, \infty)$$

denote an infinitely countable lattice such that all elements of D have distances of at least $d_0 > 0$:

$$\|(s_1, t_1, s_2, t_2) - (s_3, t_3, s_4, t_4)\| > d_0$$

for any $(s_1, t_1, s_2, t_2), (s_3, t_3, s_4, t_4) \in D$, where $\|\cdot\|$ is an arbitrary norm. Note that D describes pairs of space-time locations. Further let $\{D_n : n \in \mathbb{N}\}$ be a sequence of arbitrary finite subsets of D satisfying $|D_n| \rightarrow \infty$ as $n \rightarrow \infty$, where $|\cdot|$ denotes the cardinality. In addition the sets D_n contain only pairs of space-time locations for which $\|s_1 - s_2\| \leq r$, $|t_1 - t_2| \leq p$ and at least one of the lags $\|s_1 - s_2\|$ and $|t_1 - t_2|$ is larger than zero. The pairwise

log-likelihood function (see general definition in (4.10)) is now given by

$$PL^{(n)}(\boldsymbol{\psi}) = \sum_{(s_1, t_1, s_2, t_2) \in D_n} \log f_{\boldsymbol{\psi}}(\eta(s_1, t_1), \eta(s_2, t_2)).$$

Denote by $\mathcal{S} \times \mathcal{T}$ the sampling region with cardinality $|\mathcal{S} \times \mathcal{T}| = n$. To prove consistency of the pairwise likelihood estimates Theorems 2 and 3 in Jenish and Prucha [48] are used to show that the pairwise log-likelihood function satisfies a law of large numbers. Using the same arguments as in Theorem 4.1 in Section 4.3.2 we can show that the estimates are consistent, i.e.

$$\hat{\boldsymbol{\psi}} \xrightarrow{P} \boldsymbol{\psi}^*, \quad n \rightarrow \infty.$$

Compared to the conditions needed to prove Theorem 4.1 the stronger assumption that the pairwise log-density is uniformly $L_{1+\delta}$ integrable (for a definition see Section 3.1 in Bai et al. [4]) has to be shown. For the Brown-Resnick process (4.2) and the assumptions in Theorem 4.1 this can be verified in a similar fashion as in the derivation of the upper bound for the log-density in the proof of Theorem 4.1, (C1). To show asymptotic normality of the estimates, Bai et al. [4] use Theorem 1 in Jenish and Prucha [48] assuming eight conditions, where the first two define the setting for the space-time locations. For the Brown-Resnick process with $\delta(\mathbf{h}, u) = \theta_1 \|\mathbf{h}\|^{\alpha_1} + \theta_2 |u|^{\alpha_2}$ and together with the assumptions in Theorem 4.1 all conditions except their Assumptions (7) and (8) can be shown. For our setting, Assumptions (7) and (8) in [4] are equivalent to

$$n \text{Var}(\nabla_{\boldsymbol{\psi}} PL^n(\boldsymbol{\psi})) \rightarrow \Sigma \quad \text{and} \quad \mathbb{E} \left[\nabla_{\boldsymbol{\psi}}^2 PL^n(\boldsymbol{\psi}) \right] \rightarrow F, \quad n \rightarrow \infty, \quad (4.27)$$

where F and Σ are positive definite matrices. Note that by using Theorem 2 and 3 in Jenish and Prucha [48] together with the arguments in the proof of Theorem 4.2 we can show the first part of Assumption 8 in Bai et al. [4]:

$$\sup_{\boldsymbol{\psi} \in \mathbb{Y}} \left| \nabla_{\boldsymbol{\psi}}^2 PL^{(n)}(\boldsymbol{\psi}) - \mathbb{E} \left[\nabla_{\boldsymbol{\psi}}^2 PL^{(n)}(\boldsymbol{\psi}) \right] \right| \rightarrow 0, n \rightarrow \infty.$$

Altogether, in contrast to the regular grid case we have two additional assumptions (4.27), which seem difficult to show.

4.5.2 Random locations generated by a Poisson process

In the following we assume that observations are taken at random locations. For simplicity we consider a spatial random field and no time component. We use the ideas and results in Karr [51] and Li et al. [59] to redefine the pairwise likelihood function and to show that the resulting estimates are asymptotically normal. Let $\{\eta(\mathbf{s}), \mathbf{s} \in \mathbb{R}^d\}$ be the max-stable random field defined analogously to (4.2), where δ is now given by $\delta(\mathbf{h}) = \theta_1 \|\mathbf{h}\|^{\alpha_1}$, and let N denote a Poisson random measure with mean measure $\nu\lambda(\cdot)$, where λ is Lebesgue measure, i.e. N is a stationary homogeneous Poisson process with intensity parameter ν which is assumed to be known. As before, we denote by S_m the set of possible spatial locations where the process is observed. Suppose that the set S_m is convex and compact. Following Karr [51] we define

$$N^{(2)}(ds_1, ds_2) = N(ds_1)N(ds_2)\mathbb{1}_{\{s_1 \neq s_2\}}, \quad s_1, s_2 \in S_m.$$

The pairwise log-likelihood function is now given by

$$PL^{(m)}(\boldsymbol{\psi}) = \int_{S_m} \int_{S_m} w(\mathbf{s}_1 - \mathbf{s}_2) \log f_{\boldsymbol{\psi}}(\eta(\mathbf{s}_1), \eta(\mathbf{s}_2)) N^{(2)}(ds_1, ds_2), \quad (4.28)$$

where w is some positive weight function. We adapt Lemma A.2 from Li et al. [59] to show that the pairwise score function satisfies a central limit theorem. The variance calculation is different from Li et al. [59] in the sense that we investigate the pairwise score function instead of a kernel smoothed estimator of a covariance function, which requires different arguments.

Lemma 4.4. *Assume that locations are generated by a stationary homogeneous Poisson process N with intensity ν . Suppose further that the following conditions hold.*

1. *The sets S_m satisfy*

$$\lambda(S_m) = O(m^d), \quad \text{and} \quad \lambda(\partial S_m) = O(m^{d-1}),$$

where λ denotes the Lebesgue measure and ∂S_m is the boundary of S_m .

2. *The random field $\{\eta(\mathbf{s}), \mathbf{s} \in \mathbb{R}^d\}$ is α -mixing with mixing coefficients as in (4.25) for which hold*

$$\sup_{k \in \mathbb{N}} \frac{1}{k^2} \alpha_{k,k}(r) = O(r^{-\epsilon}), \quad \text{for some } \epsilon > 0.$$

3. Let w be a positive weight function satisfying

$$\int_{\mathbb{R}^d} w(\mathbf{u}) d\mathbf{u} < \infty \quad \text{and} \quad \int_{\mathbb{R}^d} w(\mathbf{u})^2 d\mathbf{u} < \infty.$$

4. The gradient of the bivariate log-density satisfies

$$\mathbb{E} \left[|\nabla_{\boldsymbol{\psi}} \log f_{\boldsymbol{\psi}^*}(\boldsymbol{\eta}(\mathbf{s}_1), \boldsymbol{\eta}(\mathbf{s}_2)))| \right] < \infty \quad \text{and} \quad \mathbb{E} \left[|\nabla_{\boldsymbol{\psi}} \log f_{\boldsymbol{\psi}^*}(\boldsymbol{\eta}(\mathbf{s}_1), \boldsymbol{\eta}(\mathbf{s}_2)))|^2 \right] < \infty.$$

5. Define $S_m - S_m$ as the the set of all pairwise differences in S_m . Denote further

$$A_{\boldsymbol{\psi}^*}(\mathbf{s}_1, \mathbf{s}_2) = \nabla_{\boldsymbol{\psi}} \log f_{\boldsymbol{\psi}^*}(\boldsymbol{\eta}(\mathbf{s}_1), \boldsymbol{\eta}(\mathbf{s}_2)).$$

$$\begin{aligned} & \iiint_{(S_m - S_m)^3} w(\mathbf{v}_1) w(\mathbf{v}_3 - \mathbf{v}_2) \mathbb{E} \left[A_{\boldsymbol{\psi}^*}(\mathbf{v}_1, \mathbf{0}) A_{\boldsymbol{\psi}^*}(\mathbf{v}_2, \mathbf{v}_3) \right] \\ & \quad \times \frac{\lambda(S_m \cap (S_m + \mathbf{v}_1) \cap (S_m + \mathbf{v}_2) \cap (S_m + \mathbf{v}_3))}{\lambda(S_m)} d\mathbf{v}_1 d\mathbf{v}_2 d\mathbf{v}_3 \\ & \rightarrow \iiint_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} w(\mathbf{v}_1) w(\mathbf{v}_3 - \mathbf{v}_2) \mathbb{E} \left[A_{\boldsymbol{\psi}^*}(\mathbf{v}_1, \mathbf{0}) A_{\boldsymbol{\psi}^*}(\mathbf{v}_2, \mathbf{v}_3) \right] d\mathbf{v}_1 d\mathbf{v}_2 d\mathbf{v}_3, \quad m \rightarrow \infty. \end{aligned}$$

6. There exists $\beta > 0$, such that

$$\sup_{m>0} \mathbb{E} \left[|\sqrt{\lambda(S_m)} \nabla_{\boldsymbol{\psi}} PL^{(m)}(\boldsymbol{\psi}^*)|^{2+\beta} \right] < C_\beta$$

for some constant $C_\beta > 0$.

Then,

$$\frac{1}{\sqrt{\lambda(S_m)}} \nabla_{\boldsymbol{\psi}} PL^{(m)}(\boldsymbol{\psi}^*) \xrightarrow{d} \mathcal{N}(0, \Sigma), \quad m \rightarrow \infty,$$

where

$$\Sigma = \frac{2}{v^2} \int_{\mathbb{R}^d} w^2(\mathbf{v}) \mathbb{E} \left[A_{\boldsymbol{\psi}^*}^2(\mathbf{v}, \mathbf{0}) \right] d\mathbf{v} + \frac{4}{v} \text{Var} \left(\int_{\mathbb{R}^d} w(\mathbf{u}) A_{\boldsymbol{\psi}^*}(\mathbf{u}, \mathbf{0}) d\mathbf{u} \right)$$

$$+ \iiint_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} w(\mathbf{v}_1)w(\mathbf{v}_3 - \mathbf{v}_2)\mathbf{Cov} [A_{\psi^*}(\mathbf{v}_1, \mathbf{0}), A_{\psi^*}(\mathbf{v}_2, \mathbf{v}_3)] d\mathbf{v}_1 d\mathbf{v}_2 d\mathbf{v}_3. \quad (4.29)$$

Proof. We calculate the expectation and the variance of the pairwise score function $\nabla_{\psi} PL^{(m)}(\psi^*)$. By using standard properties of the Poisson process it follows that

$$\begin{aligned} \mathbb{E} [\nabla_{\psi} PL^{(m)}(\psi^*)] &= \mathbb{E} \left[\iint_{S_m \times S_m} w(\mathbf{s}_1 - \mathbf{s}_2) A_{\psi^*}(\mathbf{s}_1, \mathbf{s}_2) N^{(2)}(d\mathbf{s}_1, d\mathbf{s}_2) \right] \\ &= \nu^2 \iint_{S_m \times S_m} w(\mathbf{s}_1 - \mathbf{s}_2) \mathbb{E} [A_{\psi^*}(\mathbf{s}_1, \mathbf{s}_2)] d\mathbf{s}_1 d\mathbf{s}_2 = 0, \end{aligned}$$

To calculate the variance note that

$$\begin{aligned} &\lambda(S_m)^{-1} \mathbb{V}ar(\nu^{-2} \nabla_{\psi} PL^{(m)}(\psi^*)) \\ &= \lambda(S_m)^{-1} \nu^{-4} \iiint\limits_{S_m^4} w(\mathbf{s}_1 - \mathbf{s}_2) w(\mathbf{s}_3 - \mathbf{s}_4) \mathbb{E} [A_{\psi^*}(\mathbf{s}_1, \mathbf{s}_2) A_{\psi^*}(\mathbf{s}_3, \mathbf{s}_4)] \\ &\quad \times \mathbb{E} [N^{(2)}(d\mathbf{s}_1, d\mathbf{s}_2) N^{(2)}(d\mathbf{s}_3, d\mathbf{s}_4)]. \end{aligned}$$

The expectation $\mathbb{E} [N^{(2)}(d\mathbf{s}_1, d\mathbf{s}_2) N^{(2)}(d\mathbf{s}_3, d\mathbf{s}_4)]$ can be calculated by using standard properties of the Poisson process leading to seven terms as stated in Karr [51] or Li et al. [59]. In the limit, some of these terms are equal. We calculate the three representative different parts. We denote by $\epsilon_s(\cdot)$ the Dirac measure.

$$\begin{aligned} &\lambda(S_m)^{-1} \nu^{-4} \iiint\limits_{S_m^4} w(\mathbf{s}_1 - \mathbf{s}_2) w(\mathbf{s}_3 - \mathbf{s}_4) \mathbb{E} [A_{\psi^*}(\mathbf{s}_1, \mathbf{s}_2) A_{\psi^*}(\mathbf{s}_3, \mathbf{s}_4)] \\ &\quad \times \nu^2 d\mathbf{s}_1 d\mathbf{s}_2 \epsilon_{s_1}(d\mathbf{s}_3) \epsilon_{s_2}(d\mathbf{s}_4) \\ &= \lambda(S_m)^{-1} \nu^{-2} \iint\limits_{S_m^2} w^2(\mathbf{s}_1 - \mathbf{s}_2) \mathbb{E} [A_{\psi^*}^2(\mathbf{s}_1 - \mathbf{s}_2, \mathbf{0})] d\mathbf{s}_1 d\mathbf{s}_2 \\ &= \nu^{-2} \int\limits_{S_m - S_m} w^2(\mathbf{u}) \mathbb{E} [A_{\psi^*}^2(\mathbf{u}, \mathbf{0})] \frac{\lambda(S_m \cap (S_m + \mathbf{u}))}{\lambda(S_m)} d\mathbf{u} = (1) \end{aligned}$$

Since $\lambda(S_m \cap (S_m + \mathbf{u})) / \lambda(S_m) \rightarrow 1$ as $m \rightarrow \infty$ for every fixed $\mathbf{u} \in \mathbb{R}^d$ (see Lemma 3.2 in

Karr [51]), and due to the fact that

$$\begin{aligned} & \int_{S_m - S_m} w^2(\mathbf{u}) \mathbb{E} \left[A_{\psi^*}^2(\mathbf{u}, \mathbf{0}) \right] \frac{\lambda(S_m \cap (S_m + \mathbf{u}))}{\lambda(S_m)} d\mathbf{u} \leq \int_{S_m - S_m} w^2(\mathbf{u}) \mathbb{E} \left[A_{\psi^*}^2(\mathbf{u}, \mathbf{0}) \right] d\mathbf{u} \\ & \rightarrow \int_{\mathbb{R}^d} w^2(\mathbf{u}) \mathbb{E} \left[A_{\psi^*}^2(\mathbf{u}, \mathbf{0}) \right] d\mathbf{u} < \infty \end{aligned}$$

it follows by dominated convergence that (1) converges to

$$v^{-2} \int_{\mathbb{R}^d} \text{Var} \left(w(\mathbf{u}) A_{\psi^*}(\mathbf{u}, \mathbf{0}) \right) d\mathbf{u}.$$

Using similar arguments,

$$\begin{aligned} & \lambda(S_m)^{-1} v^{-4} \iiint_{S_m^4} w(\mathbf{s}_1 - \mathbf{s}_2) w(\mathbf{s}_3 - \mathbf{s}_4) \mathbb{E} \left[A_{\psi^*}(\mathbf{s}_1, \mathbf{s}_2) A_{\psi^*}(\mathbf{s}_3, \mathbf{s}_4) \right] v^3 d\mathbf{s}_1 d\mathbf{s}_2 \epsilon_{s_1}(d\mathbf{s}_3) d\mathbf{s}_4 \\ & = \lambda(S_m)^{-1} v^{-1} \iiint_{S_m^3} w(\mathbf{s}_1 - \mathbf{s}_2) w(\mathbf{s}_1 - \mathbf{s}_4) \mathbb{E} \left[A_{\psi^*}(\mathbf{0}, \mathbf{s}_2 - \mathbf{s}_1) A_{\psi^*}(\mathbf{0}, \mathbf{s}_4 - \mathbf{s}_1) \right] d\mathbf{s}_1 d\mathbf{s}_2 d\mathbf{s}_4 \\ & = v^{-1} \iint_{(S_m - S_m)^2} w(\mathbf{v}_1) w(\mathbf{v}_2) \mathbb{E} \left[A_{\psi^*}(\mathbf{v}_1, \mathbf{0}) A_{\psi^*}(\mathbf{v}_2, \mathbf{0}) \right] \frac{\lambda(S_m \cap (S_m + \mathbf{v}_1) \cap (S_m + \mathbf{v}_2))}{\lambda(S_m)} d\mathbf{v}_1 d\mathbf{v}_2 \\ & \rightarrow v^{-1} \text{Var} \left(\int_{\mathbb{R}^d} w(\mathbf{u}) A_{\psi^*}(\mathbf{u}, \mathbf{0}) d\mathbf{u} \right). \end{aligned}$$

For the last term we obtain

$$\begin{aligned} & \lambda(S_m)^{-1} v^{-4} \iiint_{S_m^4} w(\mathbf{s}_1 - \mathbf{s}_2) w(\mathbf{s}_3 - \mathbf{s}_4) \mathbb{E} \left[A_{\psi^*}(\mathbf{s}_1, \mathbf{s}_2) A_{\psi^*}(\mathbf{s}_3, \mathbf{s}_4) \right] v^4 d\mathbf{s}_1 d\mathbf{s}_2 d\mathbf{s}_3 d\mathbf{s}_4 \\ & = \iiint_{(S_m - S_m)^3} w(\mathbf{v}_1) w(\mathbf{v}_3 - \mathbf{v}_2) \mathbb{E} \left[A_{\psi^*}(\mathbf{v}_1, \mathbf{0}) A_{\psi^*}(\mathbf{v}_2, \mathbf{v}_3) \right] \\ & \quad \times \frac{\lambda(S_m \cap (S_m + \mathbf{v}_1) \cap (S_m + \mathbf{v}_2) \cap (S_m + \mathbf{v}_3))}{\lambda(S_m)} d\mathbf{v}_1 d\mathbf{v}_2 d\mathbf{v}_3. \end{aligned}$$

Altogether, as $m \rightarrow \infty$,

$$\begin{aligned} \text{Var}((\lambda(S_m))^{-1/2} \nabla_{\psi} PL^{(m)}(\psi^*)) &\rightarrow \Sigma \\ &= 2\nu^{-2} \int_{\mathbb{R}^d} \text{Var}(w(\mathbf{v})A_{\psi^*}(\mathbf{v}, \mathbf{0})) d\mathbf{v} + 4\nu^{-1} \text{Var} \left(\int_{\mathbb{R}^d} w(\mathbf{u})A_{\psi^*}(\mathbf{u}, \mathbf{0}) d\mathbf{u} \right) \\ &+ \iiint_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} \text{Cov} [w(\mathbf{v}_1)A_{\psi^*}(\mathbf{v}_1, \mathbf{0}), w(\mathbf{v}_3 - \mathbf{v}_2)A_{\psi^*}(\mathbf{v}_2, \mathbf{v}_3)] d\mathbf{v}_1 d\mathbf{v}_2 d\mathbf{v}_3. \end{aligned}$$

The central limit theorem for $\nabla_{\psi} PL^{(m)}(\psi^*)$ follows in exactly the same way as in Lemma A.2. together with Lemma A.4. in Li et al. [59]. \square

In a second step we show that the estimates resulting by maximizing the pairwise log-likelihood function in (4.28) are asymptotically normal.

Theorem 4.4. *In addition to the conditions in Lemma 4.4 assume that*

$$8. \mathbb{E} [|\nabla_{\psi}^2 \log f_{\psi^*}(\eta(s_1), \eta(s_2))|] < \infty \quad \text{and} \quad \mathbb{E} [|\nabla_{\psi}^2 \log f_{\psi^*}(\eta(s_1), \eta(s_2))|^2] < \infty, \text{ and}$$

$$9. \sup_{\psi \in \Psi_{\mathbb{R}^d}} \int w(\mathbf{u}) \mathbb{E} [\nabla_{\psi}^2 \log f_{\psi}(\eta(\mathbf{u}), \eta(\mathbf{0}))] d\mathbf{u} < \infty, \text{ and}$$

10. further, as $m \rightarrow \infty$ and for fixed $\psi \in \Psi$,

$$\begin{aligned} &\iiint_{(S_m - S_m)^3} w(\mathbf{v}_1)w(\mathbf{v}_3 - \mathbf{v}_2) \mathbb{E} [\nabla_{\psi} A_{\psi}(\mathbf{v}_1, \mathbf{0}) \nabla_{\psi} A_{\psi}(\mathbf{v}_2, \mathbf{v}_3)] \\ &\quad \times \frac{\lambda(S_m \cap (S_m + \mathbf{v}_1) \cap (S_m + \mathbf{v}_2) \cap (S_m + \mathbf{v}_3))}{\lambda(S_m)} d\mathbf{v}_1 d\mathbf{v}_2 d\mathbf{v}_3 \\ &\rightarrow \iiint_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} w(\mathbf{v}_1)w(\mathbf{v}_3 - \mathbf{v}_2) \mathbb{E} [\nabla_{\psi} A_{\psi}(\mathbf{v}_1, \mathbf{0}) \nabla_{\psi} A_{\psi}(\mathbf{v}_2, \mathbf{v}_3)] d\mathbf{v}_1 d\mathbf{v}_2 d\mathbf{v}_3. \end{aligned}$$

Then, the pairwise likelihood estimate $\hat{\psi}$ is asymptotically normal:

$$\sqrt{\lambda(S_m)}(\hat{\psi} - \psi^*) \xrightarrow{d} \mathcal{N}(0, F^{-1}\Sigma(F^{-1})^{\top}), \quad m \rightarrow \infty,$$

where Σ is defined in (4.29) and

$$F = \int_{\mathbb{R}^d} w(\mathbf{u}) \mathbb{E} \left[\nabla_{\psi}^2 \log f_{\psi^*}(\eta(\mathbf{u}), \eta(\mathbf{0})) \right] d\mathbf{u}.$$

Proof. For the second derivative of the pairwise log-likelihood function (4.28) we obtain for fixed $\psi \in \Psi$

$$\begin{aligned} \mathbb{E} \left[\lambda(S_m)^{-1} \nu^{-2} \nabla^2 PL^{(m)}(\psi) \right] &= \lambda(S_m)^{-1} \iint_{S_m \times S_m} w(s_1 - s_2) \mathbb{E} \left[\nabla_{\psi} A_{\psi}(s_1, s_2) \right] ds_1 ds_2 \\ &= \int_{(S_m - S_m)} w(\mathbf{u}) \mathbb{E} \left[\nabla_{\psi} A_{\psi}(s_1 - s_2, \mathbf{0}) \right] \frac{\lambda(S_m \cap (S_m + \mathbf{u}))}{\lambda(S_m)} d\mathbf{u} \\ &\rightarrow \int_{\mathbb{R}^d} w(\mathbf{u}) \mathbb{E} \left[\nabla_{\psi} A_{\psi}(\mathbf{u}, \mathbf{0}) \right] d\mathbf{u}, \quad m \rightarrow \infty. \end{aligned}$$

Using the same argument as for the pairwise score function it follows that

$\text{Var}(\lambda(S_m)^{-1} \nu^{-2} \nabla_{\psi}^2 PL^{(m)}(\psi)) \rightarrow 0$. This shows pointwise convergence of $\nabla_{\psi}^2 PL^{(m)}(\tilde{\psi})$ to $\int_{\mathbb{R}^d} w(\mathbf{u}) \mathbb{E} \left[\nabla_{\psi}^2 \log f_{\psi}(\eta(\mathbf{u}), \eta(\mathbf{0})) \right] d\mathbf{u}$. The uniform convergence is implied by Assumption (i). Therefore,

$$\lambda(S_m)^{-1} \nu^{-2} \nabla_{\psi}^2 PL^{(m)}(\psi) \rightarrow F.$$

Using a Taylor expansion

$$0 = \nabla_{\psi} PL^{(m)}(\hat{\psi}) = \frac{1}{\lambda(S_m) \nu^2} \nabla_{\psi} PL^{(m)}(\psi^*) + \left(\frac{1}{\lambda(S_m) \nu^2} \nabla_{\psi}^2 PL^{(m)}(\tilde{\psi}) \right) (\hat{\psi} - \psi^*),$$

we obtain

$$\sqrt{\lambda(S_m)} (\hat{\psi} - \psi^*) = - \left(\frac{1}{\lambda(S_m) \nu^2} \nabla_{\psi}^2 PL^{(m)}(\tilde{\psi}) \right)^{-1} \left(\frac{1}{\sqrt{\lambda(S_m)} \nu^2} \nabla_{\psi} PL^{(m)}(\psi^*) \right).$$

Together with the central limit theorem for the pairwise score function (Lemma 4.4) it follows that

$$\sqrt{\lambda(S_m)} (\hat{\psi} - \psi^*) \xrightarrow{d} \mathcal{N}(0, \Sigma), \quad m \rightarrow \infty.$$

□

Remark 4.3. First note that the rate of convergence here is $\sqrt{\lambda(S_m)} = O(m^{d/2})$ (see Assumption (a)) which is the same as for regular grids. For the max-stable random field in (4.2) satisfying Condition 4.1, Assumption (d) was shown in Lemma 4.2. Assumptions (f) and (h) can be shown in the same way as Lemma 4.2. The condition (b) on the α -mixing coefficients is easily verified using Lemma 4.3, from which follows that

$$\frac{1}{k^2} \alpha_{k,k}(r) \leq \exp\{-\theta_1 r_1^\alpha / 2\} \leq C r_1^{-\alpha},$$

where $C > 0$ is some constant.

4.6 Proof of Lemma 4.2

In the following, we use the same abbreviations as in the proof of Theorem 4.1. The gradient of the bivariate log-density with respect to the parameter vector $\boldsymbol{\psi}$ is given by

$$\nabla_{\boldsymbol{\psi}} \log f(x_1, x_2) = \frac{\partial \log f(x_1, x_2)}{\partial \delta} \nabla_{\boldsymbol{\psi}} \delta.$$

Assume in the following that all parameters $\theta_1, \alpha_1, \theta_2$ and α_2 are identifiable. Since all partial derivatives

$$\frac{\partial \delta}{\partial \theta_1} = \|\mathbf{h}\|^{\alpha_1}, \quad \frac{\partial \delta}{\partial \theta_2} = |u|^{\alpha_2}, \quad \frac{\partial \delta}{\partial \alpha_1} = \theta_1 \alpha_1 \|\mathbf{h}\|^{\alpha_1-1}, \quad \frac{\partial \delta}{\partial \alpha_2} = \theta_2 \alpha_2 |u|^{\alpha_2-1},$$

as well as all second order partial derivatives can be bounded from below and above for $0 < \min\{\|\mathbf{h}\|, |u|\}, \max\{\|\mathbf{h}\|, |u|\} < \infty$ using assumption (4.19) and, independently of the parameters $\theta_1, \theta_2, \alpha_1$ and α_2 , it suffices to show that

$$\mathbb{E}_{\boldsymbol{\psi}^*} \left[\left| \frac{\partial \log f_{\boldsymbol{\psi}}(\eta(s_1, t_1), \eta(s_2, t_2))}{\partial \delta} \right|^3 \right] < \infty$$

and

$$\mathbb{E}_{\boldsymbol{\psi}^*} \left[\sup_{\boldsymbol{\psi} \in \Psi} \left| \frac{\partial^2 \log f_{\boldsymbol{\psi}}(\eta(s_1, t_1), \eta(s_2, t_2))}{\partial \delta} \right| \right] < \infty.$$

Since δ can be bounded away from zero using assumption (4.19), we can treat δ as a constant. For simplification we drop the argument in the following equalities. Define

$$V_1 = \frac{\partial V}{\partial x_1}, \quad V_2 = \frac{\partial V}{\partial x_2}, \quad \text{and} \quad V_{12} = \frac{\partial^2 V}{\partial x_1 \partial x_2}.$$

The partial derivative of the bivariate log-density with respect to δ has the following form

$$\frac{\partial \log f_\psi}{\partial \delta} = -\frac{\partial V}{\partial \delta} + (V_1 V_2 - V_{12})^{-1} \left(\frac{\partial V_1}{\partial \delta} V_2 + V_1 \frac{\partial V_2}{\partial \delta} - \frac{\partial V_{12}}{\partial \delta} \right).$$

We identify stepwise the ‘‘critical’’ terms, where ‘‘critical’’ means higher order terms of functions of x_1 and x_2 . To give an idea on how to handle the components in the derivatives, we describe one such step. Note that $(V_1 V_2 - V_{12})^{-1}$ can be written as

$$(V_1 V_2 - V_{12})^{-1} = \frac{x_1 x_2}{g_1 \left(\frac{1}{x_1}, \frac{1}{x_2}, \frac{1}{x_1 x_2}, \frac{1}{x_1^2}, \frac{1}{x_2^2} \right)},$$

where g_1 describes the sum of the components together with additional multiplicative factors. By using

$$\frac{\partial \Phi(q_\psi^{(1)})}{\partial \delta} = \frac{q_\psi^{(1)}}{2\delta} \varphi(q_\psi^{(1)}) \quad \text{and} \quad \frac{\partial \varphi(q_\psi^{(1)})}{\partial \delta} = -\frac{(q_\psi^{(1)})^2}{2\delta} \varphi(q_\psi^{(1)}),$$

where $q_\psi^{(1)} = \log(x_2/x_1)/(2\sqrt{\delta}) + \sqrt{\delta}$, we have

$$\frac{\partial V_1}{\partial \delta} V_2 = g_2 \left(\frac{1}{x_1^2}, \frac{q_1}{x_2}, \frac{q_1^2}{x_1^2 x_2^2}, \frac{1}{x_1^3 x_2}, \frac{q_1}{x_1^3 x_2}, \frac{q_1^2}{x_1^3 x_2}, \frac{1}{x_1 x_2^3}, \frac{q_1}{x_1 x_2^3} \right),$$

where g_2 is a linear function of the components. By combining the two representations above, we obtain that all terms in

$(V_1 V_2 - V_{12})^{-1} (\partial V_1 / \partial \delta) V_2$ are of the form

$$\frac{|\log x_1|^{k_1} |\log x_2|^{k_2}}{x_1^{k_3} x_2^{k_4}}, \quad k_1, k_2, k_3, k_4 \geq 0. \quad (4.30)$$

The second derivative of the bivariate log-density with respect to δ is given by

$$\begin{aligned} \frac{\partial^2 \log f_\psi}{(\partial \delta)^2} &= -\frac{\partial^2 V}{(\partial \delta)^2} - (V_1 V_2 - V_{12})^{-2} \left(\frac{\partial V_1}{\partial \delta} V_2 + V_1 \frac{\partial V_2}{\partial \delta} - \frac{\partial V_{12}}{\partial \delta} \right)^2 \\ &\quad + (V_1 V_2 - V_{12})^{-1} \left(\frac{\partial^2 V_1}{(\partial \delta)^2} V_2 + 2 \frac{\partial V_1}{\partial \delta} \frac{\partial V_2}{\partial \delta} + V_1 \left(\frac{\partial^2 V_2}{(\partial \delta)^2} - \frac{\partial^2 V_{12}}{(\partial \delta)^2} \right) \right) \end{aligned}$$

Stepwise calculation of the single components shows that all terms are also of form (4.30).

