MAX-STABLE PROCESSES AND SPATIAL EXTREMES

RICHARD L. SMITH

Department of Mathematics

University of Surrey, Guildford GU2 5XH, England

October 29 1990

Summary.

Max-stable processes arise from an infinite-dimensional generalisation of extreme value theory. They form a natural class of processes when sample maxima are observed at each site of a spatial process, a problem of particular interest in connection with regional estimation methods in hydrology. A general representation of max-stable processes due to de Haan and Vatan is discussed, and examples are given to show how it may be used to generate explicit examples of max-stable process. As a side-product, it is possible to generate a number of known multivariate extreme value families in this way, and in one case this suggests an extension of the family. The main contribution of the paper, however, is to define two new max-stable stochastic processes, related to the multivariate normal and multivariate t distributions. Statistical estimation and model checking are discussed, and the concepts illustrated by being applied to rainfall data.

Keywords. Gaussian extreme value process, Hydrological extremes, Max-stable process, Multivariate extreme value theory.

1. INTRODUCTION.

Extreme value theory has traditionally been developed for univariate series in which the extremes of a single variable are being considered. The recent papers of Smith (1989) and Davison and Smith (1990) include regression models which allow for dependence between the variable of interest and relevant covariates, but still in the context of estimating extreme values for a single series. *Multivariate* extreme value theory has been developed to study the joint distribution of extremes in several series. The probabilistic limit theory of multivariate extremes is reasonably well established, and has been reviewed in the books of Resnick (1987) and Galambos (1987), but statistical theory is still in rapid development. Tawn (1988, 1990a) and Smith, Tawn and Yuen (1990) have proposed methods based on multivariate extreme value distributions, which therefore generalise the classical approach to univariate extremes based on the limiting extreme value distributions (Gumbel 1958), while Coles and Tawn (1991) and Joe, Smith and Weissman (1991) proposed methods extending the threshold-exceedances approaches developed in Smith (1989) and Davison and Smith (1990).

The present paper extends these ideas to *spatial* extremes, in which data such as rainfall or sea levels are collected on a grid of points in space, and the joint distribution of extreme values at different places is of interest. For a finite grid of sampling points, it is possible to think of this problem as merely another application of multivariate extreme value theory, but our approach here is based on a general class of models for extremes in stochastic processes, known as *max-stable processes*. There are two principal advantages of this approach over the one based on general multivariate extreme value distributions. First, it generates a tractable family of multivariate extreme value distributions even when the number of grid points becomes large, whereas the general family of p-variate multivariate extreme value distributions involves an arbitrary positive measure on a (p-1)-dimensional simplex or sphere, which is intractable when p is large. Secondly, by developing a

general spatial model it is possible to answer questions, for instance about spatial aggregation or interpolation, which a simple multivariate model cannot.

A particular area where this question is of interest is that of regional methods in hydrology. The traditional approach to hydrological extremes, as represented for example by the Flood Studies Report (NERC 1975), has recognised that data at a single site are generally too sparse for satisfactory estimates of extremal properties to be based on that site alone. Consequently, hydrologists have grouped data into regions based on geographical or catchment characteristics, and have assumed the data within a region to be drawn from a common distribution except for a scaling factor at each site. Most methods in current use ignore inter-site dependence. Stedinger (1983) examined the effect of inter-site dependence, and concluded that the main effect was to increase the variance of estimates of margin parameters, compared with the independent case. Hosking and Wallis (1988) conducted a simulation study which confirmed Stedinger's results, and also concluded that the effect of inter-site dependence is less important than that of heterogeneity of the margin parameters from site to site. They also concluded that even when both dependence and heterogeneity are present, it is still better to employ a regional method (assuming independence and homogeneity) than to estimate separately at each site. The bulk of their study, however, assumed a dependence structure derived by pointwise transformation from a multivariate normal distribution. In view of the classical result of Sibuya (1960), on the asymptotic independence of extreme values from correlated normal samples, this may not be adequate to reveal the full extent of the problem. Moreover, both Stedinger and Hosking and Wallis were primarily interested in the effects of dependence on estimates of the margin parameters, rather than for its own interest. Buishand (1984) proposed a method for estimating the dependence of extreme values between two sites, but his method has several *ad hoc* features and is not easily extended to multiple

