Biometrika (2003), **90,** 1, *pp.* 139–156

Printed in Great Britain

A dependence measure for multivariate and spatial extreme values: Properties and inference

By MARTIN SCHLATHER

Soil Physics Group, University of Bayreuth, 95440 Bayreuth, Germany

martin.schlather@uni-bayreuth.de

AND JONATHAN A. TAWN

Department of Mathematics and Statistics, Lancaster University, Lancaster LA1 4YF, U.K.

j.tawn@lancaster.ac.uk

Summary

We present properties of a dependence measure that arises in the study of extreme values in multivariate and spatial problems. For multivariate problems the dependence measure characterises dependence at the bivariate level, for all pairs and all higher orders up to and including the dimension of the variable. Necessary and sufficient conditions are given for subsets of dependence measures to be self-consistent, that is to guarantee the existence of a distribution with such a subset of values for the dependence measure. For pairwise dependence, these conditions are given in terms of positive semidefinite matrices and non-differentiable, positive definite functions. We construct new nonparametric estimators for the dependence measure which, unlike all naive nonparametric estimators, impose these self-consistency properties. As the new estimators provide an improvement on the naive methods, both in terms of the inferential and interpretability properties, their use in exploratory extreme value analyses should aid the identification of appropriate dependence models. The methods are illustrated through an analysis of simulated multivariate data, which shows that a lack of self-consistency is frequently a problem with the existing estimators, and by a spatial analysis of daily rainfall extremes in south-west England, which finds a smooth decay in extremal dependence with distance.

Some key words: Dependence measure; Extreme value; Max-stable process; Multivariate extreme-value distribution

1. Introduction

Techniques for modelling and estimating the dependence structure of multivariate and spatial processes are well established when the variables or processes are Gaussian. For example, for Gaussian vector variables the dependence is characterised by a positive semidefinite covariance matrix. Similarly, for stationary d-dimensional spatial processes dependence is described by a correlation function, $\rho(h)$, of pairwise spatial separation $h \in \mathbb{R}^d$, which must be a positive definite function (Chilès & Delfiner, 1999, p. 60). Beyond the Gaussian family very broad but complex forms of dependence structure are possible, which makes model building difficult in problems of three or more variables. Consequently, in multivariate and spatial problems attention has often focused on obtaining dependence measures that capture the main characteristics of the dependence structure; see for example Cox & Wermuth (1996, Ch. 2), Joe (1997), Hutchinson & Lai (1991, Ch. 5) and Ekholm et al. (1995, 2000).

We focus on a particularly rich family of multivariate distributions and spatial processes that arise in the study of extreme values. We investigate the properties of a natural dependence measure for this family. First consider the dependence measure in the multivariate case, the spatial setting being considered in § 3. Let $X=(X_1, \ldots, X_m)$ be an m-dimensional vector variable. To simplify the presentation, throughout the paper we take all the univariate marginal distributions of X to be unit Fréchet, so that $\mathrm{pr}(X_i<z)=\exp(-1/z)$, for $z>0$ and $i=1, \ldots, m$. There is no loss of generality in this assumption as our interest is only in the dependence structure of the variables, that is the copula, and in practice the variables can be marginally standardised before dependence is considered. In § 4 we discuss this further and § 5 contains such an example.

To study the extremal dependence of the variable X it is standard to consider the joint distribution of the componentwise maximum of n independent replicates of the variable (de Haan & Resnick, 1977; Tawn, 1990; Coles & Tawn, 1994). Under weak conditions, see Resnick (1987, Ch. 5) for details and Schlather (2001a) for a family of counterexamples in a similar context, for all $A \subseteq M_m=\{1, \ldots, m\}$ there exists a real number θ_A, with $1 \leqslant \theta_A \leqslant |A|$, such that the normalised maximum of all the variables indexed by the set A converges to a Fréchet-distributed variable with parameter θ_A; that is

$$\lim_{n\to\infty} \mathrm{pr}\left(\max_{i\in A} \max_{j=1,\ldots,n} X_i^{(j)}/n<z\right)=\lim_{n\to\infty}\left\{\mathrm{pr}\left(\max_{j=1,\ldots,n} X_1^{(j)}/n<z\right)\right\}^{\theta_A}=\exp(-\theta_A/z), \tag{1}$$

for all $z>0$, where $X_i^{(j)}$ is the jth replicate of variable X_i. Here θ_A measures the extremal dependence between the variables indexed by set A and, following the unpublished University of Surrey technical report on 'Max-stable processes and spatial extremes' by R. L. Smith, see Coles & Tawn (1996), we term θ_A the extremal coefficient of these variables. The simple interpretation of θ_A as the effective number of independent variables in the set A from which the maximum is drawn has led to its use as a dependence measure in a range of practical applications (Buishand, 1984; Tawn, 1988, 1990).

A collection of extremal coefficients that characterise dependence at bivariate level and all higher orders up to and including all m variables is determined by the distribution of X. To be specific, the set of all statistically relevant extremal coefficients for X is $\{\theta_A: A\in C_m\}$, where $C_m=2^{M_m}\backslash\{\varnothing\}$ is the ensemble of all nonempty subsets of M_m and $\theta_{\{i\}}=1$, for all $i=1, \ldots, m$. We simplify the notation using $\theta_i=\theta_{\{i\}}$, $\theta_{ij}=\theta_{\{i,j\}}$, $\theta_i=\theta_{ii}$, and so on, where no confusion can arise.

To illustrate the type of property that the collection $\{\theta_A: A\in C_m\}$ must satisfy, and the connections between the extremal coefficients and similar known Gaussian conditions, first consider pairwise dependence measures and examine the conditions these must satisfy for Gaussian and extreme-value dependence problems. If $Z=(Z_1, \ldots, Z_m)$ follows a multivariate normal distribution, with ρ_{ij} being the (i,j)th element of the correlation matrix Σ, then a necessary and sufficient condition for a set of ρ_{ij} values to be self-consistent is that all the eigenvalues of Σ be nonnegative. This condition can easily be tested for numerically, but is analytically simple only in the trivariate case, where the conditions simplify to $-1\leqslant\rho_{12}\leqslant 1$, $-1\leqslant\rho_{13}\leqslant 1$, and

$$\max\{b_L, -1\}\leqslant\rho_{23}\leqslant\min\{b_U, 1\},$$

where

$$b_L = \rho_{12}\rho_{13} - (1-\rho_{12}^2)^{\frac{1}{2}}(1-\rho_{13}^2)^{\frac{1}{2}}, \quad b_U = \rho_{12}\rho_{13} + (1-\rho_{12}^2)^{\frac{1}{2}}(1-\rho_{13}^2)^{\frac{1}{2}}.$$

The same issues arise in different ways for the extremal coefficients. Although a general tight solution to this problem in the extreme-value case is not tractable, some partial conditions may be given (Schlather & Tawn, 2002). In the trivariate case the pairwise extremal coefficients are self-consistent provided $1 \leqslant \theta_{12} \leqslant 2$, $1 \leqslant \theta_{13} \leqslant 2$ and

$$1 + |\theta_{12} - \theta_{13}| \leqslant \theta_{23} \leqslant \min\{2, \theta_{12} + \theta_{13} - 1\}. \tag{2}$$

For non-Gaussian dependence structures, higher-order dependence generally is not determined by pairwise dependence. Here, we will explore the associations between dependence of different orders for extreme values. This topic has been considered generally; for example, the Fréchet bounds impose structure on joint distributions given the lower-order structure of the distribution; see Joe (1997) for recent developments. In the extreme-value context only Tiago de Oliveira (1962/63) has studied this feature, showing that pairwise independence for all pairs implies independence for all higher-order dependencies.

Benefits from obtaining this detailed understanding of dependence are as follows: it enables the incorporation of a self-consistency property into the estimation of the set of extremal coefficients so that the existence of a distribution corresponding to the estimated set of extremal coefficients is guaranteed; the contradictions in interpretation are avoided that arise when estimators of dependence are not self-consistent; and they provide insight for subsequent model building. We shall illustrate the use of the former in multivariate problems in proposing a joint estimation scheme for a set of extremal coefficients which improves on independent estimation of each, and the use of the latter in the context of building a spatial model for extremes. Furthermore, ensuring self-consistency of the estimates for the extremal coefficients is desirable when obtaining inferences on the extremal behaviour of $\max_{i \in A} X_i$ if various subsets A of M_m are considered simultaneously.