This implies that for both statements it suffices to show that for all $k_1, k_2, k_3, k_4 \geq 0$

$$\mathbb{E} \left[\frac{(\log \eta(\mathbf{s}, t))^{k_1} (\log \eta(\mathbf{s}, t))^{k_2}}{|\eta(\mathbf{s}, t)|^{k_3} |\eta(\mathbf{s}, t)|^{k_4}} \right] < \infty.$$

Since $\eta(\mathbf{s}, t)$ is standard Fréchet $\log(\eta(\mathbf{s}, t))$ is standard Gumbel and $1/\eta(\mathbf{s}, t)$ is standard exponential. Using Hölder's inequality, we obtain

$$\begin{aligned} &\mathbb{E} \left[\frac{|\log(\eta(\mathbf{s}, t))|^{k_1} |\log(\eta(\mathbf{s}, t))|^{k_2}}{|\eta(\mathbf{s}, t)|^{k_3} |\eta(\mathbf{s}, t)|^{k_4}} \right] \\ &< \left(\mathbb{E} [|\log(\eta(\mathbf{s}, t))|^{4k_1}] \mathbb{E} [|\log(\eta(\mathbf{s}, t))|^{4k_2}] \right)^{1/2} \left(\mathbb{E} \left[\left| \frac{1}{\eta(\mathbf{s}, t)} \right|^{4k_3} \right] \mathbb{E} \left[\left| \frac{1}{\eta(\mathbf{s}, t)} \right|^{4k_4} \right] \right)^{1/2} < \infty, \end{aligned}$$

since all moments of the exponential and the Gumbel distributions are finite.

CHAPTER 5

A SEMIPARAMETRIC ESTIMATION PROCEDURE FOR THE PARAMETERS IN A MAX-STABLE SPACE-TIME PROCESS

In this chapter, we introduce a new semiparametric estimation procedure as an alternative to the proposed pairwise likelihood estimation in Chapter 4. The results are based on Davis, Klüppelberg and Steinkohl [28]. The extremogram is the natural extreme analogue of the correlation function for stationary processes. It was introduced in Davis and Mikosch [23] and extended to a spatial setting in Cho, Davis and Ghosh [15]. A special case of the extremogram was also considered in Fasen et al. [37]. Based on a closed form expression of the extremogram containing the parameters of interest, we estimate the extremogram empirically and apply constrained weighted linear regression to estimate the parameters.

The chapter is organized as follows. Section 5.1 describes the semiparametric estimation procedure in full detail. Asymptotic normality of the parameter estimates is established in Section 5.2. In Section 5.3 bootstrap methods for the construction of pointwise confidence intervals are introduced and analysed theoretically.

5.1 Derivation of the semiparametric estimation procedure

As before, we consider the Brown-Resnick process in (4.2) with δ given by

$$\delta(\mathbf{h}, u) = \theta_1 \|\mathbf{h}\|^{\alpha_1} + \theta_2 |u|^{\alpha_2},$$

where $\boldsymbol{\psi} = (\theta_1, \alpha_1, \theta_2, \alpha_2)$ denotes the parameter vector. As described in Chapter 4 composite likelihood methods can be used to estimate the parameters. Unfortunately, parameter estimation using composite likelihood methods can be laborious since the computation and subsequent optimization of the objective function is time-consuming. The choice of good initial values for the optimization of the composite likelihood is essential.

In the following, we introduce an alternative estimation approach, which is based on the closed form expression for the tail dependence coefficient given by

$$\chi = \lim_{x \rightarrow \infty} P(\eta(\mathbf{s}_1, t_1) > x \mid \eta(\mathbf{s}_2, t_2) > x).$$

This can be viewed as a special case of the space-time extremogram, which is an adaption of the temporal extremogram introduced in Davis and Mikosch [23].

Definition 5.1 (The extremogram). *For strictly stationary space-time processes $\{X(\mathbf{s}, t), (\mathbf{s}, t) \in \mathbb{R}^d \times [0, \infty)\}$ the space-time extremogram is defined for two Borel-sets A and B bounded away from 0 by*

$$\rho_{AB}(r, u) = \lim_{z \rightarrow \infty} \frac{P(z^{-1}X(\mathbf{s}, t) \in A, z^{-1}X(\mathbf{s} + \mathbf{h}, t + u) \in B)}{P(z^{-1}X(\mathbf{s}, t) \in A)}, \quad \min\{r = \|\mathbf{h}\|, u\} \geq 0, \quad (5.1)$$

provided the limit exists.

In the following, we denote by $\chi(r, u)$ the tail dependence coefficient as special case of the extremogram with $A = B = (1, \infty)$, i.e. $\chi(r, u) = \rho_{(1, \infty)(1, \infty)}(r, u)$. For the Brown-Resnick process in (4.2) we obtain a closed form expression for χ which builds the basis for our estimation procedure.

Proposition 5.1. *The extremogram for $A = B = (1, \infty)$ (tail dependence coefficient) of the*

space-time Brown-Resnick process in (4.2), and $\delta(\mathbf{h}, u) = \theta_1 \|\mathbf{h}\|^{\alpha_1} + \theta_2 |u|^{\alpha_2}$, is given by

$$\chi(r, u) = 2\left(1 - \Phi\left(\sqrt{\delta(\mathbf{h}, u)}\right)\right) = 2\left(1 - \Phi\left(\sqrt{\theta_1 r^{\alpha_1} + \theta_2 u^{\alpha_2}}\right)\right), \quad \min\{r = \|\mathbf{h}\|, u\} \geq 0. \quad (5.2)$$

We describe the idea behind the new estimation procedure. Solving equation (5.2) for $\delta(\mathbf{h}, u)$ leads to

$$\delta(\mathbf{h}, u) = \left(\Phi^{-1}\left(1 - \frac{1}{2}\chi(r, u)\right)\right)^2. \quad (5.3)$$

Using a temporal lag equal to zero and taking the logarithm on both sides gives

$$2 \log\left(\Phi^{-1}\left(1 - \frac{1}{2}\chi(r, 0)\right)\right) = \log(\theta_1) + \alpha_1 \log(r).$$

In the same way, we obtain

$$2 \log\left(\Phi^{-1}\left(1 - \frac{1}{2}\chi(0, u)\right)\right) = \log(\theta_2) + \alpha_2 \log(u).$$

These equations are the basis for estimating the parameters. We replace the extremogram on the left hand side in both of these equations with a nonparametric estimate at several lags. Then we use constrained weighted least squares in a regression framework to yields estimates of the parameters. The parameter estimation is based on the following observation scheme for the space-time data.

Condition 5.1. *We assume that the locations lie on a regular d -dimensional lattice,*

$$S_m = \{s_j, j = 1, \dots, M = m^d\} = \{(i_1, \dots, i_d), i_1, \dots, i_d \in \{1, \dots, m\}\}.$$

Further assume that the time points are equidistant, given by the set $\{t_1, \dots, t_T\} = \{1, \dots, T\}$.

In the following scheme we illustrate the two step procedure to estimate the parameters. Let \mathcal{H} and \mathcal{U} denote sets of spatial and temporal lags, respectively, which are included in the estimation.

(1) Nonparametric estimates for the extremogram:

(1a) For all $t \in \{t_1, \dots, t_T\}$ calculate

$$\hat{\chi}^{(t)}(r, 0) = \frac{\sum_{i=1}^M \sum_{\substack{j=1 \\ \|s_i - s_j\|=r}}^M \mathbb{1}_{\{\eta(s_i, t) > q, \eta(s_j, t) > q\}}}{\sum_{i=1}^M \mathbb{1}_{\{\eta(s_i, t) > q\}}}, \quad r \in \mathcal{H}, \quad (5.4)$$

where q is a high empirical quantile of $\{\eta(s_i, t), i = 1, \dots, M\}$.

(1b) For all $s \in S_m$ estimate the extremogram by

$$\hat{\chi}^{(s)}(0, u) = \frac{\sum_{k=1}^{T-u} \mathbb{1}_{\{\eta(s, t_k) > q, \eta(s, t_k + u) > q\}}}{\sum_{k=1}^T \mathbb{1}_{\{\eta(s, t_k) > q\}}}, \quad u \in \mathcal{U}, \quad (5.5)$$

where q is a high empirical quantile of $\{\eta(s, t_k), k = 1, \dots, T\}$.

(2) The respective ‘‘spatial’’ and ‘‘temporal’’ extremograms are defined by computing averages over the temporal and spatial locations, i.e.,

$$(2a) \quad \hat{\chi}(r, 0) = \frac{1}{T} \sum_{k=1}^T \hat{\chi}^{(t_k)}(r, 0), \quad r \in \mathcal{H}.$$

$$(2b) \quad \hat{\chi}(0, u) = \frac{1}{M} \sum_{i=1}^M \hat{\chi}^{(s_i)}(0, u), \quad u \in \mathcal{U}.$$

(3) Solve the following optimization problems to obtain estimates for $\theta_1, \alpha_1, \theta_2$ and α_2 .

(3a)

$$\min_{\substack{\theta_1, \alpha_1 \\ \alpha_1 \in (0, 2]}} \sum_{r \in \mathcal{H}} w_r \left(2 \log \left(\Phi^{-1} \left(1 - \frac{1}{2} \hat{\chi}(r, 0) \right) \right) - (\log(\theta_1) + \alpha_1 \log(r)) \right)^2, \quad (5.6)$$

(3b)

$$\min_{\substack{\theta_2, \alpha_2 \\ \alpha_2 \in (0, 2]}} \sum_{u \in \mathcal{U}} w_u \left(2 \log \left(\Phi^{-1} \left(1 - \frac{1}{2} \hat{\chi}(0, u) \right) \right) - (\log(\theta_2) + \alpha_2 \log(u)) \right)^2, \quad (5.7)$$

where $w_u > 0$ and $w_r > 0$ are certain weights which control the emphasis of spatial or temporal lags included in the estimation. In the following sections we show asymptotic properties

of the estimates.

5.2 Estimation for the space-time Brown-Resnick process

In this section we analyse asymptotic properties for estimation in the context of the Brown-Resnick process in (4.2) and $d = 2$. The proof is done in three steps. We first show α -mixing for the model, which is needed to show asymptotic normality of the spatial extremogram. Using the results we show asymptotic normality of the semiparametric estimates. Some general theory for the spatial extremogram is discussed in Section 5.4.

5.2.1 Asymptotics of the spatial extremogram

For a strictly stationary max-stable random field recent work in Dombry and Eyi-Minko [34] shows that the α -mixing coefficient can be related to the tail dependence coefficient of the max-stable process. The following proposition is a direct application of Corollary 2.2 in Dombry and Eyi-Minko [34].

Proposition 5.2. *For all fixed time points $t \in \mathbb{N}$ the random field $\{\eta(s, t), s \in \mathbb{Z}^d\}$ defined in (4.2) is α -mixing with mixing coefficients satisfying*

$$\alpha_{k,l}(j) \leq 2kl \sup_{r \geq j} \chi(r, 0) \leq 4kl \exp\left\{-\frac{\theta_1 j^{\alpha_1}}{2}\right\}, \quad k, l, j \geq 0. \quad (5.8)$$

For all fixed locations $s \in \mathbb{Z}^d$ the time series $\{\eta(s, t), t \in \mathbb{N}\}$ in (4.2) is α -mixing with mixing coefficients

$$\alpha(j) \leq 4 \sum_{u=j}^{\infty} \chi(0, u) \leq 8 \sum_{u=j}^{\infty} \exp\left\{-\frac{\theta_2 u^{\alpha_2}}{2}\right\}, \quad j \geq 0. \quad (5.9)$$

In the following sections we will make use of the following simple result.

Lemma 5.1. *Let z denote an arbitrary integer and $x > 0$. For $(\theta, \alpha) = (\theta_1, \alpha_1), (\theta_2, \alpha_2)$ the function*

$$g_z(x) = \sum_{u=x}^{\infty} u^z \exp\{-\theta u^\alpha / 2\}$$

satisfies

$$g_z(x) \leq ce^{-\theta x^\alpha/2} x^{z+1},$$

for some constant $c > 0$.

Proof. It holds

$$\begin{aligned} g_z(x) &= \sum_{u=x}^{\infty} u^z \exp\{-\theta u^\alpha/2\} \leq \int_x^{\infty} u^z e^{-\theta u^\alpha/2} du \\ &= \left(\frac{2}{\theta}\right)^{z/\alpha+1/\alpha-1} \frac{2}{\theta\alpha} \int_{\theta x^\alpha/2}^{\infty} t^{\alpha^{-1}(z+1)-1} e^{-t} dt \leq c_1 \Gamma([\alpha^{-1}(z+1)], \theta x^\alpha/2) \\ &= c_1 ([\alpha^{-1}(z+1)] - 1)! e^{-\theta x^\alpha/2} \sum_{k=0}^{[\alpha^{-1}(z+1)]-1} \frac{\theta^k x^{\alpha k}}{2^k k!} \leq ce^{-\theta x^\alpha/2} x^{\alpha([\alpha^{-1}(z+1)]-1)} \\ &\leq ce^{-\theta x^\alpha/2} x^{z+1}, \end{aligned}$$

where $\Gamma(s, x) = \int_x^{\infty} t^{s-1} e^{-t} dt$ is the incomplete Gamma function and $c_1, c > 0$ are some constants. □

In Section 5.4 we explain and extend fundamental theory for the spatial extremogram. By showing Conditions (1) - (4) in Proposition 5.4 we prove a central limit theorem for the empirical spatial extremogram of the Brown-Resnick process (4.2) centred by the preasymptotic extremogram, which is defined for $A = B = (1, \infty)$ and $u = 0$ by

$$\chi_m(r, 0) = \frac{P(\eta(\mathbf{0}, 0) > a_{n_m}, \eta(\mathbf{h}, 0) > a_{n_m})}{P(\eta(\mathbf{0}, 0) > a_{n_m})}, \quad r \in H[1, p], \quad (5.10)$$

where the sequence (a_{n_m}) is chosen such that $P(\|\text{vec}\{\eta(\mathbf{h}), \mathbf{h} \in B(\mathbf{0}, p)\}\| > a_{n_m}) \sim n_m^{-1}$, which implies that $a_{n_m} = n_m$ since the marginal distributions are standard Fréchet. The sequence (n_m) will be specified in Theorem 5.1. For the estimation of the parameters we use all spatial lags smaller or equals to $p > 1$, i.e.

$$\mathcal{H} = H[1, p] = \{r = \|s_1 - s_2\| \in [1, p] : s_1, s_2 \in S_m\}.$$

The preasymptotic extremogram in the central limit theorem can be replaced by the theoretical one, if it converges to the theoretical extremogram with the same convergence rate as

the empirical extremogram to the preasymptotic extremogram (see Remark 5.2). We show that for the spatial setting the preasymptotic extremogram cannot be directly replaced by the theoretical one. Instead we propose a bias correction, which gives us the desired central limit theorem. The maximum norm is used throughout.

Theorem 5.1. *Consider the random field $\{\eta(\mathbf{s}, t), \mathbf{s} \in \mathbb{R}^d\}$ defined in (4.2) for fixed time point $t \in \{t_1, \dots, t_T\}$. Let further $n_m = m^\beta$ with $0 < \beta < 2/3$. The empirical spatial extremogram $\hat{\chi}^{(t)}(r, 0)$ defined in (5.4) satisfies*

$$\left(\frac{m^2}{n_m}\right)^{1/2} \left(\hat{\chi}^{(t)}(r, 0) - \chi_m(r, 0)\right)_{r \in H[1, p]} \xrightarrow{d} \mathcal{N}(0, \Pi_1^{(space)}), \quad m \rightarrow \infty,$$

where $\Pi_1^{(space)}$ is the covariance matrix given explicitly in (5.29), and χ_m is the preasymptotic extremogram in (5.10).

Proof. We use the following notation in the proof.

$$\begin{aligned} H[a, b] &= \{r = \|\mathbf{s}_1 - \mathbf{s}_2\| \in [a, b] : \mathbf{s}_1, \mathbf{s}_2 \in S_m\} \\ p(r) &= \#\{(\mathbf{s}, \mathbf{s}_j) : \|\mathbf{s} - \mathbf{s}_j\| = r, j = 1, \dots, m^2\} \\ c(p) &= \sum_{r \in H[1, p]} p(r) = |H[1, p]| \\ B(\mathbf{h}, p) &= \{\mathbf{s} \in S_m : \|\mathbf{s} - \mathbf{h}\| \leq p\} \end{aligned}$$

For fixed time point $t \in \{t_1, \dots, t_T\}$ the random field $\{\eta(\mathbf{s}, t), \mathbf{s} \in \mathbb{R}^d\}$ is stationary and all finite-dimensional distributions are multivariate extreme value distributions with Fréchet marginals and, thus, are multivariate regularly varying. In addition, these distribution functions only depend on the absolute spatial lag between two locations, which implies that the random field is isotropic. We apply Proposition 5.4, and show that the assumptions are satisfied for the Brown-Resnick process (4.2). By Proposition 5.2 it directly follows that the random field is α -mixing. In addition to the sequence (n_m) we define

$$r_m = \left(2 \log(m^5) / \theta_1\right)^{1/\alpha_1} \rightarrow \infty, \quad m \rightarrow \infty.$$

The sequences satisfy the basic assumptions $n_m/m \rightarrow 0$, $r_m/n_m \rightarrow 0$ and $r_m^3/n_m^2 \rightarrow 0$ as $m \rightarrow \infty$. We verify Assumption (3a) by using (5.8), Lemma 5.1 and the fact that the number

of pairs $p(r)$ with distance r can be bounded by $8r$.

$$\begin{aligned}
 n_m \sum_{r \in H(r_m, \infty)} p(r) \alpha_{c(p), c(p)}(r) &\leq n_m \sum_{r \in H(r_m, \infty)} 8r 4c(p)^2 e^{-\theta_1 r^{\alpha_1}/2} \\
 &\leq 32c(p)^2 n_m \int_{r_m}^{\infty} r e^{-\theta_1 r^{\alpha_1}/2} dr = 32c(p)^2 n_m g_1(r_m) \\
 &\leq 32c(p)^2 c_2 e^{-\theta_1 r_m^{\alpha_1}/2} r_m^2 \leq 32c(p)^2 c_2 \frac{n_m}{m^5} \left(\frac{2}{\theta_1} \log(m^5) \right)^{2/\alpha_1} \rightarrow 0, \quad m \rightarrow \infty.
 \end{aligned}$$

Furthermore,

$$\begin{aligned}
 n_m \sum_{r \in H(k, r_m)} p(r) P \left(\max_{s_1 \in B(\mathbf{0}, p)} \eta(s_1, t) > \frac{\epsilon a_{n_m}}{c(p)}, \max_{s_2 \in B(\mathbf{h}, p)} \eta(s_2, t) > \frac{\epsilon a_{n_m}}{c(p)} \right) \\
 &\leq n_m \sum_{r \in H(k, \infty)} \sum_{s_1 \in B(\mathbf{0}, p)} \sum_{s_2 \in B(\mathbf{h}, p)} 8r P \left(\eta(s_1, t) > \frac{\epsilon a_{n_m}}{c(p)}, \eta(s_2, t) > \frac{\epsilon a_{n_m}}{c(p)} \right) \\
 &\leq \frac{2n_m c(p)}{\epsilon a_{n_m}} \sum_{r \in H(k, \infty)} \sum_{s_1 \in B(\mathbf{0}, p)} \sum_{s_2 \in B(\mathbf{h}, p)} \exp \left\{ -\frac{\theta_1 \|s_1 - s_2\|^{\alpha_1}}{2} \right\} \\
 &= \frac{2n_m c(p)}{\epsilon a_{n_m}} \sum_{r \in H(k, \infty)} \sum_{s_1 \in B(\mathbf{0}, p)} \sum_{s_2 \in B(\mathbf{0}, p)} \exp \left\{ -\frac{\theta_1 \|\mathbf{h} + s_1 - s_2\|^{\alpha_1}}{2} \right\} < \infty.
 \end{aligned}$$

Since the marginal distributions are standard Fréchet, we have $a_{n_m} \sim n_m$ (see Section 2.1.3 for a proof) and, thus, Condition (3b) holds. In addition, it holds that

$$\begin{aligned}
 \sum_{l=1}^{(m/n_m)^2} 4\alpha_{(n_m - r_m)^2, \ln_m^2}(r_m) &\leq 4 \sum_{l=1}^{(m/n_m)^2} (n_m - r_m)^2 \ln_m^2 \frac{1}{m^5} \\
 &= 2(n_m - r_m)^2 \frac{n_m^2}{m^5} \frac{m^2}{n_m^2} \left(\frac{m^2}{n_m^2} + 1 \right) = \frac{2(n_m - r_m)^2 \left(\frac{m^2}{n_m^2} + 1 \right)}{m^3} \rightarrow 0, \quad m \rightarrow \infty,
 \end{aligned}$$

which implies (3c). We consider Condition (4b). As before, we obtain

$$\frac{n_m^7}{m^2} \sum_{r \in H[r_m, \infty)} p(r) \alpha_{1,1}(r) \leq 8c_2 \frac{n_m^7}{m^7} \left(\frac{2}{\theta_1} \log(m^5) \right)^{2/\alpha_1} \rightarrow 0 \quad m \rightarrow \infty,$$

since $\beta < 1$. In addition, $n_m^3 r_m^2 / m^2 \rightarrow 0$ if $\beta < 2/3$. By Proposition 5.4 it follows that the empirical extremogram centred by the preasymptotic extremogram is asymptotically normal. As a last step we analyse Assumption (5) in Remark 5.2.

$$\begin{aligned}
 & \left(\frac{m^2}{n_m} \right)^{1/2} \left(\frac{P(a_{n_m}^{-1} \eta(\mathbf{s}, t) \in A, a_{n_m}^{-1} \eta(\mathbf{s} + \mathbf{h}, t) \in B)}{P(a_{n_m}^{-1} \eta(\mathbf{s}, t) \in A)} - \chi(\|\mathbf{h}\|, 0) \right) \\
 &= \left(\frac{m^2}{n_m} \right)^{1/2} \left(\frac{P(\eta(\mathbf{s}, t) > a_{n_m}, \eta(\mathbf{s} + \mathbf{h}, t) > a_{n_m})}{P(\eta(\mathbf{s}, t) > a_{n_m})} - 2(1 - \Phi(\sqrt{\theta_1 \|\mathbf{h}\|^{\alpha_1}})) \right) \\
 &\sim \left(\frac{m^2}{n_m} \right)^{1/2} \frac{1}{a_{n_m}} (\Phi(\sqrt{\theta_1 \|\mathbf{h}\|^{\alpha_1}})^2 - 1) \sim \left(\frac{m^2}{m^{3\beta}} \right)^{1/2} (\Phi(\sqrt{\theta_1 \|\mathbf{h}\|^{\alpha_1}})^2 - 1) \\
 &\rightarrow 0 \quad m \rightarrow \infty, \text{ if } \beta > 2/3.
 \end{aligned}$$

The last condition is in contrary to (4b) and prevents us from directly replacing the preasymptotic extremogram by the theoretical one. \square

Bias correction for the spatial empirical extremogram

Since it is not possible to find a sequence (n_m) such that Condition (5) in Remark 5.2 is satisfied, we apply a bias correction to obtain a central limit theorem for the empirical extremogram centred by the theoretical extremogram. We calculate the preasymptotic extremogram in the case $A = B = (1, \infty)$.

$$\begin{aligned}
 \chi_{n_m}(r, 0) &= \frac{P(\eta(\mathbf{s}, t) > a_{n_m}, \eta(\mathbf{s} + \mathbf{h}, t) > a_{n_m})}{P(\eta(\mathbf{s}, t) > a_{n_m})} \\
 &\sim a_{n_m} \left(1 - 2 \left(1 - \frac{1}{a_{n_m}} + \frac{1}{2a_{n_m}^2} \right) + \left(1 - \frac{2}{a_{n_m}} \Phi(\sqrt{\theta_1 r^{\alpha_1}}) + \frac{1}{a_{n_m}^2} \Phi(\sqrt{\theta_1 r^{\alpha_1}})^2 \right) \right) \\
 &\sim 2 - \frac{1}{n_m} - 2\Phi(\sqrt{\theta_1 r^{\alpha_1}}) + \frac{1}{n_m} \Phi(\sqrt{\theta_1 r^{\alpha_1}})^2 \\
 &= \chi(r, 0) - \frac{1}{n_m} (1 - \Phi(\sqrt{\theta_1 r^{\alpha_1}})) (1 + \Phi(\sqrt{\theta_1 r^{\alpha_1}})) \\
 &= \chi(r, 0) + \frac{1}{n_m} \frac{1}{2} \chi(r, 0) (1 - \Phi(\sqrt{\theta_1 r^{\alpha_1}}) - 2) \\
 &= \chi(r, 0) + \frac{1}{4n_m} (\chi(r, 0)^2 - \chi(r, 0)) = \chi(r, 0) + \frac{1}{4n_m} v(r, 0),
 \end{aligned}$$

where $v(r, 0) := \chi(r, 0)^2 - \chi(r, 0)$. We propose to use the bias corrected estimate

$$\hat{\chi}(r, 0) = \hat{\chi}(r, 0) - \frac{1}{4n_m}(\hat{\chi}(r, 0)^2 - \hat{\chi}(r, 0)). \quad (5.11)$$

Theorem 5.2. *For fixed time point $t \in \{t_1, \dots, t_T\}$ consider the random field $\{\eta(s, t), s \in \mathbb{R}^2\}$ defined in (4.2). Then, the bias corrected empirical spatial extremogram in (5.11) satisfies*

$$\left(\frac{m^2}{n_m}\right)^{1/2} \left(\hat{\chi}^{(t)}(r, 0) - \chi(r, 0)\right)_{r \in H[1, p]} \xrightarrow{d} \mathcal{N}(0, \Pi_1^{(space)}), \quad m \rightarrow \infty,$$

where $\Pi_1^{(space)}$ is the covariance matrix in (5.29) and $n_m = m^\beta$ with $\frac{2}{5} < \beta < \frac{2}{3}$. Furthermore,

$$\left(\frac{m^2}{n_m}\right)^{1/2} \left(\frac{1}{T} \sum_{k=1}^T \hat{\chi}^{(t_k)}(r, 0) - \chi(r, 0)\right)_{r \in H[1, p]} \xrightarrow{d} \mathcal{N}(0, \Pi_2^{(space)}), \quad m \rightarrow \infty,$$

with covariance matrix $\Pi_2^{(space)}$ in (5.32), Corollary 5.4.

Proof. For simplicity we leave out the time point t in the notation. Since

$$\frac{m}{\sqrt{n_m}}(\hat{\chi}(r, 0) - \chi(r, 0)) = \frac{m}{\sqrt{n_m}}(\hat{\chi}(r, 0) - \chi_m(r, 0)) + \frac{m}{4n_m \sqrt{n_m}}(\hat{v}(r, 0) - v(r, 0))$$

it suffices to show that $(m/4n_m \sqrt{n_m})(\hat{v}(r, 0) - v(r, 0)) \xrightarrow{P} 0, m \rightarrow \infty$.

$$\begin{aligned} & \frac{m}{4n_m \sqrt{n_m}}(\hat{v}(r, 0) - v(r, 0)) \\ &= \frac{m}{4n_m \sqrt{n_m}}(\hat{v}(r) - v_m(r, 0)) + \frac{m}{4n_m \sqrt{n_m}}(v_m(r, 0) - v(r, 0)) \\ &=: A_1 + A_2. \end{aligned}$$

Using the continuous mapping theorem (delta method) together with the fact that $\chi_m(r, 0) \xrightarrow{P} \chi(r, 0)$ it holds for fixed $r \in H[1, p]$ that

$$\begin{aligned} & \frac{m}{\sqrt{n_m}(2\chi(r, 0) - 1)}(\hat{v}(r, 0) - v_m(r, 0)) \\ &= \frac{m}{\sqrt{n_m}(2\chi(r, 0) - 1)}(\hat{\chi}(r, 0)^2 - \hat{\chi}(r, 0) - (\chi_m(r, 0)^2 - \chi_m(r, 0))) \end{aligned}$$

converges weakly to a normal distribution. Since

$$\frac{2\chi(r,0) - 1}{4n_m} \rightarrow 0, \quad m \rightarrow \infty,$$

it follows that $A_1 \xrightarrow{P} 0$. Further note that

$$\begin{aligned} v_m(r,0) &= \chi_m(r,0)^2 - \chi_m(r,0) \\ &\sim \left(\chi(r,0) + \frac{1}{4n_m}(\chi(r,0)^2 - \chi(r,0)) \right)^2 - \chi(r,0) - \frac{1}{4n_m}(\chi(r,0)^2 - \chi(r,0)) \\ &= \chi(r,0)^2 + \frac{1}{2n_m}\chi(r,0)(\chi(r,0)^2 - \chi(r,0)) + \frac{1}{16n_m^2}(\chi(r,0)^2 - \chi(r,0))^2 \\ &\quad - \chi(r,0) - \frac{1}{4n_m}(\chi(r,0)^2 - \chi(r,0)) \\ &= v(r,0) + \frac{v(r,0)}{n_m} \left(\frac{1}{4} - \frac{1}{2}\chi(r,0) + \frac{1}{16n_m}v(r,0) \right). \end{aligned}$$

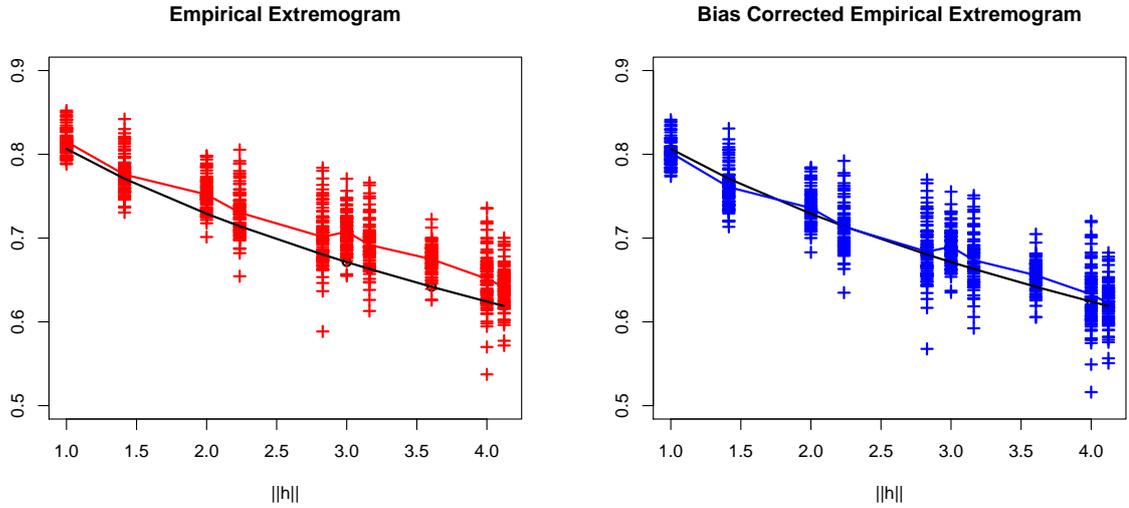


Figure 5.2.1: Empirical extremogram (left) and bias corrected version (right) for one hundred simulated max-stable random fields. The middle black line represents the theoretical extremogram and the middle blue or red line is the mean over all estimates.

