SPATIAL DEPENDENCE AND SPACE-TIME TREND IN EXTREME EVENTS

BY JOHN H.J. EINMAHL¹, ANA FERREIRA², LAURENS DE HAAN³, CLÁUDIA NEVES⁴ AND CHEN ZHOU⁵

¹Department of Econometrics and Operations Research, Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands, j.h.j.einmahl@uvt.nl
²Instituto Superior Técnico (IST), University of Lisbon, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal, anafh@tecnico.ulisboa.pt
³Department of Statistics, University of Lisbon, , Bloco C6 Piso 4, Campo Grande, 1749-016 Lisboa, Portugal; Department of Economics, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands, ldehaan@ese.eur.nl
⁴Department of Mathematics and Statistics, Whiteknights Campus, University of Reading, Reading RG6 6AX, United Kingdom, c.neves@reading.ac.uk
⁵Econometric Institute, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands; Bank of The Netherlands, PO Box 98, 1000 AB Amsterdam, The Netherlands, zhou@ese.eur.nl

The statistical theory of extremes is extended to independent multivariate observations that are non-stationary both over time and across space. The non-stationarity over time and space is controlled via the scedasis (tail scale) in the marginal distributions. Spatial dependence stems from multivariate extreme value theory. We establish asymptotic theory for both the weighted sequential tail empirical process and the weighted tail quantile process based on all observations, taken over time and space. The results yield two statistical tests for homoscedasticity in the tail, one in space and one in time. Further, we show that the common extreme value index can be estimated via a pseudo-maximum likelihood procedure based on pooling all (non-stationary and dependent) observations. Our leading example and application is rainfall in Northern Germany.

Keywords and phrases: multivariate extreme value statistics, non-identical distributions, sequential tail empirical process, testing.
Consider independent random vectors \((X_{i,1}, X_{i,2}, \ldots, X_{i,m}), i = 1, 2, \ldots, n\). In the rainfall context, “\(m\)” is the number of stations and “\(n\)” is the number of time points (in days). A key assumption is the existence of scedasis: for some continuous distribution function \(F_0\) in the domain of attraction of an extreme value distribution

\[
\lim_{x \to x^*} \frac{1 - F_{i,j}(x)}{1 - F_0(x)} = c \left( \frac{i}{n}, j \right) \in (0, \infty),
\]

holds for \(i = 1, 2, \ldots, n; j = 1, 2, \ldots, m\), where \(F_{i,j}\) is the distribution function of \(X_{i,j}\) and where for each \(j\) the scedasis \(c(\cdot, j)\) is a positive continuous function on \([0,1]\). In order to ensure that the function \(c\) is uniquely defined we impose the condition

\[
\sum_{j=1}^{m} C_j(1) = 1,
\]

where for \(0 \leq t \leq 1\) and \(j = 1, 2, \ldots, m\),

\[
C_j(t) := \frac{1}{m} \int_0^t c(u,j) \, du;
\]

\(C_j\) is called the integrated scedasis. The scedasis \(c(i/n, j)\) can be interpreted as the relative frequency of extremes at time \(i\) and location \(j\). We assume \(F_0 \in \mathcal{D}(G_\gamma)\), \(\gamma \in \mathbb{R}\), i.e., \(F_0\) is in the max-domain of attraction of \(G_\gamma(x) = \exp\{-(1 + \gamma x)^{-1/\gamma}\}, 1 + \gamma x > 0\). As a consequence of (1.1), \(\gamma\) is now the common extreme value index: \(F_{i,j} \in \mathcal{D}(G_\gamma), i = 1, \ldots, n; j = 1, \ldots, m\).

Estimators \(\hat{C}_j\) for \(C_j\) will be introduced and their joint asymptotic distribution derived. This will enable us to perform various tests. For each station \(j\) we test whether the scedasis is changing over time. We also test whether the \(C_j(1)\) are different i.e. if there are real differences in extreme rainfall over space.

Let \(F_i\) be the distribution function of \((X_{i,1}, X_{i,2}, \ldots, X_{i,m})\). Assume that the distribution function \(F_{i,j}\) of \(X_{i,j}\) is continuous and let \(U_{i,j}(t) := F_{i,j}^{-1}(1 - 1/t)\), where the arrow indicates the generalized inverse function. To model the dependence, we further assume that

\[
\bar{F}(x_1, x_2, \ldots, x_m) := F_i(U_{i,1}(x_1), U_{i,2}(x_2), \ldots, U_{i,m}(x_m))
\]

does not depend on \(i\) and is in the domain of attraction of a multivariate extreme value distribution (de Haan and Ferreira, 2006, Chapter 6). As a consequence, the multivariate tail dependence structure does not depend on \(i\). Let \(R_{j_1,j_2}\) denote the tail copula of the components \(j_1\) and \(j_2\):

\[
R_{j_1,j_2}(x,y) = \lim_{t \downarrow 0} \, \frac{1}{t} P \left( 1 - F_{i,j_1}(X_{i,j_1}) \leq tx, \, 1 - F_{i,j_2}(X_{i,j_2}) \leq ty \right),
\]

for \((x, y) \in [0, \infty]^2 \setminus \{(\infty, \infty)\}\). Note that a tail copula is continuous, non-decreasing, and homogeneous of order one.

As in Einmahl et al. (2016), the estimator of \(C_j\) could be the number of exceedances over a high empirical quantile at station \(j\). But, since we want to compare the \(C_j\)’s, we want to use the same threshold for all rain stations. Consequently the common threshold will be a high empirical quantile of all \(N := n \times m\) observations taken together. Let \(X_{N-k:N}\) be the \((N - k)\)-th order statistic of the observations \(\{X_{i,j}\}_{i=1, j=1}^{n} m\). We define the estimator

\[
\hat{C}_j(t) := \frac{1}{k} \sum_{i=1}^{nt} \mathbb{1}\{X_{i,j} > x_{N-k:N}\}
\]

where \(k\) is an intermediate sequence: \(k = k(n) \to \infty, k(n)/n \to 0\), as \(n \to \infty\); the sum is over all integers \(1 \leq i \leq nt\).

In this paper we make the following four contributions.
1. We establish the joint asymptotic behavior of \( \{ \hat{C}_j(t) \}_{t \in [0,1]}, j = 1, \ldots, m \).

2. We test \( H_0 : C_j(1) = \frac{1}{n} \) for all \( j = 1, \ldots, m \), i.e., the total integrated scedasis is constant over the various locations. We perform the test by checking whether the limit vector (in distribution) of

\[
\left( \sqrt{k} \left( \hat{C}_1(1) - \frac{1}{m} \right), \sqrt{k} \left( \hat{C}_2(1) - \frac{1}{m} \right), \ldots, \sqrt{k} \left( \hat{C}_m(1) - \frac{1}{m} \right) \right)
\]

has mean zero. This will be done via an adapted \( \chi^2 \)-test.

3. We test \( H_{0,j} : C_j(t) = tC_j(1) \) for \( 0 \leq t \leq 1 \) and some given \( j \in \{ 1, \ldots, m \} \), i.e., the scedasis \( c(\cdot, j) \) is constant over time. Since, under \( H_{0,j} \), the limit in distribution of the process \( \{ \sqrt{k}(\hat{C}_j(t) - t\hat{C}_j(1)) \}_{0 \leq t \leq 1} \) is essentially a Brownian bridge, we can use, e.g., a Kolmogorov-Smirnov-type statistic.

4. We establish the asymptotic behavior of the pseudo-maximum likelihood estimator of \( \gamma \) based on all \( n \times m \) observations.

Crucial for these results is a joint weighted Gaussian approximation of the \( m \) sequential tail empirical processes as well as one for the tail quantile process based on all \( n \times m \) observations, for general \( \gamma \in \mathbb{R} \). The main challenge is to establish these results in the above setting of observations that are dependent and have different distributions. For the tail empirical processes this is achieved by disentangling these two complications, whereas for the tail quantile process, after aggregating the \( m \) tail empirical processes, a delicate proof is required to deal with the weight functions.

An early paper where in the univariate case a linear trend in the parameters of the limit distribution is studied is Davison and Smith (1990). Various models for spatial extremes, all quite different from the setup in the present paper, are reviewed in Davison et al. (2012), see also Coles and Tawn (1996) for a specific rainfall model. The assumption of constant tail dependence, cf. (1.2), is tested in Bücher et al. (2015).

The outline of the paper is as follows. Section 2 gives a detailed account of the conditions and the ensuing results. These results are applied to the mentioned rainfall data in Section 3. Proofs are collected in Section 4 and partly deferred to the Supplementary Material, along with a simulation study showing the performance of the proposed estimation and testing procedures and a validation of the assumptions for the data application.

### 2. Results

Throughout the paper we assume the following conditions.

(i) Multivariate dependence: Assume that \( \mathbf{F} \), defined in (1.2), does not depend on \( i \) and is in the domain of attraction of a multivariate extreme value distribution.

(ii) Sharpening of the scedasis condition (1.1): Assume that there exists an eventually decreasing function \( A_1 \) with \( \lim_{t \to \infty} A_1(t) = 0 \) such that

\[
\sup_{n \geq 1} \max_{1 \leq i, j} \left| \frac{1 - F_{i,j}(x)}{1 - F_0(x)} - c \frac{i}{n^\gamma} \right| = O \left( A_1 \left( \frac{1}{1 - F_0(x)} \right) \right), \quad x \uparrow x^*.
\]

(iii) Second order condition for \( F_0 \): Write \( U_0(t) := F_0^{-1}(1 - 1/t) \). There exists \( \gamma \in \mathbb{R}, \rho < 0 \) and functions \( \bar{a}_0 \), positive, and \( A_0 \) not changing sign eventually satisfying \( \lim_{t \to \infty} A_0(t) = 0 \) such that for all \( x > 0 \),

\[
\lim_{t \to \infty} \frac{U_0(t)x - U_0(t)}{\bar{a}_0(t)} A_0(t) = \frac{x^{\gamma-1}}{\gamma} = \Psi_{\gamma, \rho}(x) := \begin{cases} \frac{x^{\gamma+\rho-1}}{\gamma+\rho}, & \gamma + \rho \neq 0, \\ \log x, & \gamma + \rho = 0, \end{cases}
\]

cf. Corollary 2.3.5 in de Haan and Ferreira (2006).
(iv) Conditions on the intermediate sequence \( k \): Assume, as \( n \to \infty \),
\[
k \to \infty, \quad k/n \to 0,
\]
\[
\sqrt{k} A_1\left( q \frac{n}{k} \right) \to 0, \quad \text{for all } q > 0, \quad \sqrt{k} A_0\left( \frac{n}{k} \right) \to 0
\]
\[
\sqrt{k} \sup_{|u-v| \leq \frac{t}{k}} |c(u,j) - c(v,j)| \to 0, \quad \text{for } j = 1, 2, \ldots, m.
\]

Throughout the paper, quantities regarding \( \gamma = 0 \) should be read as the limit as \( \gamma \) tends to zero. In particular, in (2.1) if \( \gamma = 0 \), \((x^\gamma - 1)/\gamma \) should be read as \( \log x \). Likewise, \( (1 + \gamma x)^{1/\gamma} \) is meant to be \( e^x \). Write \( \gamma_+ = \gamma \vee 0 \) and \( \gamma_- = \gamma \wedge 0 \). If \( \gamma \leq 0 \), \(-1/\gamma_+ \) means \( -\infty \); if \( \gamma \geq 0 \), \( 1/(\gamma_-) \) means \( \infty \).