In § 2 we review properties for θ_A in the multivariate case; we illustrate these with theoretical examples, discuss numerical issues and make comparisons with the correlation coefficients of Gaussian distributions. In § 3 the concept of the extremal coefficient is extended to the extremal coefficient function, a pairwise measure of spatial dependence in extremes, and properties and examples are given. In § 4 an inference scheme for the extremal coefficient is developed and this is illustrated through a simulation study. Finally, in § 5 the methods are applied to daily rainfall extremes from a network of sites in the south-west of England for which we have a small number of sites of continuous time series of daily aggregated rainfalls, and many sites with only annual maximum daily rainfalls.

2. Multivariate case

2·1. *Multivariate extreme-value distributions*

The defining representation (1) of θ_A arises as a special property of the complete class of limiting distributions for componentwise maxima of independent and identically distributed variables $X^{(1)}, \ldots, X^{(n)}$ with the same distribution as X. Provided that weak conditions in Resnick (1987, Ch. 5) hold, then

$$\lim_{n\to\infty} \mathrm{pr}\left(\max_{j=1,\ldots,n} X_i^{(j)}/n < z_i;\ i = 1, \ldots, m\right) = G_m(z_1, \ldots, z_m),$$

where the limit G_m is a member of the class of multivariate extreme-value distributions with Fréchet margins. The complete class is characterised by the function V, where

$$V(z_1, \ldots, z_m) = -\log G_m(z_1, \ldots, z_m) = \int_{S_m} \max_{i \in M_m} \left(\frac{w_i}{z_i}\right) dH_m(w_1, \ldots, w_m). \tag{3}$$

Here S_m is the $(m-1)$-dimensional unit simplex and H_m is a measure which has all its marginal expectations equal to 1; compare Pickands (1981). Combining representations (1) and (3) gives

$$\theta_A = \int_{S_m} \max_{i \in A} w_i \, dH_m(w_1, \ldots, w_m) \quad (A \in C_m). \tag{4}$$

Thus it is clear that the ensemble of the θ_A, for $A \in C_m$, satisfies a range of properties including strong relationships between the values of θ_A, for different A, that are imposed through the integral representation (4) by the range of possible structures for H_m. Furthermore, if X follows a multivariate extreme-value distribution then, for all $A \in C_m$, θ_A is an exact dependence measure and does not rely on asymptotic arguments.

2·2. *Properties of extremal coefficients*

The theoretical properties of the collection of extremal coefficients to be used as dependence measures need to be established so that estimates of the dependence measure can be constructed so as to possess these properties. In this section we present a summary of relevant marginal and joint properties satisfied by a collection of extremal coefficients.

Whatever continuous marginal distributions $X = (X_1, \ldots, X_m)$ has, the first equality in equation (1) holds provided the normalisation by n is replaced by an appropriate location and scale form such that the limit is nondegenerate. It follows that the extremal coefficient is invariant to strictly increasing transformations of the marginal distributions. This is an important feature of a dependence measure, despite not being satisfied by the correlation coefficient.

From equation (1), it is seen that θ_A is invariant to the labelling of the variables in A. Also, from the convexity of representation (4), we have that $1 \leqslant \theta_A \leqslant |A|$ for any $A \in C_m$ with these boundary values corresponding to the variables in A being completely dependent or independent respectively.

Let $G_m^{(1)}$ and $G_m^{(2)}$ be two extreme-value distributions with unit Fréchet margins that have the same set of extremal coefficients. If $Y^{(i)}$ follows $G_m^{(i)}$ for $i = 1, 2$ then

$$\text{pr}\{Y_i^{(1)} < z;\ i \in A\} = \text{pr}\{Y_i^{(2)} < z;\ i \in A\}$$

for any $A \in C_m$ and $z \in \mathbb{R}$. However, the set of extremal coefficients does not characterise G_m completely. To illustrate this fact we present a range of distributions for which there exist parameter values so that $\theta_{12} = \theta$ for any $\theta \in [1, 2]$.

Example 1. The distribution of the asymmetric logistic (Tawn, 1990) is given by

$$V(z_1, \ldots, z_m) = \sum_{c \in C_m} \left\{ \sum_{i \in c} (\xi_{c,i}/z_i)^{1/r_c} \right\}^{1/r_c}, \tag{5}$$

where $r_c \geqslant 1$ for all $c \in C_m$. For $i = 1, \ldots, m$ we have $\xi_{c,i} = 0$ if $i \notin c$, $\xi_{c,i} \geqslant 0$ if $i \in c$, and $\sum_{c \in C_m} \xi_{c,i} = 1$. Then, for $A \in C_m$,

$$\theta_A = \sum_{c \in C_m} \left(\sum_{i \in c \cap A} \xi_{c,i}^{r_c} \right)^{1/r_c}.$$

This model is dimension-specific, in that θ_A is determined by parameters that describe a higher order of structure than among the variables indexed by A. For $m=2$, $\xi_{1,1}=\xi_{2,2}=0$ and $\xi_{12,1}=\xi_{12,2}=1$ we obtain the logistic dependence model

$$V(y,z)=(y^{-1/\alpha}+z^{-1/\alpha})^{\alpha} \quad (0<\alpha=r_{12}^{-1}\leqslant 1)$$

and pairwise extremal coefficient $\theta_{12}=2^{\alpha}$.

Example 2. The distribution of the asymmetric negative logistic (Joe, 1990) is given by

$$V(z_1,\ldots,z_m)=\sum_{k=1}^{m}\frac{1}{z_k}-\sum_{c\in C_m:|c|\geqslant 2}(-1)^{|c|}\left\{\sum_{i\in c}(\xi_{c,i}/z_i)^r\right\}^{1/r},$$

where $r\leqslant 0$ for all $c\in C_m$. For $i=1,\ldots,m$ we have $\xi_{c,i}=0$ if $i\notin c$, $\xi_{c,i}\geqslant 0$ if $i\in c$, and $\sum_{c\in C_m}(-1)^{|c|}\xi_{c,i}\leqslant 1$. Then, for $A\in C_m$,

$$\theta_A=|A|-\sum_{c\in A:|c|\geqslant 2}(-1)^{|c|}\left(\sum_{i\in c}\xi_{c,i}^r\right)^{1/r}.$$

Unlike for the asymmetric logistic model, the parameters that determine θ_A only relate to dependence of order $|A|$ and lower. For $m=2$ we have

$$V(y,z)=y^{-1}+z^{-1}-\{(\xi_{12,1}/y)^r+(\xi_{12,2}/z)^r\}^{1/r}, \quad \theta_{12}=2-(\xi_{12,1}^r+\xi_{12,2}^r)^{1/r},$$

where $\xi_{12,1}$ and $\xi_{12,2}$ are in $[0,1]$ and $r\leqslant 0$.

Example 3. A family of distributions that is completely determined by the set of all pairwise extremal coefficients θ_{ik} can be obtained as a multivariate version of a spatial process developed in Schlather (2002); see also § 3·4. The distribution arises as the limit distribution of componentwise maxima of a heteroscedastic multivariate series of variables. Let $Z^{(j)}=(Z_1^{(j)},\ldots,Z_m^{(j)})$ be the jth independent copy of the standard Gaussian random vector with correlation matrix (ξ_{ik}), and let $\tau^{(1)},\tau^{(2)},\ldots$ be a sequence of identically distributed volatility variables with heavy-tailed distribution $F_\tau: 1-F_\tau(x)\sim cx^{-1}$. If $c=(2\pi)^{\frac{1}{2}}$ and $X^{(j)}=\tau^{(j)}Z^{(j)}$ then it follows that $\theta_{ik}=1+2^{-\frac{1}{2}}(1-\xi_{ik})^{\frac{1}{2}}$, at least if the $\tau^{(i)}$ are independent. Consequently, any set of pairwise extremal coefficients is self-consistent provided the matrix with (i,k)th element $1-2(\theta_{ik}-1)^2$ is positive definite.