This implies

$$A_2 = \frac{m}{4n_m \sqrt{n_m}} \frac{v(r,0)}{n_m} \left(\frac{1}{4} - \frac{\chi(r,0)}{2} - \frac{v(r,0)}{16n_m} \right) \xrightarrow{P} 0,$$

if

$$\frac{m}{n_m^2 \sqrt{n_m}} \rightarrow 0, \quad m \rightarrow \infty.$$

With $n_m = m^\beta$, it follows that $\beta > \frac{2}{5}$. Finally, with $\frac{2}{5} < \beta < \frac{2}{3}$ the statement follows together with Theorem 5.1. \square

To get an intuition on how the bias corrected empirical extremogram behaves, we simulate one hundred max-stable random fields, estimate the extremogram and compare the estimates $\hat{\chi}$ and $\hat{\chi}$. Figure 5.2.1 shows the empirical extremogram on the left hand side compared to the bias corrected one on the right hand side. We clearly see, that the bias corrected extremogram is closer to the true value. In the following, we replace the nonparametric estimate for the extremogram in (5.4) by the bias corrected extremogram.

5.2.2 Spatial parameter estimates and their properties

In this section we prove asymptotic normality for the semiparametric estimates introduced in Section 5.1. We use the following notation. Set

$$y_r = 2 \log \left(\Phi^{-1} \left(1 - \frac{1}{2} \hat{\chi}(r, 0) \right) \right), \quad x_r = \log(r), \quad r \in H[1, p],$$

where $\hat{\chi}(r, 0) = \frac{1}{T} \sum_{k=1}^T \hat{\chi}^{(t_k)}(r, 0)$. To show asymptotic normality of the constrained weighted least-squares (WLS) estimates, the design matrix X and weight matrix W are defined by

$$X = [\mathbf{1}, \text{vec}(x_r, r \in H[1, p])]^\top \quad \text{and} \quad W = \text{diag}\{w_r, r \in H[1, p]\},$$

respectively. Let $\boldsymbol{\psi}_1 = (\log(\theta_1), \alpha_1)$ be the parameter vector with parameter space $\Psi = \mathbb{R} \times (0, 2]$. Denote the WLS estimator

$$\hat{\boldsymbol{\psi}}_1 = (X^\top W X)^{-1} X^\top W \text{vec}(y_r, r \in H[1, p]).$$

Without any constraints $\hat{\boldsymbol{\psi}}_1$ may produce estimates of α_1 outside the parameter space. In such cases the parameter estimate is set equal to 2 and we denote the resulting estimate by $\hat{\boldsymbol{\psi}}_1^c = (\widehat{\log(\theta_1)}^c, \hat{\alpha}_1^c)^\top$.

Theorem 5.3. Let $\hat{\psi}_1^c = (\log(\widehat{\theta}_1)^c, \widehat{\alpha}_1^c)$ denote the constrained WLS estimate resulting from the minimization problem in (5.6) and let $\psi_1^* \in \Psi$ denote the true parameter vector. Further, let $\varphi_{\mu, \Sigma}(\beta_1, \beta_2)$ denote the density of a bivariate normal distribution with mean μ and covariance matrix Σ . The following asymptotic properties hold.

$$\left(\frac{m^2}{n_m}\right)^{1/2} \left(\hat{\psi}_1^c - \psi_1^*\right) \xrightarrow{d} \begin{cases} Z_1, & \alpha_1^* < 2, \\ Z_2, & \alpha_1^* = 2, \end{cases} \quad m \rightarrow \infty, \quad (5.12)$$

where $Z_1 \sim \mathcal{N}(0, \Pi_3^{(space)})$ and

$$\begin{aligned} P(Z_2 \in B) &= \int_{B \cap \{(b_1, b_2) \in \mathbb{R}^2, b_2 < 0\}} \varphi_{\mathbf{0}, \Pi_3^{(space)}}(\beta_1, \beta_2) d\beta_1 d\beta_2 \\ &+ \int_0^\infty \int_{\{b_1 \in \mathbb{R}: (b_1, 0) \in B\}} \varphi_{\mathbf{0}, \Pi_3^{(space)}}\left(\beta_1 - \left(\sum_{r \in H[1, p]} w_r x_r\right) / \left(\sum_{r \in H[1, p]} w_r\right), \beta_2, \beta_2\right) d\beta_1 d\beta_2 \end{aligned} \quad (5.13)$$

with

$$\Pi_3^{(space)} = Q_x^{(w)} G \Pi_2^{(space)} G^\top Q_x^{(w)\top}, \quad (5.14)$$

where $\Pi_2^{(space)}$ is the covariance matrix in (5.31), $n_m = m^\beta$, $2/5 < \beta < 2/3$,

$$Q_x^{(w)} = (X^\top W X)^{-1} X^\top W \quad \text{and} \quad (5.15)$$

$$G = \text{diag} \left\{ \left(\frac{-\sqrt{2\pi} \exp\{\theta_1^* r^{\alpha_1^*} / 2\}}{\sqrt{\theta_1^* u^{\alpha_1^*}}} \right), r \in H[1, p] \right\}. \quad (5.16)$$

Proof. For $r \in H[1, p]$ it holds

$$y_r = g(\hat{\chi}(r, 0)), \quad \text{with} \quad g(x) = 2 \log(\Phi^{-1}(1 - x/2)).$$

The derivative of g is given by

$$g'(x) = -(\Phi^{-1}(1 - \frac{1}{2}x) \varphi(\Phi^{-1}(1 - \frac{1}{2}x)))^{-1}$$

and, thus,

$$g'(\chi(r, 0)) = -\left(\sqrt{\theta_1^* r^{\alpha_1^*}} \varphi(\sqrt{\theta_1^* r^{\alpha_1^*}})\right)^{-1} = \frac{-\sqrt{2\pi} \exp\{\theta_1^* r^{\alpha_1^*}/2\}}{\sqrt{\theta_1^* r^{\alpha_1^*}}}.$$

Using the multivariate delta method together with Theorem 5.2 it follows that

$$\left(\frac{m^2}{n_m}\right)^{1/2} (y_r - g(\chi(r, 0)))_{r \in H[1, p]} \xrightarrow{d} \mathcal{N}(\mathbf{0}, G \Pi_2^{(\text{space})} G^\top), \quad m \rightarrow \infty,$$

where G is defined in (5.16). Since

$$Q_x^{(w)}(g(\chi(r, 0)))_{r \in H[1, p]} = (\log(\theta_1^*), \alpha_1^*) = \boldsymbol{\psi}_1^{*\top},$$

it follows by Cramér Wold's device that

$$\left(\frac{m^2}{n_m}\right)^{1/2} \left(\hat{\boldsymbol{\psi}}_1 - \boldsymbol{\psi}_1^*\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, Q_x^{(w)} G \Pi_2^{(\text{space})} G^\top Q_x^{(w)\top}\right), \quad m \rightarrow \infty.$$

We now turn to the constraints on α_1 . Since the objective function is quadratic, it follows by standard arguments from optimization theory, that if the unconstrained estimate exceeds two, the constraints $\alpha_2 \in (0, 2]$ result in an estimate equal to two. We consider the two cases $\alpha_1^* < 2$ and $\alpha_1^* = 2$, i.e. the parameter lies either in the interior of the parameter space or on the boundary of the parameter space. The constrained estimator, denoted by $\hat{\boldsymbol{\psi}}_1^c$, can be written in the following way.

$$\hat{\boldsymbol{\psi}}_1^c = \hat{\boldsymbol{\psi}}_1 \mathbb{1}_{\{\hat{\alpha}_1 \leq 2\}} + (\hat{\theta}_1, 2)^\top \mathbb{1}_{\{\hat{\alpha}_1 > 2\}}.$$

We calculate the asymptotic probabilities for $\hat{\alpha}_1 \leq 2$ and $\hat{\alpha}_1 > 2$,

$$P(\hat{\alpha}_1 \leq 2) = P\left(\left(\frac{m^2}{n_m}\right)^{1/2} (\hat{\alpha}_1 - \alpha_1^*) \leq \left(\frac{m^2}{n_m}\right)^{1/2} (2 - \alpha_1^*)\right).$$

Since $(m^2/n_m)^{1/2}(\hat{\alpha}_1 - \alpha_1^*) \xrightarrow{d} \mathcal{N}(0, \Pi_3^{(\text{space})}[2, 2])$, $m \rightarrow \infty$ and $(m^2/n_m)^{1/2}(2 - \alpha_1^*) \rightarrow$

∞ , $m \rightarrow \infty$ ($\alpha_1^* < 2$), it follows that

$$P(\hat{\alpha}_1 \leq 2) \rightarrow 1, \quad \text{and} \quad P(\hat{\alpha}_1 > 2) \rightarrow 0, \quad m \rightarrow \infty.$$

Therefore, for $\alpha_1^* < 2$

$$\left(\frac{m^2}{n_m}\right)^{1/2} (\hat{\psi}_1^c - \psi_1^*) \xrightarrow{d} \mathcal{N}(0, \Pi_3^{(\text{space})}), \quad m \rightarrow \infty.$$

We now consider the case $\alpha_1^* = 2$ and $\hat{\alpha}_1 > 2$ (the unconstrained estimate exceeds 2). In this case the optimization problem

$$\min_{\psi_1} (W^{1/2}(\mathbf{y} - X\psi_1))^\top (W^{1/2}(\mathbf{y} - X\psi_1)), \quad \text{s.t.} \quad (0, 1)\psi_1 = 2,$$

has to be solved. To obtain asymptotic results for the optimal solution, the vector $\hat{\psi}_1 - \psi_1^*$ is projected onto the line $\Lambda = \{\psi \in \mathbb{R}^2, (0, 1)\psi = 0\}$, i.e. the projection matrix is given by $P_\Lambda = I_2 - (X^\top WX)^{-1}(0, 1)^\top ((0, 1)(X^\top WX)^{-1}(0, 1)^\top)^{-1}(0, 1)$. For simplicity we use the abbreviation $p_{wx} = \sum_{r \in H[1, p]} w_r x_r / \sum_{r \in H[1, p]} w_r$. It follows,

$$\begin{aligned} (\hat{\psi}_1^c - \psi_1^*) \mathbb{1}_{\{\hat{\alpha}_1 > 2\}} &= P_\Lambda (\hat{\psi}_1 - \psi_1^*) \mathbb{1}_{\{\hat{\alpha}_1 > 2\}} \\ &= (\hat{\psi}_1 - \psi_1^*) \mathbb{1}_{\{\hat{\alpha}_1 > 2\}} - (X^\top WX)^{-1}(0, 1)^\top \left((0, 1)(X^\top WX)^{-1}(0, 1)^\top \right)^{-1} (\hat{\alpha}_1 - 2) \mathbb{1}_{\{\hat{\alpha}_1 > 2\}} \\ &= \hat{\psi}_1 \mathbb{1}_{\{\hat{\alpha}_1 > 2\}} + \begin{pmatrix} -p_{wx} \\ 1 \end{pmatrix} (\hat{\alpha}_1 - 2) \mathbb{1}_{\{\hat{\alpha}_1 > 2\}}. \end{aligned}$$

For the joint constrained estimator we obtain

$$\begin{aligned} \hat{\psi}_1^c - \psi_1^* &= (\hat{\psi}_1^c - \psi_1^*) \mathbb{1}_{\{\hat{\alpha}_1 \leq 2\}} + (\hat{\psi}_1^c - \psi_1^*) \mathbb{1}_{\{\hat{\alpha}_1 > 2\}} \\ &= (\hat{\psi}_1 - \psi_1^*) \mathbb{1}_{\{\hat{\alpha}_1 \leq 2\}} + (\hat{\psi}_1 - \psi_1^*) \mathbb{1}_{\{\hat{\alpha}_1 > 2\}} + \begin{pmatrix} p_{wx} \\ -1 \end{pmatrix} (\hat{\alpha}_1 - 2) \mathbb{1}_{\{\hat{\alpha}_1 > 2\}} \end{aligned}$$

This implies

$$\left(\frac{m^2}{n_m}\right)^{1/2} (\hat{\psi}_1^c - \psi_1^*) = \left(\frac{m^2}{n_m}\right)^{1/2} \begin{pmatrix} (\widehat{\log(\theta_1)} - \log(\theta_1^*)) + p_{wx}(\hat{\alpha}_1 - 2) \mathbb{1}_{\{\hat{\alpha}_1 > 2\}} \\ (\hat{\alpha}_1 - \alpha_1^*) - (\hat{\alpha}_1 - 2) \mathbb{1}_{\{\hat{\alpha}_1 > 2\}} \end{pmatrix}.$$

For the joint asymptotic distribution of the constrained estimator it follows with

$f(x_1, x_2) = (x_1 + p_{wx}x_2\mathbb{1}_{\{x_2>0\}}, x_2 - x_2\mathbb{1}_{\{x_2>0\}})^\top$, that

$$\begin{aligned}
 & P\left(\left(\frac{m^2}{n_m}\right)^{1/2} (\hat{\psi}_1^c - \psi_1^*) \in B\right) = P(f(\hat{\psi}_1 - \psi_1^*) \in B) \\
 & = P(\hat{\psi}_1 - \psi_1^* \in f^{-1}(B \cap \{(b_1, b_2) \in \mathbb{R}^2 : b_2 < 0\}) \cup f^{-1}(B \cap \{(b_1, 0), b_1 \in \mathbb{R}\})) \\
 & = P(\hat{\psi}_1 - \psi_1^* \in (B \cap \{(b_1, b_2) \in \mathbb{R}^2 : b_2 < 0\}) \cup (\{(b_1 - cb_2, b_2), b_2 \geq 0, b_1 \in B\})) \\
 & \rightarrow \int_{B \cap \{(b_1, b_2) \in \mathbb{R}^2, b_2 < 0\}} \varphi_{\mathbf{0}, \Pi_3^{(\text{space})}}(\beta_1, \beta_2) d\beta_1 d\beta_2 \\
 & \quad + \int_0^\infty \int_{\{b_1 \in \mathbb{R} : (b_1, 0) \in B\}} \varphi_{\mathbf{0}, \Pi_3^{(\text{space})}}(\beta_1 - p_{wx}\beta_2, \beta_2) d\beta_1 d\beta_2, \quad m \rightarrow \infty.
 \end{aligned}$$

□

Remark 5.1. The derivation of the asymptotic properties for the constrained estimate is in fact a special case of Corollary 1 in Andrews [2], who shows asymptotic properties of parameter estimates in a very general setting when the true parameter is on the boundary of the parameter space. The asymptotic distribution of the estimates in the case $\alpha_1^* = 2$ is driven by the fact that approximately half of the estimates lie above the true value and are therefore equal to two, which is reflected by the second term in the asymptotic distribution of the estimates.

5.2.3 Temporal parameter estimates and their properties

The estimation of the temporal parameters θ_2 and α_2 is analogous to the estimation of the spatial parameters as described in Sections 5.2.1 and 5.2.2 with one exception. Since the empirical temporal extremogram centered by the theoretical one is asymptotically normal, a bias correction is not needed. The set of temporal lags used in the estimation is given by

$$\mathcal{U} = \{1, \dots, p\}.$$

Theorem 5.4. For fixed location $s \in S_m$ consider the Brown-Resnick process $\{\eta(s, t), t \in [0, \infty)\}$ defined in (4.2). The empirical extremogram in (5.5) for $A = B = (1, \infty)$ satisfies

$$\left(\frac{T}{nT}\right)^{1/2} \left(\hat{\chi}^{(s)}(0, u) - \chi(0, u)\right)_{u \in \{1, \dots, p\}} \xrightarrow{d} \mathcal{N}(0, \Pi_1^{(\text{time})}), \quad T \rightarrow \infty,$$

where $n_T = T^\beta$ with $1/3 < \beta < 1/2$, and $\Pi_1^{(time)}$ is the covariance matrix defined in a similar fashion to $\Pi_1^{(space)}$ in Proposition (5.4). For the exact definition see Davis and Mikosch [23]. Furthermore,

$$\left(\frac{T}{n_T}\right)^{1/2} \left(\frac{1}{M} \sum_{i=1}^M \hat{\chi}^{(s_i)}(0, u) - \chi(0, u)\right)_{u \in \{1, \dots, p\}} \xrightarrow{d} \mathcal{N}(0, \Pi_2^{(time)}), \quad T \rightarrow \infty, \quad (5.17)$$

with covariance matrix $\Pi_2^{(time)}$ in (5.18).

Proof. We verify the assumptions for the central limit theorem of the temporal extremogram in Davis and Mikosch [23], Corollary 3.4. Define the sequences $n_T = T^\beta \rightarrow \infty$, $\beta < 1$ and $r_T = (2 \log T / \theta_2)^{1/\alpha_2} \rightarrow \infty$ as $T \rightarrow \infty$. The sequences satisfy the basic assumptions $n_T/T \rightarrow 0$ and $r_T/n_T \rightarrow 0$. As for the spatial setting, all finite-dimensional distributions are multivariate extreme value distributions and, thus, are regularly varying. Since the marginal distributions are standard Fréchet, we choose $a_T \sim T$. From Lemma 5.2 the time series $\{\eta(s, t), t \in [0, \infty)\}$ is α -mixing. By Lemma 5.1 it holds that

$$n_T \sum_{u=r_T}^{\infty} \alpha(u) \leq 8c_1 n_T g_1(r_T) \leq 8cT^{\beta-1} (2 \log T / \theta_2)^{2/\alpha_2} \rightarrow 0, \quad T \rightarrow \infty.$$

In addition,

$$\begin{aligned} & n_T \sum_{u=k}^{r_T} P(\|(\eta(s, t_{1+u}), \dots, \eta(s, t_{1+u+p}))\| > \epsilon a_{n_T}, \|(\eta(s, t_1), \dots, \eta(s, t_{1+p}))\| > \epsilon a_{n_T}) \\ & \leq n_T \sum_{u=k}^{\infty} \sum_{i=u}^{u+p} \sum_{j=0}^p P(\eta(s, t_{i+1}) > \epsilon a_{n_T}, \eta(s, t_{j+1}) > \epsilon a_{n_T}) \\ & = n_T \sum_{u=k}^{\infty} \sum_{i=u}^{u+p} \sum_{j=0}^p \left(\frac{2}{\epsilon a_{n_T}} - \frac{2\Phi(\sqrt{\delta(\mathbf{0}, t_{i+1} - t_{j+1})})}{\epsilon a_{n_T}} + O\left(\frac{1}{\epsilon^2 a_{n_T}^2}\right) \right. \\ & \quad \left. + O\left(\frac{4\Phi(\sqrt{\delta(\mathbf{0}, t_{i+1} - t_{j+1})})^2}{\epsilon^2 a_{n_T}^2}\right) \right) \\ & \sim n_T \sum_{u=k}^{\infty} \sum_{i=u}^{u+p} \sum_{j=0}^p \frac{2}{\epsilon a_{n_T}} \bar{\Phi}(\sqrt{\delta(\mathbf{0}, t_{i+1} - t_{j+1})}) \end{aligned}$$

$$\begin{aligned} &\leq n_T \sum_{u=k}^{\infty} \sum_{i=u}^{u+p} \sum_{j=0}^p \frac{2}{\epsilon a_{n_T}} \exp \left\{ \frac{-\delta(\mathbf{0}, t_{i+1} - t_{j+1})}{2} \right\} \\ &\leq \frac{2n_T}{a_{n_T}} \sum_{u=k}^{\infty} \sum_{i=0}^p \sum_{j=0}^p \exp \left\{ -\frac{\theta_2 |u + t_{i+1} - t_{j+1}|^{\alpha_2}}{2} \right\} < \infty. \end{aligned}$$

Since $a_{n_T} \sim n_T$ this shows Condition (M) in Davis and Mikosch [23]. As before, we have

$$\frac{n_T^4}{T} \sum_{u=r_T}^{n_T} \alpha(u) \leq 8c \frac{1}{T^{4\beta-2}} \left(\frac{2}{\theta_2} \log T \right)^{2/\alpha_2} \rightarrow 0, \quad T \rightarrow \infty,$$

which is satisfied for $\beta < 1/2$. Furthermore, $n_T r_T^3 / T \rightarrow 0$, $T \rightarrow \infty$. As a last step we verify the convergence rate of the preasymptotic to the theoretical extremogram.

$$\begin{aligned} &\left(\frac{T}{n_T} \right)^{1/2} \left(\frac{P(a_{n_T}^{-1} \eta(\mathbf{s}, t_1) \in A, a_{n_T}^{-1} \eta(\mathbf{s}, t_1+u) \in B)}{P(a_{n_T}^{-1} \eta(\mathbf{s}, t_1) \in A)} - \chi(0, u) \right) \\ &= \left(\frac{T}{n_T} \right)^{1/2} \left(\frac{1 - 2 \exp\{-1/a_{n_T}\} + \exp\left\{-\frac{2}{a_{n_T}} \Phi(\sqrt{\theta_2 |u|^{\alpha_2}})\right\}}{1 - \exp\{-1/a_{n_T}\}} - 2(1 - \Phi(\sqrt{\theta_2 |u|^{\alpha_2}})) \right) \\ &\sim \left(\frac{T}{n_T} \right)^{1/2} \left(-\frac{1}{a_{n_T}} + \frac{1}{a_{n_T}} \Phi(\sqrt{\theta_2 |u|^{\alpha_2}})^2 \right) = \left(\frac{T}{n_T} \right)^{1/2} \frac{1}{a_{n_T}} \left(\Phi(\sqrt{\theta_2 |u|^{\alpha_2}})^2 - 1 \right) \\ &\sim \frac{T^{1/2}}{T^{3\beta/2}} \left(\Phi(\sqrt{\theta_2 |u|^{\alpha_2}})^2 - 1 \right) = \frac{1}{T^{3\beta/2-1/2}} \left(\Phi(\sqrt{\theta_2 |u|^{\alpha_2}})^2 - 1 \right) \rightarrow 0, \quad T \rightarrow \infty, \end{aligned}$$

if $\beta > 1/3$. Altogether, we require $n_T = T^\beta$ with $1/3 < \beta < 1/2$. By Corollary 3.4 in Davis and Mikosch [23] it follows that

$$\left(\frac{T}{n_T} \right)^{1/2} \left(\hat{\chi}^{(s)}(0, u) - \chi(0, u) \right)_{u \in \{1, \dots, p\}} \xrightarrow{d} \mathcal{N}(0, \Pi_1^{(\text{time})}), \quad T \rightarrow \infty,$$

for each fixed location $\mathbf{s} \in S_m$. The extension of the statement to spatial means of extremograms follows in the same way as in Corollary 5.4 by using the vectorized process

$$(\mathbf{Y}_t) = \text{vec}\{\eta(\mathbf{s}, t), \dots, \eta(\mathbf{s}, t+p)\}, \mathbf{s} \in S_m\}$$

and defining the sets $D_{u,k}$ and C_k for $u = 1, \dots, p$ and $k = 1, \dots, M = m^2$ properly to extend

the covariance matrix. This leads to the statement in (5.17), where

$$\Pi_2^{(time)} = \frac{1}{M^2} \begin{pmatrix} 1 \cdots 1 & 0 \cdots 0 & \cdots & 0 \cdots 0 \\ 0 \cdots 0 & 1 \cdots 1 & \cdots & 0 \cdots 0 \\ & & \ddots & \\ 0 \cdots 0 & \cdots & \cdots & 1 \cdots 1 \end{pmatrix} F \Sigma F^\top \begin{pmatrix} 1 \cdots 1 & 0 \cdots 0 & \cdots & 0 \cdots 0 \\ 0 \cdots 0 & 1 \cdots 1 & \cdots & 0 \cdots 0 \\ & & \ddots & \\ 0 \cdots 0 & \cdots & \cdots & 1 \cdots 1 \end{pmatrix}^\top \quad (5.18)$$

and F and Σ are defined in a similar fashion as the matrices in Corollary 5.4. \square

The asymptotic properties of the semiparametric estimates $\hat{\theta}_2$ and $\hat{\alpha}_2$ can be derived in exactly the same way as for the spatial parameters θ_1 and α_1 . Accordingly, we define for $u \in \{1, \dots, p\}$,

$$y_u = 2 \log \left(\Phi^{-1} \left(1 - \frac{1}{2} \hat{\chi}(0, u) \right) \right), \quad x_u = \log(u),$$

and

$$X = [\mathbf{1}, \text{vec}(x_u, u \in \{1, \dots, p\})], \quad W = \text{diag}\{w_u, u \in \{1, \dots, p\}\}.$$

We state the theorem for asymptotic normality.

Theorem 5.5. *Let $\hat{\psi}_2^c = (\log(\widehat{\theta}_2)^c, \widehat{\alpha}_2^c)$ be the constrained weighted estimate resulting from the minimization problem in (5.7) and let $\psi_2^* \in \Psi$ be the true parameter vector. The following asymptotic properties hold.*

$$\left(\frac{T}{n_T} \right)^{1/2} \left(\hat{\psi}_2^c - \psi_2^* \right) \xrightarrow{d} \begin{cases} Z_1, & \alpha_2^* < 2, \\ Z_2, & \alpha_2^* = 2, \end{cases} \quad T \rightarrow \infty, \quad (5.19)$$

where $Z_1 \sim \mathcal{N}(0, \Pi_3^{(time)})$ and

$$\begin{aligned} P(Z_2 \in B) &= \int_{B \cap \{(b_1, b_2) \in \mathbb{R}^2, b_2 < 0\}} \varphi_{\mathbf{0}, \Pi_3^{(time)}}(\beta_1, \beta_2) d\beta_1 d\beta_2 \\ &+ \int_0^\infty \int_{\{b_1 \in \mathbb{R}: (b_1, 0) \in B\}} \varphi_{\mathbf{0}, \Pi_3^{(time)}} \left(\beta_1 - \left(\sum_{u=1}^p w_u x_u \right) / \left(\sum_{u=1}^p w_u \right), \beta_2, \beta_2 \right) d\beta_1 d\beta_2 \end{aligned} \quad (5.20)$$

with

$$\Pi_3^{(time)} = Q_x^{(w)} G \Pi_2^{(time)} G^\top Q_x^{(w)\top}, \quad (5.21)$$

where $\Pi_2^{(time)}$ is the covariance matrix in (5.18), $n_T = T^\beta$, $1/3 < \beta < 1/2$,

$$Q_x^{(w)} = (X^\top WX)^{-1} X^\top W \quad \text{and} \quad (5.22)$$

$$G = \text{diag} \left\{ \left(\frac{-\sqrt{2\pi} \exp\{\theta_2^* u^{\alpha_2^*}/2\}}{\sqrt{\theta_2^* u^{\alpha_2^*}}} \right), u \in \{1, \dots, p\} \right\}. \quad (5.23)$$

5.3 Bootstrapping parameter estimates for the Brown-Resnick process

Since the variance of the estimates is computationally intractable, we use bootstrap methods to construct pointwise confidence intervals for the estimates. The circular bootstrap, introduced in Politis and Romano [68], for the spatial empirical extremogram and observations on a 2-dimensional regular grid is described in Cho et al. [15]. As before, the set of locations is given by $\{(i_1, i_2), i_1, i_2 \in \{1, \dots, m\}\}$. The data is first wrapped to obtain observations outside the actual set of locations. In particular, for some $l > 0$ large enough, such that all blocks fall within the new set of observations,

$$\tilde{X}((i, j)) = \begin{cases} X((i, j-m)), & i = 1, \dots, m, j = m+1, \dots, m+l-1, \\ X((i-m, j)), & i = m+1, \dots, m+l-1, j = 1, \dots, m, \\ X((i-m, j-m)), & i = m+1, \dots, m+l-1, j = m+1, \dots, m+l-1. \end{cases}$$

The blocks with side length $L = L_m$ and random starting point (I_1, I_2) are defined by

$$B_{(I_1, I_2), L} = \{\tilde{X}((I_1, I_2)), \dots, \tilde{X}((I_1 + L, I_2 + L))\}.$$

Now, let $(I_1^{(1)}, I_2^{(1)}), (I_1^{(2)}, I_2^{(2)}), \dots$ be an iid sequence of uniform random variables on $\{1, \dots, m\}^2$. The bootstrap sample $X^*((1, 1)), \dots, X^*((m, m))$ is then constructed from the blocks

$$B_{(I_1^{(1)}, I_2^{(1)}), L}, B_{(I_1^{(2)}, I_2^{(2)}), L}, \dots$$

Figure 5.3.1 visualizes the procedure. On the left hand side the coloured squares show the blocks for two random starting points with block length $L = 2$, which build the first two blocks in the bootstrap sample on the right hand side.

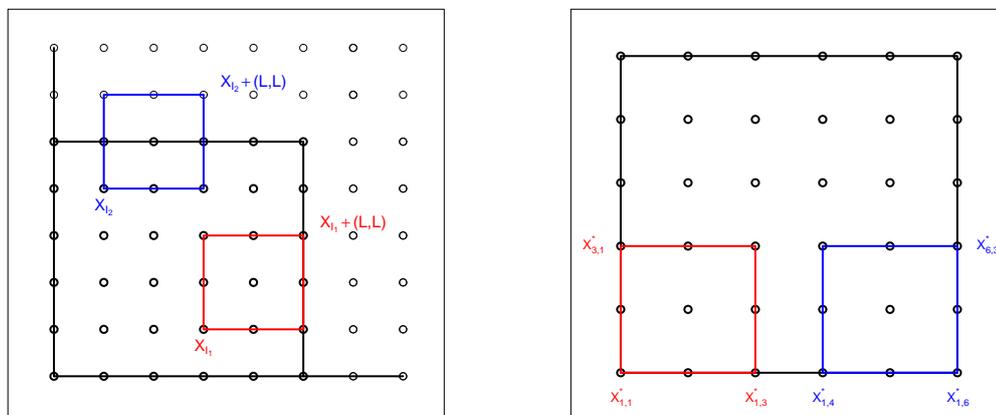


Figure 5.3.1: Visualization of the circular bootstrap with block length $L = 2$ for a two-dimensional grid. The red squared grid on the left hand side builds the first block for the bootstrap sample on the right hand side.

5.3.1 Spatial parameters: the unconstrained case

For the unconstrained estimates, theoretical results can be obtained using the same techniques and calculations as in previous sections. The following proposition is an extension of the results in Davis, Mikosch and Cribben [25] and states a central limit theorem for the bootstrapped spatial extremogram.

Proposition 5.3 (Cho et al. [15]). *Assume that the conditions of Proposition 5.4 are satisfied for the random field $\{X(s), s \in \mathbb{R}^2\}$. In addition, suppose that*

$$(1) \quad n_m^2 \sum_{r \in H[L_m, \infty)} p(r) \alpha_{2p(r), 2p(r)}(r) \rightarrow 0, \quad m \rightarrow \infty \text{ and}$$

$$(2) \quad L_m^2/n_m \rightarrow 0, \quad m \rightarrow \infty.$$

Then, conditionally on $(X(s))$ the bootstrapped spatial empirical extremogram $\hat{\rho}_{AB}^{(BS)}$ satisfies

$$\left(\frac{m^2}{n_m}\right)^{1/2} (\hat{\rho}_{AB}^{(BS)}(r) - \hat{\rho}_{AB}(r))_{r \in H[1, p]} \xrightarrow{d} \mathcal{N}(0, \Pi_1^{(space)}), \quad m \rightarrow \infty, \quad (5.24)$$

where $\Pi_1^{(space)}$ is the covariance matrix defined in Proposition 5.4. The same result holds for the bias corrected extremogram defined in (5.11) for the case $A = B = (1, \infty)$.