We begin with presenting two fundamental approximations, which are the basis for the main results (Theorem 2.3, Corollaries 2.4 and 2.5, Theorem 2.6), but they are also of independent interest. For the following theorem we need to bound the difference between the fraction to the left in (2.1) and its limit by a uniform inequality. For details, see the inequality (2.3.9) in Corollary 2.3.7, obtained from (B.3.19) in Theorem B.3.10 of de Haan and Ferreira (2006). Such an inequality holds provided that the functions \( U, a_0 \) and \( \Psi, \gamma, \rho \) given in Corollary 2.3.7.

**THEOREM 2.1.** Assume conditions (i)-(iv). Let \( x_0 > -1/\gamma_+ \); set \( x_1 := 1/(\gamma_-) \).

a) Tail empirical distribution functions

Using a Skorokhod construction, for \( 0 \leq \eta < 1/2 \), as \( n \to \infty \), it holds almost surely,
\[
\max_{1 \leq j \leq m} \sup_{x_0 \leq x_j < x_0 + t_j \leq 1} \left( 1 + \gamma x_j \right)^{n/\gamma} \left| \sqrt{k} \sum_{i=1}^m \left\{ \frac{x_j - x_0}{a_0}\left( \frac{x_j}{\gamma} \right) \right\} - (1 + \gamma x_j)^{-1/\gamma} C_j(t_j) - W_j\left( (1 + \gamma x_j)^{-1/\gamma}, C_j(t_j) \right) \right| \to 0,
\]
where \((W_1, \ldots, W_m)\) is a mean zero Gaussian vector of bivariate Wiener processes \( W_j \). Its covariances follow from the covariance matrix \( \Sigma = \Sigma(s_1, s_2, t_1, t_2) \) with entries
\[
\sigma_{j_1, j_2}(s_1, s_2, t_1, t_2) := \text{Cov}\left( W_{j_1}\left( s_1, C_{j_1}(t_1) \right), W_{j_2}\left( s_2, C_{j_2}(t_2) \right) \right) = \frac{1}{m} \int_0^{t_1 \wedge t_2} R_{j_1, j_2}\left( s_1 c(u, j_1), s_2 c(u, j_2) \right) du,
\]
for \( 1 \leq j_1, j_2 \leq m \).

b) Tail empirical quantile function

With any \( \varepsilon > 0 \) and \( X_{1:N} \leq \ldots \leq X_{N:N} \) the order statistics of the sample \( \{X_{i,j}\}_{i,j} \) of all \( N \) observations, we have, for \( T > 0 \), as \( n \to \infty \),
\[
\sup_{\frac{s}{n} \leq s \leq T} s^{-1/2+\varepsilon} \left| \sqrt{s}^\gamma \left( \frac{X_{N-[k]s}N - b_0(N)}{a_0}\left( \frac{s}{n} \right)^{1/\gamma} \right) - s^{-\gamma - 1/\gamma} - \sum_{j=1}^m W_j\left( s, C_j(1) \right) \right| \to 0.
\]

We consider the two boundary cases for the dependence structure. In case of tail independence, that is \( R_{j_1, j_2} \equiv 0 \), for all pairs \((j_1, j_2), j_1 \neq j_2\), the bivariate Wiener processes \( W_1, \ldots, W_m \) are independent, and (2.2) for \( \gamma > 0 \) specializes to a generalization of the results in Einmahl et al. (2016) where independence of the data is assumed. Further specializing the situation to having \( N \) i.i.d. data, Theorem 2.1 recovers Theorem 2.4.2 and generalizes Theorem 5.1.2 in de Haan and Ferreira (2006), on tail quantile and tail empirical
processes, respectively. In case of complete tail dependence between two locations, that is \( R_{j_1,j_2}(x,y) = x \wedge y \), we obtain \( \sigma_{j_1,j_2}(s_1, s_2, t_1, t_2) = \frac{1}{m} \int_{0}^{t_1} \int_{0}^{t_2} (s_1 c(u, j_1)) \wedge (s_2 c(u, j_2)) \, du \). In case \( c(\cdot, j_1) = c(\cdot, j_2) \), the covariance simplifies to \( (s_1 \wedge s_2) C_{j_1}(t_1 \wedge t_2) \), which means that the processes \( W_{j_1} \) and \( W_{j_2} \) are the same.

**Corollary 2.2.** Under the conditions and in the setup of Theorem 2.1, with \( \varepsilon > 0 \),

\[
\sup_{\frac{1}{\varepsilon} \leq s \leq T} \left( 1 + s^{\gamma + 1/2 + \varepsilon} \right) \sqrt{k} \left( \frac{X_{N-[k\varepsilon]:N} - X_{N-k\varepsilon:N}}{a_0(N/k)} - s^{-\gamma} - 1 \right) - \sum_{j=1}^{m} \left( s^{-\gamma} W_j(s, C_j(1)) - W_j(1, C_j(1)) \right) \xrightarrow{P} 0.
\]

As a result we get the joint asymptotic behavior of the \( \hat{C}_j, j = 1, \ldots, m \).

**Theorem 2.3.** Under the conditions and in the setup of Theorem 2.1, as \( n \to \infty \),

\[
\max_{1 \leq j \leq m} \sup_{0 \leq t \leq 1} \left| \sqrt{k} (\hat{C}_j(t) - C_j(t)) - \left\{ W_j(1, C_j(1)) - C_j(t) \sum_{r=1}^{m} W_r(1, C_r(1)) \right\} \right| \xrightarrow{P} 0,
\]

Moreover, we have the uniform consistency of the estimator of the covariance matrix \( \Sigma \) as follows. For \( T > 0 \) and \( j_1 \neq j_2 \), as \( n \to \infty \),

\[
\sup_{0 \leq t \leq 1} \sup_{0 \leq s_1, s_2 \leq T} \left| \frac{1}{k} \sum_{i=1}^{n} I \{ X_{i,j_1} > X_{N-[k\varepsilon]:N}, X_{i,j_2} > X_{N-[k\varepsilon]:N} \} - \frac{1}{m} \int_{0}^{t} R_{j_1,j_2} \left( s_1 c(u, j_1), s_2 c(u, j_2) \right) \, du \right| \xrightarrow{P} 0.
\]

Now we proceed with the two aforementioned tests. First we discuss the testing problem

\[
\begin{cases}
H_0 : C_j(1) = \frac{1}{m}, & \text{for all } j = 1, \ldots, m, \\
H_1 : C_j(1) \neq \frac{1}{m}, & \text{for some } j = 1, \ldots, m.
\end{cases}
\]

Let \( I_m \) be the \( m \)-unit vector, \( I_m \) the identity matrix of dimension \( m \) and define \( M := I_m - \frac{1}{m} I_m \). From Theorem 2.3 and under \( H_0 \), \( D = \sqrt{k} \left( \hat{C}_1(1) - \frac{1}{m}, \ldots, \hat{C}_m(1) - \frac{1}{m} \right)^t \) is asymptotically \( m \)-variate normal with zero mean vector and covariance matrix \( M \Sigma_1 M' \), where \( \Sigma_1 = 1 \). Assume that \( \Sigma_1 \) is invertible. Then \( \text{rank}(M \Sigma_1 M') = \text{rank}(M) = m - 1 \). We therefore confine attention to the first \( m - 1 \) components of \( D \) denoted by \( D_{m-1} \), which has an asymptotic covariance matrix \( (M \Sigma_1 M')_{m-1} \). Here for an \( m \times m \) matrix \( A \), the notation \( A_{m-1} \) refers to the matrix consisting of the first \( m - 1 \) rows and \( m - 1 \) columns of \( A \). Finally, we define the test statistic

\[
T_n := D_{m-1}' \left( (M \hat{\Sigma}_1 M')_{m-1} \right)^{-1} D_{m-1},
\]

with \( \Sigma_1 \) estimated via the empirical counterpart given in (2.7). From Theorem 2.3 we immediately get the asymptotic behavior of \( T_n \) under \( H_0 \).

**Corollary 2.4.** Assume that \( \Sigma_1 \) is invertible. Then under \( H_0 \), \( T_n \buildrel d \over \to \chi_{m-1}^2 \), as \( n \to \infty \).
Next we consider, for \( j \in \{1, \ldots, m\} \), the testing problem \( H_{0,j} : C_j(t) = tC_j(1), 0 \leq t \leq 1 \), and \( H_{1,j} \) that this is not the case. This null hypothesis means that the scedasis for station \( j \), \( c(\cdot, j) \), is constant over time. We can use test statistics of the Kolmogorov-Smirnov-type or Cramér-von Mises-type based on the process \( \sqrt{k}(\hat{C}_j(t) - t\hat{C}_j(1))/\sqrt{\hat{C}_j(1)}, 0 \leq t \leq 1 \).

**Corollary 2.5.** Fix \( j \in \{1, \ldots, m\} \). Under the hypothesis that \( C_j(t) = tC_j(1) \), for \( 0 \leq t \leq 1 \),

\[
\left\{ \sqrt{k}\hat{C}_j(1) \left( \frac{\hat{C}_j(t)}{\hat{C}_j(1)} - t \right) \right\}_{t \in [0,1]} \xrightarrow{d} \{ B(t) \}_{t \in [0,1]},
\]

with \( B \) a Brownian bridge.