sites. Reed and Dales (1989) developed a technique for spatial extremes based on the notion of an *equivalent number of sites*. Suppose there are N sites from which data have a common marginal distribution function G. They assumed that the distribution function of the maximum over the N sites is of the form G^{N_e} where N_e is the equivalent number of independent sites. That is, the maximum over N dependent sites is assumed to have the same distribution as the maximum over N_e independent sites, where $1 \leq N_e \leq N$, the extreme cases $N_e = 1$, $N_e = N$ corresponding respectively to the cases of perfect dependence (same observations at all sites) and independence. The method is a convenient method of summarising dependence, but it lacks a firm theoretical basis and this causes a number of difficulties in its application.

In a recent paper, Coles and Tawn (1990) have surveyed these and a number of other aspects of regional methods, and have proposed a model for time-series dependence based on a Markov chain model for consecutive sites. This is a simple form of spatial model, though as with most time series models it misses important features of spatial data.

The approach based on max-stable processes, which is developed in this paper, may be thought of as an infinite-dimensional extension of multivariate extreme value theory. Some theory for such processes was developed by de Haan (1984), Vatan (1985) and de Haan and Pickands (1986), who obtained amongst other things a *spectral representation* for such processes. In Section 2 we review this theory and show how special cases of this representation lead to alternative derivations of some existing multivariate extreme value distributions. Section 3 is devoted to a new class of stochastic processes, termed *Gaussian extreme-value processes*, which have close connections with the multivariate normal distribution as well as being max-stable. It also contains an extension based on the multivariate t distribution instead of the normal.

2. MAX-STABLE PROCESSES

A stochastic process $\{Y_t,\ t \in T\}$, where T is an arbitrary index set, is called *max-stable* if there exist constants $A_{Nt} > 0$, B_{Nt} (for $N \geq 1,\ t \in T$) with the following property: if $Y_t^{(1)}, ..., Y_t^{(N)}$ are N independent copies of the process and

$$Y_t^* = \left(\max_{1 \leq n \leq N} Y_t^{(n)} - B_{Nt} \right) / A_{Nt}, \quad t \in T,$$

then $\{Y_t^*,\ t \in T\}$ is identical in law to $\{Y_t,\ t \in T\}$.

In the case where T is a finite set, this is precisely the definition of a multivariate extreme value distribution (for maxima) and of course, if $|T| = 1$, this reduces further to the classical "three types" of Fisher and Tippett (Galambos, 1987, is one of numerous references covering classical extreme value theory). There is no loss of generality in transforming the margins to one particular extreme value distribution, and it turns out to be convenient to assume the standard Fréchet distribution

$$\Pr\{Y_t \leq y\} = e^{-1/y}, \quad \text{for all } t \tag{2.1}$$

in which case $A_{Nt} = N,\ B_{Nt} = 0$. Henceforth we assume this, though it is important to remember, with reference to any real-data application, that the models we are going to describe apply only after a pointwise transformation has already taken place.

We now consider a general method of constructing a max-stable process. Let $\{(\xi_i, s_i),\ i \geq 1\}$ denote the points of a Poisson process on $(0, \infty) \times S$ with intensity measure $\xi^{-2} d\xi \times \nu(ds)$, where S is an arbitrary measurable set and ν a positive measure on S. Let $\{f(s,t),\ s \in S,\ t \in T\}$ denote a non-negative function for which

$$\int_S f(s,t) \nu(ds) = 1, \quad \text{for all } t \in T \tag{2.2}$$

and define

$$Y_t = \max_i \{\xi_i f(s_i, t)\}, \quad t \in T. \tag{2.3}$$

The following "rainfall-storms" interpretation may help to motivate this. Think of S as a space of "storm centres", and ν as a measure which represents the distribution of storms over S. Each ξ_i represents the magnitude of a storm, and $\xi_i f(s_i, t)$ represents the amount of rainfall at position t from a storm of size ξ_i centred at s_i; the function f represents the "shape" of the storm. The max operation in (2.3) represents the notion that the observed maximum rainfall Y_t is a maximum over independent storms.

This construction is similar to one that has been employed in earlier rainfall modelling work (e.g. by Rodriguez-Iturbe, Cox and Isham 1987, 1988, Cox and Isham 1988), but both the motivation and the interpretation here, where we are concerned specifically with extremes, are different from earlier applications of this nature.