Example 4. A special case of the asymmetric logistic distribution arises as a limit by letting $\xi_{c,i}=\tau_c$, for all $i\in c$, and $r_c\to\infty$, for all $c\in C_m$, to give

$$V(z_1,\ldots,z_m)=\sum_{c\in C_m}t_c\max_{i\in c}z_i^{-1}.$$

This corresponds to H_m being a discrete measure. Here $\sum_{c\ni i}\tau_c=1$ for all $i\in M_m$, and, for $A\in C_m$,

$$\theta_A=\sum_{c\in C_m:c\cap A\neq\varnothing}\tau_c.$$

The class of distributions given in Example 4 is able to model any set of extremal coefficients for any multivariate extreme-value distribution with unit Fréchet margins, a property made precise by Theorem 1.

THEOREM 1. *Assume that the distribution of the random variable $Y=(Y_1,\ldots,Y_m)$ is a member of the class of multivariate extreme-value distributions with unit Fréchet margins. Suppose also that the set of extremal coefficients of Y is $\{\theta_A: A\in C_m\}$. Then there exist*

$2^m - 1$ *independent, unit Fréchet-distributed random variables* T_c *and constants* $0 \leqslant \tau_c$ *for* $c \in C_m$, *where* $\sum_{c,c \ni i} \tau_c = 1$, *for* $i = 1, \ldots, m$, *such that the random variables* $U = \{U_1, \ldots, U_m\}$ *given by*

$$U_i = \max_{c,c \ni i} \tau_c T_c \tag{6}$$

have the same set of the extremal coefficients, of any order, as Y. Furthermore

$$\theta_A = \sum_{c \in C_m, c \cap A \neq \varnothing} \tau_c \quad (A \in C_m),$$

and the τ_c *are uniquely given by*

$$\tau_c = \sum_{A \in C_m, A \supseteq M_m \backslash c} (-1)^{|A \cap c| + 1} \theta_A \quad (c \in C_m). \tag{7}$$

Schlather & Tawn (2002) give a generalisation and a proof of Theorem 1. It can be easily seen that the $2^m - 1$ independent variables used in representation (6) correspond to the minimum number that are required so that the extremal coefficients of U and Y are identical for all possible dependence structures of Y.

In the estimation of extremal coefficients, the problem of finding whether or not a set of extremal coefficients is consistent is solved from a theoretical point of view: if all τ_c in expression (7) are in $[0, 1]$, then the set is self-consistent. The difficulty with these conditions is that the number of conditions to be checked equals $2^m - 1$ for m variables; that is $O(2^{2m})$ computational operations have to be performed. The equivalent calculation for the multivariate Gaussian case involves $O(m^3)$ calculations to check the required inequalities on the eigenvalues of the correlation matrix. The practical applicability of inequalities obtained from Theorem 1, in its full generality, is therefore rather limited.

In general, bounds on a specific extremal coefficient given a self-consistent subset of extremal coefficients can be obtained by linear programs, where the extremal coefficient of interest is the target function and the bounds are derived from the constraint that $\tau_c \geqslant 0$ in expression (7). Such a linear program can be solved numerically only for relatively small m since the number of auxiliary variables involved is $O(2^m)$. Schlather & Tawn (2002) show that in some special cases explicit conditions on the extremal coefficients can be given to ensure that the set is, or can be expanded to, a set of extremal coefficients that satisfies property (7). For example, if the extremal coefficients θ_A for $A \in C_m \backslash \{M_m\}$ are given and the set is self-consistent, then θ_{M_m} is bounded by the sharp bounds

$$\max \left\{ \sum_{\substack{A \in C_m \backslash \{M_m\} \\ A \supseteq c}} (-1)^{|A \backslash c|} \theta_A \right\} \leqslant \theta_{M_m} \leqslant \min \left\{ \sum_{\substack{A \in C_m \backslash \{M_m\} \\ A \supseteq c}} (-1)^{|A \backslash c| + 1} \theta_A \right\},$$

where the maximum and minimum are both over the set $c \in C_m$ such that $|M_m \backslash c| \bmod 2 = 1$ and 0 respectively. To clarify notation and to provide an understanding of these bounds consider the case of $m = 3$, in which case the bounds are

$$\max\{\theta_{12}, \theta_{13}, \theta_{23}, \theta_{12} + \theta_{13} + \theta_{23} - 3\} \leqslant \theta_{123} \leqslant \theta_{12} + \theta_{13} + \theta_{23} - 1 - \max\{\theta_{12}, \theta_{13}, \theta_{23}\}. \tag{8}$$

The term $\max\{\theta_{12}, \theta_{13}, \theta_{23}\}$ is a sharp lower bound if there is strong dependence, whereas $\theta_{12} + \theta_{13} + \theta_{23} - 3$ is a sharp lower bound if the random variables are nearly independent. In practice the pairwise dependence measures are the most useful, and a range of results exists for these including conditions (2), which are a consequence of Theorem 1, and the following new conditions given by Theorem 2.

Theorem 2. *A necessary condition for a set of pairwise extremal coefficients θ_{ij} to be self-consistent is that the matrix $(2-\theta_{ij})_{i,j=1,\ldots,m}$ be positive semidefinite. The positive definiteness of the matrix $\{1-2(\theta_{ij}-1)^2\}_{i,j=1,\ldots,m}$ is a sufficient condition.*

The necessary condition is shown in the Appendix, and the sufficient condition is given by Example 3 above. These self-consistency constraints will be used in § 4 to improve the estimation of the extremal coefficients.

3. Spatial case

3·1. *Extremal coefficient function*

Let $X(x)$ be a stationary random field for $x \in \mathbb{R}^d$ with unit Fréchet margins. Let $X^{(j)}(x)$ be independent replicates of the random field $X(x)$ for $j=1,\ldots,n$. To study the extremal dependence of the process $X(.)$ we focus on the componentwise maximum of the n replicates. Under weak conditions (de Haan, 1984), there exists a real-valued function $\theta(.)$ such that the asymptotic distribution of the normalised maximum at a pair of locations separated by h is Fréchet-distributed with scale parameter $\theta(h)$; that is, for $z>0$,

$$\lim_{n\to\infty} \operatorname{pr}\left[\max_{j=1,\ldots,n} \max\{X^{(j)}(h), X^{(j)}(o)\}/n \leqslant z\right] = \exp\{-\theta(h)/z\}$$

for all $h \in \mathbb{R}^d$, where o denotes the origin. We term $\theta(.)$ the extremal coefficient function. We focus on this function as it provides sufficient information about extremal dependence for many problems even though it measures the pairwise structure only.

3·2. *Max-stable processes*

By definition a max-stable process Y in $\mathbb{R}^d$ with unit Fréchet margins is a random field that has all its higher-order marginal distributions belonging to the class of multivariate extreme-value distributions with unit Fréchet margins. An important class is defined by

$$Y(x) = \max_{j=1,2,\ldots} \zeta_j g(s_j, x),$$

where ζ and s can be interpreted as the magnitude and type of a storm respectively, and g as determining the shape of the event over the region, so that $\zeta_j g(s_j, x)$ is the size of the jth storm at location x. The sizes and the types of the storms are all assumed to be independent. Technically, $\{(\zeta_j, s_j): j=1,2,\ldots\}$ are the points of a Poisson process Π on the space $(0,\infty)\times S$, for some parameter space S of the types of storm, with the intensity of Π given by $\zeta^{-2}\, d\zeta v(ds)$, and g is a nonnegative function which is measurable in the first argument and upper semicontinuous in the second; see de Haan (1984) and Giné et al. (1990). For $Y(x)$ to have unit Fréchet margins we require that

$$\int_S g(s,x) v(ds) = 1$$

for all $x \in \mathbb{R}^d$. It follows that, for any semicontinuous function y,

$$\operatorname{pr}\{Y(x) \leqslant y(x) \text{ for all } x \in \mathbb{R}^d\} = \exp\left\{-\int_S \max_{x\in\mathbb{R}^d} \frac{g(s,x)}{y(x)} v(ds)\right\}.$$

If $Y(.)$ is stationary then, for all $z>0$ and $h \in \mathbb{R}^d$,

$$\mathrm{pr}\{Y(h) \leqslant z,\, Y(o) \leqslant z\} = \exp\{-\theta(h)/z\},$$

where

$$\theta(h) = \int_S \max\{g(s, o), g(s, h)\}\, \nu(ds). \tag{9}$$

Clearly, the right-hand side of relationship (9) restricts the functions $\theta(.)$ can take. For example, if g is an indicator function then $2-\theta(.)$ is a geometric covariance function (Matheron, 1987) and, conversely, any geometric covariance function for sets with unit Lebesgue measure is a model for $2-\theta(.)$.