Finally, we introduce the pseudo-maximum likelihood estimator (MLE) of \( (\gamma,a_0(\frac{N}{k})) \). The estimator is based on the \( N = n \times m \) dependent and non-identically distributed observations (as described in Section 1) and in particular on the order statistics of excesses \( \{X_{N \times i+1:N} - X_{N \times k:N}\}_{i=1}^k \). The (pseudo) ML procedure is based on the assumption that these order statistics are taken from the (limiting) generalized Pareto distribution \( 1 - (1 + \gamma x/\sigma)^{-1/\gamma}, \gamma \in \mathbb{R}, \sigma > 0 \), where \( \sigma \) represents the scale component \( a_0(\frac{N}{k}) \); for more details see de Haan and Ferreira (2006), Section 3.4. This leads to the log-likelihood

\[
\ell(\gamma,\sigma,x) = -\log \sigma - \left( 1 + \frac{1}{\gamma} \right) \log \left( 1 + \frac{\gamma x}{\sigma} \right), \quad 0 < x < \frac{\sigma}{-\gamma} - 1
\]

(2.8)

(for \( \gamma = 0 \) the formula is interpreted as \( -\log \sigma - x/\sigma \)). The pseudo log-likelihood based on the above sample can be written as, with parameter space \( (\gamma,\sigma) \in \mathbb{R} \times (0,\infty) \),

\[
L_{N,k}(\gamma,\sigma_{N/k}) = \sum_{i=1}^{k} \ell(\gamma,\sigma_{N/k}, X_{N \times i+1:N} - X_{N \times k:N})
\]

\[
= k \int_{0}^{1} \ell(\gamma,\sigma_{N/k}, X_{N \times \lfloor ks\rfloor:N} - X_{N \times k:N}) \, ds
\]

(2.9)

\[
= k \int_{0}^{1} \ell \left( \gamma, \sigma_{N/k}, \frac{X_{N \times \lfloor ks\rfloor:N} - X_{N \times k:N}}{a_0(N/k)} \right) \, ds - k \log a_0(N/k).
\]

Generally \( (\hat{\gamma},\hat{\sigma}_{N/k}) \) is an MLE if it is a local maximizer of \( L_{N,k}(\gamma,\sigma_{N/k}) \) solving the score equations,

\[
\left\{ \frac{\partial}{\partial \gamma} L_{N,k}(\gamma,\sigma_{N/k}) = 0 \right. \left. \frac{\partial}{\partial \sigma} L_{N,k}(\gamma,\sigma_{N/k}) = 0 \right. \]

In the following we highlight \( \gamma_0 \) as the true unknown parameter value.

**Theorem 2.6.** Under conditions (i)-(iv) with \( \gamma_0 > -1/2 \), with probability tending to 1, there exists a unique sequence of estimators \( (\hat{\gamma}_n,\hat{a}_0(N/k)) \), maximizing (2.9), for which

\[
\sqrt{k} \left( \hat{\gamma}_n - \gamma_0, \left( \frac{\hat{a}_0(N/k)}{a_0(N/k)} - 1 \right) \right) \xrightarrow{d} N(0, I_{\gamma_0}^{-1} \Sigma_{\gamma_0} I_{\gamma_0}^{-1})
\]

where

\[
I_{\gamma_0}^{-1} = \left[ \begin{array}{cc}
(\gamma_0 + 1)^2 & - (\gamma_0 + 1) \\
- (\gamma_0 + 1) & 2(\gamma_0 + 1)
\end{array} \right]^{-1}
\]
and \( \Sigma_{\gamma_0} \) is the covariance matrix of the random vector

\[
\begin{bmatrix}
\sum_{j=1}^{m} \int_{0}^{1} \left( s^{\gamma_0+1} - (1 + \gamma_0) s^{2\gamma_0} \right) \left\{ s^{-\gamma_0-1} W_j(s, C_j(1)) - W_j(1, C_j(1)) \right\} ds \\
\sum_{j=1}^{m} (1 + \gamma_0) \int_{0}^{1} s^{2\gamma_0} \left\{ s^{-\gamma_0-1} W_j(s, C_j(1)) - W_j(1, C_j(1)) \right\} ds
\end{bmatrix}
\]

with \( \{W_j\}_{j=1}^{m} \) from Theorem 2.1.

**Remark 1.** The covariance matrix \( \Sigma_{\gamma_0} \) can be calculated as follows: let \( U \) be a \((1 \times 2m)\) vector with the first \( i = 1, \ldots, m \), components as

\[
\gamma_0^{-1} \int_{0}^{1} \left( s^{\gamma_0} - (1 + \gamma_0) s^{2\gamma_0} \right) \left\{ s^{-\gamma_0-1} W_i(s, C_i(1)) - W_i(1, C_i(1)) \right\} ds,
\]

and the remaining \( i = m+1, \ldots, 2m \) components as

\[
(1 + \gamma_0) \int_{0}^{1} s^{2\gamma_0} \left\{ s^{-\gamma_0-1} W_i(s, C_i(1)) - W_i(1, C_i(1)) \right\} ds.
\]

The covariance matrix of \( U \) has entries \( \tau_{ij}, i = 1, \ldots, 2m, j = 1, \ldots, 2m \), given by,

\[
\tau_{ii} = \frac{2 + 6\gamma_0 + 5\gamma_0^2}{(1 + \gamma_0)^2(1 + 2\gamma_0)^2} C_i(1), \quad i = 1, \ldots, m;
\]

\[
\tau_{ii} = \left( \frac{1 + \gamma_0}{1 + 2\gamma_0} \right)^2 C_{i-m}(1), \quad i = m + 1, \ldots, 2m;
\]

\[
\tau_{i,i+m} = \tau_{i+m,i} = \frac{1 + \gamma_0}{(1 + 2\gamma_0)^2} C_i(1), \quad i = 1, \ldots, m;
\]

\[
\tau_{ij} = \int_{0}^{1} \int_{0}^{1} f(s) f(t) r_{ij}(s, t) - 2f(s)g(t) r_{ij}(s, 1) + g(s) g(t) r_{ij}(1, 1) ds dt,
\]

\( i \neq j, i, j = 1, \ldots, m; \)

\[
\tau_{ij} = (1 + \gamma_0)^2 \left\{ \int_{0}^{1} \int_{0}^{1} s^{\gamma_0+1} t^{\gamma_0+1} r_{i-j-m}(s, t) - 2s^{\gamma_0+1} t^{\gamma_0+1} r_{i-j-m}(s, 1)
\]

\[+ s^{\gamma_0+1} t^{\gamma_0+1} r_{i-j-m}(1, 1) ds dt \right\},
\]

\( i \neq j, i = m + 1, \ldots, 2m, j = m + 1, \ldots, 2m; \)

\[
\tau_{ij} = (1 + \gamma_0) \left\{ \int_{0}^{1} \int_{0}^{1} f(s) t^{\gamma_0+1} r_{i-j-m}(s, t) - f(s) t^{\gamma_0+1} r_{i-j-m}(s, 1)
\]

\[- g(s) t^{\gamma_0+1} r_{i-j-m}(1, t) + g(s) t^{\gamma_0+1} r_{i-j-m}(1, 1) ds dt \right\},
\]

\( j \neq i + m, i = 1, \ldots, m, j = m + 1, \ldots, 2m; \)

\[
\tau_{ij} = (1 + \gamma_0) \left\{ \int_{0}^{1} \int_{0}^{1} f(s) t^{\gamma_0+1} r_{i-m-j}(s, t) - f(s) t^{\gamma_0+1} r_{i-m-j}(s, 1)
\]

\[- g(s) t^{\gamma_0+1} r_{i-m-j}(1, t) + g(s) t^{\gamma_0+1} r_{i-m-j}(1, 1) ds dt \right\},
\]

\( i \neq j + m, j = 1, \ldots, m, i = m + 1, \ldots, 2m; \)

with

\[
f(s) = s^{-1} \gamma_0^{-1} (1 - (1 + \gamma_0)s^{\gamma_0}), \quad g(s) = s^{\gamma_0+1} \gamma_0^{-1} (1 - s^{\gamma_0}) - s^{2\gamma_0}.
\]
\[ r_{ij}(s,t) = \sigma_{i,j}(s,t,1,1) = EW_i(s,C_i(1))W_j(t,C_j(1)). \]

Then \( \Sigma_{\gamma_0} = \text{Cov}(I \times U) = I \text{Cov}(U) I^T \) where \( I = I_{2 \times 2m} = \begin{bmatrix} 1 \cdots 1 & 0 \cdots 0 \\ 0 \cdots 0 & 1 \cdots 1 \end{bmatrix} \).

We consider the two boundary cases for the dependence structure again. In case of tail independence for all pairs \((j_1, j_2)\), the matrix \( I_{\gamma_0}^{-1} \Sigma_{\gamma_0} I_{\gamma_0}^{-1} \) reduces to

\[
\Sigma_{\text{iid}} = \begin{bmatrix} (\gamma_0 + 1)^2 & -(\gamma_0 + 1) \\ -(\gamma_0 + 1) & 1 + (\gamma_0 + 1)^2 \end{bmatrix},
\]

which is the same covariance matrix as that for \( N \) i.i.d. data. In case of complete tail dependence for all pairs \((j_1, j_2)\) and if all \( m \) scedases \( c(\cdot, j) \) are the same, the matrix \( I_{\gamma_0}^{-1} \Sigma_{\gamma_0} I_{\gamma_0}^{-1} \) becomes \( m \Sigma_{\text{iid}} \), which is the same as that for \( n \) i.i.d. data with \( k \) replaced by \( k/m \).

As an example, consider \( N \) data such that all \( m \) scedases \( c(\cdot, j) \) are the same and the tail dependence structure is described by the well-known logistic stable tail dependence function

\[
l_{\theta}(x_1, \ldots, x_m) = \left( \sum_{j=1}^{m} x_j^{1/\theta} \right)^{\theta}, \quad x_j \geq 0;\]

cf. Einmahl et al. (2012). Here the parameter \( \theta \in (0, 1] \) reflects the degree of tail dependence, with \( \theta = 1 \) corresponding to tail independence and the limiting case \( \theta = 0 \) corresponding to complete tail dependence. Then for all \( j_1 \neq j_2 \) the pairwise tail copula becomes

\[
R_{j_1,j_2}(x,y) = x + y - \left( x^{1/\theta} + y^{1/\theta} \right)^{\theta} = mr_{j_1,j_2}(x,y), \quad x,y \geq 0.
\]

**Figure 1.** Asymptotic variance of the MLE of \( \gamma_0 \) for the logistic stable tail dependence function in dimension \( m = 20 \), as a function of the dependence parameter \( \theta \) and the extreme value index \( \gamma_0 \).

With the latter formula we can numerically compute the \( \tau_{ij} \) in Remark 1 and obtain \( \Sigma_{\gamma_0}(\theta) \).

Figure 1 shows, for \( m = 20 \), the asymptotic variance of \( \sqrt{k} (\hat{\gamma}_0 - \gamma_0) \) as a function of \( \theta \) and \( \gamma_0 \). The variance increases smoothly with dependence and with \( \gamma_0 \). For fixed \( \gamma_0 \), it increases from \((\gamma_0 + 1)^2\) to \(20(\gamma_0 + 1)^2\), as \( \theta \) decreases from 1 to 0. Clearly, tail independence yields the most accurate estimation of \( \gamma_0 \).
3. Application. This section is devoted to illustrating the testing methods for detecting a trend in extreme rainfalls, both across stations and over time. We use a subset of rainfall data from the German national meteorological service, which consists of daily rainfall amounts recorded in 49 stations ($m = 49$) in three regions of North-West Germany: Bremen, Niedersachsen and Hamburg. The data set comprises nearly complete time series records over 84 years (1931-2014). We divide the data into two seasons: winter from November to March, and summer from May to September, excluding the transitional months April and October.