Fix $y_t > 0$ for each t, and consider the set

$$B = \{(\xi, s): \quad \xi f(s,t) > y_t \text{ for at least one } t \in T\}.$$

The event $\{Y_t \leq y_t \text{ for all } t\}$ occurs if and only if no points of the Poisson process lie in B. However, the Poisson measure of the set B is

$$\int_S \int_0^\infty I\left\{\xi > \min \frac{y_t}{(f(s,t)}\right\} \xi^{-2} d\xi \; \nu(ds) = \int_S \max_t \left\{\frac{f(s,t)}{y_t}\right\} \nu(ds)$$

where I is the indicator function, and consequently

$$\Pr\{Y_t \leq y_t \;\text{ for all } t\} = \exp\left[-\int_S \max_t \left\{\frac{f(s,t)}{y_t}\right\} \nu(ds)\right]. \tag{2.4}$$

It follows from (2.4) and (2.2) that the marginal distribution of Y_t, for any fixed t, is of the Fréchet form (2.1). Moreover, the process is max-stable: this may be verified either from (2.4), or more simply by observing

that the superposition of N independent, identical Poisson processes is itself a Poisson process with its intensity multiplied by N.

A number of results in the literature point towards a converse of this result: if Y is max-stable, then it may be represented by a process of the form (2.3). Examples of such results are Theorem 3 of de Haan (1984), theorem 4.5 of Vatan (1985) and some as yet unpublished work of Giné, Hahn and Vatan. These results provide strong motivation for considering (2.3) as a characterisation of max-stable processes.

For the reminder of this section, we show how a number of well-known multivariate extreme value families may be represented in the form (2.4). In at least one case, this suggests an extension of the family.

Example 1. Let $T = \{1, 2\},\ S = [0, 1],\ \nu$ be Lebesgue measure and let

$$f(s,t) = \begin{cases} 2(\gamma - s)/\gamma^2, & 0 < s < \gamma,\ t = 1, \\ 2(s - 1 + \gamma)/\gamma^2, & 1 - \gamma < s < 1,\ t = 2, \\ 0, & \text{otherwise}, \end{cases}$$

where $1/2 \leq \gamma \leq 1$. Then

$$\int_0^1 \max\left\{\frac{f(s,1)}{y_1}, \frac{f(s,2)}{y_2}\right\} ds = \frac{1}{y_1} + \frac{1}{y_2} - \frac{1}{y_1 + y_2}\left(2 - \frac{1}{\gamma}\right)^2$$

and hence

$$\Pr\{Y_1 \leq y_1,\ Y_2 \leq y_2\} = \exp\left\{-\left(\frac{1}{y_1} + \frac{1}{y_2}\right) A\left(\frac{y_1}{y_1 + y_2}\right)\right\} \tag{2.5}$$

where for $0 \leq w \leq 1$,

$$A(w) = 1 - \phi w(1 - w), \quad \phi = (2 - 1/\gamma)^2. \tag{2.6}$$

In the terminology of Tawn (1988) or Smith, Tawn and Yuen (1990), the function A is called the *dependence function* of the bivariate extreme value distribution function, and the specific form (2.6) is known as the *mixed*

model. Thus we have provided a derivation of a known bivariate family as a special case of a max-stable process.

Example 2. Let $T,\ S$ and ν be as in Example 1 and

$$f(s,t) = \begin{cases} (1-\alpha)s^{-\alpha}, & t=1, \\ (1-\alpha)(1-s)^{-\alpha}, & t=2, \end{cases}$$

where $0 < \alpha < 1$. Then

$$\int_0^1 \max\left\{\frac{f(s,1)}{y_1}, \frac{f(s,2)}{y_2}\right\} ds = \left(y_1^{-1/\alpha} + y_2^{-1/\alpha}\right)^{\alpha}.$$

This is the *logistic* model of bivariate extreme value theory (Tawn, 1988), for which the dependence function is $A(w) = \left\{w^{-1/\alpha} + (1-w)^{-1/\alpha}\right\}^{\alpha}$.