3·3. *Properties of the extremal coefficient function*

From the connections with the extremal coefficient of two variables, a number of properties of the extremal coefficient function are immediate: invariance to marginal distribution; symmetry about the origin; $\theta(o)=1$; $1 \leqslant \theta(h) \leqslant 2$ for $h \in \mathbb{R}^d$ with the lower and upper bounds corresponding to complete dependence and independence at separation h respectively; if Y is a max-stable process the extremal coefficient function is an exact dependence measure; and any finite convex combination of extremal coefficient functions is an extremal coefficient function. Additional properties of the extremal coefficient function are presented in Theorem 3, whose proof is given in the Appendix.

THEOREM 3. *Let $\theta(h)$, for $h \in \mathbb{R}^d$, be an extremal coefficient function belonging to a stationary max-stable process in $\mathbb{R}^d$. Then the following assertions hold.*

(i) *The function $2-\theta(h)$ is positive semidefinite.*
(ii) *Function $\theta(.)$ is not differentiable at the origin unless $\theta(h)=1$ for all $h \in \mathbb{R}^d$.*
(iii) *If the dimension of the random field is greater than or equal to two and if the random field is isotropic, then $\theta(.)$ has at most a jump at the origin and is continuous elsewhere.*

Theorem 3 shows that the dependence structure of a stationary random field can be characterised by a positive definite function, even if the random field does not have a variance or expectation. Only a subset of all positive definite functions are valid extremal coefficient functions; for example the Gaussian correlation model $\exp(-\|h\|^2)$ is not allowed since it is differentiable at the origin. This feature is shared with other classes of distribution; see Matheron (1987) for restrictions on random fields with multivariate log-normal marginal distributions, for instance. If the assumptions of Theorem 3(iii) are not satisfied then $\theta(.)$ is not necessarily continuous anywhere; for example $\theta(h)=2-I(h \in \mathbb{Q})$ is a valid extremal coefficient function in one dimension, where $I(h \in \mathbb{Q})$ is the indicator of the rationals.

3·4. *Examples*

To illustrate the properties derived in § 3·3 for the extremal coefficient function we consider two max-stable models. The previously mentioned technical report by R. L. Smith, see Coles (1993), gives a range of models with deterministic shape functions of which probably the most useful is the Gaussian storm shape model, illustrated in Coles & Tawn (1996). In two-dimensional space, the model is formulated as follows: storm types are determined by the positions of their centres $s \in \mathbb{R}^2$, with $\nu(ds)$ being Lebesgue measure, and the storm intensity at location x is given by $g(s, x) = \varphi_2(x-s, \Sigma_2)$ where φ_2 is the density function of a zero-mean bivariate normal vector with covariance matrix Σ_2. Smith's report shows that the joint distribution G_h of $\{Y(o), Y(h)\}$ satisfies

$$-\log G_h(y, z) = y^{-1}\Phi\left(\frac{1}{2}\beta_h + \beta_h^{-1}\log\frac{z}{y}\right) + z^{-1}\Phi\left(\frac{1}{2}\beta_h + \beta_h^{-1}\log\frac{y}{z}\right), \tag{10}$$

where $\beta_h = (h^{\mathrm{T}}\Sigma_2^{-1}h)^{\frac{1}{2}}$ and Φ is the standard normal distribution function. The extremal coefficient function is therefore $\theta(h) = 2\Phi(\beta_h/2)$. The bivariate distribution (10) and also the higher-order joint distributions for this process correspond to multivariate distributions derived by Hüsler & Reiss (1989).

Models with stochastic shape functions are proposed by Schlather (2002). The extremal Gaussian model Y is defined as

$$Y(x) = \max_{t_j \in \Pi} t_j \max\{0, Z^{(j)}(x)\},$$

where Π is the Poisson process on the positive real axis with intensity $\lambda(t) = (2\pi)^{\frac{1}{2}}t^{-2}$ for $t > 0$, and the $Z^{(j)}(x)$ are independent stationary Gaussian random fields in $\mathbb{R}^d$ with standard marginals and correlation function $\rho(h)$, for $h \in \mathbb{R}^d$.

In this context t_j represents the size of the storm and $\max\{0, Z^{(j)}(x)\}$ is the shape of the jth storm. Thus the shape is stochastic, but with a common latent structure of dependence. For this process it is shown that the bivariate distribution for two locations a distance h apart is

$$\mathrm{pr}\{Y(o) \leqslant y, Y(h) \leqslant z\} = \exp\left\{-(z^{-1} + y^{-1})a_h\left(\frac{y}{y+z}\right)\right\},$$

where

$$a_h(\xi) = \frac{1}{2} + \left[\frac{1}{4} - \frac{1}{2}\{\rho(h) + 1\}\xi(1-\xi)\right]^{\frac{1}{2}}.$$

Hence, the extremal coefficient equals

$$\theta(h) = 1 + 2^{-\frac{1}{2}}\{1 - \rho(h)\}^{\frac{1}{2}} \quad (h \in \mathbb{R}^d). \tag{11}$$

This class is completely characterised by $\theta(.)$ or $\rho(.)$ as (11) defines a 1–1 relationship. Although $\theta(.)$ can take values in $[1, 2]$, for practically relevant, stationary and isotropic, random fields of dimension 2 or higher, there are restrictions on $\theta(.)$; for example, $\lim_{\|h\| \to \infty} \theta(h) \leqslant 1 + 2^{-\frac{1}{2}}$, $\theta(.) \leqslant 1{\cdot}838$ in $\mathbb{R}^2$, and $\theta(.) \leqslant 1{\cdot}781$ in $\mathbb{R}^3$. That is, independence cannot be obtained for stationary and isotropic extremal Gaussian processes. The restrictions follow from the properties of isotropic positive definite functions; see Matérn (1986, p. 16) and Schoenberg (1938).

4. Estimation

4·1. *Introduction*

We study flexible nonparametric estimators which are ideal for use in exploratory studies. A number of such naive estimators for the extremal coefficient are easily constructed, but none of these yields self-consistent sets of coefficients in general. In the multivariate case, we present different strategies for achieving self-consistent estimators, for which numerical solutions can be given in relatively small dimensions. For the spatial case we aim to provide diagnostics only, so self-consistency of estimators is less important.

Whatever the marginal distributions of the data, the values are first transformed to Fréchet values. We assume that the data are observations on independent and identically distributed variables so we can use the empirical probability integral transformation separately for each margin; that is only the ranks of the marginal values are used. We presume this transformation has been made to all Fréchet marginal data, even when the data are generated from a Fréchet distribution explicitly. For data which are nonidentically distributed in its marginal structure over observations, but have a dependence structure which is invariant over observations, some modelling assumptions are required to construct an appropriate transformation to Fréchet marginal distributions. A natural starting point is to use covariate models for the marginal tail behaviour; see Smith (1989) and Hall & Tajvidi (2000a).