Although the raw data set comprises a tally of $49 \times 84 \times 150$ rainfall amounts within each season, the actual number of observations we use at each station, $n$, will be determined by a declustering procedure. This has been designed to remove the effect of temporal dependence and is viewed as a key step to ensure that after pre-processing, the data set can be regarded as having no temporal dependence. The idea of our pre-processing procedure is to create gaps between consecutive observations by removing some days in the data set. The detailed procedure we have employed is outlined in the next paragraph.

The raw data set consists of daily rainfall amounts (in mm) at each gauging station $j = 1, \ldots, m$, including zero rainfall. From this data set we will use the daily maximum rainfall amount across the $m$ stations, henceforth referred to as station-wise maxima, for eliciting potential serial dependence. We order all station-wise maxima from high to low. The declustering procedure is initiated by picking up the pair of calendar days with the largest and second largest station-wise maxima. If this second maximum was recorded within two consecutive days of the first station-wise maximum, then all $m$ observations on its corresponding day are removed; otherwise both days are kept. This procedure then rolls out to the subsequent ordered station-wise maxima: for each station-wise maxima, we remove the corresponding day if it is recorded within two consecutive days of any of the previously kept days. This results in the declustered data sets with sample sizes $n = 3561$ and $n = 3552$ for winter and summer respectively. The two data sets are used for testing the presence of scedasis over time and/or across space, whilst accounting for the spatial dependence.\footnote{We make the declustered data sets and the codes for application available at https://github.com/zhouchen0527/rainscedasis}

In addition, we check the assumptions of having a constant extreme value index over time and over space (cf. Sections 3.2 and 3.3 in the Supplementary Material) and find no evidence...
against making such assumptions. In relation to the spatial dependence, we present in Figure 2 two plots displaying estimates for the pairwise spatial dependence through

$$m \tilde{\sigma}_{j_1,j_2}(1,1,1,1) := \frac{m}{k} \sum_{i=1}^{n} \mathbb{I} \{ X_{i,j_1} > X_{N-k:N}, X_{i,j_2} > X_{N-k:N} \},$$

as in (2.7), for $k = 300, \ldots, 1500$. Note that the choice of $k$ in this estimation differs from that in the estimation of tail dependence coefficient by a factor of $m$, i.e. $k = 300$ corresponds to choosing $k = 6$ at each station. In addition, we choose to estimate $m \sigma_{j_1,j_2}(1,1,1,1)$, since by construction, a factor $1/m$ appears in the expression for $\sigma_{j_1,j_2}(1,1,1,1)$, see (2.3). The plots in Figure 2 illustrate estimation of pairwise spatial dependence for one single pair of stations for winter and for summer. These suggest that spatial dependence is stronger in the winter than in the summer, a finding consistent with both the convective nature of extreme rainfall in the summer and widespread persistent precipitation in the winter.

**FIG 3.** Obtained $p$-values through the test $T_n$ for homogeneity across space, all plotted as a function of $k = 300, \ldots, 1500$. 
First, we test whether the total integrated scedasis of extreme rainfall is constant across $m = 49$ stations by adopting the test statistic $T_n$ in Corollary 2.4. We reject the null hypothesis of having constant total integrated scedasis across all stations for large values of $T_n$. We plot the $p$-values against $k$ the number of upper observations used in the test in the two plots of Figure 3, for winter and summer seasons respectively. The $p$-values obtained for the winter season stay below 5% for all $k > 350$. Therefore, we conclude that for the winter season, the total integrated scedasis of extreme rainfall is not constant across stations. In other words, the frequencies of having extreme rainfall differ across stations. In contrast, there is no statistical evidence of a trend in the space-domain over the summer. This finding holds for almost all values of $k$ in the lower panel of Figure 3.

![Winter](image1)

![Summer](image2)

**Figure 4.** Obtained $p$-values for the test of the null hypothesis $H_{0,j}$ of a non-existent time trend at each station $j = 1, 2, \ldots, 49$. The Kolmogorov-Smirnov statistic is applied with $k = 1000$ higher observations. The numbers next to the station-marks indicate elevation in meters; one station on the east side falls outside the map.
Next, we investigate a possible temporal trend in the extreme rainfall process for each station mapped in Figure 4 by means of a Kolmogorov-Smirnov (KS) type test based on the left hand side of the limit relation in Corollary 2.5. For each season, we apply this test at each station \( j \) with \( k = 1000 \), and plot the \( p \)-values of the test in the two plots in Figure 4, for winter and summer seasons respectively. The sharper the red in the renderings, the lower the estimated \( p \)-values, and the more evidence for rejecting the null hypothesis. The brighter the green marks, the higher the \( p \)-values. We find that the \( p \)-values vary widely across the selected region, and more so in the winter.

Overall, we find that \( p \)-values plunge in the winter but soar in the summer at many locations. In the winter season, the KS type test highlights two stations with \( p \)-values below the nominal level \( \alpha = 5\% \): \( p = 0.044 \) for station Steinau, Kr. Cuxhaven, with elevation 1m, and \( p = 0.05 \) for Bramsche at 58m high. Nevertheless, we need to interpret such \( p \)-values with caution. Given that these are the lower \( p \)-values across all 49 stations, we are encountering a potential multiple test problem. One potential solution is to consider the Bonferroni correction: the corrected nominal level is \( \alpha^* = 5\%/49 \approx 0.1\% \). Since these low \( p \)-values do not breach the corrected nominal level, we find no temporal trend over the winter, at the usual significance levels. Similarly, for the summer, the KS type test identifies one significant individual \( p \)-value of 0.048 for station Uetze, standing at 60m of elevation. Again this individual \( p \)-value is not in the vicinity of the Bonferroni’s corrected critical barrier \( \alpha^* = 0.1\% \). To summarize, there seems to be no temporal trend in extreme rainfalls in the winter or in the summer.

Finally, we report the estimated extreme value index \( \gamma \) using all data from all stations in one season, using the maximum likelihood estimator (MLE) in Theorem 2.6. Figure 5 shows the estimates against various values of \( k \), for the winter and summer seasons respectively. We observe that, for \( k \) ranging between 900 and 1100, both estimates paths seem to consolidate a plateau of stability. For the purpose of point estimation, we fix \( k = 1000 \), highlighted in both plots with a vertical gray line. The estimated extreme value indices are \( \hat{\gamma} = 0.041 \) and 0.078 for winter and summer seasons, respectively, with corresponding 95\% confidence intervals \((-0.0546, 0.1365)\) and \((0.0159, 0.1391)\). We find that the estimated standard deviation for the winter, based on the asymptotic distribution of the maximum likelihood estimator for \( \gamma \), is greater than \((1 + \hat{\gamma})/\sqrt{k}\), the estimated standard deviation assuming tail independence. This indicates that ignoring spatial dependence results in underestimation of the asymptotic variance. The analogous results for the summer season lead to estimated standard deviations nearly matching the values \((1 + \hat{\gamma})/\sqrt{k}\), which suggests weaker spatial tail dependence during the summer months.

4. Proofs. Write for convenience \( X_{i,j} = U_{i,j}(Y_{i,j}) \), where \((Y_{i,1}, Y_{i,2}, \ldots, Y_{i,m})\) follows the distribution function \( \tilde{F} \) with standard Pareto marginals. Let \( Y_{i,m}^{(j)} \) be the \( i \)-th order statistic from \((Y_{1,j}, Y_{2,j}, \ldots, Y_{n,j})\), for all \( j \).

**Proof of Theorem 2.1 a) Tail empirical distribution functions**

Consider one station \( j \) for the time being and define \( C_{j,n}(t) := \frac{1}{N} \sum_{i=1}^{nt} c \left( \frac{i}{nt}, j \right) \). (Recall \( N = nm \).) According to Proposition 1 in Einmahl et al. (2016) we have under a Skorokhod construction for any \( t_0 > 0 \) and \( 0 \leq \eta < 1/2 \), almost surely,

\[
\sup_{0 < v \leq t_0, 0 \leq t \leq 1} v^{-\eta} \sqrt{k} \left\{ \frac{1}{k} \sum_{i=1}^{nt} \mathbb{I} \left\{ Y_{i,j} > \frac{mv C_{j,n}(t)}{C_{j,n}(1)} \right\} - \frac{v C_{j}(t)}{C_{j,n}(1)} \right\} - W_j \left( v, \frac{C_{j}(t)}{C_{j,n}(1)} \right) \rightarrow 0,
\]

as \( n \rightarrow \infty \). After some rearrangement we get, almost surely,

\[
(4.1) \quad \sup_{0 < v \leq t_0, 0 \leq t \leq 1} v^{-\eta} \sqrt{k} \left\{ \frac{1}{k} \sum_{i=1}^{nt} \mathbb{I} \left\{ Y_{i,j} > \frac{mv C_{j}(t)}{C_{j,n}(1)} \right\} - mv C_{j}(t) \right\} - W_j \left( mv, C_{j}(t) \right) \rightarrow 0.
\]
We are going to transform this result in several steps. First replace $v$ with $\frac{n}{ke(i_n,j)}\left(1 - F_{i,j}(U_0(N_{ku}^i))\right)$, $0 < u \leq t_0$. Note that, as $n \to \infty$, by condition (ii),

$$\frac{n}{ke(i_n,j)}\left(1 - F_{i,j}(U_0(N_{ku}^i))\right) = \frac{n}{k}\left(1 - F_0(U_0(N_{ku}^i))\right)\left\{1 + O(A_1(N_{ku}^i))\right\}$$

$$= \frac{u}{m}\left\{1 + O(A_1(N_{ku}^i))\right\},$$
and by condition (iv), \( \sqrt{k} A_1 \left( \frac{N}{ku} \right) \to 0 \) uniformly for \( 0 < u \leq t_0 \). Hence we have, almost surely,

\[
\sup_{0 < u \leq t_0, 0 \leq t \leq 1} u^{-q} \left\{ \sqrt{k} \left\{ \frac{1}{k} \sum_{i=1}^{nt} \mathbb{1} \{ U_{i,j}(Y_{i,j}) > u a_0(\frac{N}{k}) \} - u C_j(t) \right\} - W_j(u, C_j(t)) \right\} \to 0.
\]

Next replace \( u \) with \( \frac{N}{k} \left( 1 - F_0 \left( b_0(\frac{N}{k}) + x a_0(\frac{N}{k}) \right) \right) \) and note that by condition (iii) and Proposition 3.2 (equation 3.2) of Drees et al. (2006)

\[
\frac{N}{k} \left( 1 - F_0 \left( b_0(\frac{N}{k}) + x a_0(\frac{N}{k}) \right) \right) = (1 + \gamma x)^{-1/\gamma} \left\{ 1 + O \left( A_0(\frac{N}{k}) \right) \right\}
\]

uniformly for \( x \geq x_0 > -1/\gamma \) and \( \sqrt{k} A_0(\frac{N}{k}) \to 0 \) by condition (iv). Recall \( X_{i,j} = U_{i,j}(Y_{i,j}) \). Hence we have for each \( j \),

\[
(4.2) \quad \sup_{x_0 \leq x_j < 1, 0 \leq t \leq 1} (1 + \gamma x_j)^{1/\gamma} \left\{ \sqrt{k} \left\{ \frac{1}{k} \sum_{i=1}^{nt} \mathbb{1} \{ \frac{x_{i,j} - b_0(\frac{N}{k})}{a_0(\frac{N}{k})} > x_j \} - (1 + \gamma x_j)^{-1/\gamma} C_j(t) \right\} - W_j((1 + \gamma x_j)^{-1/\gamma}, C_j(t)) \right\} \to 0,
\]

which yields the weak convergence of the weighted process on the left to the weighted Wiener process on the right.