Example 3. An extension suggested by Example 2 is to let

$$f(s,t) = \begin{cases} (1-\alpha)s^{-\alpha}, & t=1, \\ (1-\beta)(1-s)^{-\beta}, & t=2, \end{cases}$$

where $0 < \alpha < 1,\ 0 < \beta < 1$. The resulting integral from (2.4) must be evaluated numerically but this is easily done. The point of this *bilogistic model* is that it provides an asymmetric extension of the logistic model which seems more natural than Tawn's (1988) asymmetric logistic model. See also Smith (1990) and Joe, Smith and Weissman (1991).

Example 4. Let $S = \{(s_1, ..., s_p) :\ s_i \geq 0 \text{ for all } i,\ \sum_i s_i = 1\}$ denote the unit simplex in $\Re^p$, let ν be Lebesgue measure on S, $T = \{1, 2, ..., p\}$ and

$$f(s_1, ..., s_p, i) = C s_i^{-\alpha}, \quad 0 < \alpha < 1,\ (s_1, ..., s_p) \in S,\ i \in \{1, ..., p\}$$

where C is a normalising constant which turns out to be

$$C = \prod_{i=1}^{p-1} (i - \alpha)$$

and

$$\int_S \max_{1\le i\le p} \frac{f(s_1,...,s_p,i)}{y_i} ds = \left(\sum_{i=1}^{p} y_i^{-1/\alpha}\right)^{\alpha} \tag{2.7}$$

which is Gumbel's (1960) p-dimensional logistic model (Tawn 1990a). The derivation of (2.7) is considerably more complicated than the preceding examples, and is given in full in Appendix 1.

To summarise, the max-stable viewpoint provides a new derivation of several known families of bivariate and multivariate extreme value distributions, and has in at least one case (Example 3) suggested an extension of an existing family.

3. THE GAUSSIAN AND t EXTREME VALUE PROCESSES.

Now consider a max-stable process in which $S = T = \Re^d$, ν is Lebesgue measure, and

$$f(s,t) = f_0(s-t) = (2\pi)^{-d/2}|\Sigma|^{-1}\exp\left\{-\frac{1}{2}(s-t)^T\Sigma^{-1}(s-t)\right\},$$

so that the function f (as a function of t for fixed s) is a multivariate normal density with mean s and covariance matrix Σ.

The joint distribution at two sites may be calculated (Appendix 2) in the form

$$\Pr\{Y_{t_1}\le y_1, Y_{t_2}\le y_2\} = \exp\left\{-\frac{1}{y_1}\Phi\left(\frac{a}{2}+\frac{1}{a}\log\frac{y_2}{y_1}\right) - \frac{1}{y_2}\Phi\left(\frac{a}{2}+\frac{1}{a}\log\frac{y_1}{y_2}\right)\right\} \tag{3.1}$$

where Φ is the standard normal distribution function and

$$a^2 = (t_1-t_2)^T\Sigma^{-1}(t_1-t_2). \tag{3.2}$$

Equation (3.1) represents a new family of bivariate extreme value distribution functions with dependence function

$$A(w) = (1-w)\Phi\left(\frac{a}{2} + \frac{1}{a}\log\frac{1-w}{w}\right) + w\Phi\left(\frac{a}{2} + \frac{1}{a}\log\frac{w}{1-w}\right) \qquad (3.3)$$

with the dependence parameter a representing a generalised distance between the points t_1 and t_2. The limits $a \to 0$ and $a \to \infty$ in (3.3) become the extreme cases $A(w) = \max(w, 1-w)$ and $A(w) = 1$ representing, respectively, perfect dependence and independence.

No closed-form expression has been found for the k-dimensional distributions of the process when $k > 2$. The following procedure, however, shows how they can be simulated. The method to be described is valid whenever $f(s,t)$ is of the form $f_0(s-t)$ for some f_0.

Suppose $y_1 > 0, ..., y_k > 0$ are given, corresponding to sites $t_1, ..., t_k$. Then

$$\begin{aligned}
& \int_S \max_i \left\{\frac{f_0(s-t_i)}{y_i}\right\} ds \\
= & \int_S \sum_i \frac{f_0(s-t_i)}{y_i} I\left\{\frac{f_0(s-t_i)}{y_i} > \max_{j\neq i}\frac{f_0(s-t_j)}{y_j}\right\} ds \\
= & \sum_i \int_S \frac{f_0(s)}{y_i} I\left\{\frac{f_0(s)}{y_i} > \max_{j\neq i}\frac{f_0(s-t_j+t_i)}{y_j}\right\} ds \\
= & E\left[\sum_i \frac{1}{y_i} I\left\{\frac{f_0(X)}{y_i} > \max_{j\neq i}\frac{f_0(X-t_j+t_i)}{y_j}\right\}\right] \qquad (3.4)
\end{aligned}$$

where the random variable X has density f_0. The final expression in (3.4) is in a form that can easily be evaluated by simulation.