For the space-time Brown-Resnick process in (4.2), we extend the procedure as follows. For each fixed time point $t \in \{t_1, \dots, t_T\}$ we construct a sample $\eta^*(s_1, t), \dots, \eta^*(s_M, t)$ by using the steps described above and estimate the parameters as described in Section 4.1. Note, that we use the same sequence (I_i) for all time points t_1, \dots, t_T to keep the dependence structure in time. The result in Proposition 5.3 can also be extended to the temporal mean of spatial extremograms as done in the Corollary 5.4. In particular,

$$\left(\frac{m^2}{n_m}\right)^{1/2} \left(\frac{1}{T} \sum_{k=1}^T \hat{\rho}_{AB}^{(BS), (t_k)}(r, 0) - \frac{1}{T} \sum_{k=1}^T \hat{\rho}_{AB}^{(t_k)}(r, 0)\right) \xrightarrow{d} \mathcal{N}(0, \Pi_2^{(space)}), \quad m \rightarrow \infty,$$

conditional on $\{\eta(s, t)\}$, where $\hat{\rho}_{AB}^{(BS), (t)}(r, 0)$ denotes the bootstrapped spatial extremogram for fixed time point $t \in \{t_1, \dots, t_T\}$. We obtain bootstrap estimates $\hat{\theta}_1^{(BS)}$ and $\hat{\alpha}_1^{(BS)}$. Showing asymptotic normality of the unconstrained estimates is straight forward using Proposition 5.3 together with the steps in the proofs of Theorem 5.1, 5.2 and 5.3.

Theorem 5.6. Assume that the block size L_m is given by $L_m = \left(\frac{2}{\theta_1} \log(Cm^2)\right)^{1/\alpha_1}$ for some constant $C > 0$. Then, conditionally on $(\eta(s, t))$ the bootstrapped unconstrained semiparametric estimates satisfy

$$\left(\frac{m^2}{n_m}\right)^{1/2} \left(\hat{\psi}_1^{(BS)} - \hat{\psi}_1\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \Pi_3^{(space)}\right), \quad m \rightarrow \infty, \quad (5.25)$$

where $\Pi_3^{(space)}$ is the covariance matrix in (5.14).

Proof. Using equation (5.8) it holds

$$\begin{aligned} n_m^2 \sum_{r \in H[L_m, \infty)} p(r) \alpha_{2p(r), 2p(r)}(r) &\leq 32n_m^2 \int_{L_m}^{\infty} r^3 e^{-\theta_1 r^{\alpha_1}/2} dr \\ &= 32n_m^2 g_3(L_m) \leq 32c_2 n_m^2 e^{-\theta_1 L_m^{\alpha_1}/2} L_m^4 = 32c_2 \frac{n_m^2}{m^2} L_m^4 \rightarrow 0, \quad m \rightarrow \infty. \end{aligned}$$

Obviously, $L_m^2/n_m \rightarrow 0$ as $m \rightarrow \infty$. This shows Assumptions (1) and (2) in Proposition 5.3. \square

5.3.2 Spatial parameters: the constrained case

In the literature many examples can be found in which the bootstrap fails to be consistent, see Bickel, Götze and Zwet [9]. In particular, Andrews [3] shows that bootstrap estimates can be inconsistent when parameters are on the boundary of the parameter space which is the case in our procedure. Since $P(\hat{\alpha}_1 \geq 2) \rightarrow 0, m \rightarrow \infty$, if the true value $\alpha_1^* < 2$ (take for example $\alpha_1^* = 2 - (n_m/m^2)^{1/2}$) we cannot apply the same arguments as in the proof of Theorem 5.3 to show that the asymptotic distribution of $(m^2/n_m)^{1/2}(\hat{\psi}_1 - \psi_1^*)$ can be approximated by the conditional asymptotic distribution of $(m^2/n_m)^{1/2}(\hat{\psi}_1^{(BS)} - \hat{\psi}_1)$. As suggested by several authors, including for example Andrews [3], a possible solution to this problem is to use a smaller sample size $U_m < m$ for the bootstrap such that $U_m n_m / m^2 n_{U_m} \rightarrow 0$. Let $\hat{\chi}_{U_m}^{(BS)}$ denote the temporal mean of the bootstrapped bias corrected empirical extremograms based on a spatial subsample with side length U_m , and let $\hat{\chi}$ denote the empirical extremogram based on the whole set of locations S_m . In particular,

$$\hat{\chi}_{U_m}^{(BS)}(r, 0) = \frac{1}{T} \sum_{k=1}^T \hat{\chi}_{U_m}^{(BS), (t_k)}(r, 0), \quad \hat{\chi}(r, 0) = \frac{1}{T} \sum_{k=1}^T \hat{\chi}^{(t_k)}(r, 0),$$

where $\hat{\chi}^{(t_k)}(r, 0)$ is defined in (5.11). Then,

$$\begin{aligned} & \left(\frac{U_m^2}{n_{U_m}} \right)^{1/2} (\hat{\chi}_{U_m}^{(BS)}(r, 0) - \hat{\chi}(r, 0))_{r \in H[1, p]} \\ &= \left(\frac{U_m^2}{n_{U_m}} \right)^{1/2} (\hat{\chi}_{U_m}^{(BS)}(r, 0) - \chi(r, 0))_{r \in H[1, p]} - \left(\frac{U_m^2 n_m}{m^2 n_{U_m}} \right)^{1/2} \left(\frac{m^2}{n_m} \right)^{1/2} (\hat{\chi}(r, 0) - \chi(r, 0))_{r \in H[1, p]} \\ &= A_1 - A_2. \end{aligned}$$

Since the bias corrected empirical extremogram is asymptotically normal it holds that $A_2 \xrightarrow{P} 0$ as $m \rightarrow \infty$. Using the same arguments as in the proof of Theorem 2.1 in Davis et al. [25] one can show that conditionally on $(\eta(s, t))$,

$$A_1 = \left(\frac{m^2}{n_m} \right)^{1/2} (\hat{\chi}_{U_m}^{(BS)}(r, 0) - \chi(r, 0))_{r \in H[1, p]} \xrightarrow{d} \mathcal{N}(0, \Pi_2^{(\text{space})}), \quad m \rightarrow \infty.$$

Since $\hat{\chi}$ does not appear in the first sum, we can now use the same steps as in the proof of Theorem 5.3 to prove the asymptotic properties of the bootstrap estimates.

Corollary 5.1. *Assume the bootstrap sample is based on a subsample grid with side length $U_m < m$ satisfying $U_m^2 n_m / m^2 n_{U_m} \rightarrow 0$, $m \rightarrow \infty$. Let $\hat{\psi}_{1,U_m}^{(BS),c}$ denote the constrained weighted least-squares estimate resulting from a bootstrap sample based on a grid of side length U_m and let $\hat{\psi}_1^c$ be the parameter estimate based on the whole sample grid with side length m . Then, conditional on $(\eta(s, t))$,*

$$\left(\frac{U_m^2}{n_{U_m}}\right)^{1/2} \left(\hat{\psi}_{1,U_m}^{(BS),c} - \hat{\psi}_1^c\right) \xrightarrow{d} \begin{cases} Z_1 & \alpha_1^* < 2, \\ Z_2 & \alpha_1^* = 2, \end{cases} \quad m \rightarrow \infty,$$

where $Z_1 \sim \mathcal{N}(0, \Pi_3^{(space)})$ with $\Pi_3^{(space)}$ defined in (5.14) and the distribution function of Z_2 is given in (5.13).

5.3.3 Temporal parameters

As described in Davis, Mikosch and Cribben [25] for the extremogram, we use the stationary bootstrap method proposed by Politis and Romano [69] and reviewed by Lahiri [54] for the temporal parameters. The procedure is as follows. In a first step the observations are periodically extended.

$$\tilde{X}_i = \tilde{X}_j, \quad i = lT + j, \quad 1 \leq j \leq T,$$

for some $l > 0$. Let (L_j) denote a sequence of block sizes used to construct the bootstrap sample. It is assumed that (L_j) is geometrically distributed with mean $1/q$. The blocks of length L_j are defined as

$$B_{I,L_j} = \{\tilde{X}_I, \dots, \tilde{X}_{I+L_j-1}\}, \quad j = 1, 2, \dots$$

for some start index I . In addition, let I_1, I_2, \dots be a sequence of iid random variables with uniform distribution on $\{1, \dots, T\}$. The bootstrap sample X_1^*, \dots, X_T^* is given by the first T observations from the blocks $B_{I_1, L_1}, B_{I_2, L_2}, \dots$,

$$\tilde{X}_{I_1}, \dots, \tilde{X}_{I_1+L_1-1}, \tilde{X}_{I_2}, \dots, \tilde{X}_{I_2+L_2-1}, \dots$$

For the resulting bootstrap sample X_1^*, \dots, X_T^* , the empirical extremogram is calculated for $u = 1, \dots, p$, leading to bootstrap estimates $\hat{\rho}_{AB}^{(BS)}(u)$, $u \in \{1, \dots, p\}$. As for the spatial parameters we use the following procedure for the space-time model. For each fixed location

$\mathbf{s} \in \{\mathbf{s}_1, \dots, \mathbf{s}_M\}$ construct a sample $\eta^*(\mathbf{s}, t_1), \dots, \eta^*(\mathbf{s}, t_T)$ by using the stationary bootstrap and estimate the parameters as described in Section 4.1, i.e. we estimate the extremogram for all $u \in \{1, \dots, p\}$ and solve the minimization problem in (5.7). Note, that we use the same sequences (I_i) and (L_j) for all locations $\mathbf{s}_1, \dots, \mathbf{s}_M$ to keep the dependence structure in space. This results in bootstrap estimates $\hat{\theta}_2^{(BS)}$ and $\hat{\alpha}_2^{(BS)}$.

The asymptotic properties can be obtained in the same way as for the spatial parameters. We only state the result when the constraints on α_2 are ignored. It follows directly from Corollary 3.3 in Davis et al. [25].

Corollary 5.2. *Assume that the sequence (q_T) for the generation of the block sizes (L_i) satisfies $q = q_T \rightarrow 0$ and $Tq_T^2/n_T \rightarrow 0$, as $T \rightarrow \infty$. Then, the bootstrapped unconstrained semiparametric estimates (ignoring the constraints) satisfy conditionally on $(\eta(\mathbf{s}, t))$*

$$\left(\frac{T}{n_T}\right)^{1/2} \left(\hat{\boldsymbol{\psi}}_2^{(BS)} - \hat{\boldsymbol{\psi}}_2\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \Pi_3^{(time)}\right), \quad T \rightarrow \infty, \quad (5.26)$$

where $\Pi_3^{(time)}$ is the covariance matrix in (5.21).

The theoretical result including the constraints in the estimation can be obtained in exactly the same way as for the spatial parameters. For completeness we state the result for the bootstrapped temporal parameter estimates when subsamples of size $U_T < T$ are used for the bootstrap samples.

Corollary 5.3. *Assume the bootstrap sample is based on a subsample of size $U_T < T$ satisfying $U_T n_T / T n_T \rightarrow 0$, $T \rightarrow \infty$. Let $\hat{\boldsymbol{\psi}}_{2, U_T}^{(BS), c}$ denote the constrained weighted least-squares estimate resulting from a bootstrap sample based on a time series of length U_T and let $\hat{\boldsymbol{\psi}}_2^c$ be the parameter estimate based on the whole sample of size T . Then, conditional on the data $\{\eta(\mathbf{s}, t), \mathbf{s} \in \{1, \dots, m\}^2, t \in \{t_1, \dots, t_T\}\}$,*

$$\left(\frac{U_T^2}{n_{U_T}}\right)^{1/2} \left(\hat{\boldsymbol{\psi}}_{2, U_T}^{(BS), c} - \hat{\boldsymbol{\psi}}_2^c\right) \xrightarrow{d} \begin{cases} Z_1 & \alpha_2^* < 2, \\ Z_2 & \alpha_2^* = 2, \end{cases} \quad m \rightarrow \infty,$$

where $Z_1 \sim \mathcal{N}(0, \Pi_3^{(time)})$ with $\Pi_3^{(time)}$ defined in (5.21) and the distribution function of Z_2 is given in (5.20).

5.4 Some general theory for the spatial extremogram

In this section we state results for the spatial extremogram, which are used to show asymptotic properties of the semiparametric estimates introduced in Section 5.1. First, we introduce α -mixing coefficients for random fields which can be found for instance in Doukhan [35] or Bolthausen [11].

Definition 5.2. Consider a strictly stationary random field $\{X(\mathbf{s}), \mathbf{s} \in \mathbb{R}^d\}$ and define the distances

$$d(\mathbf{s}_1, \mathbf{s}_2) = \max_{1 \leq i \leq d} |s_1(i) - s_2(i)|, \quad \mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d,$$

$$d(\Lambda_1, \Lambda_2) = \inf \{d(\mathbf{s}_1, \mathbf{s}_2), \mathbf{s}_1 \in \Lambda_1, \mathbf{s}_2 \in \Lambda_2\}, \quad \Lambda_1, \Lambda_2 \subset \mathbb{R}^d.$$

Let further $\mathcal{F}_{\Lambda_i} = \sigma\{X(\mathbf{s}), \mathbf{s} \in \Lambda_i\}$, $i = 1, 2$ denote the σ -algebra generated by $X(\mathbf{s})$, $\mathbf{s} \in \Lambda_i$ for $i = 1, 2$. The α -mixing coefficients are defined by

$$\alpha_{k,l}(j) = \sup \left\{ |P(A_1 \cap A_2) - P(A_1)P(A_2)| : A_i \in \mathcal{F}_{\Lambda_i}, |\Lambda_1| \leq k, |\Lambda_2| \leq l, d(\Lambda_1, \Lambda_2) \geq j \right\}, \quad (5.27)$$

where $|\Lambda_i|$ is the cardinality of the set Λ_i , $i = 1, 2$. The random field is called α -mixing, if $\alpha_{k,l}(j) \rightarrow 0$ as $j \rightarrow \infty$ for all $k, l \geq 0$.

Asymptotic normality of the spatial empirical extremogram is proved in Cho et al. [15], where also extensions to the d -dimensional case can be found. We only summarize the case $d = 2$. In accordance we use the following notations.

$$S_m = \{1, \dots, m\}^2$$

$$D_m = \{r = \|\mathbf{s}_1 - \mathbf{s}_2\| : \mathbf{s}_1, \mathbf{s}_2 \in S_m\}$$

$$H[a, b] = \{r \in D_m, a \leq r < b\}, \quad D_\infty = H[0, \infty)$$

$$p(r) = \#\{(s, s_j) : \|s - s_j\| = r, j = 1, \dots, M\}$$

$$c(p) = \sum_{r \in H[1, p]} p(r)$$

$$B(\mathbf{h}, p) = \{s \in S_m : \|s - \mathbf{h}\| \leq p\}$$

$$X(s_r) = \text{one of the } p(r) \text{ neighbours from a fixed point with distance } r.$$

For simplified notation, we enumerate the distances $r \in H[1, p]$ by $r_1, \dots, r_{h(p)}$, where

$h(p) = |H[1, p]|$ denotes the cardinality of $H[1, p]$. Accordingly, we define the sets $D_1, \dots, D_{h(p)}$ by the property,

$$\{\text{vec}\{X(\mathbf{s} + \mathbf{h}), \mathbf{h} \in B(\mathbf{0}, p)\} \in D_i\} = \{X(\mathbf{s}) \in A, X(\mathbf{s}') \in B; \|\mathbf{s} - \mathbf{s}'\| = r_i\},$$

for $i = 1, \dots, h(p)$. We further define the preasymptotic extremogram $\rho_{AB,m}$ for $m \in \mathbb{N}$ by

$$\rho_{AB,m}(r) = \frac{P(a_{n_m}^{-1}X(\mathbf{s}) \in A, a_{n_m}^{-1}X(\mathbf{s} + \mathbf{h}) \in B)}{P(a_{n_m}^{-1}X(\mathbf{s}) \in A)}. \quad (5.28)$$

Proposition 5.4 (Cho et al. [15]). *Suppose the following conditions are satisfied.*

(1) *The vectorized random field $\{\mathbf{Y}(\mathbf{s}), \mathbf{s} \in \mathbb{R}^2\} = \{\text{vec}(X(\mathbf{s} + \mathbf{h}), \mathbf{h} \in B(\mathbf{0}, p)), \mathbf{s} \in \mathbb{R}^2\}$ is strictly stationary, isotropic and all finite dimensional distributions are regularly varying.*

(2) *$\{\mathbf{Y}(\mathbf{s}), \mathbf{s} \in \mathbb{R}^2\}$ is spatially α -mixing with mixing coefficients $\alpha_{k,l}(j)$ defined in (5.27).*

(3) *There exist sequences $n_m \rightarrow \infty$ and $r_m \rightarrow \infty$ with $n_m/m \rightarrow 0$, $r_m/n_m \rightarrow 0$ and $r_m^3/n_m^2 \rightarrow 0$ as $m \rightarrow \infty$, such that*

(3a)

$$n_m \sum_{r \in H(r_m, \infty)} p(r) \alpha_{c(p), c(p)}(r) \rightarrow 0, \quad m \rightarrow \infty.$$

(3b) *For all $\epsilon > 0$ and $k \rightarrow \infty$,*

$$\limsup_{m \rightarrow \infty} n_m \sum_{r \in H(k, r_m)} p(r) P\left(\max_{\mathbf{s} \in B(\mathbf{0}, p)} |X(\mathbf{s})| > \frac{\epsilon a_{n_m}}{c(p)}, \max_{\mathbf{s}' \in B(\mathbf{h}, p)} |X(\mathbf{s}')| > \frac{\epsilon a_{n_m}}{c(p)}\right) \rightarrow 0.$$

$$(3c) \text{ and } \sum_{l=1}^{(m/n_m)^2} 4\alpha_{(n_m-r_m)^2, l n_m^2}(r_m) \rightarrow 0, \quad m \rightarrow \infty.$$

(4) *In addition,*

$$(4a) \quad n_m^5/m^2 \rightarrow 0, \quad m \rightarrow \infty \text{ or}$$

$$(4b) \quad \frac{n_m^7}{m^2} \sum_{r \in H(r_m, \infty)} p(r) \alpha_{1,1}(r) \rightarrow 0 \text{ and } n_m^3 r_m^2/m^2 \rightarrow 0, \quad m \rightarrow \infty.$$

Then, the spatial empirical extremogram centred by the preasymptotic extremogram is asymptotically normal, i.e.

$$\left(\frac{m^2}{n_m}\right)^{1/2} \left(\hat{\rho}_{AB}(r) - \rho_{AB,m}(r)\right)_{r \in H[1,p]} \xrightarrow{d} \mathcal{N}(0, \Pi_1^{(space)}), \quad m \rightarrow \infty, \quad (5.29)$$

where $\rho_{AB,m}$ is defined in (5.28), $\Pi_1^{(space)} = \mu(A)^{-4} F \Sigma F^\top$ with

$$\begin{aligned} \Sigma[i, i] &= \sigma(D_i) = \mu_{B(\mathbf{0}, p)}(D_i) + 2 \sum_{r \in D_\infty \setminus \{0\}} p(r) \tau_{\{B(\mathbf{0}, p) \times B(s_r, p)\}}(D_i \times D_i), \\ \Sigma[i, j] &= r_{D_i, D_j} = \mu_{B(\mathbf{0}, p)}(D_i \cap D_j) + \sum_{r \in D_\infty \setminus \{0\}} p(r) \tau_{\{B(\mathbf{0}, p) \times B(s_r, p)\}}(D_i \times D_j), \\ F_1 &= \mu(A) I_{|H[1,p]|}, \quad F_2 = \text{vec}(-\mu_{B(\mathbf{0}, p)}(D_1), \dots, -\mu_{B(\mathbf{0}, p)}(D_{h(p)})), \\ F &= [F_1, F_2], \end{aligned}$$

and

- $\mu(A)$ is the limit distribution resulting from the relation

$$\mu(A) = \lim_{z \rightarrow \infty} \frac{P(z^{-1} X(s) \in A)}{P(\|Y(s)\| > z)}.$$

- $\mu_{B(\mathbf{0}, p)}$ is defined as limit distribution of the vectorized process Y , i.e.

$$\mu_{B(\mathbf{0}, p)}(D_i) = \lim_{z \rightarrow \infty} \frac{P(z^{-1} Y(s) \in D_i)}{P(\|Y(s)\| > z)}.$$

- $\tau_{\{B(\mathbf{0}, p) \times B(s_r, p)\}}$ is given by

$$\tau_{\{B(\mathbf{0}, p) \times B(s_r, p)\}}(D_i \times D_j) = \lim_{z \rightarrow \infty} \frac{P(Y(s) \in D_i, Y(s_r) \in D_j)}{P(\|\text{vec}\{Y(s), s \in B(\mathbf{0}, r)\}\| > z)}.$$

Remark 5.2. Consider the following condition.

$$(5) \quad \lim_{m \rightarrow \infty} \left(\frac{m^2}{n_m}\right)^{1/2} \left(\frac{P(a_{nm}^{-1} X(s) \in A, a_{nm}^{-1} X(s+\mathbf{h}) \in B)}{P(a_{nm}^{-1} X(s) \in A)} - \rho_{AB}(\|\mathbf{h}\|) \right) = 0,$$

for all spatial distances $\|\mathbf{h}\| = r \in H[1, p]$.

If (5) is satisfied in addition to conditions (1)-(4), then the empirical extremogram centred by the theoretical extremogram is asymptotically normal, i.e.

$$\left(\frac{m^2}{n_m}\right)^{1/2} \left(\hat{\rho}_{AB}(r) - \rho_{AB}(r)\right)_{r \in H[1,p]} \xrightarrow{d} \mathcal{N}(0, \mu(A)^{-4} F \Sigma F^\top), \quad m \rightarrow \infty.$$

We now turn to space-time processes. In the following corollary we show that the temporal mean of empirical spatial extremograms satisfies a central limit theorem under certain conditions. The proof is a simple extension of Corollary 3.3 and Corollary 3.4 in Davis and Mikosch [23].

Corollary 5.4. *Consider the space-time process $\{X(\mathbf{s}, t), \mathbf{s} \in \mathbb{R}^d, t \in [0, \infty)\}$. Assume we observe the process at time points $t \in \{t_1, \dots, t_T\}$. If the vectorized process*

$$(\mathbf{Y}(\mathbf{s})) = \text{vec}\{(X(\mathbf{s} + \mathbf{h}, t_1), \dots, X(\mathbf{s} + \mathbf{h}, t_T)), \mathbf{h} \in B(\mathbf{0}, p)\}, \quad \mathbf{s} \in \mathbb{R}^2 \quad (5.30)$$

satisfies the assumptions as in Proposition 5.4, then

$$\left(\frac{m^2}{n_m}\right)^{1/2} \left(\frac{1}{T} \sum_{k=1}^T \hat{\rho}_{AB}^{(t_k)}(r, 0) - \rho_{AB,m}(r, 0)\right)_{r \in H[1,p]} \xrightarrow{d} \mathcal{N}(0, \Pi_2^{(space)}), \quad m \rightarrow \infty, \quad (5.31)$$

where $\hat{\rho}_{AB}^{(t_k)}(r, 0)$ is the empirical spatial extremogram based on $\{\eta(\mathbf{s}_i, t_k), i = 1, \dots, M\}$ for fixed time point $t_k, k = 1, \dots, T$ and $\Pi_2^{(space)}$ is the covariance matrix specified in (5.32).

Proof. We redefine the sets D_r for each spatial distance $r \in H[1, p]$ by including the time component. In particular, define $D_{r,k}$ for $r = 1, \dots, h(p)$ and $k = 1, \dots, T$ by

$$\{\mathbf{Y}(\mathbf{s}) \in D_{r,k}\} = \text{vec}\{X(\mathbf{s}, t_k) \in A, X(\mathbf{s}_1, t_k) \in B; \|\mathbf{s} - \mathbf{s}_1\| = r\},$$

and $C_k = D_{h(p)+1,k}$ by

$$\{\mathbf{Y}(\mathbf{s}) \in D_{h(p)+1,k}\} = \{X(\mathbf{s}, t_k) \in A\}, \quad k = 1, \dots, T.$$

Define the matrices $\Sigma \in \mathbb{R}^{T(h(p)+1) \times T(h(p)+1)}$ by

$$\Sigma = \begin{pmatrix} \sigma^2(D_{1,1}) & \cdots & r_{D_{1,1}, D_{1,T}} & \cdots & r_{D_{1,1}, D_{h(p)+1,1}} & \cdots & r_{D_{1,1}, D_{h(p)+1,T}} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ r_{D_{1,1}, D_{h(p)+1,T}} & \cdots & r_{D_{1,T}, D_{h(p)+1,T}} & \cdots & r_{D_{h(p)+1,1}, D_{h(p)+1,T}} & \cdots & \sigma^2(D_{h(p)+1,T}) \end{pmatrix},$$

and $F \in \mathbb{R}^{T(h(p)+1) \times T(h(p)+1)}$ by

$$F = \begin{pmatrix} F_1 & \mathbf{0} & \cdots & \mathbf{0} & F_2^{(1)} \\ \mathbf{0} & F_1 & \cdots & \mathbf{0} & F_2^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & F_1 & F_2^{(h(p))} \end{pmatrix},$$

where

$$F_1 = \begin{pmatrix} \frac{1}{\mu(C_1)} & 0 & \cdots & 0 \\ 0 & \frac{1}{\mu(C_2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\mu(C_T)} \end{pmatrix}, \quad F_2^{(i)} = \begin{pmatrix} -\frac{\mu(D_{i,1})}{\mu(C_1)^2} & 0 & \cdots & 0 \\ 0 & -\frac{\mu(D_{i,2})}{\mu(C_2)^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -\frac{\mu(D_{i,T})}{\mu(C_T)^2} \end{pmatrix},$$

for $i = 1, \dots, h(p)$. The covariance terms $r_{D_{i,j}, D_{k,l}}$ and the limit distributions μ are defined as in Proposition 5.4 but using the extended vector process $\{Y(s)\}$ in (5.30). Using the proof of Corollary 3.3 in Davis and Mikosch [23], it follows that

$$\left(\frac{m^2}{n_m} \right)^{1/2} \begin{pmatrix} \hat{\rho}_{AB}^{(t_1)}(r, 0) - \rho_{AB,m}(r, 0) \\ \vdots \\ \hat{\rho}_{AB}^{(t_T)}(r, 0) - \rho_{AB,m}(r, 0) \end{pmatrix}_{r \in H[1,p]} \xrightarrow{d} \mathcal{N}(0, F \Sigma F^\top) \quad m \rightarrow \infty,$$

and by using the Cramér Wold device the statement in (5.31) follows with

$$\Pi_2^{(\text{space})} = \frac{1}{T^2} \begin{pmatrix} 1 \cdots 1 & 0 \cdots 0 & \cdots & 0 \cdots 0 \\ 0 \cdots 0 & 1 \cdots 1 & \cdots & 0 \cdots 0 \\ & & \ddots & \\ 0 \cdots 0 & \cdots & \cdots & 1 \cdots 1 \end{pmatrix} F \Sigma F^\top \begin{pmatrix} 1 \cdots 1 & 0 \cdots 0 & \cdots & 0 \cdots 0 \\ 0 \cdots 0 & 1 \cdots 1 & \cdots & 0 \cdots 0 \\ & & \ddots & \\ 0 \cdots 0 & \cdots & \cdots & 1 \cdots 1 \end{pmatrix}^\top. \quad (5.32)$$

□

CHAPTER 6

SIMULATION STUDY

In this chapter we evaluate pairwise likelihood estimation and the semiparametric estimation procedure for the Brown-Resnick process introduced in Chapter 3. The construction of the max-stable process was introduced in Theorem 3.1, and is based on taking pointwise maxima of independent replications of Gaussian space-time processes. This allows for an easy simulation setup, provided that we can simulate from the Gaussian processes.

6.1 Setup for simulation study

We illustrate the small sample behaviour of the two proposed estimation methods for spatial dimension $d = 2$ in a simulation experiment. The setup for this study is:

1. The spatial locations consisted of a 10×10 grid

$$S_{10} = \{s_{(i_1, i_2)} = (i_1, i_2), i_1, i_2 \in \{1, \dots, 10\}\}.$$

The time points are chosen equidistantly, $\{1, \dots, T = 200\}$.

2. One thousand independent Gaussian space-time processes $Z_j(s_n s, t_n t)$, $j = 1, \dots, n = 1000$ were generated using the R-package `RandomFields` with covariance function $\gamma(s_n \mathbf{h}, t_n u)$, where the circulant embedding procedure introduced in Section 2.2.2 is implemented. We use the following correlation function for the underlying Gaussian

random field.

$$\gamma(\mathbf{h}, u) = (1 + \theta_1 \|\mathbf{h}\|^{\alpha_1} + \theta_2 |u|^{\alpha_2})^{-3/2}.$$

Assumption 4.1 is fulfilled and the limit function δ is given by

$$\lim_{n \rightarrow \infty} (\log n)(1 - \gamma(s_n \mathbf{h}, t_n u)) = \delta(\mathbf{h}, u) = \frac{3}{2} \theta_1 \|\mathbf{h}\|^{\alpha_1} + \frac{3}{2} \theta_2 |u|^{\alpha_2},$$

where $s_n = (\log n)^{-1/\alpha_1}$ and $t_n = (\log n)^{-1/\alpha_2}$.

3. The simulated processes were transformed to standard Fréchet margins using the transformation $-1 / \log(\Phi(Z_j(\mathbf{s}, t)))$ for $\mathbf{s} \in S_{10}$ and $t \in \{1, \dots, T\}$.
4. The pointwise maximum of the transformed Gaussian random fields was computed and rescaled by $1/n$ to obtain an approximation of a max-stable random field, i.e.

$$\eta(\mathbf{s}, t) = \frac{1}{1000} \bigvee_{j=1}^{1000} \frac{1}{\log(\Phi(Z_j(s_n \mathbf{s}, t_n t)))}, \quad \mathbf{s} \in S_{10}, \quad t \in \{1, \dots, T\}.$$

5. The parameters $\theta_1, \alpha_1, \theta_2$ and α_2 were estimated by using pairwise likelihood estimation and the semiparametric procedure.
6. Steps (a)-(e) were repeated 100 times.

Note first, that we only get an approximation of a Brown-Resnick process since we cannot choose $n = \infty$. The true parameter vector for the simulation study equals

$$\boldsymbol{\psi}^* = (\theta_1^*, \alpha_1^*, \theta_2^*, \alpha_2^*) = (0.06, 1, 0.04, 1).$$

6.2 Results for pairwise likelihood estimation

We start by presenting the results for the pairwise likelihood estimation. The computation of the pairwise likelihood function is implemented in C and used as function in R. The code is adapted such that it takes care of identifiability issues, when some of the parameters are not identifiable, cf. Remark 4.1. Figures 6.2.1 and 6.2.2 show the resulting estimates as a function of (r, p) . Figure 6.2.1 shows boxplots of the resulting estimates for the spatial parameters θ_1 and α_1 . The horizontal axis shows the different maximal space-time lags included in the

pairwise likelihood function from (4.11). We also show qq-plots against a normal distribution for all parameters and different combinations of r and p in Figure 6.2.3. In addition to the graphical output we calculate the root mean square error (RMSE) and the mean absolute error (MAE) to see how the choice of (r, p) influences the estimation.