4·2. *Naive estimators*

Assume that we have observations $X^{(1)}, \ldots, X^{(n)}$ which are independent m-variate random vectors with unit Fréchet margins and are in the domain of max-attraction of a multivariate extreme-value distribution. It follows that the univariate variable $\max_{i \in A} X_i$ is in the domain of attraction of the extreme-value distribution function $\exp(-\theta_A/x)$ for $x>0$. Following standard univariate threshold methods, we assume that the distribution of $\max_{i \in A} X_i$ is identical to $\exp(-\theta_A/x)$ for $x>0$ above some high threshold level, z, and we estimate the θ_A using a censored likelihood approach; see Smith (1989) and Nadarajah et al. (1998). We refer to the threshold by the marginal probability t of not exceeding z, that is $t=\exp(-1/z)$. Then the loglikelihood for θ_A, for $A \in C_m$, is given by

$$l_A(\theta_A) = \text{card}\left\{ j : \max_{i \in A} (X_i^{(j)} \bar{X}_i) > z \right\} \log \theta_A - \theta_A \sum_{j=1}^{n} \left[\max\left\{ z, \max_{i \in A} (X_i^{(j)} \bar{X}_i) \right\} \right]^{-1},$$

where $\bar{X}_i = n^{-1} \sum_{j=1}^{n} 1/X_i^{(j)}$ and the corresponding maximum likelihood estimator is denoted by $\hat{\theta}_{A,t}$. The Fréchet variables are scaled by $\bar{X}_i$ to ensure that $\hat{\theta}_i = 1$ when $t=0$. For $t=0$, $\hat{\theta}_{A,t}$ reduces to a variant of the Hall & Tajvidi (2000b) estimator that has been introduced to estimate the V function of a multivariate extreme-value distribution. In the following we denote $\hat{\theta}_{A,0}$ by $\hat{\theta}_{A,\mathrm{HT}}$. As $\hat{\theta}_{A,\mathrm{HT}}$ is not restricted to $[1, |A|]$, we also consider the estimator $\hat{\theta}_{A,\mathrm{b}}$, which truncates $\hat{\theta}_{A,\mathrm{HT}}$ to $[1, |A|]$.

4·3. *Self-consistent estimators*

Two different strategies for building self-consistent estimators are proposed, namely sequential correction of naive estimators and direct estimation of the set of extremal coefficients under consideration of the inequalities in Theorem 1. The choice of approach depends on the size of the problem, with the sequential-correction method being much the faster algorithm but requiring an ordering of the extremal coefficients in which the values are corrected and which in general is arbitrary.

For sequential correction, we correct coefficients in the order of increasing values of $|A|$. For equal values of $|A|$ the correction is made in a randomly selected order unless some coefficients are estimated with more certainty than others, in which case the better estimated coefficients are corrected first. The ensemble of all consistent sets of extremal coefficients is convex, and the admissible set of values for a specific coefficient, given a consistent subset of $\{\theta_A : A \in C_m\}$, is an interval to which an estimate is truncated. We denote by $\hat{\theta}_{A,\mathrm{s}}$ the sequentially-corrected Hall & Tajvidi (2000b) estimator.

The direct approach is based on joint estimation of the set of extremal coefficients subject to the estimates being constrained to be self-consistent. We construct a joint loglikelihood function l of $\{\theta_A : A \in C_m\}$ by falsely assuming independence between the observations of $\max_{i \in A} X_i$ for all different A to give the pseudo-loglikelihood function

$$l = \sum_{A \in C_m, |A| \geqslant 2} l_A(\theta_A). \tag{12}$$

It is easily shown that the maximum pseudolikelihood estimator is consistent; see Liang & Self (1996). We maximise the pseudolikelihood subject to the parameters $\{\theta_A : A \in C_m\}$ satisfying the condition that $\tau_c \geqslant 0$ in expression (7). We denote the resulting estimator of θ_A by $\hat{\theta}_{A,m}$.

4·4. *Simulation study: Multivariate case*

We investigate the logistic distribution and restrict ourselves to the trivariate case for simplicity. To reduce further complexity we choose a pairwise symmetric case, that is

$$\xi_1 = \xi_{1,1} = \xi_{2,2} = \xi_{3,3}, \quad \xi_2 = \xi_{12,1} = \xi_{12,2} = \ldots = \xi_{23,3}, \quad \xi_3 = \xi_{123,1} = \xi_{123,2} = \xi_{123,3}$$

in (5), where $\xi_1 + 2\xi_2 + \xi_3 = 1$. Let r_2 and r_3 be the exponents corresponding to ξ_2 and ξ_3, respectively. Then

$$\theta_{ik} = 2\xi_1 + (2 + 2^{1/r_2})\xi_2 + 2^{1/r_3}\xi_3, \quad \theta_{123} = 3\xi_1 + 2^{1/r_2}3\xi_2 + 3^{1/r_3}\xi_3.$$

The sequence of correction for $\hat{\theta}_{A,\mathrm{s}}$ that we use is $\theta_{12}, \theta_{23}, \theta_{13}, \theta_{123}$; hence $\hat{\theta}_{A,\mathrm{s}}$ is identical to $\hat{\theta}_{A,\mathrm{b}}$ for $A = \{1, 2\}, \{2, 3\}$. Numerical maximisation of l in (12) is performed with the simulated annealing algorithm implemented in R (Ihaka & Gentleman, 1996). In order to compare the performance of the estimators, we choose as a measure

$$\alpha_A = \theta_A^{-1}[E\{(\hat{\theta}_A - \theta_A)^2\}]^{\frac{1}{2}}.$$

The number of replicates n on which a single estimation is based is taken to be 6, 10, 20, 40, 100 and 1000, and the estimation of α_A is based on 500 simulations for each parameter specification.

Figure 1 presents the simulation results when the parameter vector $(\xi_1, \xi_2, \xi_3, r_2, r_3)$ equals $(0{\cdot}2, 0{\cdot}3, 0{\cdot}2, 2, 2)$ and $(0, 0, 1, 1, \frac{10}{3})$, so that $(\theta_{ik}, \theta_{123})$ equals $(1{\cdot}71, 2{\cdot}21)$ and $(1{\cdot}23, 1{\cdot}39)$, respectively. Figures 1(a)–(f) show that improvement of $\hat{\theta}_{\mathrm{HT}}$ is obtained by marginal correction, $\hat{\theta}_{\mathrm{b}}$, and a further improvement is obtained by sequential correction, $\hat{\theta}_{\mathrm{s}}$. Depending on the true parameter and the extremal coefficient being estimated, the performance of the pseudolikelihood estimator is comparable to or better than $\hat{\theta}_{\mathrm{HT}}$. All $\hat{\theta}_{\mathrm{s}}$ and $\hat{\theta}_m$ estimates are self-consistent, whereas the proportion of self-consistent estimates of $\hat{\theta}_{\mathrm{HT}}$ and $\hat{\theta}_{\mathrm{b}}$ is low for relatively small n, with $\hat{\theta}_{\mathrm{b}}$ being marginally the better of the two; see Figs 1(c) and (f).

4·5. *Empirical extremal coefficient function*

Analysing and modelling spatial extremes raises additional difficulties, as the data are both high-dimensional and sparse. 'High-dimensional' means 'of order 50', which is far from what can be treated in multivariate extreme-value statistics. On the other hand, in classical geostatistics, several hundred observations are necessary to estimate reliably the dependence structure of a Gaussian random field, even when stationarity and isotropy in one form or the other are assumed. In contrast we have many replications in time for each site, so the estimates of dependence for any pair of sites should have relatively small