An extension of the Gaussian extreme-value process, to allow for storm profiles with fatter tails than a Gaussian density, is to replace the Gaussian density f_0 with a multivariate t density:

$$f_0(x) = |\Sigma|^{-1/2}(\pi v)^{-d/2}\frac{\Gamma(v/2)}{\Gamma((v-d)/2)}\left(1+\frac{x^T\Sigma^{-1}x}{v}\right)^{-v/2}, \qquad (3.5)$$

valid for all $x \in \Re^d$, where Σ is again a positive definite covariance matrix and $v > d$. The resulting max-stable process will be referred to as the *t-extreme value process*. The calculation of joint distributions in this case is even harder than for the Gaussian extreme value process, but the simulation formula (3.4) is still valid, and we also have, for the joint distribution at two sites t_1 and t_2,

$$A\left(\frac{1}{2}\right) = \frac{1}{2}\left\{1 + B\left(\frac{a^2}{a^2+4v^2}; \frac{1}{2}, \frac{v-d}{2}\right)\right\}, \qquad (3.6)$$

where a is given by (3.2) and B is the incomplete beta function,

$$B(y;\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^y u^{\alpha-1}(1-u)^{\beta-1}du, \quad 0 \le y \le 1.$$

A derivation of (3.6) is given in Appendix 3. Note that, as $v \to \infty$, $A(1/2) \to \Phi(a/2)$ which is consistent with (3.3).

The dependence function (3.3) has previously been derived in a different context by Hüsler and Reiss (1989). They obtained it as a limiting form for the joint distribution of bivariate extremes from a bivariate normal distribution with correlation ρ_n varying with sample size n, in such a way that $n \to \infty$, $\rho_n \to 1$ and $(1-\rho_n)\log n \to a^2/4$. They also gave a multivariate extension.

4. APPLICATION TO SPATIAL DATA.

We now consider the application of these ideas to spatial data of the kind described in the introduction. Attention is restricted to the "traditional" approach to extreme value based on annual maxima, i.e. we assume that the data consist of annual maxima at a set of sites. Clearly there is much

scope for the development of threshold methods in this context, but this will not be considered here.

The purpose of the present section is to propose a statistical procedure for fitting the models that have been described. In many respects, it is a highly *ad hoc* procedure, and no claim to any form of optimality is made; this also is a question for further research. Our main purpose is to demonstrate that the models that have been developed are at least potentially applicable to real data. In the next section, their performance on some real data will be considered.

The proposed procedure is as follows:

Step 1. The Generalised Extreme Value distribution

$$F(x;\mu,\sigma,\xi) = \exp\left[-\left\{1-\frac{\xi(x-\mu)}{\sigma}\right\}_+^{1/\xi}\right] \tag{4.1}$$

is fitted to the data at each site, and then the probability integral transformation

$$y = \left\{1-\frac{\xi(x-\mu)}{\sigma}\right\}^{-1/\xi}$$

used to transform the data at each site so that they have a standard Fréchet distribution function ($F(y) = e^{-1/y}$). For the present study this is done separately at each site though, in the context of regional methods in hydrology, it would be sensible also to consider procedures in which some or all of the parameters μ, σ, ξ are common from site to site. Fitting is by maximum likelihood, though again there are other procedures such as the probability weighted moments method (Hosking, Wallis and Wood 1985), which could be considered.

From now on, we assume that the data at site t_i $(i = 1, ..., p)$ consist of standardised observations $\{Y_{ni},\ 1 \le n \le N\}$ with a standard Fréchet marginal distribution.

Step 2. Estimate the *extremal coefficient* between each pair of sites.