We make the following observations. As already pointed out by Davis and Yau [24] and Huser and Davison [46], there might be a loss in efficiency if too many pairs are included in the estimation. This can be explained by the fact that pairs get more and more independent as the space-time lag increases. Adding more pairs to the pairwise log-likelihood function can introduce some noise which decreases the efficiency. This is evident in Figure 6.2.2 for the temporal parameter α_2 , where the estimates vary more around the mean as more pairs are included in the estimation.

An interesting observation for our model is that using a maximal spatial lag of 0 or a maximal temporal lag 0, respectively, leads to very good results. For the spatial parameters, the space-time lags which lead to the lowest RMSE and MAE are $(2, 0)$ for θ_1 and $(2, 0)$ (RMSE) or $(3, 0)$ (MAE) for α_2 (see Table 6.1), i.e. we use all pairs within a spatial distance of 2 or 3 at the same time point. Basically, this suggests that we could also estimate the spatial parameters based on each individual random field for fixed time points and then take the mean over all estimates in time. The same holds for the time parameters θ_2 and α_2 , where the best results in the sense of the lowest RMSE and MAE are obtained for the space-time lags $(0, 3)$, i.e. if we use all pairwise densities corresponding to the space-time pairs (s, t_1) and (s, t_2) , where $|t_2 - t_1| \leq 3$ (see Table 6.2). The reason for this observation is that the parameters of the underlying space-time correlation function get “separated” in the extremal setting in the sense that for example a spatial lag equal to zero does not affect the temporal parameters θ_1 and α_1 and vice versa.

θ_1	(1,0)	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(2,0)	(2,1)	(2,2)
RMSE	0.010	0.012	0.012	0.012	0.012	0.012	0.010	0.010	0.011
MAE	0.007	0.009	0.009	0.009	0.009	0.010	0.008	0.008	0.008
	(2,3)	(2,4)	(2,5)	(3,0)	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)
RMSE	0.010	0.010	0.010	0.011	0.011	0.011	0.011	0.011	0.011
MAE	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008
α_1	(2,0)	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(3,0)	(3,1)	(3,2)
RMSE	0.134	0.140	0.153	0.149	0.154	0.157	0.135	0.141	0.158
MAE	0.108	0.112	0.115	0.114	0.123	0.125	0.105	0.111	0.113
	(3,3)	(3,4)	(3,5)	(4,0)	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)
RMSE	0.160	0.164	0.165	0.142	0.148	0.161	0.167	0.174	0.175
MAE	0.123	0.129	0.130	0.112	0.118	0.111	0.128	0.137	0.139

Table 6.1: RMSE and MAE based on 100 simulations for the spatial estimates θ_1 and α_1 for different combinations of maximal space-time lags (r, p) .

$\hat{\theta}_2$	(0,1)	(0,2)	(0,3)	(0,4)	(0,5)	(1,1)	(1,2)	(1,3)	(1,4)
RMSE	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018
MAE	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017
	(1,5)	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(3,1)	(3,2)	(3,3)
RMSE	0.018	0.019	0.019	0.019	0.019	0.019	0.019	0.019	0.019
MAE	0.017	0.018	0.017	0.017	0.017	0.017	0.018	0.018	0.017
$\hat{\alpha}_2$	(0,2)	(0,3)	(0,4)	(0,5)	(1,2)	(1,3)	(1,4)	(1,5)	(2,2)
RMSE	0.132	0.127	0.128	0.129	0.144	0.140	0.143	0.144	0.147
MAE	0.110	0.099	0.102	0.104	0.109	0.108	0.114	0.115	0.118
	(2,3)	(2,4)	(2,5)	(3,2)	(3,3)	(3,4)	(3,5)	(4,2)	(4,3)
RMSE	0.153	0.158	0.162	0.147	0.153	0.159	0.164	0.155	0.161
MAE	0.122	0.128	0.129	0.117	0.122	0.127	0.132	0.123	0.128

Table 6.2: RMSE and MAE based on 100 simulations for the spatial estimates θ_2 and α_2 for different combinations of maximal space-time lags (r, p) .

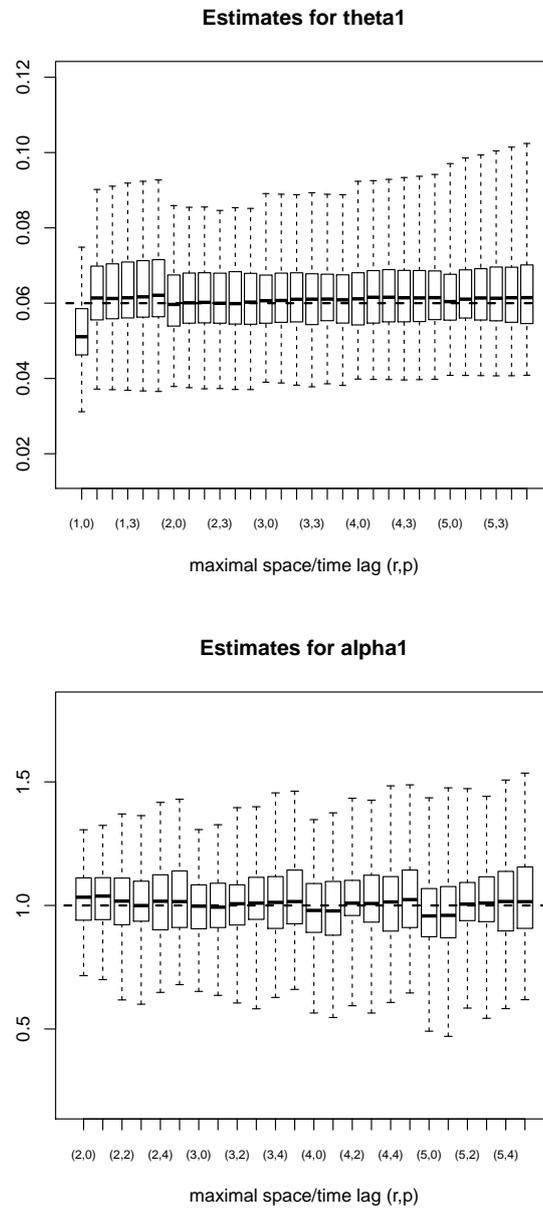


Figure 6.2.1: Pairwise likelihood estimates for θ_1 and α_1 (spatial parameters) as a function of maximal space-time lags (r, p) . Each boxplot represents the estimates for 100 simulations. The dashed line is the true value.

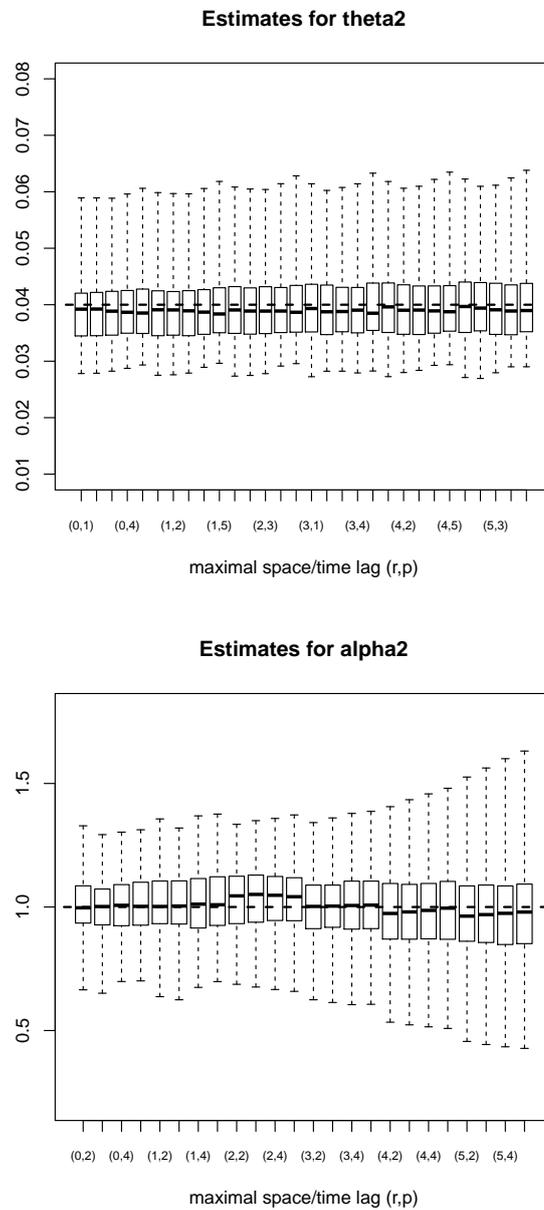


Figure 6.2.2: Pairwise likelihood estimates for θ_2 and α_2 (spatial parameters) as a function of maximal space-time lags (r, p) . Each boxplot represents the estimates for 100 simulations. The dashed line is the true value.

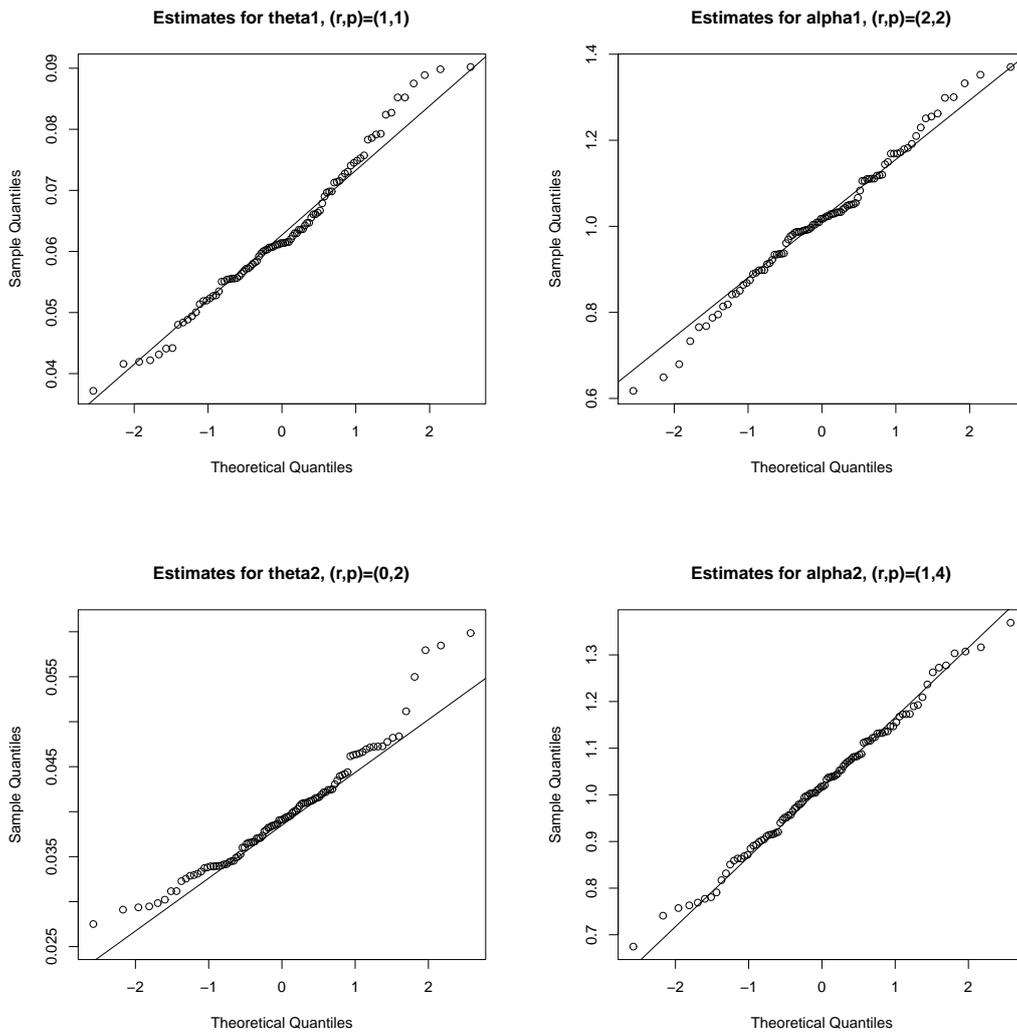


Figure 6.2.3: QQ-plots of the pairwise likelihood estimates against the normal distribution, where for each parameter we chose a random combination of r and p .

6.3 Results for the semiparametric estimation procedure using the extremogram

We present results for the semiparametric estimation procedure introduced and analysed in Chapter 5. For the empirical extremogram we choose the 70%-quantile, which allows for enough observations included in the estimation. Since we assume that data arises from a max-stable process, i.e. that we have extreme data, we choose a threshold which is lower than in the usual case. The weights in the constrained weighted linear regression problem (see (5.6) and (5.7)) are chosen as $w_i = \exp\{-i^2\}$ for $i = u \in \mathcal{U}$ or $i = r \in \mathcal{H}$. For the spatial and temporal parameters 500 bootstrap samples are constructed separately. In Section 5.3.2 we showed theoretical results for a bootstrap procedure which takes care of the constraints on α_1 or α_2 . Here, we ignore the constraints since it is not possible to simulate the max-stable space-time processes with a reasonable number of observations to use subsamples for the bootstrap samples. Figure 6.3.1 shows the estimates for the spatial parameters θ_1 and α_1 and Figure 6.3.2 visualizes the temporal estimates for θ_2 and α_2 . Each dot represents one estimate based on one simulated max-stable space-time process. The dotted lines above and below are pointwise confidence intervals based on the bootstrap sample. They were calculated by using quantiles, i.e. for $\psi = \theta_1, \alpha_1, \theta_2, \alpha_2$,

$$CI = \left[\hat{\psi} - \hat{q}\left(1 - \frac{\beta}{2}\right), \hat{\psi} - \hat{q}\left(\frac{\beta}{2}\right) \right],$$

where \hat{q} is the empirical quantile of the bootstrap sample and β is the level of significance. Table 6.3 shows the mean, RMSE and MAE of the simulations. Altogether, we observe that the estimates are close to the true values, and that the bootstrap method gives accurate pointwise confidence intervals.

	MEAN	RMSE	MAE
θ_1	0.0584	0.0106	0.0088
α_1	1.1016	0.2375	0.1897
θ_2	0.0423	0.0166	0.0139
α_2	0.9509	0.1154	0.0876

Table 6.3: Mean, root mean squared error and mean absolute squared error for the semiparametric estimates

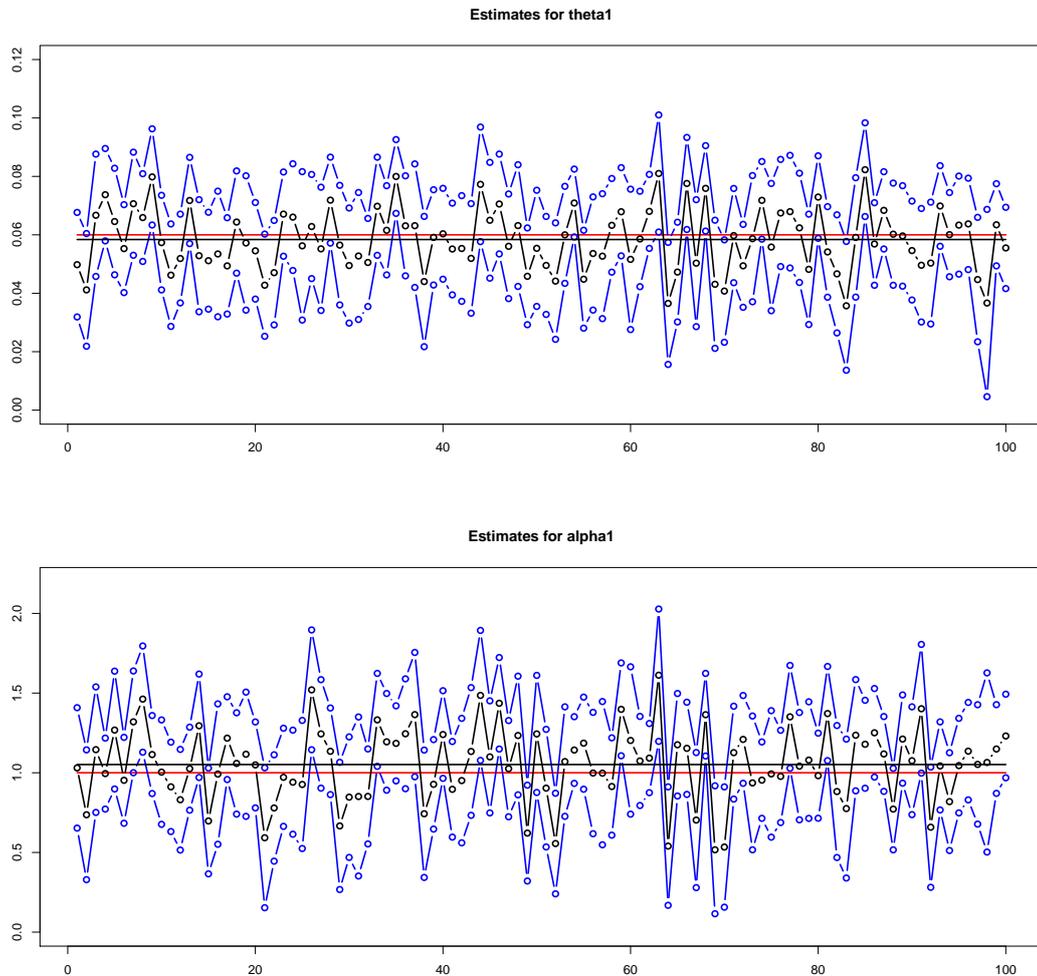


Figure 6.3.1: Semiparametric estimates for θ_1 (top) and α_1 (bottom) based on 100 simulated max-stable processes together with pointwise 95%- bootstrap confidence intervals (blue). The middle red line is the true value and the middle black line represents the mean over all estimates.

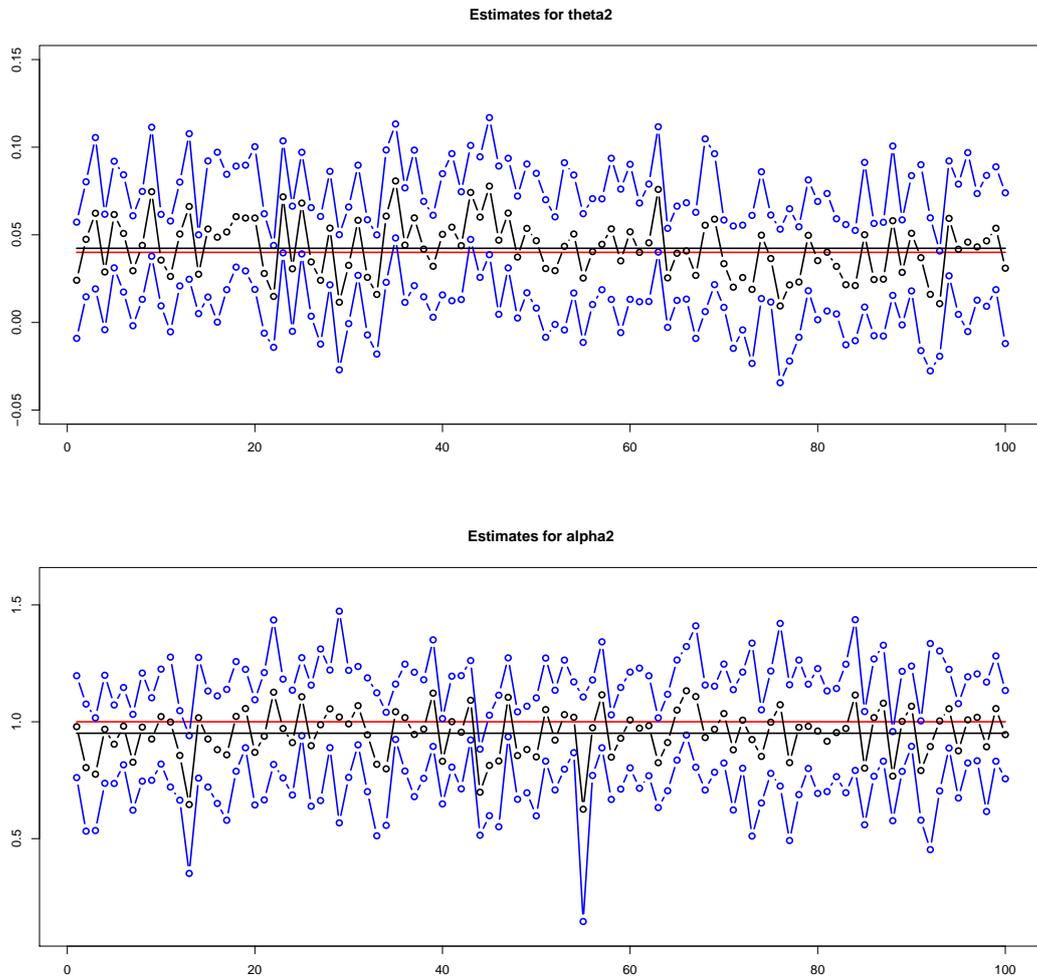


Figure 6.3.2: Semiparametric estimates for θ_2 (top) and α_2 (bottom) based on 100 simulated max-stable processes together with pointwise 95%- bootstrap confidence intervals (blue). The middle red line is the true value and the middle black line represents the mean over all estimates.

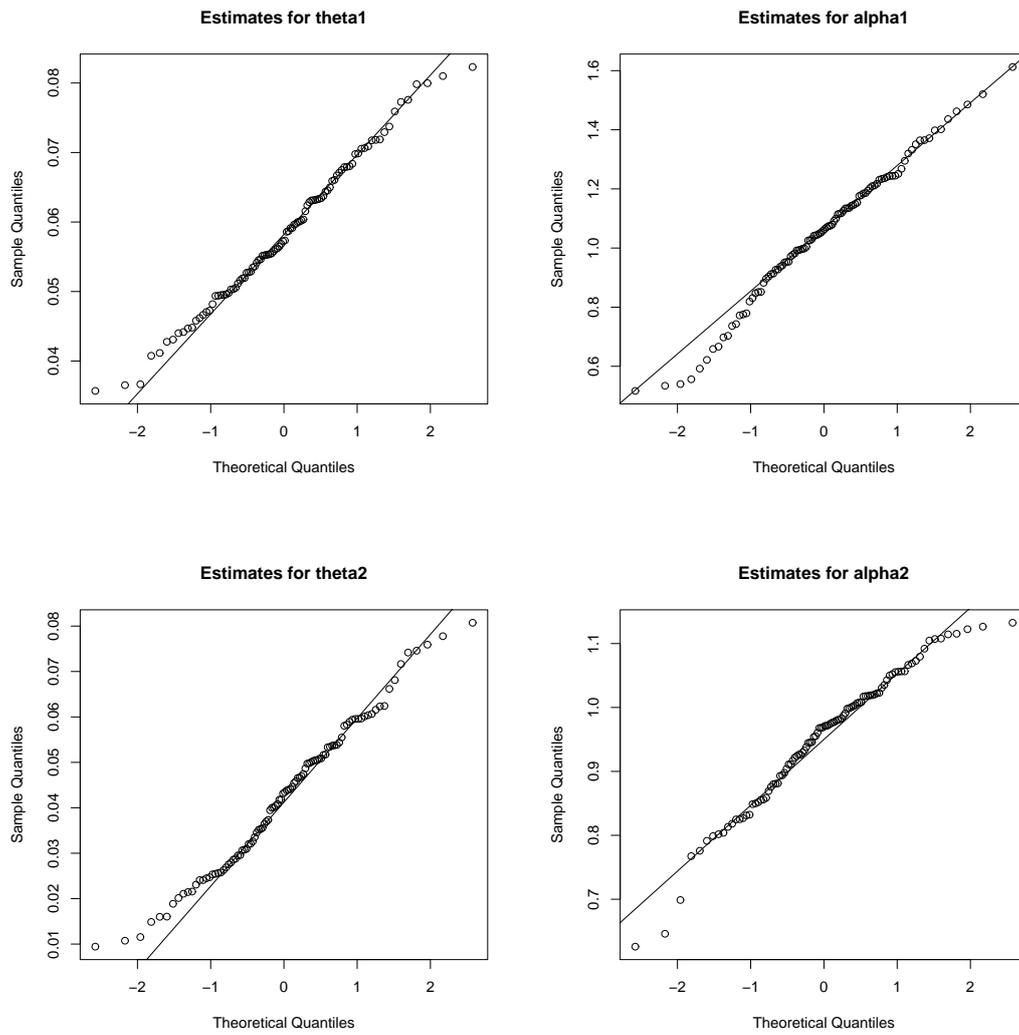


Figure 6.3.3: QQ-plots of the semiparametric estimates against the normal distribution.

6.4 Comparison of pairwise likelihood and semiparametric estimation

As a last step in the simulation study, we compare the estimates resulting from the two different methods. Figures 6.4.1 and 6.4.2 contrast the pairwise likelihood estimates with the semiparametric estimates. On the left hand side, the estimates (blue: pairwise likelihood, red: semiparametric) are plotted for each of the 100 simulations together with the true value as solid line in the middle. On the right hand side, boxplots resulting from the simulations are shown. We see that the semiparametric estimates differ more from the true value than the pairwise likelihood estimates, which can be explained from the bias in the spatial and temporal empirical extremogram, which cannot be cancelled completely, since the threshold used in the estimation cannot be set equal to infinity. The variability around the true value differs for the different parameters, and no particular pattern can be detected.

The semiparametric estimates rely on the choice of the threshold for the empirical extremogram, where a tradeoff between a high threshold and enough observations has to be made. We measured the time needed to estimate one set of parameters in the simplest case, i.e. for the pairwise likelihood estimation we use $(r, p) = (0, 2)$ to estimate θ_2 and α_2 and $(r, p) = (2, 0)$ for θ_1 and α_1 . Table 6.4 compares the computation time in seconds measured for the calculation on a Laptop (Windows 7, 64 bit, Intel(R) Core(TM) i5 CPU, 2.53 GHz) using the statistic software R. We see that the semiparametric estimation is much faster than the pairwise likelihood estimation.

	Pairwise likelihood $(r, p) = (2, 0)$ and $(r, p) = (0, 2)$	Semiparametric
Computation time	34.3	2.2

Table 6.4: Computation time in seconds for the estimation of the parameters $\theta_1, \alpha_1, \theta_2$ and α_2 .

To conclude this section, we make some summarizing comments.

Pairwise likelihood estimation:

- The pairwise likelihood estimation serves as reliable method to estimate parameters in max-stable process.

- Accurate starting values are needed for the optimization routine, which maximizes the log-likelihood function.
- The computation time is rather high compared to the semiparametric estimation.

Semiparametric estimation:

- The semiparametric estimates show a larger bias than the pairwise likelihood estimates and are sensible to the choice of the threshold used for the extremogram.
- The computation time is around 15 times lower than for the pairwise likelihood estimation.
- The implementation of bootstrap confidence intervals is feasible and gives regions of confidence in a reasonable amount of time.
- The semiparametric estimates could serve as starting values for the optimization routine used to maximize the pairwise log-likelihood function.

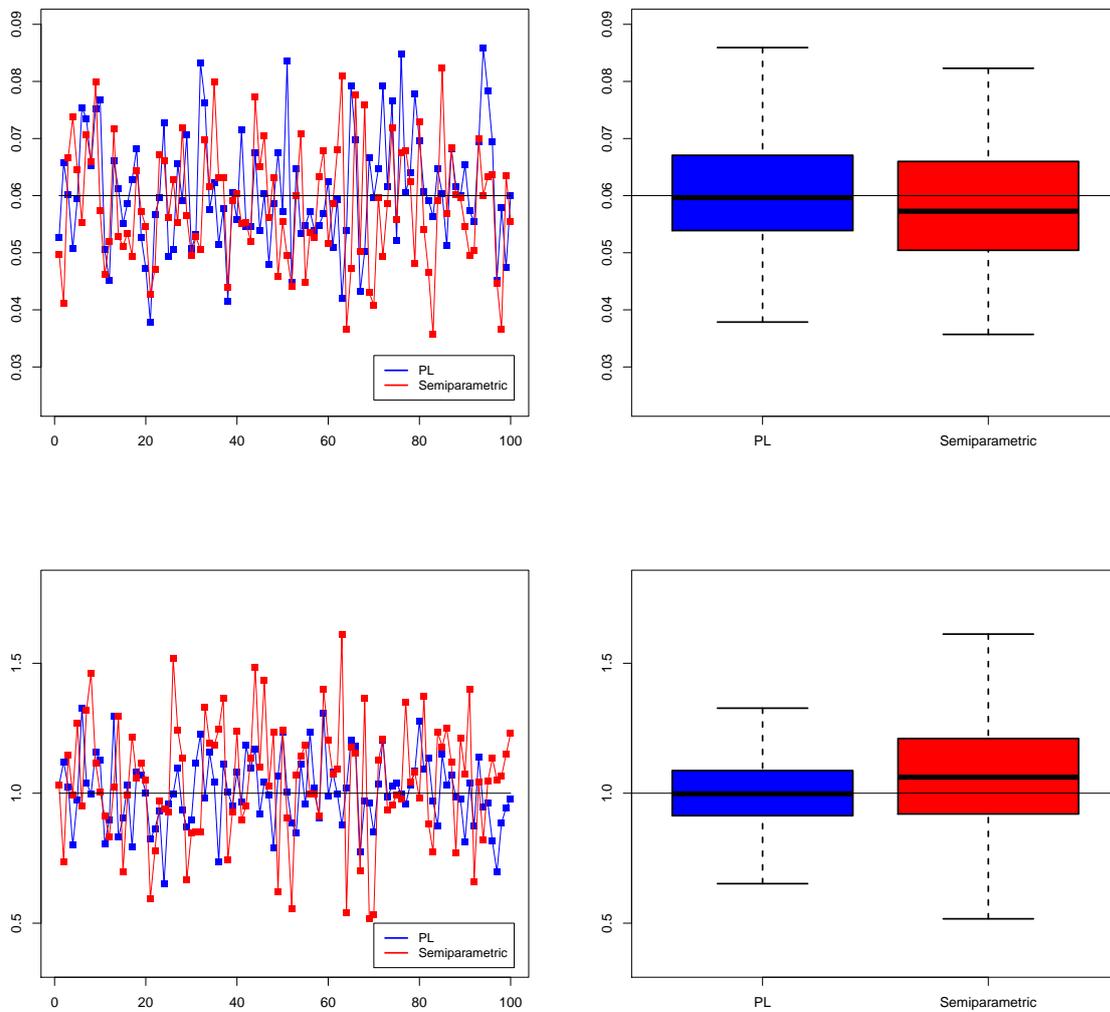


Figure 6.4.1: Comparison of estimates for θ_1 (top) and α_1 (bottom), where in the left plots each dot (blue: pairwise likelihood, red: semiparametric) corresponds to one of the 100 simulations. The plots on the right hand side show the corresponding boxplots resulting from the simulations.