It remains to prove the joint convergence of the processes at different locations. First we deal with convergence of the finite dimensional distributions. After that we consider tightness. For ease of writing we confine ourselves to the first two dimensions, i.e. \( \{(X_{i,1}, X_{i,2})\}_{i=1}^n \) and one point \((s_1, t_1)\) and \((s_2, t_2)\) at each dimension. According to the Cramèr-Wold device, we look first at all linear combinations \( (u, v) \in \mathbb{R}, x_0 \leq x, y < x_1, 0 \leq t_1, t_2 \leq 1 \)

\[
(4.3) \quad \sqrt{k} \left[ u \frac{1}{k} \sum_{i=1}^{nt_1} \mathbb{1} \{ \frac{x_{i,1} - b_0(\frac{N}{k})}{a_0(\frac{N}{k})} > x \} - P \left\{ \frac{X_{i,1} - b_0(\frac{N}{k})}{a_0(\frac{N}{k})} > x \right\} \right] \\
+ v \frac{1}{k} \sum_{i=1}^{nt_2} \mathbb{1} \{ \frac{x_{i,2} - b_0(\frac{N}{k})}{a_0(\frac{N}{k})} > y \} - P \left\{ \frac{X_{i,2} - b_0(\frac{N}{k})}{a_0(\frac{N}{k})} > y \right\} \right]\].
\]

The Lindeberg-Feller central limit theorem applies to this expression since the summands are indicators and hence bounded. As a consequence we obtain that

\[
\sqrt{k} \left\{ \frac{1}{k} \sum_{i=1}^{nt_1} \mathbb{1} \{ \frac{x_{i,1} - b_0(\frac{N}{k})}{a_0(\frac{N}{k})} > x \} - P \left\{ \frac{X_{i,1} - b_0(\frac{N}{k})}{a_0(\frac{N}{k})} > x \right\}, \right. \\
\frac{1}{k} \sum_{i=1}^{nt_2} \mathbb{1} \{ \frac{x_{i,2} - b_0(\frac{N}{k})}{a_0(\frac{N}{k})} > y \} - P \left\{ \frac{X_{i,2} - b_0(\frac{N}{k})}{a_0(\frac{N}{k})} > y \right\} \right\}
\]

converges in distribution to \( \left( W_1((1 + \gamma x)^{-1/\gamma}, C_1(t_1)), W_2((1 + \gamma y)^{-1/\gamma}, C_2(t_2)) \right) \). Note that, e.g. for the first dimension, by assumptions (ii) and (iv),

\[
\frac{1}{mn} \sum_{i=1}^{nt_1} N P \left\{ \frac{X_{i,1} - b_0(\frac{N}{k})}{a_0(\frac{N}{k})} > x \right\} = \frac{1}{mn} \sum_{i=1}^{nt_1} \left( 1 + \gamma x \right)^{-1/\gamma} c \left( \frac{i}{n}, 1 \right) + o \left( \frac{1}{\sqrt{k}} \right) \\
= (1 + \gamma x)^{-1/\gamma} C_1(t_1) + o \left( \frac{1}{\sqrt{k}} \right).
\]
It follows that
\[
(4.4) \quad \sqrt{k} \left( \frac{1}{n} \sum_{i=1}^{nt_1} \mathbb{I} \left\{ \frac{x_{i,1} - b_0 \left( \frac{N}{k} \right)}{a_0 \left( \frac{N}{k} \right)} > x \right\} - (1 + \gamma x)^{-1/\gamma} C_1(t_1) \right),
\]
\[
\frac{1}{k} \sum_{i=1}^{nt_2} \mathbb{I} \left\{ \frac{x_{i,2} - b_0 \left( \frac{N}{k} \right)}{a_0 \left( \frac{N}{k} \right)} > y \right\} - (1 + \gamma y)^{-1/\gamma} C_2(t_2)
\]
converges in distribution to \( W_1 \left( (1 + \gamma x)^{-1/\gamma} C_1(t_1) \right), W_2 \left( (1 + \gamma y)^{-1/\gamma} C_2(t_2) \right). \)

Next we prove tightness of the process in (4.4) in the space \( D([0, 1]^2 \times [x_0, x_1]) \) with index \((t_1, t_2, x, y)\). We know from (4.2) that for \( \eta \leq 1/2 \) the sequence of processes
\[
(1 + \gamma x_j)^{\eta/\gamma} \sqrt{k} \left( \frac{1}{n} \sum_{i=1}^{nt_j} \mathbb{I} \left\{ \frac{x_{i,j} - b_0 \left( \frac{N}{k} \right)}{a_0 \left( \frac{N}{k} \right)} > x_j \right\} - (1 + \gamma x_j)^{-1/\gamma} C_j(t_j) \right)
\]
is tight in the space \( D([0, 1] \times [x_0, x_1]) \), for \( j = 1 \) and \( j = 2 \). It then follows from Ferger and Vogel (2015) that the joint process (4.4) is also tight. Hence the weak convergence is established. A Skorokhod construction yields the result.

For the proof of (2.3) consider the covariance of the two components in (4.4). It suffices to show that the leading term in this covariance
\[
(4.5) \quad \to \frac{1}{m} \int_{t_1 \wedge t_2} \int_0^s R_{1,2}(s_1 c(u, 1), s_2 c(u, 2)) du, \quad n \to \infty,
\]
for fixed \((t_1, t_2, s_1, s_2) \in [0, 1]^2 \times [0, T]^2\), where \( T = (1 + \gamma x_0)^{-1/\gamma} > 0 \). Since \( X_{i,j} = U_{i,j}(Y_{i,j}) \) with \( Y_{i,j} \) standard Pareto distributed random variables, given any \( \varepsilon > 0 \), for sufficiently large \( n \),
\[
P \left( \frac{X_{i,1} - b_0 \left( \frac{N}{k} \right)}{a_0 \left( \frac{N}{k} \right)} > \frac{s_1^\gamma - 1}{\gamma}, \frac{X_{i,2} - b_0 \left( \frac{N}{k} \right)}{a_0 \left( \frac{N}{k} \right)} > \frac{s_2^\gamma - 1}{\gamma} \right) = \frac{1}{N} \left[ 1 - F_{i,1} \left( \frac{N}{k} \right) + a_0 \left( \frac{N}{k} \right) \frac{s_1^\gamma - 1}{\gamma} \right]^{m},
\]
\[
\frac{k}{n} \frac{1}{Y_{i,2}} > \frac{N}{k} \left[ 1 - F_{i,2} \left( \frac{N}{k} \right) + a_0 \left( \frac{N}{k} \right) \frac{s_2^\gamma - 1}{\gamma} \right]^{m},
\]
\[
\leq P \left( \frac{1}{s_1 c \left( \frac{k}{n}, 1 \right)} (1 + \varepsilon)^{s_1 c \left( \frac{k}{n}, 1 \right)}, \frac{1}{s_2 c \left( \frac{k}{n}, 2 \right)} (1 + \varepsilon) \right).
\]
The definition of \( R_{1,2} \) implies that, as \( n \to \infty \),
\[
n \frac{1}{k} P \left( \frac{k}{n} Y_{i,1} > \frac{1}{v_1}, \frac{k}{n} Y_{i,2} > \frac{1}{v_2} \right) \to R_{1,2}(v_1, v_2).
\]
uniformly for \((v_1, v_2)\) \(\in [0, V]^2\) with any fixed \(V > 0\). By the continuity and boundedness of \(c\) and of \(R_{1,2}\), as \(n \to \infty\),

\[
\frac{n}{k} P \left( \frac{k}{n} Y_{i,1} > \frac{m}{s_1 c \left( \frac{i}{n}, 1 \right) (1 + \varepsilon)}, \frac{k}{n} Y_{i,2} > \frac{m}{s_2 c \left( \frac{i}{n}, 2 \right) (1 + \varepsilon)} \right) - R_{1,2} \left( \frac{s_1 c \left( \frac{i}{n}, 1 \right) (1 + \varepsilon)}{m}, \frac{s_2 c \left( \frac{i}{n}, 2 \right) (1 + \varepsilon)}{m} \right) \to 0,
\]

uniformly in \(i\), \(s_1\) and \(s_2\). Hence, uniformly in \((t, s_1, s_2)\) \(\in [0, 1] \times [0, T]^2\),

\[
\frac{1}{k} \sum_{i=1}^{nt} P \left( \frac{k}{n} Y_{i,1} > \frac{m}{s_1 c \left( \frac{i}{n}, 1 \right) (1 + \varepsilon)}, \frac{k}{n} Y_{i,2} > \frac{m}{s_2 c \left( \frac{i}{n}, 2 \right) (1 + \varepsilon)} \right) \leq \frac{1}{k} \sum_{i=1}^{nt} P \left( \frac{k}{n} Y_{i,1} > \frac{m}{s_1 c \left( \frac{i}{n}, 1 \right) (1 + \varepsilon)}, \frac{k}{n} Y_{i,2} > \frac{m}{s_2 c \left( \frac{i}{n}, 2 \right) (1 + \varepsilon)} \right)
\]

\[
= \int_{1/n}^{(nt+1)/n} \frac{n}{k} P \left( \frac{k}{n} Y_{i,1} > \frac{m}{s_1 c \left( \frac{[nu]}{n}, 1 \right) (1 + \varepsilon)}, \frac{k}{n} Y_{i,2} > \frac{m}{s_2 c \left( \frac{[nu]}{n}, 2 \right) (1 + \varepsilon)} \right) du
\]

\[
\to \int_0^t R_{1,2} \left( \frac{s_1 c(u, 1)(1 + \varepsilon)}{m}, \frac{s_2 c(u, 2)(1 + \varepsilon)}{m} \right) du, \quad \text{as} \ n \to \infty.
\]

The lower bound follows similarly by replacing \(1 + \varepsilon\) with \(1 - \varepsilon\) and reversing the inequality. Now by the homogeneity of \(R\), (4.5) follows by letting \(\varepsilon \to 0\) and taking \(t = t_1 \wedge t_2\).

The following lemma is useful for the proof of Theorem 2.1 b).