The only thing new about the concept of extremal coefficient is its name. The idea has been proposed by Tiago de Oliveira (see de Haan 1985) as an index of the extreme-value dependence between two variables, and has also appeared in various guises in the hydrology literature. The name chosen here is intended as a link between the correlation coefficient (of which it is, in some sense, an extreme values analogue) and the extremal index (Leadbetter 1983), which is a similar concept for measuring the effect of dependence in a stationary stochastic process. Direct use of the correlation coefficient in extreme value theory is less desirable because the correlation coefficient lacks the property of invariance under marginal transformations, and some of the cases we consider (including the Fréchet) do not even have finite variance.

Suppose (X_1, X_2) have a bivariate extreme value distribution with common marginal distribution function F; the extremal coefficient θ between X_1 and X_2 is defined by the relation $\Pr\{\max(X_1, X_2) \leq x\} = F^{\theta}(x)$. In terms of the dependence function A (see (2.5)), we have $\theta = 2A(1/2)$. In terms of the "equivalent number of independent sites" concept described in Section 1, it can be seen that the extremal coefficient is just that when there are only two sites. Finally, it should be noted that Buishand (1984) defined a function $\theta(x)$ by $\Pr\{\max(X_1, X_2) \leq x\} = F^{\theta(x)}(x)$; this is a constant when the joint distribution is indeed a bivariate extreme value distribution but in practice is often found not to be constant, a phenomenon also noted by Tawn (1990b, 1990c) in a number of real data series.

If the marginal distribution is unit Fréchet, then $1/X_1$ and $1/X_2$ have unit exponential distributions and $1/\max(X_1, X_2)$ has an exponential distribution with mean $1/\theta$. This suggests an obvious estimator. Returning to our present context in which we have transformed data $\{Y_{ni},\ 1 \leq n \leq N,\ 1 \leq i \leq p\}$, and noting that the maximum likelihood fitting of the marginal distributions will have ensured that $\sum_n Y_{ni}^{-1} = \sum_n Y_{nj}^{-1} = N$, the natural estimator of the extremal coefficient θ_{ij} between sites i and j is

$$\tilde{\theta}_{ij} = N / \left\{ \sum_{n=1}^{N} \min(Y_{ni}^{-1}, Y_{nj}^{-1}) \right\}. \tag{4.2}$$

These quantities are called *raw estimates* of the extremal coefficients, because they are not based on any model.

The theoretical properties of bivariate extremes show that $1 \leq \theta_{ij} \leq 2$, the extreme cases 1 and 2 corresponding respectively to complete dependence and independence. The estimator $\tilde{\theta}_{ij}$ could take any value between 1 and $+\infty$, the former only if $Y_{ni} = Y_{nj}$ for all n, but cases for which $\tilde{\theta}_{ij} > 2$ correspond to negative dependence between the sites and are comparatively rare in an extreme values context.

It should be noted that the concept of extremal coefficient has an obvious extension to more than 2 variables: the extremal coefficient of k variables $X_1, ..., X_k$, assumed multivariate extreme with common marginal distribution function F, is defined by the relation $\Pr\{\max(X_1, ..., X_k) \leq x\} = F^{\theta}(x)$. In this case the theoretical range is $1 \leq \theta \leq k$.

Step 3. Estimate the standard error of each $\tilde{\theta}_{ij}$.

This must take into account Step 1 as well as Step 2, since the marginal transformation appreciably affects the standard error. To see this, consider the extreme case in which $Y_{in} = Y_{jn}$ for all n. The marginal transformation ensures $\sum_n Y_{ni}^{-1} = \sum_n Y_{nj}^{-1} = N$ and hence $\tilde{\theta}_{ij} = 1$; in other words, the standard error is 0. Without the marginal transformation, $1/\tilde{\theta}_{ij}$ would be the mean of N independent unit exponential variates and hence have variance $1/N$.

In view of this difficulty, there is no simple approximate formula for the standard error. The most reasonable procedure would appear to be one based on resampling, i.e. jackknife or bootstrap. For the present study the jackknife estimator of standard error (Efron 1982) was used; this leads to the formula for the standard error

$$\left\{ \frac{N-1}{N} \sum_{n=1}^{N} \left(\tilde{\theta}_{ij}^{(n)} - \tilde{\theta}_{ij} \right)^2 \right\}^{1/2},$$

in which $\tilde{\theta}_{ij}^{(n)}$ is the estimate obtained from the n'th jackknife sample in which the obervations from year n are omitted. Of course, the marginal parameters must be re-estimated for each jackknife sample.