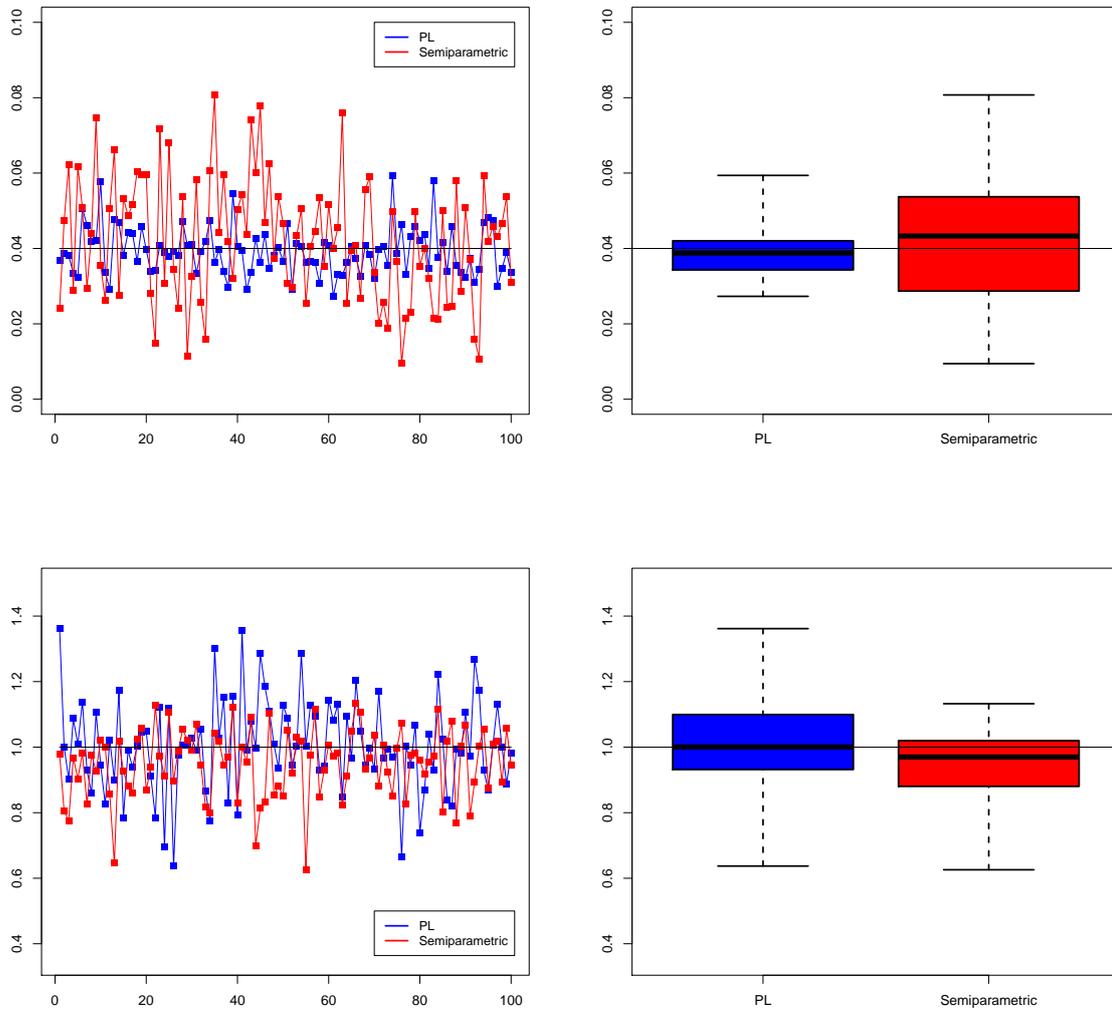


Figure 6.4.2: Comparison of estimates for θ_2 (top) and α_2 (bottom), where in the left plots each dot (blue: pairwise likelihood, red: semiparametric) corresponds to one of the 100 simulations. The plots on the right hand side show the corresponding boxplots resulting from the simulations.

CHAPTER 7

ANALYSIS OF RADAR RAINFALL MEASUREMENTS IN FLORIDA

The final step in this thesis is the application of our introduced model and methods to real data. In particular, we quantify the extremal behaviour of radar rainfall data in a region in Florida by using spatial block maxima and two different time domains, namely daily maxima and hourly measurements.

7.1 Description of data set

The rainfall data we use in this study was collected by the Southwest Florida Water Management District (SWFWMD) ¹. The data base consists of radar values measured on a grid with size 2 km covering a region in Florida. A map of the area in Florida is shown in Figure 7.1.1. Radar rainfall observations are given in inches as 15-minutes increments from 1999 to 2004. To present our methods we choose 60×60 locations in the middle of the considered region (see red square in Figure 7.1.1) resulting in a $120 \text{ km} \times 120 \text{ km}$ squared region. We calculate the accumulated hourly rainfall measurements and take the maximum over 25 locations lying on a square with side length of 10 km. In that way we obtain 12×12 locations with spatial maxima of rainfall observations. The corresponding time series for each spatial

¹<http://www.swfwmd.state.fl.us/>

maxima is used for further analysis. In the following two sections we consider two different settings for the time domain taken into account. In both settings we only consider the wet season in Florida (June to September). For simplified notation we denote the set of locations by $S = \{(i_1, i_2), i_1, i_2 \in \{1, \dots, 12\}\}$.

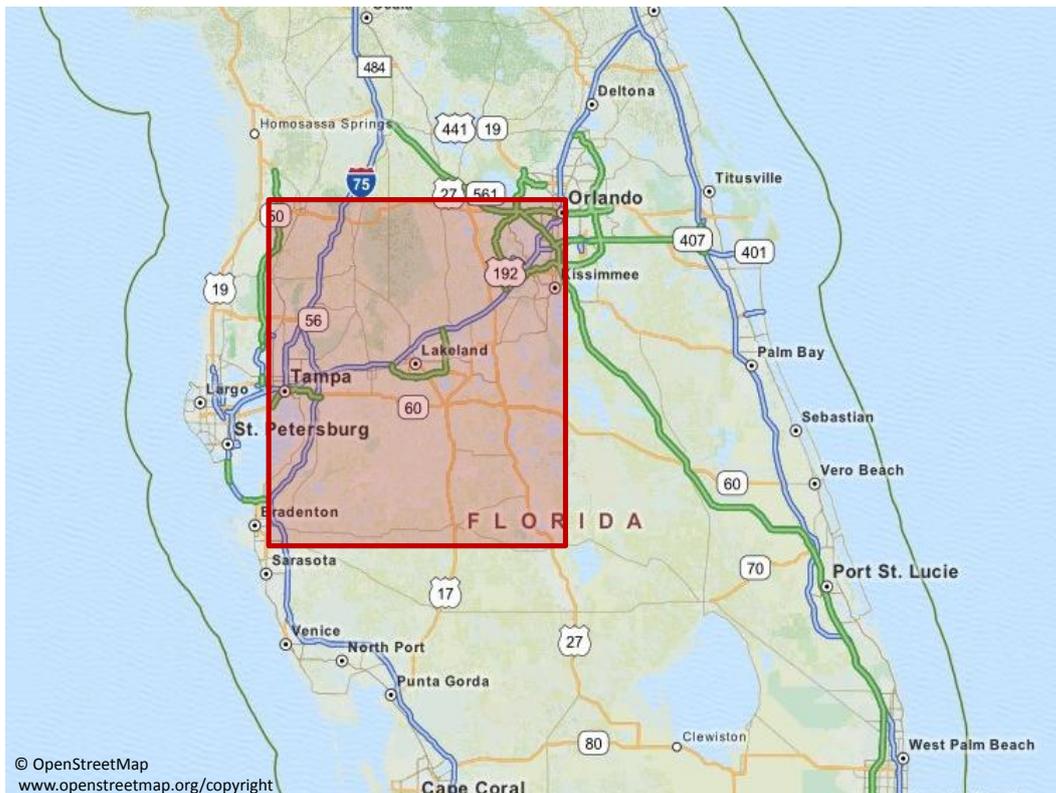


Figure 7.1.1: Map of Florida with region of observations.

7.2 Daily maxima of rainfall measurements

We start by analysing daily maxima of rainfall measurements. Since we take maxima in space and time, the assumption of a max-stable space-time process becomes reasonable. As a first step we remove a seasonal component by using a simple moving average with a period of 122 days, which is the number of days in the wet season considered in one particular year. By inspecting autocorrelation plots for the daily maxima with fixed location, we conclude that there is no temporal dependence in the time series. As described in Section 2.1.1 we fit the generalized extreme value distribution to the block maxima for each fixed location. Since the estimates for the shape parameter are close to zero with confidence intervals containing

zero, we directly fit a Gumbel distribution (GEV with $\xi = 0$) to the data. The observations are then transformed to standard Fréchet margins by using the probability integral transform. Afterwards, we estimate the parameters of the max-stable Brown-Resnick process to obtain the extremal dependence structure. We explain the procedure in more detail.

1. Marginal modelling: For each fixed location $\mathbf{s} \in S$ we follow the steps below.

a) The time series $\{\eta(\mathbf{s}, t), t \in \{t_1, \dots, t_{732}\}\}$ is deseasonalized by calculating

$$\tilde{\eta}(\mathbf{s}, t_{k+122(j-1)}) = \eta(\mathbf{s}, t_{k+122(j-1)}) - \frac{1}{6} \sum_{j=1}^6 \eta(\mathbf{s}, t_{k+122(j-1)}) \quad (7.1)$$

for $k = 1, \dots, 122$.

b) The deseasonalized observations are fitted to a Gumbel distribution with distribution function

$$GEV_{\mu(\mathbf{s}), \sigma(\mathbf{s}), 0}(x) = \exp\left\{-\exp\left\{-\frac{x - \mu(\mathbf{s})}{\sigma(\mathbf{s})}\right\}\right\}. \quad (7.2)$$

c) The data are transformed according to

$$\tilde{\tilde{\eta}}(\mathbf{s}, t) = -\frac{1}{\log(GEV_{\hat{\mu}(\mathbf{s}), \hat{\sigma}(\mathbf{s}), 0}(\tilde{\eta}(\mathbf{s}, t)))}, \quad t \in \{t_1, \dots, t_{732}\},$$

where $\hat{\mu}(\mathbf{s})$ and $\hat{\sigma}(\mathbf{s})$ are the parameter estimates resulting from fitting the Gumbel distribution.

2. Estimating the extremal dependence structure: We assume that $\tilde{\tilde{\eta}}(\mathbf{s}, t), \mathbf{s} \in S, t \in \{t_1, \dots, t_{732}\}$ are realizations from the Brown-Resnick process (cf. (4.2))

$$\tilde{\tilde{\eta}}(\mathbf{s}, t) = \bigvee_{j=1}^{\infty} \xi_j \exp\{W_j(\mathbf{s}, t) - \delta(\mathbf{s}, t)\},$$

where W_j are independent replications of a Gaussian process with stationary increments and correlation function $\delta(\mathbf{s}_1, t_1) + \delta(\mathbf{s}_2, t_2) - \delta(\mathbf{s}_1 - \mathbf{s}_2, t_1 - t_2)$. The δ function is assumed to equal $\delta(\mathbf{s}, t) = \theta_1 \|\mathbf{s}\|^{\alpha_1} + \theta_2 |t|^{\alpha_2}$ and contains the extremal dependence parameters $\theta_1, \alpha_1, \theta_2$ and α_2 , which are estimated using the following two procedures.

- a) We use the semiparametric estimation procedure introduced in Section 5.1 to estimate the parameters and construct 500 bootstrap samples to obtain pointwise confidence intervals for the estimated parameters.
- b) In addition, we estimate the parameters using the pairwise likelihood estimation procedure developed in Section 4.2.

The results are summarized in Figures 7.2.2, 7.2.3, 7.2.4, 7.2.5, 7.2.6, Tables 7.1 and 7.2. Figure 7.2.2 shows time series for the daily maxima of four representative locations, indicated in Figure 7.2.1. Three of the locations are safely inside the study region and one is on the boundary. The fit of the Gumbel distribution is supported by qq-plots in Figure 7.2.3, which show a straight line pattern. In Figure 7.2.4 plots of the original rainfall fields (left) and the standard Fréchet transformations are shown for two arbitrary fixed time points t_{166} (July 14th 2000) and t_{434} (August 7th 2002). The parameter estimates are given in Tables 7.1 and 7.2. The temporal estimates $\hat{\theta}_2$ and $\hat{\alpha}_2$ indicate that there is little or no temporal extremal dependence. In particular, recall from (5.2) that the extremogram for a spatial lag equal to zero is given by

$$\chi(0, u) = 2(1 - \Phi(\sqrt{\theta_2|u|^{\alpha_2}}))$$

for temporal lags $u > 0$. If α_2 is near zero then $\chi(0, u)$ is approximately constant indicating that the extremal dependence is the same for all temporal lags $u > 0$. If, in addition, θ_2 is large $\chi(0, u)$ is close to zero indicating asymptotic independence. So, the combination of α_2 small and θ_2 large implies asymptotic independence. Figure 7.2.5 shows the empirical spatial (left) and temporal (right) extremogram, which are used to estimate the parameters. Note, that the plots show the temporal mean of spatial extremograms (left) and the spatial mean of temporal extremograms (right) as described in (2a) and (2b) in Section 5.1. In addition, we include pointwise 95%-bootstrap confidence intervals (red lines) calculated based on 500 bootstrap samples. Recall from (5.32) that the asymptotic variance for the temporal mean of spatial extremograms includes the factor $1/T^2$ which reduces the standard error. Since we have 732 time points, the confidence intervals for the spatial extremogram are narrow. To check whether the extremal dependence is significant, we permute the space-time data and calculate empirical extremograms as before. For the temporal extremogram, we fix the one location s and sample from the corresponding time series without replacement. Using the same indices as for s we sample the other time series accordingly to keep the spatial dependence structure, i.e. for all locations $s \in S$ the permuted space-time observations are given by $\{(s, \kappa(t_1)), \dots, (s, \kappa(t_{732}))\}$, where κ describes one permutation. We calculate the empiri-

cal temporal extremogram as before and repeat the procedure 1000 times. 97.5% and 2.5% empirical quantiles are calculated from the resulting temporal extremogram sample which gives a 95% confidence region for temporal extremal independence. To see how the extremal dependence lasts in space, we do the same for the spatial extremogram. In particular, for each fixed time point we permute the spatial locations and estimate the empirical spatial extremogram for the permuted space-time data. Figure 7.2.6 shows the estimates together with the independence confidence intervals. Since the estimate for the temporal extremogram lies within the obtained confidence region, we conclude that there is no temporal extremal dependence for the daily maxima of rainfall. We also conclude that there is no spatial extremal dependence for spatial lags larger than 4.

Estimate	$\hat{\theta}_1$	0.3496	$\hat{\alpha}_1$	0.9040
Bootstrap-CI		[0.3254,0.3640]		[0.8183,0.9897]
Estimate	$\hat{\theta}_2$	2.2803	$\hat{\alpha}_2$	0.0951
Bootstrap-CI		[1.8810,2.6796]		[0.000,0.2161]

Table 7.1: Semiparametric estimates for the spatial parameters θ_1 and α_1 and the temporal parameters θ_2 and α_2 for the Brown-Resnick process in (4.2) together with 95% bootstrap confidence intervals.

PL estimates	$\hat{\theta}_1$	$\hat{\alpha}_1$	$\hat{\theta}_2$	$\hat{\alpha}_2$
	0.3485	0.8858	2.419	0.1973

Table 7.2: Pairwise likelihood estimates for the parameters in the Brown-Resnick process in (4.2).

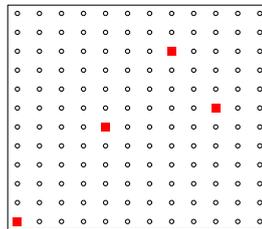


Figure 7.2.1: Reference locations for time series plots in Figure 7.2.2.

7 Analysis of radar rainfall measurements in Florida

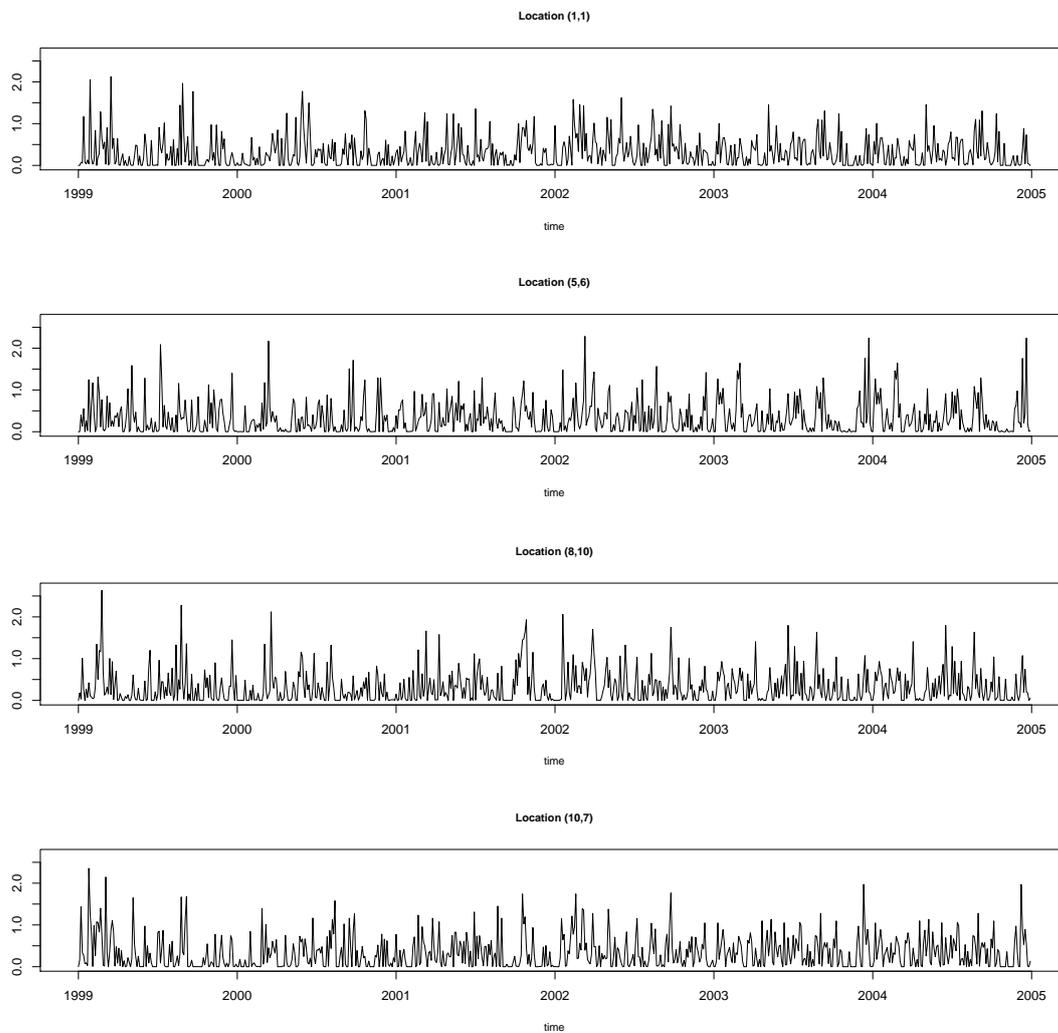


Figure 7.2.2: Plot of daily maximal rainfall (calculated based on the 15 minutes measurements) in inches for four different grid locations with simplified coordinates (1,1), (5,6), (8,10) and (10,7).

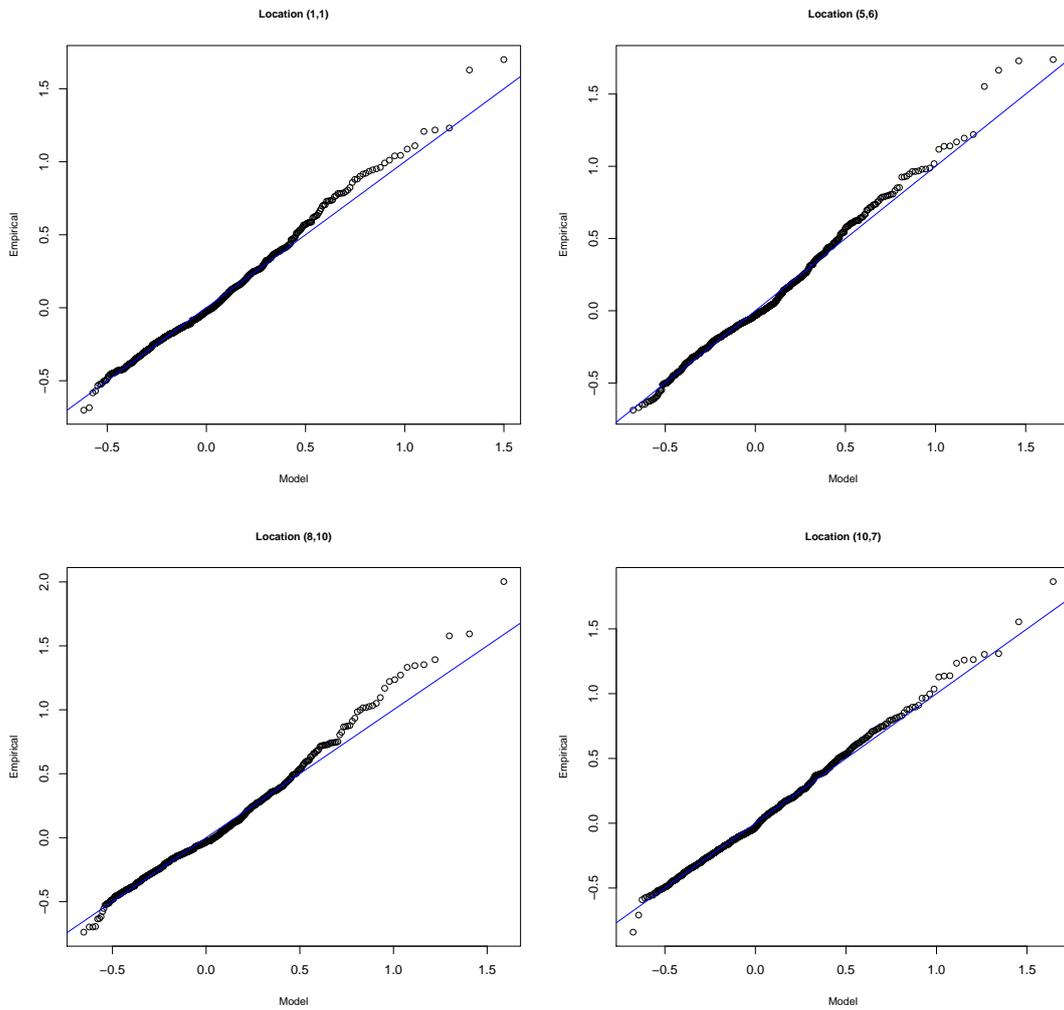


Figure 7.2.3: QQ-Plot of deseasonalized rain series against the fitted Gumbel distribution (GEV with $\hat{\mu}(s)$, $\hat{\sigma}(s)$ and 0) based on the time series corresponding to the grid locations shown in 7.2.1.

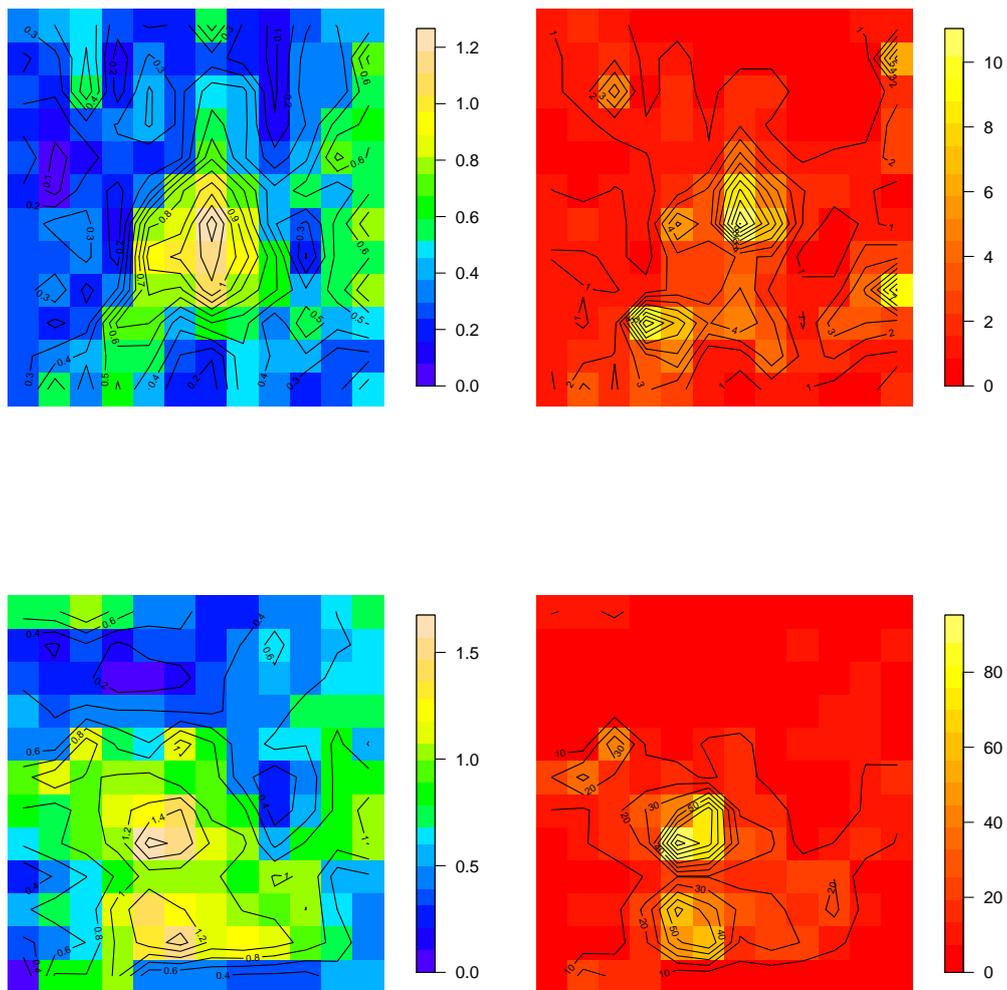


Figure 7.2.4: Plots of spatial maxima of rainfall measurements (left) and to standard Fréchet margins transformed rain fields for fixed time points t_{166} (July 14th 2000) and t_{434} (August 7th 2002).

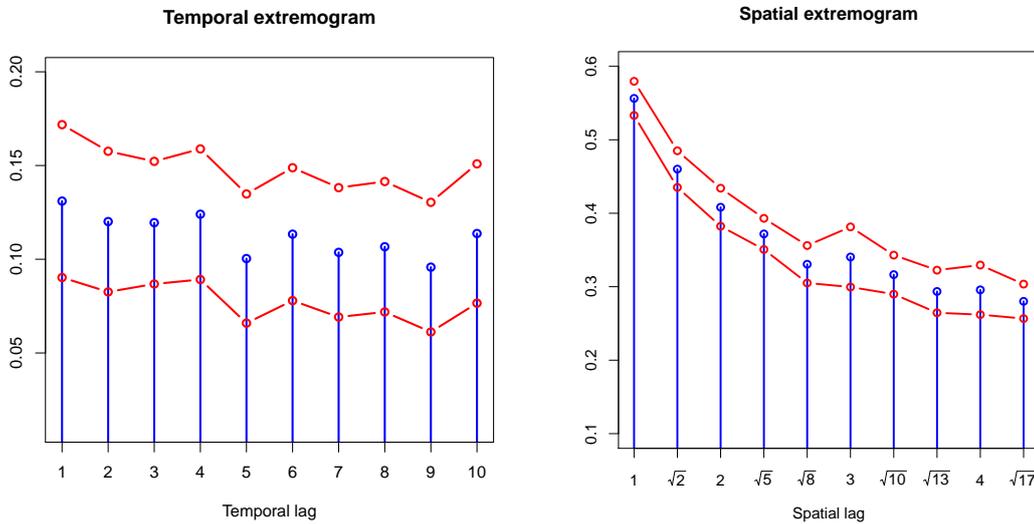


Figure 7.2.5: Temporal (left) and spatial (right) empirical extremogram as introduced (2a) and (2b) of Section 5.1 together with 95% bootstrap confidence intervals. Note, that the scale of the temporal extremogram is much smaller than the scale of the spatial extremogram.

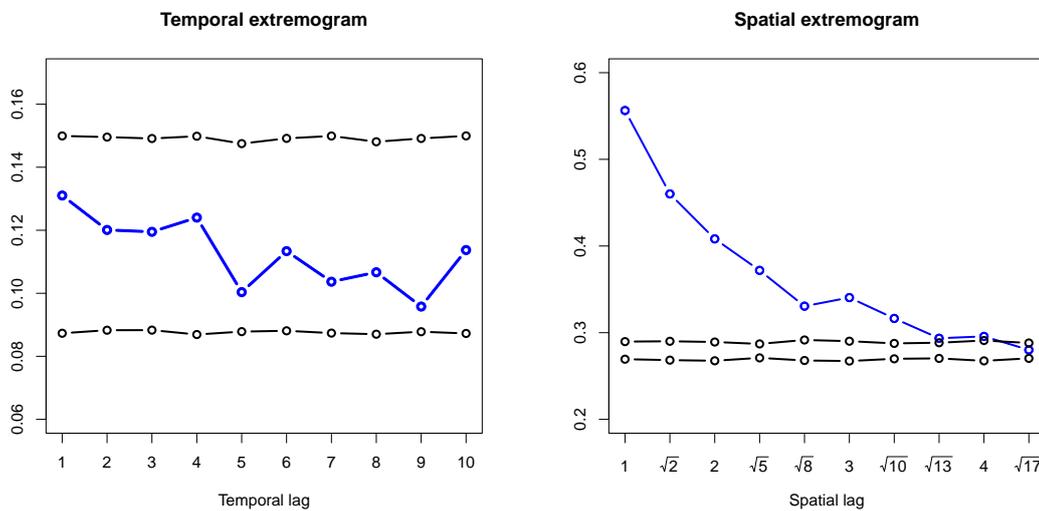


Figure 7.2.6: Permutation test for extremal independence in time (left) and space (right). The upper and lower black lines show 97.5% and 2.5% empirical quantiles of empirical extremograms for 1000 temporal (left) and spatial (right) permutations of the space-time observations.