**Lemma 4.1.** For all \(\delta > 0\), there exists \(0 < a < 1\) such that for large \(n\),

\[
P \left\{ U_0 \left( \frac{aN}{ks} \right) \leq X_{N-[ks]:N} \leq U_0 \left( \frac{N}{ak \varepsilon} \right) \text{ for } \frac{1}{2k} \leq s \leq 1 \right\} > 1 - \delta.
\]

**Proof.** Condition (1.1) implies, by inversion, that there exist \(M > 1\) and \(t_0 > 0\) such that for all \(t \geq t_0\),

\[
U_0 \left( \frac{t}{M} \right) \leq U_{i,j}(t) \leq U_0(M t)
\]

Note that from Inequality on page 419 of Shorack and Wellner (1986), we get that for every \(\delta > 0\) there exists \(0 < b < 1\) such that

\[
P \left\{ \frac{bn}{ks} \leq Y^{(j)}_{n-[ks]:n} \text{ for } \frac{1}{2k} \leq s \leq 1 \text{ and } Y^{(j)}_{n-[ks]:n} \leq \frac{n}{kbs} \right\} > 1 - \delta/2.
\]

Note also that

\[
Y^{(1)}_{n-[ks]:n} \leq Y^{(j)}_{n-[ks]:n} \leq \max_{1 \leq j \leq m} Y^{(j)}_{n-[ks]:n}
\]

Next we show, using (4.6), that with probability tending to 1,

\[
U_0 \left( \frac{1}{M} Y_{N-[ks]:N} \right) \leq X_{N-[ks]:N} \leq U_0 \left( M Y_{N-[ks]:N} \right).
\]
By definition \([ks] \) of the \(Y_{i,j} \) are not less than \(Y_{N-[ks];N} \). Hence at least \([ks] \) of the \(U_{i,j}(Y_{i,j})\) are not less than \(U_0(\frac{1}{M} Y_{N-[ks];N}) \). The right-hand inequality is similar. The result follows combining \((4.7), (4.8)\) and \((4.9)\).

**Proof of Theorem 2.1 b)** Tail empirical quantile function

We start from \((2.2)\) in Theorem 2.1a). By taking \(t_j = 1\) and aggregating over \(1 \leq j \leq m\), we get that for any \(0 \leq \eta < 1/2\), as \(n \to \infty\), almost surely,

\[
\sup_{x_0 \leq x < x_1} (1 + \gamma x)^{\eta/\gamma} \left| \frac{\sqrt{k}}{\gamma} \left( \mathbb{P}_n(x) - (1 + \gamma x)^{-1/\gamma} \right) - \sum_{j=1}^m W_j((1 + \gamma x)^{-1/\gamma}, C_j(1)) \right| \to 0,
\]

where

\[
\mathbb{P}_n(x) = \frac{1}{k} \sum_{i=1}^n \sum_{j=1}^m \mathbb{I}\{x_{i,j} - b_0(\frac{s}{\gamma}) > x\},
\]

and \(x_0 > -1/\gamma+\) and \(x_1 = 1/(-\gamma-)\). As a consequence, for \(\delta_1 > 0\),

\[
P\left\{ \sup_{x_0 \leq x < x_1} (1 + \gamma x)^{\eta/\gamma} \left| \frac{\sqrt{k}}{\gamma} \left( \mathbb{P}_n(x) - (1 + \gamma x)^{-1/\gamma} \right) - \sum_{j=1}^m W_j((1 + \gamma x)^{-1/\gamma}, C_j(1)) \right| > \delta_1 \right\} \to 0.
\]

We remark that the region \(x_0 \leq x < x_1\) has different implications for \(\gamma > 0\), \(\gamma < 0\), and \(\gamma = 0\). For \(\gamma > 0\), it implies that \(1 + \gamma x \geq 1 + \gamma x_0 > 0\), i.e. \(1 + \gamma x\) is bounded away from zero. For \(\gamma < 0\), \(1 + \gamma x > 1 + \gamma x_1 = 0\). Hence \(1 + \gamma x > 0\) but not necessarily bounded away from zero. On the other hand, \(1 + \gamma x \leq 1 + \gamma x_0\), i.e. \(1 + \gamma x\) is bounded away from \(\infty\). For \(\gamma = 0\), \(1 + \gamma x = 1\).

Then, split the range of \(s\) in two subintervals, \([1/(2k), s_0]\) and \([s_0, T]\), where \(s_0\) is a sufficiently small, but fixed, positive number.

For the range \([s_0, T]\) we use \((4.10)\) and Vervaat’s Lemma (cf. e.g. Appendix A of de Haan and Ferreira (2006)) with \(x_n(s) := \mathbb{P}_n(s)\) and \(x_n^-(s) := (X_{N-[ks];N} - b_0(N/k)) / a_0(N/k)\). We then obtain the statement in \((2.4)\), with the ‘sup’ taken over \([s_0, T]\).

For \(s \in [1/(2k), s_0]\), we first deal with the Gaussian processes term. Let \(W_0\) be a univariate standard Wiener process. It is well-known (and follows from the law of the iterated logarithm) that for every \(\delta > 0\) there exists an \(s_0\), such that \(P\{\sup_{0 < s \leq s_0} |W_0(s)|/s^{\eta} < \delta\} > 1 - \delta\). Now \(W_j(\cdot, C_j(1)) \overset{d}{=} \sqrt{C_j(1)} W_0\) for all \(j\). Hence for \(\delta > 0\), there exists an \(s_0(\delta)\), such that for all \(s_0 \leq s_0(\delta)\),

\[
P\left\{ \sup_{0 < s \leq s_0} s^{-\eta} \left| \sum_{j=1}^m W_j(s, C_j(1)) \right| < \delta \right\} > 1 - \delta.
\]

Hence, we shall concentrate on proving that with probability larger than \(1 - \delta\), with a proper choice of \(s_0\), for large \(n\),

\[
\sup_{\frac{1}{2k} \leq s \leq s_0} s^{\gamma + \frac{1}{2} + \varepsilon} \sqrt{k} \left( \frac{X_{N-[ks];N} - b_0(\frac{N}{k})}{a_0(\frac{N}{k})} - \frac{s^{-\gamma} - 1}{\gamma} \right) \leq \delta,
\]
and

\[
\inf_{\frac{1}{2} \leq s \leq s_0} s^{\gamma+\frac{1}{2}+\varepsilon} \sqrt{k} \left( \frac{X_{N-[ks]}-b_0 \left( \frac{N}{k} \right)}{a_0 \left( \frac{N}{k} \right)} - \frac{s^{-\gamma} - 1}{\gamma} \right) \geq -\delta.
\]

In the following we further split the range of \( s, [1/(2k), s_0] \), into two subintervals, \([1/(2k), t_n]\) and \([t_n, s_0]\), where \( t_n \) depends only on the constant \( a \) in Lemma 4.1 (eventually depending on \( \delta \)) and a sufficiently small \( \xi > 0 \), although the choice of \( t_n \) is different for proving the upper bound in (4.13) and the lower bound in (4.14):

\[
\sqrt{k t_n^{1/2+\varepsilon}} := \begin{cases} \delta \Delta_1^{-1} \text{ with } \Delta_1 := a^{-\gamma} - 1 + \xi > 0, \text{for the upper bound}, \\ \delta \Delta_2^{-1} \text{ with } \Delta_2 := a^{-\gamma} - 1 + \xi > 0, \text{for the lower bound}. \end{cases}
\]

Moreover the following technical relations provide the constants to determine an upper bound for \( s_0 \). Note that for large enough \( \eta \) there exists \( \eta' \) with \( 1 - \eta < 1 - \eta' < 1/2 + \varepsilon \), such that the following hold:

\[
\sup_{s > t_n} \frac{s^{-(1-\eta')}}{\sqrt{k}} \leq \frac{t_n^{-(1-\eta')}}{\sqrt{k}} = \frac{\eta'-1/2+\varepsilon}{\sqrt{k t_n^{1/2+\varepsilon}}} \rightarrow 0,
\]

\[
\left( 1 + \frac{\delta \gamma}{\sqrt{k}} s^{-(1-\eta')} \right)^{-1/\gamma} \leq 1 - c_I \frac{\delta}{\sqrt{k}} s^{-(1-\eta')} \text{ for some } 0 < c_I \leq 1 \text{ and large } n,
\]

\[
\left( 1 - \frac{\delta \gamma}{\sqrt{k}} s^{-(1-\eta')} \right)^{-1/\gamma} \geq 1 + c_{II} \frac{\delta}{\sqrt{k}} s^{-(1-\eta')} \text{ for some } 0 < c_{II} \leq 1 \text{ and large } n,
\]

where the last two inequalities follow from the inequalities \((1 + \gamma x)^{-1/\gamma} \leq 1 - c_I x\) and \((1 - \gamma x)^{-1/\gamma} \geq 1 + c_{II} x\) respectively, for \( 0 < x < \min(1, 1/(1-\gamma)) \) and some \( 0 < c_I, c_{II} \leq 1 \). Then, we should take

\[
s_0 \leq \min(s_0(\delta), (\delta c_I/(1 + \delta_1))^{(\eta-\eta')^{-1}}, \delta c_{II}/((m/a)^{\gamma}(1 + \xi)^{-\eta'/\gamma}(1 + \delta_1))^{(\eta-\eta')^{-1}}).
\]

(A) Upper bound and \( s \in [(2k)^{-1}, t_n] \): Assume first that \( A_0 \) is eventually positive. Corollary 2.3.7 in de Haan and Ferreira (2006) and Lemma 4.1 imply: for all \( \varepsilon, \delta, \theta > 0 \), there exists \( 0 < a < 1 \) such that for large \( n \), with probability at least \( 1 - \delta \),

\[
\frac{X_{N-[ks]}-b_0 \left( \frac{N}{k} \right)}{a_0 \left( \frac{N}{k} \right)} - \frac{s^{-\gamma} - 1}{\gamma} \leq (as)^{-\gamma} - s^{-\gamma} + \overline{\psi}_{\gamma,\rho} \left( \frac{1}{as} \right) A_0 \left( \frac{N}{k} \right) + (as)^{-\gamma-\rho-\theta} A_0 \left( \frac{N}{k} \right)
\]

\[
\leq s^{-\gamma} \left\{ \frac{a^{-\gamma} - 1}{\gamma} + s^{-\rho-\theta} K A_0 \left( \frac{N}{k} \right) \right\},
\]

for some \( K > 0 \). Hence,

\[
s^{\gamma+1/2+\varepsilon} \sqrt{k} \left( \frac{X_{N-[ks]}-U_0 \left( \frac{N}{k} \right)}{a_0 \left( \frac{N}{k} \right)} - \frac{s^{-\gamma} - 1}{\gamma} \right) \leq s^{1/2+\varepsilon} \sqrt{k} \left\{ \frac{a^{-\gamma} - 1}{\gamma} + s^{-\rho-\theta} K A_0 \left( \frac{N}{k} \right) \right\}
\]

\[
\leq \frac{\delta}{\Delta_1} \left\{ \frac{a^{-\gamma} - 1}{\gamma} + s^{-\rho-\theta} K A_0 \left( \frac{N}{k} \right) \right\} \leq \delta,
\]
for \( n \) large, uniformly in \( s \in [(2k)^{-1}, t_n] \), since \( \sup_{s \in [(2k)^{-1}, t_n]} s^{-\rho - \theta} A_0 \left( \frac{N}{k} \right) \to 0 \) choosing \( \theta < -\rho \).