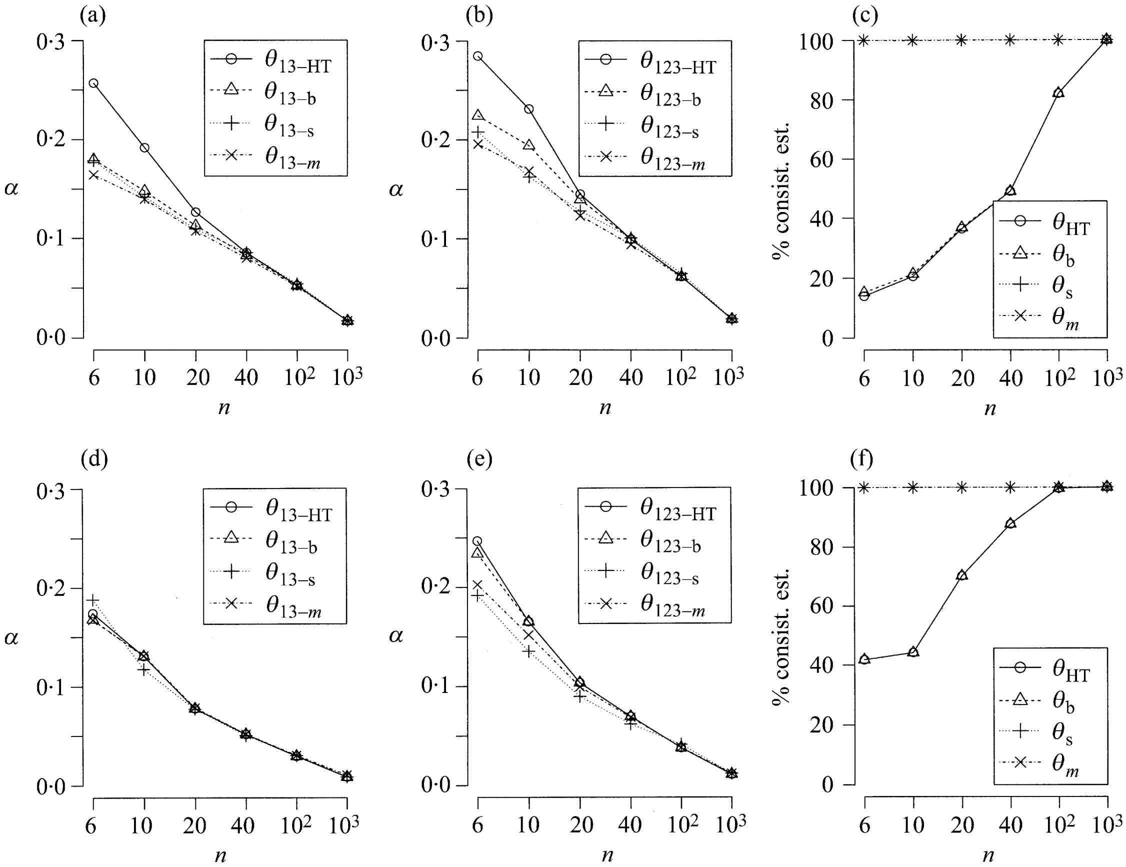

Fig. 1: Simulation study. Quality, α, of the estimators as a function of the number n of independent random vectors on which $\hat{\theta}_A$ is based; $(\xi_1, \xi_2, \xi_3, r_2, r_3) = (0{\cdot}2, 0{\cdot}3, 0{\cdot}2, 2, 2)$ in (a)–(c) corresponding to $(\theta_{ik}, \theta_{123}) = (1{\cdot}707, 2{\cdot}219)$ and $(\xi_1, \xi_2, \xi_3, r_2, r_3) = (0, 0, 1, 1, \frac{10}{3})$ in (d)–(f) corresponding to $(\theta_{ik}, \theta_{123}) = (1{\cdot}231, 1{\cdot}390)$; (c) and (f) show percentage of self-consistent estimates plotted against n.

variability by comparison to problems in classical geostatistics. Here we study the estimation of the pairwise dependence structure of stationary and isotropic max-stable random fields, that is the extremal coefficient function $\theta(h)$.

Recall that, by Theorem 3, $\theta(h) - 1$ is a conditionally negative definite function; that is, for any stationary max-stable process, $\theta(h) - 1$ is a valid variogram model for a second-order stationary random field. Consequently, for constructing an estimator of the extremal coefficient function we use a similar approach to the variogram analysis in geostatistics; see Chilès & Delfiner (1999). We introduce the extremal coefficient cloud C_θ as an intermediate step to our estimator, the empirical extremal coefficient function $\hat{\theta}(h)$, of $\theta(h)$. Assume that we have independent and identically distributed replicated data $z^{(j)}(x_i)$ from a stationary and isotropic max-stable random field at locations $x_1, \ldots, x_m$ and instances $j \in T_i$, where T_i is the index set of the observed data for site x_i. The pairwise extremal coefficients θ_{ik} can be estimated by any naive estimator $\hat{\theta}_{ik}$, see § 4·2, using all the instances where data are available simultaneously for both locations, x_i and x_k. The plot of $\hat{\theta}_{ik}$ against the distance $\|x_i - x_k\|$ for all i and k allows for a first check on the consistency of the data. We call the ensemble of the pairs $(\|x_i - x_k\|, \hat{\theta}_{ik})$ the extremal coefficient cloud C_θ. The empirical extremal coefficient function $\hat{\theta}(h)$ is the local average of the C_θ cloud, which is then truncated to the interval $[1, 2]$; that is

$$\hat{\theta}(h) = 1 \vee \hat{\theta}^*(h) \wedge 2,$$

where

$$\hat{\theta}^*(h) = \frac{\sum_{(i,k) \in S_h} |T_i \cap T_k|^{\frac{1}{2}} \hat{\theta}_{ik}}{\sum_{(i,k) \in S_h} |T_i \cap T_k|^{\frac{1}{2}}},$$

in which $S_h = \{(i, k) : \|x_i - x_k\| \in [h - \varepsilon, h + \varepsilon]; |T_i \cap T_k| > 1\}$. In contrast to standard geostatistical analyses the binning for the empirical extremal coefficient function is weighted because the $\hat{\theta}_{ik}$ estimates have unequal variances as a result of missing values causing different values in the number $|T_i \cap T_k|$ of overlapping data at the two sites.

5. Exploratory analysis of rainfall extremes

5·1. *Introduction*

The volume of rainfall and its spatial spread determine the degree of flooding at sites. Modelling of both the spatial variation in the marginal distribution of rainfall extremes and the spatial dependence structure of extreme rainfall events over a catchment is required to provide this information (Coles & Tawn, 1996). We use a slightly extended version of the rainfall data studied by Coles & Tawn (1996), and reanalyse the spatial extremal dependence in the data to provide diagnostic information about the characteristics of extreme events and for use in future model building. Our data are daily aggregates, recorded from a start time of 9:00 each day at sites over a 40 km × 40 km region in south-west England. We have two forms of data: 11 sites have 19 years of continuous data and 54 additional sites have annual maximum data with some sites having up to 100 years of data and many sites having periods of missing data relative to other sites; see Fig. 2. The number of concurrent years of annual maximum data at pairs of sites ranges between 0 and 85, with a median of 12 years and an upper quartile of 20 years. Let $\tilde{X}_i^{(j)}$ denote the rainfall at site i on day j, and let $\tilde{Y}_i^{(j)}$ denote the annual maximum rainfall at site i in year j.

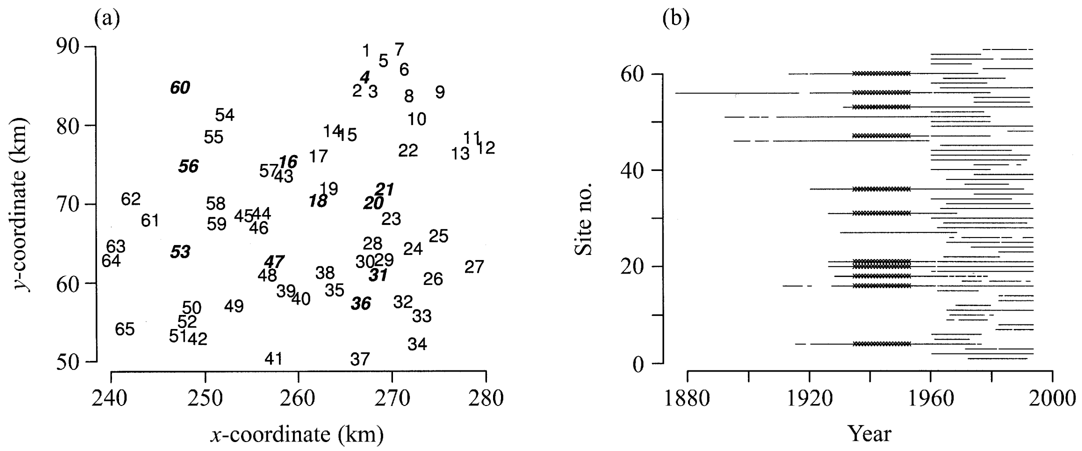

Fig. 2: Rainfall data. (a) Locations of rain gauge sites and (b) availability of annual maxima, shown by horizontal lines, and daily data, shown by crosses. In (a) the bold italic numbers correspond to sites with daily data.