Conditional probabilities and return level maps

We now return to the fundamental questions raised in the Introduction. Based on the fitted model we are interested in the prediction of conditional probabilities as well as conditional return level maps. Given a particular reference space-time location (s^*, t^*) we consider the conditional probability

$$P(\tilde{\eta}(s, t) > z_c \mid \tilde{\eta}(s^*, t^*) > z^*),$$

where $P(\tilde{\eta}(s^*, t^*) > z^*) = p^*$, and $\tilde{\eta}$ is the deseasonalized process in (7.1). In other words, we predict the conditional probability that the rainfall process exceeds the level z_c at any location (s, t) , given that the process exceeds z^* at the reference location (s^*, t^*) . Assuming that $\{\tilde{\eta}(s, t), s \in S, t \in \{t_1, \dots, t_{732}\}\}$ are realizations from a Brown-Resnick process (see (4.2)) with generalized extreme value marginal distributions $GEV_{\mu(s), \sigma(s), 0}$ fitted in (7.2), we obtain

$$\begin{aligned} & P(\tilde{\eta}(s, t) > z_c \mid \tilde{\eta}(s^*, t^*) > z^*) = \\ &= \frac{1}{p^*} \left(1 - GEV_{\mu(s), \sigma(s), 0}(z_c) - (1 - p^*) + P(\tilde{\eta}(s, t) \leq z_c, \tilde{\eta}(s^*, t^*) \leq z^*) \right) \\ &= 1 - \frac{1}{p^*} GEV_{\mu(s), \sigma(s), 0}(z_c) \\ &+ \frac{1}{p^*} P \left(\frac{-1}{\log(GEV_{\mu(s), \sigma(s), 0}(\tilde{\eta}(s, t)))} \leq \frac{-1}{\log(GEV_{\mu(s), \sigma(s), 0}(z_c))}, \right. \\ &\quad \left. \frac{-1}{\log(GEV_{\mu(s^*), \sigma(s^*), 0}(\tilde{\eta}(s^*, t^*)))} \leq \frac{-1}{\log(1 - p^*)} \right) \\ &= 1 - \frac{1}{p^*} GEV_{\mu(s), \sigma(s), 0}(z_c) + \frac{1}{p^*} F \left(\frac{-1}{\log(GEV_{\mu(s), \sigma(s), 0}(z_c))}, \frac{-1}{\log(1 - p^*)} \right), \end{aligned}$$

where F is the bivariate extreme value distribution of the Brown-Resnick process (see (3.6)). As a first step, we plot the predicted conditional probabilities with $z^* = z_c = 2.5$ for all locations s and reference locations in Figure 7.2.1, where the marginal estimates $\hat{\mu}(s)$ and $\hat{\sigma}(s)$ are used in $GEV_{\mu(s), \sigma(s), 0}(z_c)$ and the extremal dependence estimates for the Brown-Resnick process $\hat{\theta}_1, \hat{\alpha}_1, \hat{\theta}_2$ and $\hat{\alpha}_2$ are plugged into F . Since there is no temporal extremal dependence (see Figure 7.2.6), we only consider fixed time points. Figure 7.3.1 shows the predicted conditional probability fields.

The conditional return level is defined as the value z_c for which

$$P(\tilde{\eta}(\mathbf{s}, t) > z_c \mid \tilde{\eta}(\mathbf{s}^*, t^*) > z^*) = p_c,$$

where z^* is chosen as the $(1 - p^*)$ -quantile, $P(\tilde{\eta}(\mathbf{s}^*, t^*) > z^*) = p^*$. In particular, the conditional return level with return period $1/p_c$ is the value for which the probability that the process exceeds that level given that the process at a reference location exceeds z^* equals p^* . To estimate z_c we plug in the parameter estimates $\hat{\mu}(\mathbf{s})$ and $\hat{\sigma}(\mathbf{s})$ for the marginal distribution and $\hat{\theta}_1, \hat{\alpha}_1, \hat{\theta}_2$ and $\hat{\alpha}_2$ for the extremal dependence. First note, that if z^* equals the minimum of the data, i.e. $p^* = 1$, we obtain a field of unconditional quantiles which is the same as if the return levels at each location are estimated using the marginal fitted distributions. Figure 7.3.2 shows the predicted conditional return levels for $p^* = 0.01$ and $p_c = 0.01$ with the same reference locations. The conditional probability, as well as the conditional return level fields confirm the findings from estimating the extremogram, where we concluded that spatial extremal dependence is present for lags smaller than $\sqrt{17}$.

7.3 Hourly rainfall measurements June 2002-September 2002

Since there is no temporal extremal dependence for the daily maxima we turn to hourly rainfall measurements. In particular, we analyse the hourly rainfall observations in the rainfall season 2002 from June to September, leading to 2928 time points and 144 locations, denoted by

$$\{Z(\mathbf{s}, t), \mathbf{s} \in S, t \in \{t_1, \dots, t_{2928}\},$$

where $S = \{(i_1, i_2), i_1, i_2 \in \{1, \dots, 12\}\}$. As for the daily maxima, we first remove daily seasonality by using a simple moving average with block length of one day (24 time points). By inspecting autocorrelation plots we assume that the marginal time series are stationary with short-range dependence. Leadbetter [55, 56] shows that stationary time series satisfying additional mixing conditions have the same extreme value properties than the corresponding iid sequence. Therefore, we fit the generalized Pareto distribution to threshold exceedances as described in Section 2.1.1 and use the fitted distributions to transform the data to standard Fréchet marginal distributions. Afterwards, we estimate the extremal dependence parameters as before. The procedure is explained in full detail below.

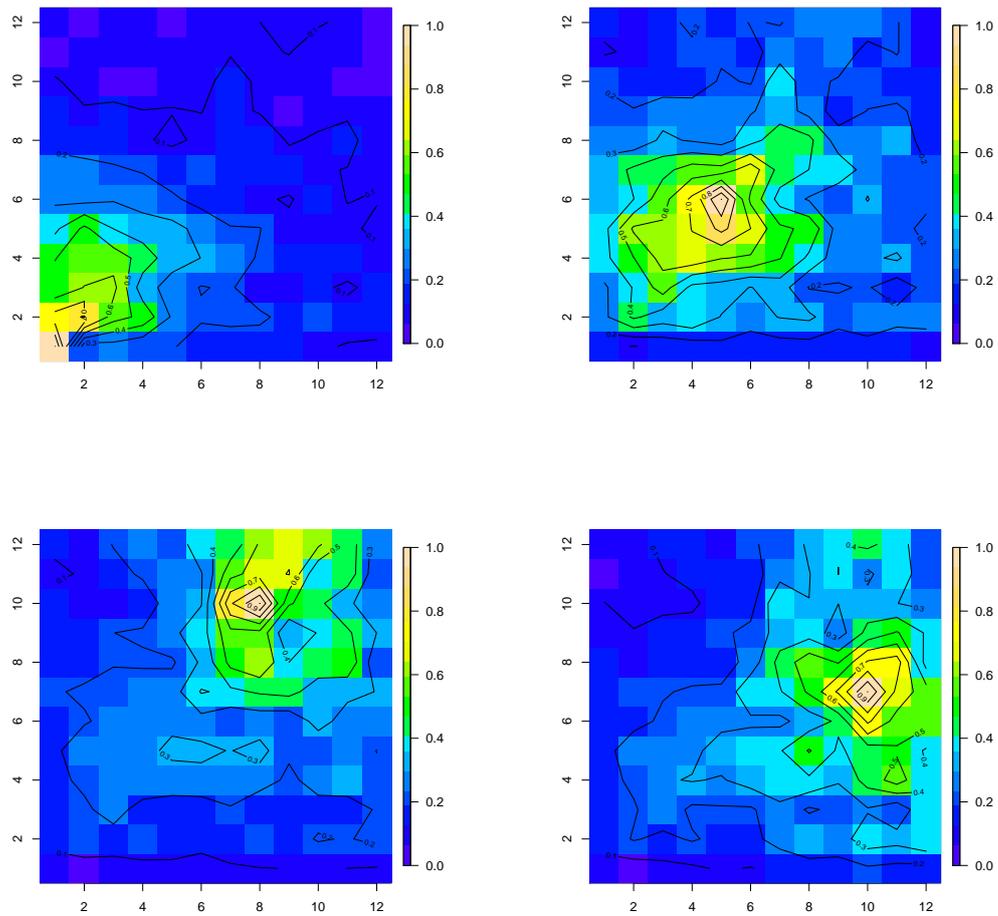


Figure 7.3.1: Predicted conditional probability fields $P(Z(s, t) > z_c \mid Z(s^*, t) > z^*)$ for fixed time point t with reference locations $(s^*, t^*) = ((1, 1), 1)$ (upper left) $(s^*, t^*) = ((5, 6), 1)$ (upper right), $(s^*, t^*) = ((8, 10), 1)$ (lower left) and $(s^*, t^*) = ((10, 7), 1)$ (lower right) together with contour lines. The values z_c and z^* equal 2.5.

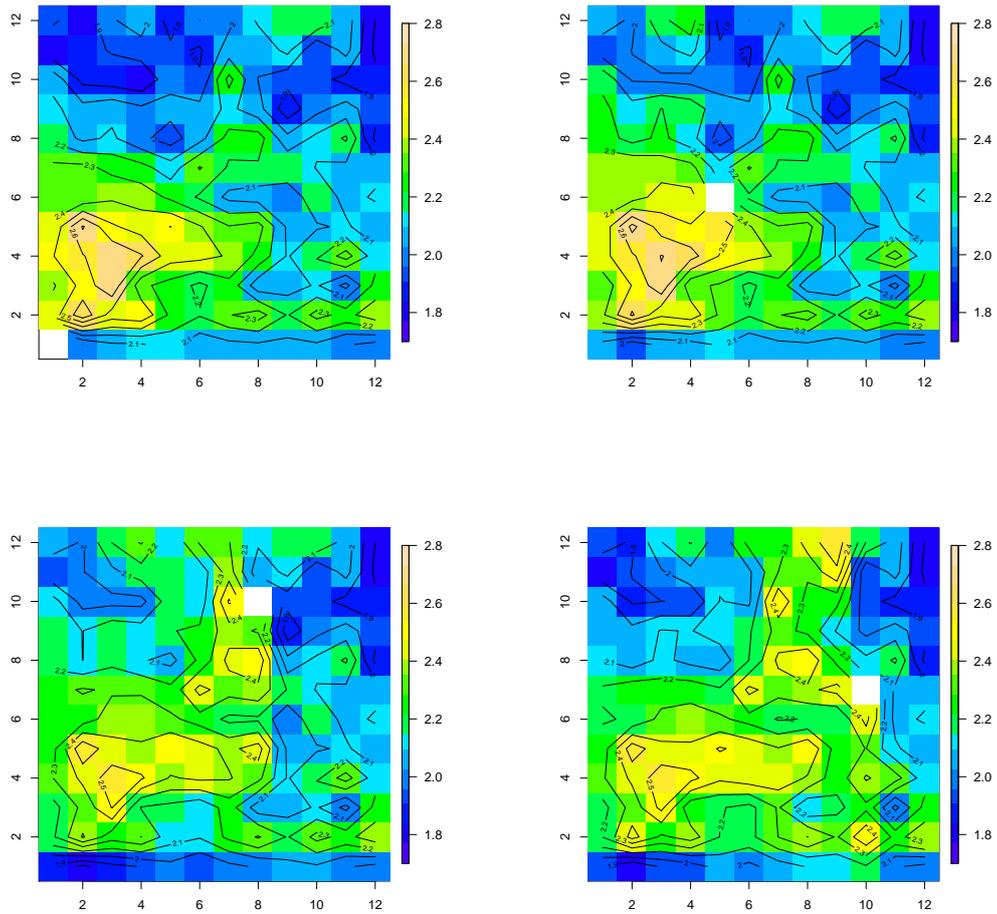


Figure 7.3.2: Conditional 100-day return levels resulting from solving $P(Z(s, t) > z_c \mid Z(\mathbf{s}^*, t^*) > z^*) - p_c = 0$ for z_c with four different reference locations $(\mathbf{s}^*, t^*) = ((1, 1), 1)$ (upper left) $(\mathbf{s}^*, t^*) = ((5, 6), 1)$ (upper right), $(\mathbf{s}^*, t^*) = ((8, 10), 1)$ (lower left) and $(\mathbf{s}^*, t^*) = ((10, 7), 1)$ (lower right), which are shown as white square.

1. Marginal modelling: For each fixed location $\mathbf{s} \in S = \{(i_1, i_2), i_1, i_2 \in \{1, \dots, 12\}\}$ we follow the steps below.

a) The time series $\{Z(\mathbf{s}, t), t \in \{t_1, \dots, t_{2928}\}\}$ is deseasonalized by calculating

$$\tilde{Z}(\mathbf{s}, t_{k+24(j-1)}) = Z(\mathbf{s}, t_{k+24(j-1)}) - \frac{1}{122} \sum_{j=1}^{122} Z(\mathbf{s}, t_{k+24(j-1)})$$

for $k = 1, \dots, 24$.

b) Threshold exceedances with a threshold $thres = 0.2$ are fitted to a Generalized Pareto distribution with distribution function

$$GPD_{\sigma(\mathbf{s}), \xi(\mathbf{s})}(x) = 1 - \left(1 + \frac{\xi(\mathbf{s})x}{\sigma(\mathbf{s})}\right)^{-1/\xi(\mathbf{s})}. \quad (7.3)$$

c) The data are transformed according to

$$\tilde{\tilde{Z}}(\mathbf{s}, t) = -\frac{1}{\log(\hat{F}_s(\tilde{Z}(\mathbf{s}, t)))}, \quad t \in \{t_1, \dots, t_{2928}\},$$

where

$$\hat{F}_s(x) = \begin{cases} 1 - \frac{N_{thres, s}}{2928} \left(1 + \hat{\xi}(\mathbf{s}) \frac{x - thres}{\hat{\sigma}(\mathbf{s})}\right)^{-1/\hat{\xi}(\mathbf{s})}, & x > thres, \\ \frac{1}{2928} \sum_{k=1}^{2928} \mathbb{1}_{\{\tilde{Z}(\mathbf{s}, t_k) \leq x\}}, & x \leq thres, \end{cases} \quad (7.4)$$

and $N_{thres, s}$ denotes the number of threshold exceedances for fixed location \mathbf{s} .

2. Estimating the extremal dependence structure: We estimate the parameters θ_1 , α_1 , θ_2 and α_2 of the Brown-Resnick process (cf. (4.2)) by using the semiparametric estimation procedure in Chapter 5 and the pairwise likelihood estimation in Chapter 4.

The results are shown in Figures 7.3.4, 7.3.5, 7.3.6, 7.3.7, 7.3.8 and Tables 7.3 and 7.4. Figure 7.3.4 visualizes the hourly accumulated rainfall series for the locations shown in 7.3.3. QQ-plots in Figure 7.3.5 of the exceedances against the fitted GPD support the marginal goodness of fit. Figure 7.3.6 shows rainfall fields for four consecutive hours on the left hand side together with the marginally transformed fields on the right hand side. Using the transformed space-time observations we estimate the parameters of the Brown-Resnick process, which are shown in Tables 7.3 and 7.4. We notice that the spatial parameter estimates

$\hat{\theta}_1$ and $\hat{\alpha}_1$ are close to the estimates for the daily maxima (compare to Tables 7.1 and 7.2). This is not unexpected since the spatial extremal dependence structure remains the same if the time domain is scaled. To see this, consider the extremogram of the Brown-Resnick process with scaled time lag u/m where $m \in \mathbb{N}$ (To get from the daily maxima to the hourly values use $m = 24$),

$$\chi(\mathbf{h}, u/m) = 2\left(1 - \Phi\left(\sqrt{\theta_1\|\mathbf{h}\|^{\alpha_1} + \theta_2|u/m|^{\alpha_2}}\right)\right).$$

For a temporal lag equal to zero, this is the same as for the original time domain.

Figure 7.3.7 shows the spatial and temporal mean of temporal (left) and spatial (right) extremograms as described in (2a) and (2b) of Section 5.1 together with 95% bootstrap confidence intervals. When we compare Figure 7.3.7 to Figure 7.2.5 we see that the spatial extremogram on the right hand side is similar to the one for the daily maxima. The temporal extremogram now shows more extremal dependence compared to the daily maxima. We repeat the permutation test for independence. The results are shown in Figure 7.3.8 and indicate that there is no temporal extremal dependence for time lags larger than 6 and no spatial extremal dependence for spatial lags larger than 4.

As for the daily maxima, we return to the fundamental questions raised in the introduction. In particular, we predict the conditional probabilities

$$P(\tilde{Z}(\mathbf{s}, t) > z_c \mid \tilde{Z}(\mathbf{s}^*, t^*) > z^*),$$

where $P(\tilde{Z}(\mathbf{s}^*, t^*) > z^*) = p^*$, by

$$1 - \frac{1}{p^*} \hat{F}_s(z_c) + \frac{1}{p^*} \hat{F}\left(\frac{-1}{\log(\hat{F}_s(z_c))}, \frac{-1}{\log(1 - p^*)}\right),$$

where \hat{F}_s is defined in (7.4) and \hat{F} is the fitted bivariate distribution function of the Brown-Resnick process (see (3.6)). In addition, we estimate the conditional return levels z_c for which

$$P(\tilde{Z}(\mathbf{s}, t) > z_c \mid \tilde{Z}(\mathbf{s}^*, t^*) > z^*) = p_c.$$

Figure 7.3.9 shows the predicted conditional probability fields for $z_c = z^* = 1$ for two reference locations $\mathbf{s}^* = (5, 6)$ (left) and $\mathbf{s}^* = (10, 7)$ (right) and $t^* = t_1$. The plots visualize the prediction for all locations and for $t \in \{t_1, t_2, t_3, t_4\}$ (from the top to the bottom). Since we assume stationarity only the time lag $t - t^*$, and not the starting reference time point t^* , is of interest. Figures 7.3.10 and 7.3.11 show the predicted conditional return level maps for time

lags $u = 0, 1, \dots, 7$ based on the reference locations $\mathbf{s}^* = (1, 1)$ (left), $\mathbf{s}^* = (5, 6)$ (middle) and $\mathbf{s}^* = (10, 7)$ (right). Both, the conditional probability fields as well as the conditional return level fields show how the dependence in space and time develops. The plots confirm the findings from the extremogram estimates, where we concluded spatial independence for lags larger than 4 and temporal independence for time lags larger than 6.

As a final step in this chapter we simulate from the fitted space-time process for extremes to obtain realizations for the space-time rainfall process. We start by simulating from the max-stable space-time process as described in Chapter 6 with the parameter estimates obtained for the rainfall measurements. Afterwards, we transform the marginal distributions according to the fitted GPD distributions for each location. Figure 7.3.12 shows the resulting realizations for four consecutive time points. By comparing the figures with the plots in the introduction (see Figure 1.1.1), we see that the simulations are similar to the rainfall measurements.

Estimate	$\hat{\theta}_1$	0.2987	$\hat{\alpha}_1$	0.9664
Bootstrap-CI		[0.2469,0.3505]		[0.8407,1.0921]
Estimate	$\hat{\theta}_2$	0.4763	$\hat{\alpha}_2$	1.0686
Bootstrap-CI		[0.2889,0.6637]		[0.8514,1.2859]

Table 7.3: Semiparametric estimates for the spatial parameters θ_1 and α_1 and the temporal parameters θ_2 and α_2 for the Brown-Resnick process in (4.2) together with 95% bootstrap confidence intervals.

PL estimates	$\hat{\theta}_1$	$\hat{\alpha}_1$	$\hat{\theta}_2$	$\hat{\alpha}_2$
	0.3353	0.9302	0.4845	1.0648

Table 7.4: Pairwise likelihood estimates for the parameters in the Brown-Resnick process in (4.2).

7.3 Hourly rainfall measurements June 2002-September 2002

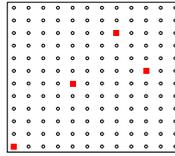


Figure 7.3.3: Reference locations for time series plots in Figure 7.2.2.

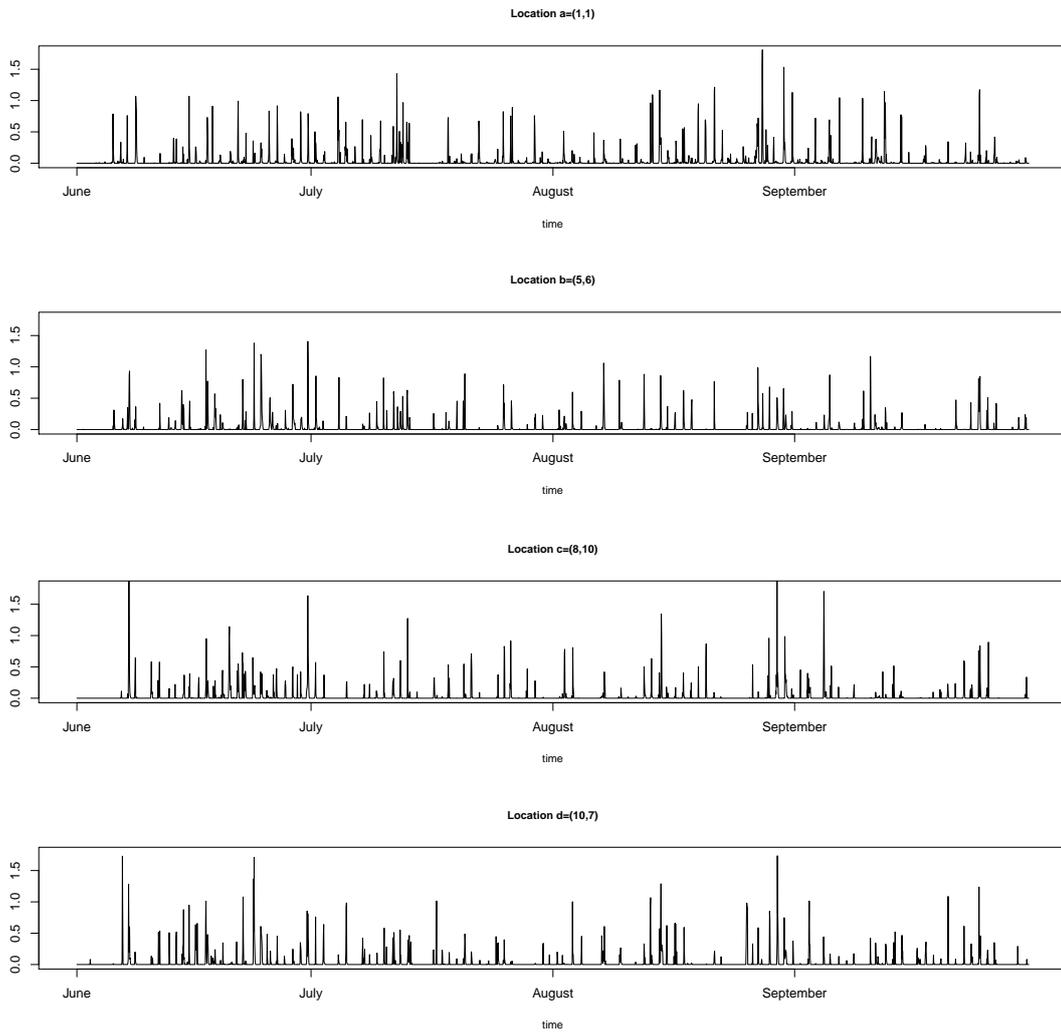


Figure 7.3.4: Accumulated hourly rainfall measurements calculated from the 15-minutes measurements (June 2002 - September 2002) in inches for the grid locations in 7.3.3.

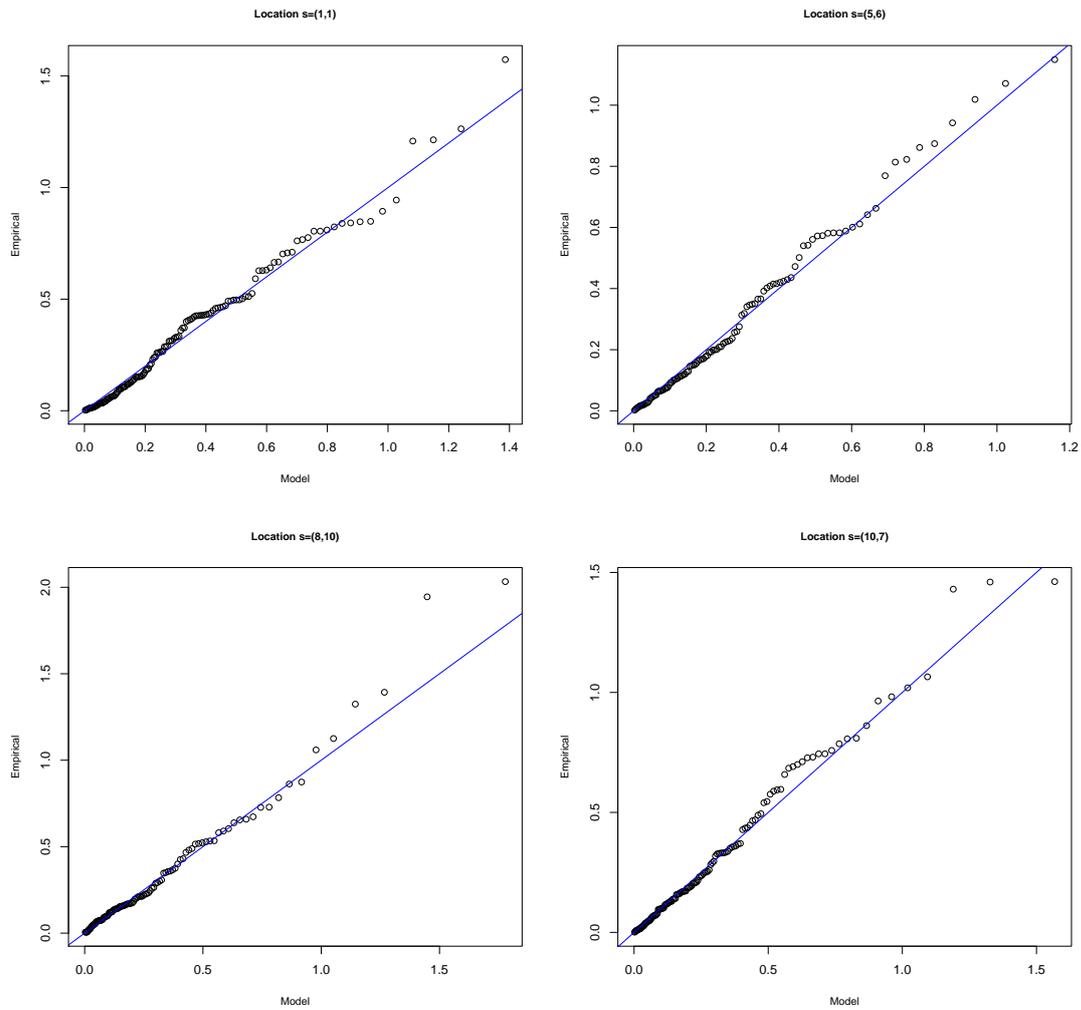


Figure 7.3.5: QQplot of threshold exceedances against fitted GPD distribution for the grid locations in 7.3.3 with a threshold equal to 0.2.

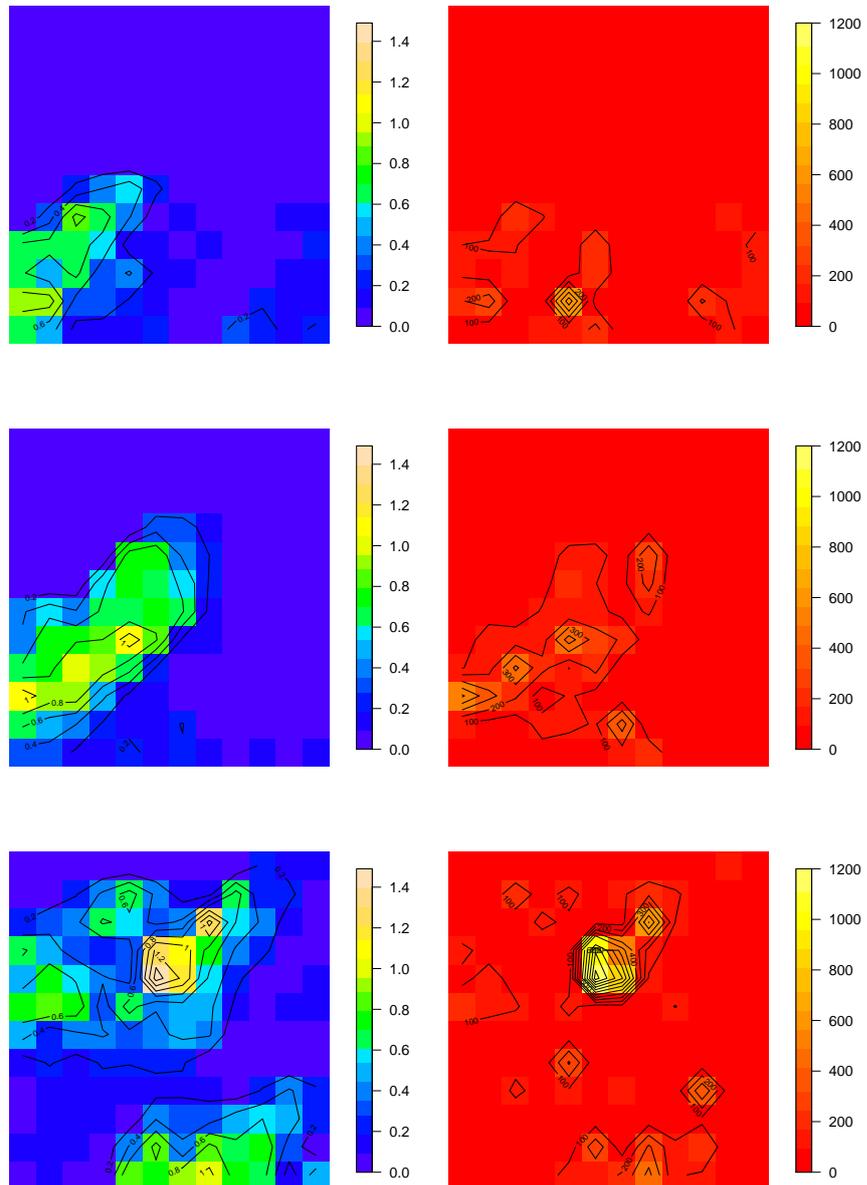


Figure 7.3.6: Plots of original and transformed (to standard Fréchet) rain fields for three subsequent time points t_{1016} , t_{1017} and t_{1018} (July 13th, 2002, 10pm - 12pm).

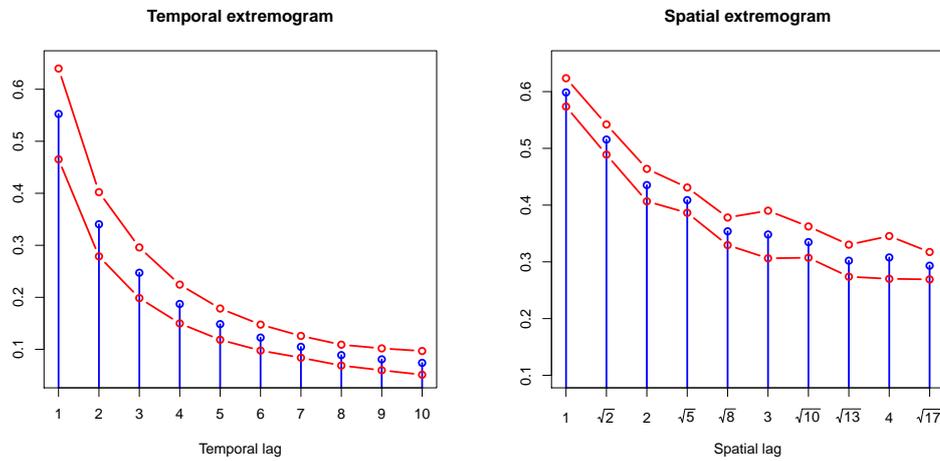


Figure 7.3.7: Temporal (left) and spatial (right) empirical extremogram based on spatial and temporal means for the space-time observations as described in (2a) and (2b) in Section 5.1 together with 95% bootstrap confidence intervals.