(B) Upper bound and \( s \in (t_n, s_0) \): We prove that, with probability at least \( 1 - \delta \), for large \( n \),

\[
\mathbb{P}_n \left( \frac{s^{-\gamma} - 1}{\gamma} + \frac{\delta}{\sqrt{k}} s^{-\gamma - 1/2 - \varepsilon} \right) \leq s, \quad \text{for all} \ s \in (t_n, s_0),
\]

which implies the upper bound in (4.13).

We intend to apply (4.11) with \( x \) replaced by \( \gamma^{-1}(s^{-\gamma} - 1) + \delta k^{-1/2} s^{-\gamma - 1/2 - \varepsilon} \). For this, note that,

\[
1 + \gamma \left( \frac{s^{-\gamma} - 1}{\gamma} + \frac{\delta}{\sqrt{k}} s^{-\gamma - 1/2 - \varepsilon} \right) = s^{-\gamma} \left( 1 + \frac{\delta \gamma}{\sqrt{k}} s^{-1/2 - \varepsilon} \right)
\]

and, for \( s \in (t_n, s_0) \) and \( \gamma > 0 \) the right-hand side is at least \( s_0^{-\gamma} \left( 1 + \delta \gamma k^{-1/2} s_0^{-1/2 - \varepsilon} \right) \), consequently bounded away from zero. For \( \gamma < 0 \) the inequality is reversed and the expression is bounded from above. Hence,

\[
\mathbb{P}_n \left( \frac{s^{-\gamma} - 1}{\gamma} + \frac{\delta}{\sqrt{k}} s^{-\gamma - 1/2 - \varepsilon} \right) \leq s \left( 1 + \frac{\delta \gamma}{\sqrt{k}} s^{-1/2 - \varepsilon} \right)^{-1/\gamma}
\]

\[
+ \frac{1}{\sqrt{k}} \sup_{s \in (t_n, s_0)} \left( s \left( 1 + \frac{\delta \gamma}{\sqrt{k}} s^{-1/2 - \varepsilon} \right)^{-1/\gamma} \right) + \frac{\delta_1}{\sqrt{k}} s^\eta \left( 1 + \frac{\delta g}{\sqrt{k}} s^{-1/2 - \varepsilon} \right)^{-\eta/\gamma}
\]

\[
\leq s \left( 1 + c_1 \frac{\delta}{\sqrt{k}} s^{(1 - \eta)} \right) + \frac{\delta_1}{\sqrt{k}} s^\eta = s - \frac{s^\eta}{\sqrt{k} c_1 \delta s^{\eta - \eta} - (1 + \delta_1)}
\]

with \( \sup_{s \in (t_n, s_0)} \left( s \left( 1 + \frac{\delta \gamma}{\sqrt{k}} s^{-1/2 - \varepsilon} \right)^{-1/\gamma} \right) \leq 1 \).

It remains to check that \( c_1 \delta s^{\eta - \eta} - (1 + \delta_1) \geq 0 \) which holds by the choice of \( s_0 \).

(C) Lower bound and \( s \in [(2k)^{-1}, t_n] \): As in (A) assume first that \( A_0 \) is eventually positive. Corollary 2.3.7 in de Haan and Ferreira (2006) and Lemma 4.1 imply: for all \( \varepsilon, \delta, \theta > 0 \), there exists \( 0 < a < 1 \) such that for large \( n \), with probability at least \( 1 - \delta \),

\[
\frac{X_{N - [ks] \colon N} - b_0 \left( \frac{N}{k} \right)}{a_0 \left( \frac{N}{k} \right) \gamma} \geq \frac{a}{s} \frac{X_{N - [ks] \colon N} - b_0 \left( \frac{N}{k} \right)}{a_0 \left( \frac{N}{k} \right) \gamma} + \frac{\varepsilon}{\gamma} A_0 \left( \frac{N}{k} \right) - \frac{a}{s} \frac{A_0 \left( \frac{N}{k} \right) \gamma + (\gamma + \theta)}{A_0 \left( \frac{N}{k} \right)}
\]

\[
\geq s^{-\gamma} \left\{ \frac{a \gamma - 1}{\gamma} - s^{-\rho - \theta} K A_0 \left( \frac{N}{k} \right) \right\},
\]

for some \( K > 0 \), hence,

\[
s^{\gamma + 1/2 + \varepsilon} \sqrt{k} \left( \frac{X_{N - [ks] \colon N} - b_0 \left( \frac{N}{k} \right)}{a_0 \left( \frac{N}{k} \right) \gamma} - \frac{s^{-\gamma} - 1}{\gamma} \right) \geq s^{1/2 + \varepsilon} \sqrt{k} \left\{ \frac{a \gamma - 1}{\gamma} - s^{-\rho - \theta} K A_0 \left( \frac{N}{k} \right) \right\}
\]

\[
\geq \frac{\delta}{\Delta_2} \left\{ \frac{a \gamma - 1}{\gamma} - s^{-\rho - \theta} K A_0 \left( \frac{N}{k} \right) \right\} \geq -\delta,
\]

for large \( n \), uniformly in \( s \in [(2k)^{-1}, t_n] \), since \( \sup_{s \in [(2k)^{-1}, t_n]} s^{-\rho - \theta} A_0 \left( \frac{N}{k} \right) \to 0 \) choosing \( \theta < -\rho \).
(D) Lower bound and \(s \in (t_n, s_0]\): We prove that, with probability at least \(1 - \delta\), for large \(n\),
\[
\mathbb{P}_n \left( \frac{s^{-\gamma} - 1}{\gamma} - \frac{\delta}{\sqrt{k}}s^{-\gamma-1/2-\varepsilon} \right) \geq s + \frac{1}{k}, \quad \text{for all } s \in (t_n, s_0],
\]
which implies the lower bound in (4.13). Similarly as in (B) apply (4.11) with \(x\) replaced by \(\gamma^{-1}(s^{-\gamma} - 1) - \delta k^{-1/2}s^{-\gamma-1/2-\varepsilon}\). Note that,
\[
1 + \gamma \left( \frac{s^{-\gamma} - 1}{\gamma} - \frac{\delta}{\sqrt{k}}s^{-\gamma-1/2-\varepsilon} \right) = s^{-\gamma} \left( 1 - \frac{\delta \gamma}{\sqrt{k}}s^{-1/2-\varepsilon} \right)
\]
and, for \(s \in (t_n, s_0]\) and \(\gamma > 0\) the right-hand side is at least \(s^{-\gamma}_0(1 - \gamma \Delta_2) > 0\) and consequently bounded away from zero. For \(\gamma < 0\) the inequality is reversed and the expression is bounded above. Hence,
\[
\mathbb{P}_n \left( \frac{s^{-\gamma} - 1}{\gamma} - \frac{\delta}{\sqrt{k}}s^{-\gamma-1/2-\varepsilon} \right) \geq s \left( 1 - \frac{\delta \gamma}{\sqrt{k}}s^{-1/2-\varepsilon} \right)^{-1/\gamma}
\]
\[
+ \frac{1}{\sqrt{k}} \frac{\tilde{W}}{s} \left( s \left( 1 - \frac{\delta \gamma}{\sqrt{k}}s^{-1/2-\varepsilon} \right)^{-1/\gamma} \right) - \frac{\delta \gamma}{\sqrt{k}}s^{\eta} \left( 1 - \frac{\delta \gamma}{\sqrt{k}}s^{-1/2-\varepsilon} \right)^{-\eta/\gamma}
\]
\[
\geq s \left( 1 + c_{II} \frac{\delta}{\sqrt{k}}s^{-(1-\eta')} \right) - a^{-\eta}(1 + \xi)^{-\eta/\gamma} (1 + \delta_1) s^{\eta}/\sqrt{k}
\]
\[
= s + s^{\eta} \left( c_{II} \delta s^{\eta'} - a^{-\eta}(1 + \xi)^{-\eta/\gamma}(1 + \delta_1) \right)
\]
where we have used in particular (4.15). It remains to check that the right-hand side of (4.16) is at least \(s + 1/k\) which is equivalent to \(s^{\eta/\sqrt{k}} \left( c_{II} \delta s^{\eta'} - a^{-\eta}(1 + \xi)^{-\eta/\gamma}(1 + \delta_1) \right) \geq 1\). This holds by the choice of \(\eta' < \eta\) and choosing \(s_0 \leq \delta c_{II}/((m/a)^{\eta}(1 + \xi)^{-\eta/\gamma}(1 + \delta_1))^{(\eta-\eta')^{-1}}\).

Finally, if \(A_0\) is eventually negative, the proofs are the same, except that the signs of the remaining terms in (A) and in (C), \((as)^{-\gamma-\rho-\theta} A_0 \left( \frac{N}{k} \right)\) and \(-a/s)^{\gamma+\rho+\theta} A_0 \left( \frac{N}{k} \right),\) should be interchanged. \(\square\)

**Proof of Theorem 2.3**

Fix \(j \in \{1, \ldots, m\}\). Replace \(x\) in (2.2) with \(\left( X_{N-k:N} - b_0(\frac{N}{k}) \right)/a_0(\frac{N}{k})\) and use condition (iii) jointly with Theorem 2.3.8 of de Haan and Ferreira (2006) to get
\[
\sqrt{k} \left\{ \frac{1}{k} \sum_{i=1}^{nt} \mathbb{1}_{\{x_{i,j} > X_{N-k:N}\}} - \left( 1 + \gamma \frac{X_{N-k:N} - b_0(\frac{N}{k})}{a_0(\frac{N}{k})} \right)^{-1/\gamma} C_j(t) \right\} - W_j \left( 1 + \gamma \frac{X_{N-k:N} - b_0(\frac{N}{k})}{a_0(\frac{N}{k})} \right)^{-1/\gamma} C_j(t) \right\} = o_p \left( 1 + \gamma \frac{X_{N-k:N} - b_0(\frac{N}{k})}{a_0(\frac{N}{k})} \right)^{-\eta/\gamma}.
\]
Now by (2.4)
\[
\sqrt{k} \frac{X_{N-k:N} - b_0(\frac{N}{k})}{a_0(\frac{N}{k})} \xrightarrow{p} \sum_{j=1}^{m} W_j(1, C_j(1))
\]
and hence
\[
\sqrt{k} \left\{ 1 - \left( 1 + \gamma \frac{X_{N-k:N} - b_0(\frac{N}{k})}{a_0(\frac{N}{k})} \right)^{-1/\gamma} \right\} - \sqrt{k} \frac{X_{N-k:N} - b_0(\frac{N}{k})}{a_0(\frac{N}{k})} \xrightarrow{p} 0.
\]
Combining this we obtain
\[
\sup_{0 \leq t \leq 1} \left| \frac{1}{k} \sum_{i=1}^{nt} \mathbb{1}_r \left\{ X_{i,j} > X_{N-k,N} \right\} - C_j(t) \right| - \{ W_j(1, C_j(t)) - C_j(t) \sum_{r=1}^{m} W_r(1, C_r(1)) \} \xrightarrow{P} 0,
\]
which yields (2.6).