As we are focusing on the dependence structure we first transform the data at each site to have a unit Fréchet marginal distribution. We let

$$X_i^{(j)} = -1/\log F_{\tilde{X}_i}(\tilde{X}_i^{(j)}), \quad Y_i^{(j)} = -1/\log F_{\tilde{Y}_i}(\tilde{Y}_i^{(j)}),$$

where $F_{\tilde{X}_i}$ and $F_{\tilde{Y}_i}$ are the empirical distribution functions of $\tilde{X}_j$ and $\tilde{Y}_i$, respectively, which are estimated under the assumption of stationarity. It is assumed that the transformed annual maxima data are multivariate extreme-value distributed and that the transformed daily data are in the domain of attraction of a multivariate extreme-value distribution, in each case with Fréchet marginal distributions.

5·2. *Preliminary analysis*

First consider an exploratory analysis of the daily rainfall data. Figure 3(a) shows that the cloud of extremal coefficients for the daily data has outliers. Here the cloud is estimated by the methods given in § 4·5, using estimator $\hat{\theta}_t$ with $t = 0{\cdot}99$, which corresponds to 3 to 4 extreme events per year. Investigation showed that the outlying $\hat{\theta}_{ik,t}$ values arise when i or k corresponds to site number 18, 20 or 21. When exploring the cause of this inconsistency, we identified that a time-shift of a day occurred for these sites on the first day of some of the years of data. We have subsequently discovered that this type of timing error is an occasional problem with daily rainfall data because of a confusion over how to register the daily rainfall as a result of the aggregation period running from 9:00 on one day to 9:00 on the next day. This feature had not been identified before for these data, but immediately illustrates the value of the diagnostic procedures we have developed. The data were subsequently edited and the remaining analysis is based on the modified data; see Fig. 3(b). The modified estimates suggest that there is a simpler structure of extremal dependence than that found by Coles & Tawn (1996), with our interpretation of the slightly weaker dependence, between sites 18, 20 and 21 and the other sites, representing additional evidence of the poorer data quality control for the three sites.

For the annual maxima data, $\hat{\theta}_{\mathrm{HT}}$ has been used to estimate the C_θ cloud. We did not find any evidence of covariate relationships for altitude or its gradient. We conclude that

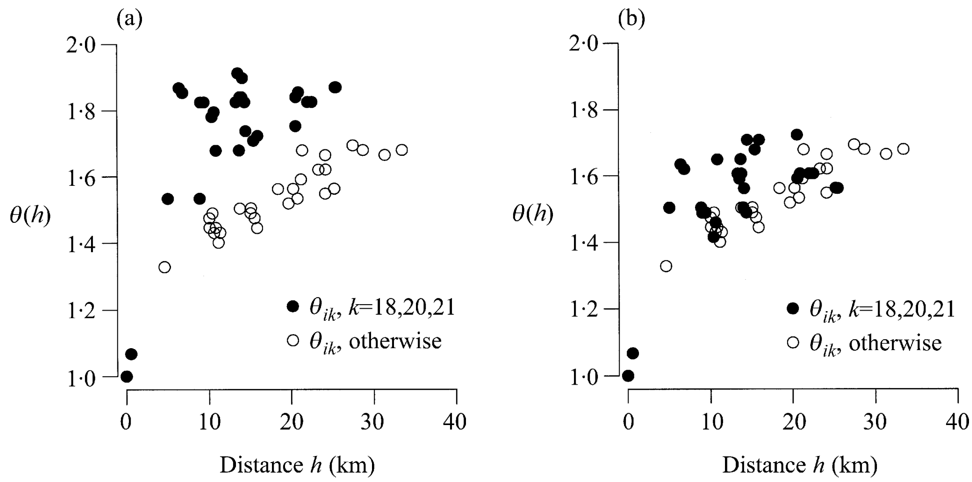

Fig. 3: Rainfall data. Extremal coefficient cloud for the daily data (a) before, (b) after correction of the timing errors

the preliminary analysis has showed that stationarity and isotropy are reasonable assumptions for the extremes of the rainfall data.

5·3. *Empirical extremal coefficient function*

We use $\hat{\theta}_{\rm HT}$ and $\hat{\theta}_t$ to estimate the cloud C_θ for annual maxima and daily data respectively, and denote the corresponding weighted spatial binned estimators by $\hat{\theta}_{\rm HT}(h)$ and $\hat{\theta}_t(h)$. Comparison between $\hat{\theta}_{\rm HT}(h)$ in Fig. 4 and the $\hat{\theta}_t$ values in Fig. 3(b) shows good agreement. However, further investigation reveals that the cloud values in Fig. 3(b) depend heavily on the chosen t-value: for $t = 0{\cdot}97$, 0·99 and 0·995 the asymptotes for distances $h \to 40$ km are approximately 1·55, 1·65 and 1·75 respectively. Coles & Tawn (1996) also note weaker dependence at higher thresholds, and regard this as an indicator of the tendency for the most extreme events to be more localised; see also Ancona-Navarrete & Tawn (2002). Nevertheless, both Figs 3(b) and 4 show strong extremal dependence, which decreases with distance, for distances less than 10 km, and a near constant level of dependence for all other separations of sites. As the region is relatively small by comparison to the spatial scale of meteorological systems, the fact that no two sites in the region are independent is not surprising.

The variability of the estimates obtained for both the daily data and the annual maxima has to be assessed. To this end we perform a Monte Carlo study using models that both satisfy the restrictions on $\theta(h)$, identified in Theorem 3, and capture the shape of $\hat{\theta}(h)$. First the extremal Gaussian process with extremal coefficient function

$$\theta(h) = 1 + 2^{-\frac{1}{2}} - 2^{-\frac{1}{2}} \exp(-h/s) \quad (h \geqslant 0) \tag{13}$$

was used, see Fig. 4, where the scale parameter s had been estimated to be 5·3, by a weighted least squares method. Model (13) is valid, since the function ρ given by equality (11) has a nonnegative Fourier inverse.

We present estimates based on the analysis of the annual maxima only. We simulate from the extremal Gaussian process assuming independence of the data for different years,

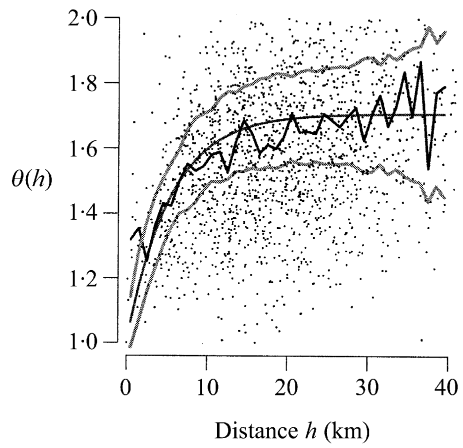

Fig. 4: Rainfall data. Extremal coefficient function $\hat{\theta}_{\rm HT}(h)$ for the annual maxima (black), exponential model $\theta(h)$ (dark grey) and 95% confidence bounds (light grey); the points build the coefficient cloud of $\hat{\theta}_{\rm HT}$. Points with values of $\hat{\theta}_{\rm HT}$ larger than 2 are omitted.

and generating the annual maxima for the same years and the same arrangement of the missing values as given in Fig. 2(b), using the algorithm of Schlather (2001b, 2002). The resulting approximate 95% confidence interval, for each bin of the empirical extremal coefficient function separately, is shown in Fig. 4. The confidence-interval envelope shows that there is significant variation in the level of dependence within the region. The envelope contains $\hat{\theta}(h)$ for all h away from 0, illustrating that the model is a good fit, though a possible nugget effect at $h=0$ in $\hat{\theta}(h)$ is identified, indicating local variations in the extreme rainfalls or inaccuracy of the measuring apparatus at storm events. Similar overall findings were obtained using data simulated from Smith's Gaussian extreme-value process with $\Sigma_2=\operatorname{diag}(64, 64)$, indicating that these results are not really dependent on the form of the underlying max-stable process.

To summarise, local averaging of a naive estimator that is based on annual maxima data is informative. Despite the additional information in the daily data, we recommend care in exploiting these data given the apparent threshold sensitivity of the corresponding estimates. However, unreported results suggest that the threshold can be selected to give estimates that are similar to those based on the annual maxima but have smaller variances, so further investigation of these methods seems to be required. For the rainfall data we have found a simple and physically realistic pattern of spatial dependence in the extreme events which is much more easily modelled than the representation found by Coles & Tawn (1996).