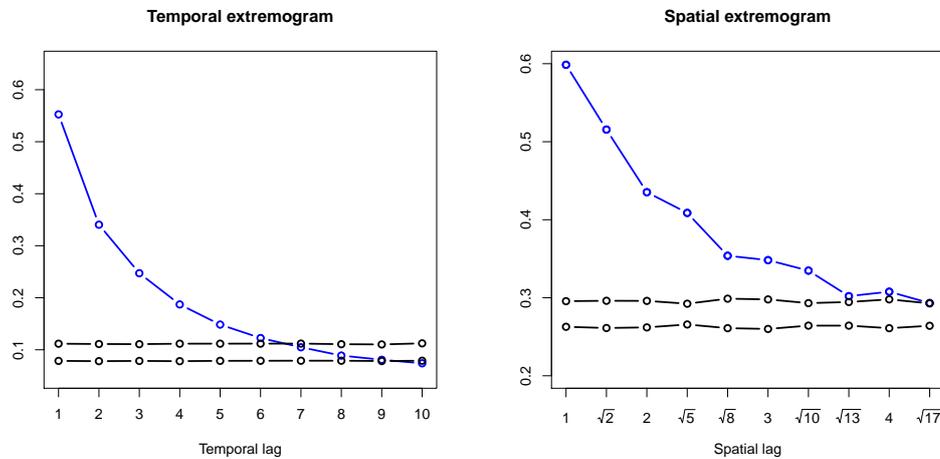


Figure 7.3.8: Permutation test for extremal independence based on 1000 space-time permutations of the observations, which destroy the extremal dependence in time (left) and space (right). The black lines show the 97.5% and 2.5% quantiles of the extremogram estimates for the 1000 permutations.

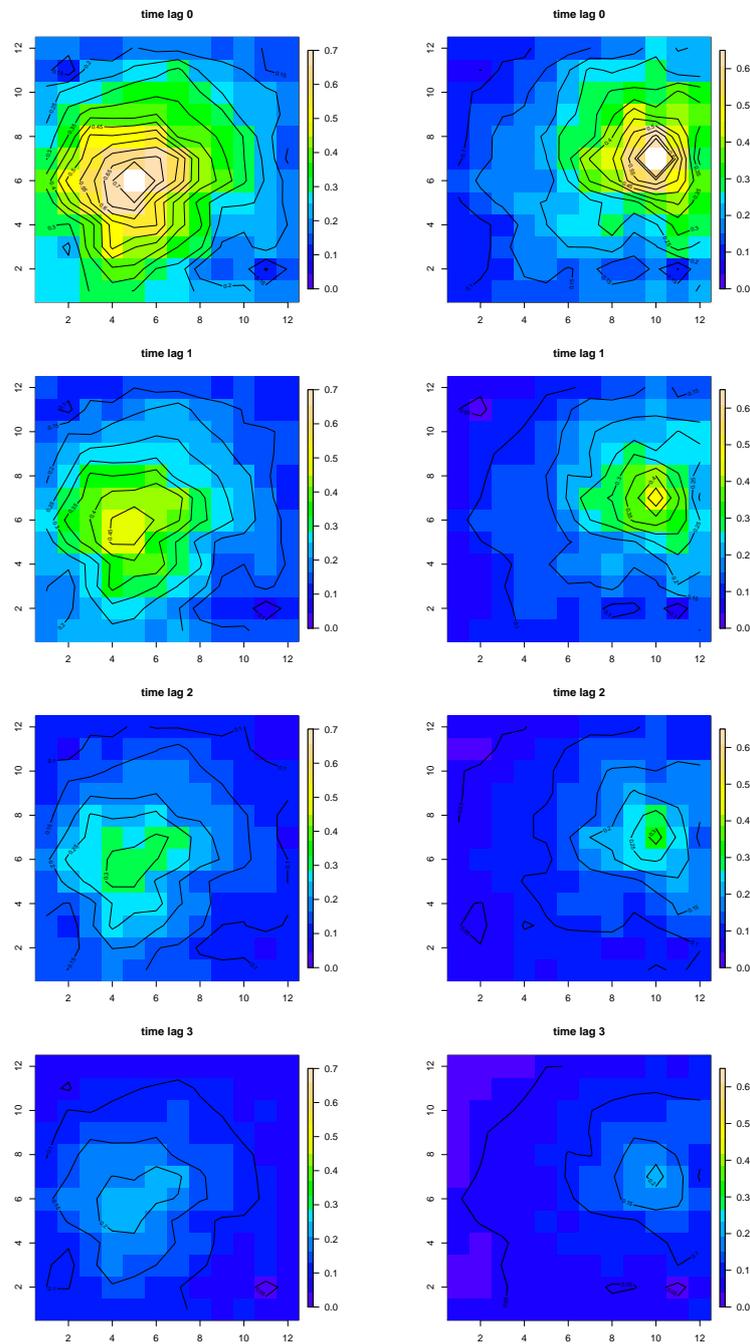


Figure 7.3.9: Predicted conditional probability fields $P(Z(s, t) > z_c | Z(s^*, t^*) > z^*)$ for $z_c = z^* = 1$, $s^* = (5, 6)$ (left) and $s^* = (10, 7)$ (right). From the top to the bottom the fields show the development in time for lags $u = 0, 1, 2, 3$.

7 Analysis of radar rainfall measurements in Florida

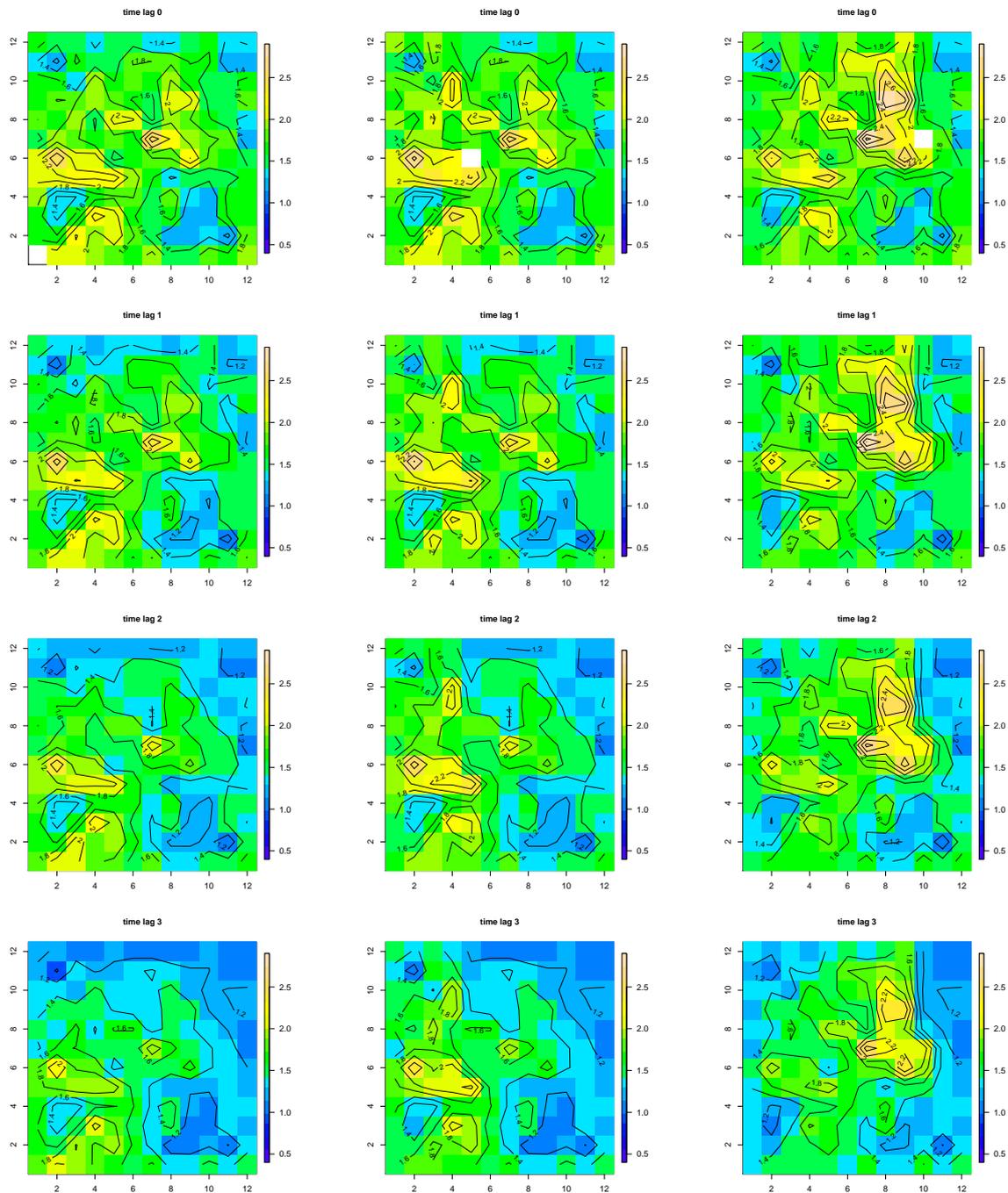


Figure 7.3.10: Conditional 100-hour return level maps for time lags $u = 0, \dots, 3$ and reference locations $s^* = (1, 1)$ (left), $s^* = (5, 6)$ (middle) and $s^* = (10, 7)$ (right).

7.3 Hourly rainfall measurements June 2002-September 2002

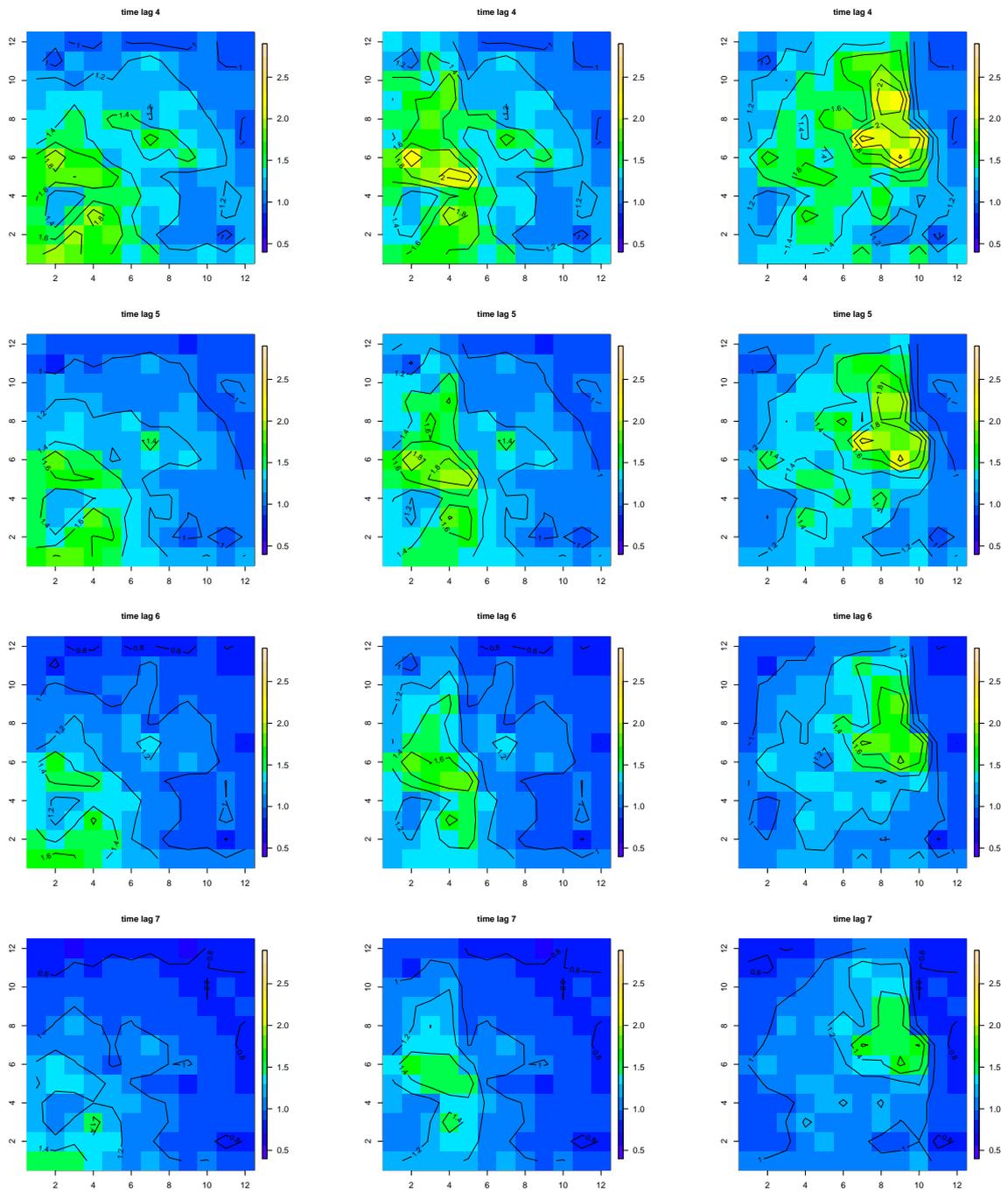


Figure 7.3.11: Conditional 100-hour return level maps for time lags $u = 4, \dots, 7$ and reference locations $s^* = (1, 1)$ (left), $s^* = (5, 6)$ (middle) and $s^* = (10, 7)$ (right).

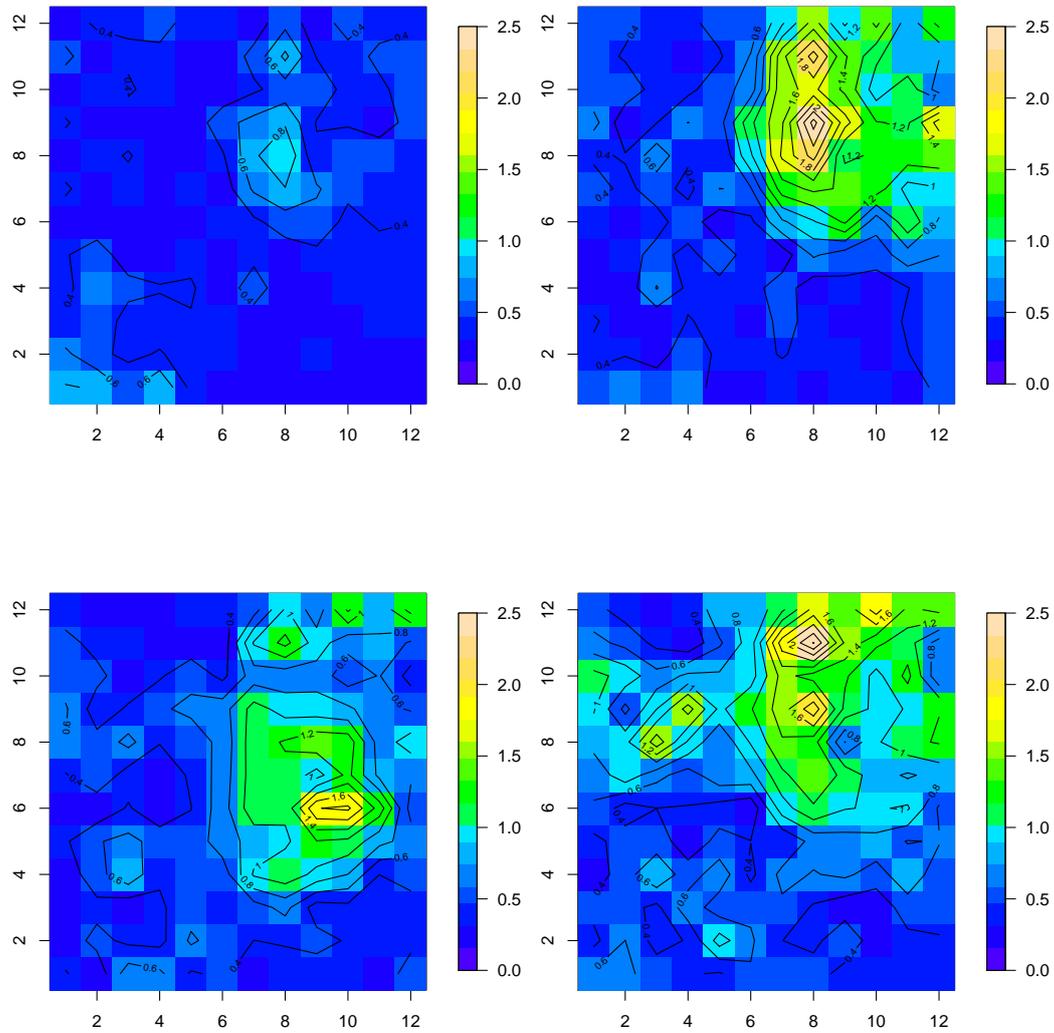


Figure 7.3.12: Extremal rainfall fields for four consecutive hours (clockwise from the top left to the bottom left) simulated from the fitted max-stable space-time process, where the marginal distributions follow the GPD with parameter estimates resulting from the marginal fit for each fixed location.

BIBLIOGRAPHY

- [1] R.J. Adler. *The Geometry of Random Fields*. John Wiley, New York, 1981.
- [2] D. W. K. Andrews. Estimation when a parameter is on a boundary. *Econometrica*, 67(6):1341–1383, 1999.
- [3] D. W. K. Andrews. Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space. *Econometrica*, 68(2):399–405, 2000.
- [4] Y. Bai, X.-K. Song, and T.E. Raghunathan. Joint composite estimating functions in spatiotemporal models. *Journal of the Royal Statistical Society B*, 75(5):799–824, 2012.
- [5] A. Baxevani, S. Caires, and I. Rychlik. Spatio-temporal stationary modelling of significant wave height. *Environmetrics*, 20(1):14–31, 2009.
- [6] A. Baxevani, K. Podgórski, and I. Rychlik. Dynamically evolving Gaussian spatial random fields. *Extremes*, 14(2):223–251, 2011.
- [7] J. Beirlant, Y. Goegebeur, J. Segers, and J. Teugels. *Statistics of Extremes, Theory and Applications*. Wiley Series in Probability and Statistics, John Wiley & Sons Ltd, Chichester, 2004.
- [8] J.E. Besag. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society B*, 34(2):192–236, 1974.

- [9] P.J. Bickel, F. Götze, and W.R. van Zwet. Resampling fewer than n observations: gains, losses, and remedies for losses. *Statistica Sinica*, 7:1–31, 1997.
- [10] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics, New York, 1999.
- [11] E. Bolthausen. On the central limit theorem for stationary mixing random fields. *The Annals of Probability*, 10(4):1047–1050, 1982.
- [12] G.E.P. Box and M.E. Muller. A note on the generation of random normal deviates. *Annals of Mathematical Statistics*, 29(2):610–611, 1958.
- [13] P.J. Brockwell and R.A. Davis. *Time Series: Theory and Methods*. Springer-Verlag, New York, 2nd edition, 1991.
- [14] B.M. Brown and S.I. Resnick. Extreme values of independent stochastic processes. *Journal of Applied Probability*, 14(4):732–739, 1977.
- [15] Y. Cho, R.A. Davis, and S. Ghosh. Asymptotic properties of the spatial empirical extremogram. 2012. In preparation.
- [16] S.G. Coles. Regional modelling of extreme storms via max-stable processes. *Journal of the Royal Statistical Society B*, 55(4):797–816, 1993.
- [17] S.G. Coles. *An Introduction to Statistical Modeling of Extreme Values*. Springer Series in Statistics, Springer, New York, 2001.
- [18] S.G. Coles and J.A Tawn. Modelling extremes of the areal rainfall process. *Journal of the Royal Statistical Society B*, 58(2):329–347, 1996.
- [19] S.G. Coles and D. Walshaw. Directional modelling of extreme wind speeds. *Journal of the Royal Statistical Society C*, 43:139–157, 1994.
- [20] D.R. Cox and N. Reid. Miscellanea: A note on pseudolikelihood constructed from marginal densities. *Biometrika*, 91(3):729–737, 2004.
- [21] N. Cressie and H.C. Huang. Classes of nonseparable spatio-temporal stationary covariance functions. *Journal of the American Statistical Association*, 94(448):1330–1340, 1999.

-
- [22] R.A. Davis and T. Mikosch. Extreme value theory for space-time processes with heavy-tailed distributions. *Stochastic Processes and their Applications*, 118(4):560–584, 2008.
- [23] R.A. Davis and T. Mikosch. The extremogram: A correlogram for extreme events. *Bernoulli*, 15(4):977–1009, 2009.
- [24] R.A. Davis and C.Y. Yau. Comments on pairwise likelihood in time series models. *Statistica Sinica*, 21(1):255–277, 2011.
- [25] R.A. Davis, T. Mikosch, and I. Cribben. Towards estimating extremal series dependence via the bootstrapped extremogram. *Journal of Econometrics*, 17(1):142–152, 2012.
- [26] R.A. Davis, C. Klüppelberg, and C. Steinkohl. Max-stable processes for extremes of processes observed in space and time. *Journal of the Korean Statistical Society*, 2013. <http://dx.doi.org/10.1016/j.jkss.2013.01.002>.
- [27] R.A. Davis, C. Klüppelberg, and C. Steinkohl. Statistical inference for max-stable processes in space and time. *Journal of the Royal Statistical Society B*, 2013. to appear.
- [28] R.A. Davis, C. Klüppelberg, and C. Steinkohl. Semiparametric estimation for max-stable space-time processes. 2013. In preparation.
- [29] L. de Haan. A spectral representation for max-stable processes. *The Annals of Probability*, 12(4):1194–1204, 1984.
- [30] L. de Haan and A. Ferreira. *Extreme Value Theory: An Introduction*. Springer Series in Operations Research and Financial Engineering, New York, 2006.
- [31] L. de Haan and J. Pickands. Stationary min-stable stochastic processes. *Probability Theory and Related Fields*, 72(4):477–492, 1986.
- [32] L. de Haan and S. Resnick. Limit theory for multivariate sample extremes. *Zeitschrift für Wahrscheinlichkeitstheorie*, 40:317–337, 1977.
- [33] P. Deheuvels. Point processes and multivariate extreme values. *Journal of Multivariate Analysis*, 13(2):257–272, 1983.
- [34] C. Dombry and F. Eyi-Minko. Strong mixing properties of max-infinitely divisible random fields. *Stochastic Processes and their Applications*, 122(11):3790–3811, 2012.

- [35] P. Doukhan. *Mixing: Properties and Examples*. Springer, New York, 1994.
- [36] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling Extremal Events*. Springer, Berlin, 1997.
- [37] V. Fasen, C. Klüppelberg, and M. Schlather. High-level dependence in time series models. *Extremes*, 13(1):1–33, 2010.
- [38] W. Feller. *An introduction to probability theory and its applications*, volume 2 of *Wiley series in probability and mathematical statistics: Probability and mathematical statistics*. Wiley, 1971.
- [39] R.A. Fisher and L.H.C. Tippett. On the estimation of the frequency distributions of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society*, 24:180–190, 1928.
- [40] J. Geffroy. Contributions à la théorie des valeurs extrême. *Publ. Inst. Stat. Univ. Paris*, 7:36–123, 1958.
- [41] J. Geffroy. Contributions à la théorie des valeurs extrême. *Publ. Inst. Stat. Univ. Paris*, 8:3–52, 1959.
- [42] B.V. Gnedenko. Sur la distribution limite du terme maximum d’une série aléatoire. *Annals of Mathematics*, 44(3):423–453, 1943.
- [43] T. Gneiting. Nonseparable, stationary covariance functions for space-time data. *Journal of the American Statistical Association*, 95(458):590–600, 2002.
- [44] V.P. Godambe. An optimum property of regular maximum likelihood estimation. *Annals of Mathematical Statistics*, 31:1208–1211, 1960.
- [45] E.J. Gumbel. Distributions de valeurs extrêmes en plusieurs dimensions. *Publications of the Institute of Statistics of the University of Paris*, 9:171–173, 1960.
- [46] R. Huser and A. Davison. Space-time modelling for extremes. *Preprint*, 2012. arXiv:1201.3245v1 [stat.ME].
- [47] J. Hüsler and R.-D. Reiss. Maxima of normal random vectors: between independence and complete dependence. *Statistics and Probability Letters*, 7(4):283–286, 1989.

-
- [48] N. Jenish and I.R. Prucha. Central limit theorems and uniform laws of large numbers for arrays of random fields. *Journal of Econometrics*, 150(1):86–98, 2009.
- [49] Z. Kabluchko. Extremes of space-time Gaussian processes. *Stochastic Processes and their Applications*, 119:3962–3980, 2009.
- [50] Z. Kabluchko, M. Schlather, and L. de Haan. Stationary max-stable fields associated to negative definite functions. *The Annals of Probability*, 37(5):2042–2065, 2009.
- [51] A. F. Karr. Inference for stationary random fields given Poisson samples. *Advances in Applied Probability*, 18(2):406–422, 1986.
- [52] B. Kozintsev. *Computations with Gaussian Random Fields*. PhD thesis, Institute for Systems Research, University of Maryland, 1999.
- [53] U. Krengel. *Ergodic Theorems*. de Gruyter, Berlin, 1985.
- [54] S. N. Lahiri. Theoretical comparisons of block bootstrap methods. *The Annals of Statistics*, 27(1):386–404, 1999.
- [55] M.R. Leadbetter. On extreme values in stationary sequences. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 28:289–303, 1974.
- [56] M.R. Leadbetter. Extremes and local dependence in stationary sequences. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 65:291–306, 1983.
- [57] M.R. Leadbetter, G. Lindgren, and H. Rootzén. *Extremes and Related Properties of Random Sequences and Processes*. Springer Verlag, New York, 1983.
- [58] M. Ledoux and M. Talagrand. *Probability in Banach Spaces: Isoperimetry and Processes*. Springer Verlag, Berlin, 1991.
- [59] B. Li, M.G. Genton, and M. Sherman. On the asymptotic joint distribution of sample space-time covariance estimators. *Bernoulli*, 14(1):208–248, 2008.
- [60] B.G. Lindsay. Composite likelihood methods. *Contemporary Mathematics*, 80:221–239, 1988.
- [61] C. Ma. Spatio-temporal covariance functions generated by mixtures. *Mathematical geology*, 34(8):965–975, 2002.

- [62] C. Ma. Spatio-temporal stationary covariance functions. *Journal of Multivariate Analysis*, 86(1):97–107, 2003.
- [63] C. Ma. Linear combinations of space-time covariance functions and variograms. *IEEE Transactions on signal processing*, 53(3):857–864, 2005.
- [64] J. Mateu, E. Porcu, and P. Gregori. Recent advances to model anisotropic space-time data. *Statistical Methods and Applications*, 17(2):209–223, 2007.
- [65] D. J. Nott and T. Rydén. Pairwise likelihood methods for inference in image models. *Biometrika*, 86(3):661–676, 1999.
- [66] S.A. Padoan, M. Ribatet, and S.A. Sisson. Likelihood-based inference for max-stable processes. *Journal of the American Statistical Association*, 105(489):263–277, 2009.
- [67] J. Pickands. Multivariate extreme value distributions. *Proceedings of the 43rd Session of the Int. Stat. Institute*, pages 859–878, 1981.
- [68] D.N. Politis and J.P. Romano. A circular block resampling procedure for stationary data. In R. Lepage and L. Billard, editors, *Exploring the Limits of Bootstrap*. Wiley, New York, 1992.
- [69] D.N. Politis and J.P. Romano. The stationary bootstrap. *Journal of the American Statistical Society*, 89(428):1303–1313, 1998.
- [70] E. Porcu, P. Gregori, and J. Mateu. La descente et la montée étendues: the spatially d-anisotropic and spatio-temporal case. *Stochastic Environmental Research and Risk Assessment*, 21(6):683–693, 2007.
- [71] S.I. Resnick. Point processes, regular variation and weak convergence. *Advances in Applied Probability*, 18(1):66–138, 1986.
- [72] S.I. Resnick. *Adventures in Stochastic Processes*. Birkhäuser, Boston, 1992.
- [73] S.I. Resnick. *Heavy-Tail Phenomena, Probabilistic and Statistical Modeling*. Springer Series in Operations Research and Financial Engineering, 2007.
- [74] H. Sang and A. Gelfand. Hierarchical modeling for extreme values observed over space and time. *Environmental and Ecological Statistics*, 16(3):407–426, 2009.

- [75] M. Schlather. Models for stationary max-stable random fields. *Extremes*, 5(1):33–44, 2002.
- [76] M. Schlather. Some covariance models based on normal scale mixtures. *Bernoulli*, 16(3):780–797, 2010.
- [77] M. Schlather and J.A. Tawn. A dependence measure for multivariate and spatial extreme values: Properties and inference. *Biometrika*, 90(1):139–156, 2003.
- [78] M. Sibuya. Bivariate extreme statistics. *Annals of the Institute of Statistical Mathematics*, 11:195–210, 1960.
- [79] R. L. Smith. Max-stable processes and spatial extremes. Unpublished manuscript, University of North California, 1990.
- [80] C. Steinkohl, R.A. Davis, and C. Klüppelberg. Extreme value analysis of multivariate high frequency wind speed data. *Journal of Statistical Theory and Practice*, 2013. to appear.
- [81] S.A. Stoev. On the ergodicity and mixing of max-stable processes. *Stochastic Processes and their Applications*, 18(9):1679–1705, 2008.
- [82] S.A. Stoev and M.S. Taqqu. Extremal stochastic integrals: a parallel between max-stable processes and α -stable processes. *Extremes*, 8:237–266, 2005.
- [83] D. Straumann. *Estimation in Conditionally Heteroscedastic Time Series Models*. Lecture Notes in Statistics, Springer, Berlin, 2004.
- [84] J.A. Tawn. Bivariate extreme value theory: models and estimation. *Biometrika*, 75(3):397–415, 1988a.
- [85] J.A. Tawn. Modelling multivariate extreme value distributions. *Biometrika*, 77(2):245–253, 1990.
- [86] C. Varin. On composite marginal likelihoods. *Advances in Statistical Analysis*, 92(1):1–28, 2008.
- [87] C. Varin and P. Vidoni. Pairwise likelihood inference and model selection. *Biometrika*, 92(3):519–528, 2005.

- [88] H. Wackernagel. *Multivariate Geostatistics*. Springer, Heidelberg, 2003.
- [89] A. Wald. Note on the consistency of the maximum likelihood estimate. *The Annals of Mathematical Statistics*, 20(4):595–601, 1946.
- [90] Y. Wang and S.A. Stoev. On the structure and representations of max-stable processes. *Advances in Applied Probability*, 42(3):855–877, 2010.
- [91] Y. Wang, P. Roy, and S.A. Stoev. Ergodic properties of sum-and max-stable stationary random fields via null and positive group actions. *Preprint*, 2011. arXiv:0911.0610v2 [math.PR].
- [92] A.T.A. Wood and G. Chan. Simulation of stationary gaussian processes on $[0, 1]^d$. *Journal of Computational and Graphical Statistics*, 3(4):409–432, 1994.
- [93] Y. Xue and Y. Xiao. Fractal and smoothness properties of space-time gaussian models. *Frontiers of Mathematics in China*, 6(6):1217–1248, 2011.