Next we prove for fixed \((t, w_1, w_2) \in [0,1] \times [0,T]^2\), as \(n \to \infty\),
\begin{equation}
\tag{4.17}
\frac{1}{k} \sum_{i=1}^{nt} \mathbb{1}_r \left\{ \frac{X_{i,j} - b_0 (\frac{N}{k})}{a_0 (\frac{N}{k})} > \frac{w_1^{-1}}{\gamma} \right\} \xrightarrow{P} \frac{1}{m} \int_0^t R_{j_1,j_2}(w_1 c(u,j_1), w_2 c(u,j_2)) du.
\end{equation}

We check the variance of the left hand side: using (4.5), as \(n \to \infty\),
\[
\begin{align*}
&\text{Var} \left( \frac{1}{k} \sum_{i=1}^{nt} \mathbb{1}_r \left\{ \frac{X_{i,j} - b_0 (\frac{N}{k})}{a_0 (\frac{N}{k})} > \frac{w_1^{-1}}{\gamma} \right\} \right) \\
&= \frac{1}{k^2} \sum_{i=1}^{nt} \text{Var} \left( \mathbb{1}_r \left\{ \frac{X_{i,j} - b_0 (\frac{N}{k})}{a_0 (\frac{N}{k})} > \frac{w_1^{-1}}{\gamma} \right\} \right) \\
&\leq \frac{1}{k^2} \sum_{i=1}^{nt} P \left( \frac{X_{i,j} - b_0 (\frac{N}{k})}{a_0 (\frac{N}{k})} > \frac{w_1^{-1}}{\gamma}, \frac{X_{i,j} - b_0 (\frac{N}{k})}{a_0 (\frac{N}{k})} > \frac{w_2^{-1}}{\gamma} \right) \\
&\leq \frac{1}{km} (1 + \varepsilon) \int_0^t R_{j_1,j_2}(w_1 c(u,j_1), w_2 c(u,j_2)) du \to 0.
\end{align*}
\]

Hence, by Chebyshev inequality, as \(n \to \infty\),
\[
\frac{1}{k} \sum_{i=1}^{nt} \mathbb{1}_r \left\{ \frac{X_{i,j} - b_0 (\frac{N}{k})}{a_0 (\frac{N}{k})} > \frac{w_1^{-1}}{\gamma} \right\} \xrightarrow{P} 0.
\]

Then, (4.5) implies (4.17).

By the continuity of the right hand side of (4.17) and the monotonicity of both sides of (4.17) in \(t, w_1\) and \(w_2\), this result holds uniformly for \((t, w_1, w_2) \in [0,1] \times [0,T]^2\).

Finally, for fixed \((s_1, s_2) \in [0, T]^2\), by (2.4), as \(n \to \infty\)
\[
\left( 1 + \gamma \frac{X_{N-[k_s]}:N - b_0 (\frac{N}{k})}{a_0 (\frac{N}{k})} \right)^{-1/\gamma} \xrightarrow{P} s_j,
\]
for \(j = 1, 2\). We can then replace \(w_j\) in (4.17) with \(\left( 1 + \gamma \frac{X_{N-[k_s]}:N - b_0 (\frac{N}{k})}{a_0 (\frac{N}{k})} \right)^{-1/\gamma}\), which yields (2.7) for fixed \((t, s_1, s_2) \in [0,1] \times [0,T]^2\). The uniformity follows as before.

**Proof of Corollary 2.5** Under the null hypothesis, using Theorem 2.3, under a Skorokhod construction,
\[
\sqrt{k} (\hat{C}_j(t) - t \hat{C}_j(1)) = \sqrt{k} (C_j(t) - C_j(t)) - t \sqrt{k} (\hat{C}_j(1) - C_j(1))
\]
converges uniformly \((0 \leq t \leq 1)\) to

\[
W_j(1, C_j(t)) - C_j(t) \sum_{r=1}^{m} W_r(1, C_j(1)) - t \left[ W_j(1, C_j(1)) - C_j(1) \sum_{r=1}^{m} W_r(1, C_r(1)) \right]
\]

\[
= W_j(1, tC_j(1)) - tW_j(1, C_j(1)) \xrightarrow{d} \sqrt{C_j(1)} B(t).
\]

Then the result follows directly via Slutsky’s theorem.

The proof of Theorem 2.6 is deferred to the Supplementary Material. Here we only present the main steps of the proof.

We use “local asymptotic normal theory”, where the local log-likelihood and local score processes are fundamental, consisting of reparametrizations of the former with local parameter \(h = (h_1, h_2) \in \mathbb{R}^2\):

\[
\begin{aligned}
&\begin{cases}
  h_1 = \sqrt{k} (\gamma_N/k - \gamma_0) \\
  h_2 = \sqrt{k} (\sigma_N/k/a_0(N/k) - 1)
\end{cases}
\iff
&\begin{cases}
  \gamma = \gamma_N/k = \gamma_0 + h_1/\sqrt{k} \\
  \sigma = \sigma_N/k = a_0\left(\frac{N}{k}\right)(1 + h_2/\sqrt{k}),
\end{cases}
\end{aligned}
\]

\[
\tilde{L}_{N,k}(h) = k \int_{0}^{1} \ell \left( \gamma_0 + \frac{h_1}{\sqrt{k}} a_0\left(\frac{N}{k}\right)(1 + \frac{h_2}{\sqrt{k}}), X_{N-[k]N} - X_{N-k,N} \right) ds
\]

\[
= k \int_{0}^{1} \ell \left( \gamma_0 + \frac{h_1}{\sqrt{k}}, 1 + \frac{h_2}{\sqrt{k}} \frac{X_{N-[k]N} - X_{N-k,N}}{a_0\left(\frac{N}{k}\right)} \right) ds - k \log a_0\left(\frac{N}{k}\right),
\]

and

\[
\begin{aligned}
&\frac{\partial \tilde{L}_{N,k}}{\partial h_1} (h_1, h_2) = \frac{1}{\sqrt{k}} \frac{\partial L_{N,k}}{\partial \gamma} (\gamma, \sigma_N/k) = 0 \\
&\frac{\partial \tilde{L}_{N,k}}{\partial h_2} (h_1, h_2) = \frac{1}{\sqrt{k}} \frac{\partial L_{N,k}}{\partial \sigma} (\gamma, \sigma_N/k) = 0.
\end{aligned}
\]

The main steps of the proof are as follows:

a) Denote \(\theta = (\gamma, \sigma) \in \mathbb{R} \times (0, \infty)\). First prove that

\[
\frac{\partial^2 \tilde{L}_{N,k}(h)}{\partial h \partial h^T} = -I_{\gamma_0} + o_P(1), \quad I_{\gamma_0} = -\int_{0}^{1} \frac{\partial^2 \ell}{\partial \theta \partial \theta^T} \left( \gamma_0, 1, \frac{s^{-\gamma_0} - 1}{\gamma_0} \right) ds,
\]

uniformly in a large enough ball \(H_n\) to ensure that it covers the true solution; \(I_{\gamma_0}\) is the Fisher information matrix related to the approximate \(GP_{\gamma_0, 1}\) model

\[
I_{\gamma_0} = \begin{pmatrix}
\frac{2}{1+3\gamma_0+2\gamma_0^2} & \frac{1}{1+3\gamma_0+2\gamma_0^2} \\
\frac{1}{1+3\gamma_0+2\gamma_0^2} & \frac{1}{1+2\gamma_0}
\end{pmatrix}.
\]

\(I_{\gamma_0}\) is positive definite, which implies that the local log-likelihood process is eventually strictly concave on \(H_n\) with probability tending to 1.

b) Then, by integration one obtains an expansion for the local log-likelihood process (holding uniformly for \(h\) in compact sets):

\[
\tilde{L}_{N,k}(h) = \tilde{L}_{N,k}(0) + h^T \frac{\partial \tilde{L}_{N,k}}{\partial h} (0) - \frac{1}{2} h^T I_{\gamma_0} h + o_P(1)
\]

where

\[
\frac{\partial \tilde{L}_{N,k}}{\partial h} (0) = \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \frac{\partial \ell}{\partial \theta} \left( \gamma_0, 1, \frac{X_{N-i+1,N} - X_{N-k,N}}{a_0\left(\frac{N}{k}\right)} \right) \xrightarrow{d} N(0, \Sigma_{\gamma_0}).
\]
c) Finally the Argmax Theorem (van der Vaart, 2000, Corollary 5.58) provides the result: let
\[
M_n(h) = \tilde{L}_{N,k}(h) - \tilde{L}_{N,k}(0),
\]
\[
M(h) = h^T N(0, \Sigma_{\gamma_0}) - \frac{1}{2} h^T I_{\gamma_0} h, \quad h \in \mathbb{R}^2.
\]
Then,
\[
\hat{h}_n = \arg\max_{h \in H_n} M_n(h) \xrightarrow{d} h = \arg\max_{h \in \mathbb{R}^2} M(h) = \frac{I_{\gamma_0}^{-1} N(0, \Sigma_{\gamma_0})}{I_{\gamma_0}^{-1}}
\]
provided \( \hat{h}_n \) is tight which holds as in Dombry and Ferreira (2019). Finally, note that
\[
\tilde{L}_{N,k}(\hat{h}_n) = L_{N,k} \left( \gamma_0 + \frac{\hat{h}_{n,1}}{\sqrt{k}}, a_0 \left( \frac{N}{k} \right) (1 + \frac{\hat{h}_{n,2}}{\sqrt{k}}) \right)
\]
and similarly for its derivatives.

Acknowledgements. We would like to thank the Editor, the Associate Editor and four anonymous referees for their detailed reviews, which helped to improve the paper substantially.

John Einmahl holds the Arie Kapteyn Chair 2019-2022 and gratefully acknowledges the corresponding research support. Ana Ferreira was partially supported by FCT-Portugal: UIDB/04621/2020, UIDB/00006/2020 and SFRH/BSAB/142912/2018; IST: P.5088. Laurens de Haan was financially supported by FCT - Fundação para a Ciência e a Tecnologia, Portugal, through the projects UIDB/00006/2020 and PTDC/MAT-STA/28649/2017. Cláudia Neves gratefully acknowledges support from EPSRC-UKRI Innovation Fellowship grant EP/S001263/1 and project FCT-UIDB/00006/2020.

SUPPLEMENTARY MATERIAL

Supplementary Material for “Spatial dependence and space-time trend in extreme events” The supplementary material consists of three sections. Section 1 provides the detailed proof of Theorem 2.6. Section 2 shows a simulation study regarding the proposed estimation and testing procedures. Section 3 validates the assumptions for the data application.

(doi: TO BE COMPLETED BY THE TYPESETTER).

REFERENCES