Acknowledgement

The authors thank the editor and the referees for valuable comments. M. Schlather has been supported by the European Union Training and Mobility of Researchers network on 'Computational and statistical methods for the analysis of spatial data' and the German Federal Ministry of Research and Technology, and J. A. Tawn by the UK Engineering and Physical Sciences Research Council.

Appendix

Proofs

Proof of Theorem 2. Theorem 1 yields suitable τ_c, for $c \in C_m$, so that

$$2-\theta_{ij}=\theta_i+\theta_j-\theta_{ij}=\sum_{c\subseteq C_m, c\supseteq\{i,j\}} \tau_c$$

for any $i \in M_m$ and $j \in M_m$. For $c \in C_m$, let $J_c=\{(i,j) \in M_m \times M_m : i \neq j, \{i,j\} \subseteq c\}$. Then, for any $a_i \in \mathbb{C}$ and its conjugate complex number $\bar{a}_i$,

$$\begin{aligned}\sum_{i,j=1}^{m} a_i(\theta_i+\theta_j-\theta_{ij})\bar{a}_j &= \sum_{i=1}^{m} |a_i|^2 + \sum_{c \in C_m, c\supseteq\{i\}} \tau_c + \sum_{(i,j)\in J_{M_m}} a_i\bar{a}_j \sum_{c\in C_m, c\supseteq\{i,j\}} \tau_c \\ &= \sum_{i=1}^{m} \tau_{\{i\}}|a_i|^2 + \sum_{c\in C_m, |c|\geqslant 2} \tau_c \sum_{i\in c} |a_i|^2 + \sum_{c\in C_m, |c|\geqslant 2} \tau_c \sum_{(i,j)\in J_c} a_i\bar{a}_j \\ &= \sum_{i=1}^{m} \tau_{\{i\}}|a_i|^2 + \sum_{c\in C_m, |c|\geqslant 2} \tau_c \left|\sum_{i\in c} a_i\right|^2 \geqslant 0.\end{aligned}$$

Proof of Theorem 3. The assertions (i) and (iii) follow directly from Theorem 2 and the representation theorem for isotropic positive definite functions of Gneiting & Sasvári (1999), respectively. In order to show part (ii) let $\zeta(.) = \theta(.) - 1$. The second inequality of (8) yields that $\zeta(x-y) \leqslant \zeta(x) + \zeta(y)$ for any $x, y \in \mathbb{R}^d$. Replacing y by $-x$ we get $\zeta(2x) \leqslant 2\zeta(x)$ and by iteration $2^{-n}\zeta(x) \leqslant \zeta(2^{-n}x)$. Since $\theta(o) = 1$ we have $\zeta(o) = 0$, and the differential quotient satisfies

$$\frac{\zeta(2^{-n}x) - \zeta(o)}{2^{-n}\|x\|} \geqslant \frac{\zeta(x)}{\|x\|} > 0 \quad (n \to \infty)$$

if $\zeta(x) \neq 0$. On the other hand, the symmetry of ζ implies that the derivative must be zero at the origin if it exists.

References

Ancona-Navarrete, M. A. & Tawn, J. A. (2002). Diagnostics for extremal dependence in spatial processes. *Extremes*. To appear.

Buishand, T. A. (1984). Bivariate extreme-value data and the station-year method. *J. Hydrol.* **69**, 77–95.

Chilès, J.-P. & Delfiner, P. (1999). *Geostatistics. Modeling Spatial Uncertainty*. New York: Wiley.

Coles, S. G. (1993). Regional modelling of extreme storms via max-stable processes. *J. R. Statist. Soc.* B **55**, 797–816.

Coles, S. G. & Tawn, J. A. (1994). Statistical methods for multivariate extremes: an application to structural design (with Discussion). *Appl. Statist.* **43**, 1–48.

Coles, S. G. & Tawn, J. A. (1996). Modelling extremes of the areal rainfall process. *J. R. Statist. Soc.* B **58**, 329–47.

Cox, D. R. & Wermuth, N. (1996). *Multivariate Dependencies*. London: Chapman and Hall.

De Haan, L. (1984). A spectral representation for max-stable processes. *Ann. Prob.* **12**, 1194–204.

De Haan, L. & Resnick, S. I. (1977). Limit theory for multivariate sample extremes. *Z. Wahr. verw. Geb.* **40**, 317–37.

Ekholm, A., Smith, P. W. F. & McDonald, J. W. (1995). Marginal regression analysis of a multivariate binary response. *Biometrika* **82**, 847–54.

Ekholm, A., McDonald, J. W. & Smith, P. W. F. (2000). Association models for a multivariate binary response. *Biometrics* **56**, 712–8.

Giné, E., Hahn, M. G. & Vatan, P. (1990). Max-infinitely divisible and max-stable sample continuous processes. *Prob. Theory Rel. Fields* **87**, 139–65.

Gneiting, T. & Sasvári, Z. (1999). The characterization problem for isotropic covariance functions. *Math. Geol.* **31**, 105–11.

Hall, P. & Tajvidi, N. (2000a). Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data. *Statist. Sci.* **15**, 153–67.

Hall, P. & Tajvidi, N. (2000b). Distribution and dependence-function estimation for bivariate extreme-value distributions. *Bernoulli* **6**, 835–44.

Hüsler, J. & Reiss, R. D. (1989). Maxima of normal random variables: Between independence and complete dependence. *Statist. Prob. Lett.* **7**, 283–6.

Hutchinson, T. P. & Lai, C. D. (1991). *The Engineering Statistician's Guide to Continuous Bivariate Distributions*. Sydney: Rumsby Scientific Publishing.

Ihaka, R. & Gentleman, R. (1996). R: A language for data analysis and graphics. *J. Comp. Graph. Statist.* **5**, 299–314.

Joe, H. (1990). Families of min-stable multivariate exponential and multivariate extreme value distributions. *Statist. Prob. Lett.* **9**, 75–81.

Joe, H. (1997). *Multivariate Models and Dependence Concepts*. London: Chapman and Hall.

Liang, K.-Y. & Self, S. G. (1996). On the asymptotic behaviour of the pseudo-likelihood ratio test statistic. *J. R. Statist. Soc.* B **58**, 785–96.

Matérn, B. (1986). *Spatial Variation*, 2nd ed. Berlin: Springer.

Matheron, G. (1987). Suffit-il, pour une covariance, d'être de type positif? *Sci. de la Terre*, *Sér. Informatique Géologique* **26**, 51–66.

Nadarajah, S., Anderson, C. W. & Tawn, J. A. (1998). Ordered multivariate extremes. *J. R. Statist. Soc.* B **60**, 473–96.

Pickands, J. (1981). Multivariate extreme value distributions. *Bull. Int. Statist. Inst.* **49**, 859–78.

Resnick, S. I. (1987). *Extreme Values, Regular Variation, and Point Processes*. New York: Springer.

Schlather, M. (2001a). Examples for the coefficient of tail dependence and the domain of attraction of an extreme value distribution. *Statist. Prob. Lett.* **53**, 325–9.

Schlather, M. (2001b). Simulation and analysis of random fields. *R News* **1**(2), 18–20.

Schlather, M. (2002). Models for stationary max-stable random fields. *Extremes* **5**, 33–44.

Schlather, M. & Tawn, J. A. (2002). Inequalities for the extremal coefficients for multivariate extreme value distributions. *Extremes* **5**, 87–102.

Schoenberg, I. J. (1938). Metric spaces and completely monotone functions. *Ann. Math.* **39**, 811–41.

Smith, R. L. (1989). Extreme value analysis of environmental time series: an application to trend detection in ground level ozone (with Discussion). *Statist. Sci.* **4**, 367–93.

Tawn, J. A. (1988). Bivariate extreme value theory: Models and estimation. *Biometrika* **75**, 397–415.

Tawn, J. A. (1990). Modelling multivariate extreme value distributions. *Biometrika* **77**, 245–53.

Tiago de Oliveira, J. (1962/63). Structure theory of bivariate extremes, extensions. *Est. Mat. Estat. Econ.* **7**, 165–95.

[*Received August* 2001. *Revised April* 2002]