Step 4. Fit a max-stable model to the $\tilde{\theta}_{ij}$'s.

The algorithm proposed here is a sum of squares of weighted residuals: defining

$$r_{ij} = (\tilde{\theta}_{ij} - \hat{\theta}_{ij})/s_{ij},$$

where $\hat{\theta}_{ij}$ is the fitted value from a model, $\tilde{\theta}_{ij}$ is the raw estimate calculated in Step 2, and s_{ij} the standard error of Step 3, the model parameters are chosen to minimise $\sum_{i,j} r_{ij}^2$. The weighted sum of squares criterion was proposed by S.J. Neil, as part of an M.Phil. thesis at the University of Surrey, after an earlier attempt based on unweighted least squares was found unsatisfactory in regions of high dependence, when $\tilde{\theta}_{ij}$ is close to 1. Nevertheless, it is still very much an *ad hoc* criterion; it is an open question whether it is possible to find some means of approximating the likelihood function.

Step 5. Test the fit of the model.

One method of testing fit is the obvious one of examining the weighted residuals of Step 4 for outlying values, or for any systematic evidence of non-random behaviour (plotting residuals against fitted values or against other variables such as the distances between the points). A second idea, noting that the estimation algorithm is based entirely on pairwise dependences, is to test the fit on higher-order dependences. This can be done by calculating the higher-order extremal coefficients mentioned at the end of Step 2, with

their standard errors, and comparing them with the fitted values based on the estimated model.

5. AN EXAMPLE

Data were supplied by the Institute of Hydrology in the form of 21 years' annual maxima of daily rainfall measurements at a network of 405 rain gauges in South-East England. For the purpose of the present study, it is not feasible to fit the proposed models to anything like 405 sites simultaneously, and two subsets of 10 sites, one a group of close neighbours (on the Isle of Wight) and the other chosen so as to be scattered throughout the network, have been used. The calculations have been repeated for a number of other subsets of similar numbers of sites, with similar results to those reported here. Figure 1 shows the grid points for our two subsets, labelled 0,1,2...,9 for subset 1 and A,B,...,J for subset 2; units of measurement are tenths of a kilometre.

We first consider subset 1, which cover a total area approximately 20 km. by 15 km. There are 45 raw extremal coefficients calculated from all pairs of points, and these are plotted against the distance between the points in Figure 2. All the raw extremal coefficients lie between 1.1 and 1.9 and, as is to be expected, they tend to increase with distance.

The Gaussian model is defined by the covariance matrix Σ; in the present (2×2) case this is most conveniently parametrised by writing $\Sigma^{-1} = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}$. The model was fitted to the raw extremal coefficients, producing estimates (with standard errors in parentheses) of $\hat{\alpha} = 1.7 \times 10^{-4}(0.3 \times 10^{-4})$, $\hat{\beta} = -0.5 \times 10^{-4}(0.3 \times 10^{-4})$, $\hat{\gamma} = 1.4 \times 10^{-4}(0.3 \times 10^{-4})$. The fitted values $\hat{\theta}_{ij}$ are plotted against distance in Figure 3; the figure gives an indication of how much the variability of the extremal coefficients has been reduced by the model fit. The residuals defined in Step 4 of Section 4 have been plotted against a variety of related variables, including both the fitted values

of extremal coefficient and a variable indicating the direction from one site to the other, but the only plot to show any interesting feature is Figure 5, which is a plot of residual against distance. In this figure it can be seen that there is a tendency for residual to decrease with distance. A negative residual indicates that the fitted model exhibits less dependence than the raw data, so the plot suggests that the model leads to too much dependence at short distances and too little at larger distances. This feature was also observed in several other data sets tried. A possible explanation is that the assumed f_0 function is too short-tailed, and it was in fact this feature that motivated the introduction of the t model in Section 3, as an alternative to the Gaussian model.

Consequently, a t-extreme value process was also fitted to the data, resulting in $\hat{v}$=2.5 and a reduction in $\sum r_{ij}^2$ from 15.2 under the Gaussian model to 14.0 under the t model. In this case we find $\hat{\alpha} = 5.9 \times 10^{-2}(2.1 \times 10^{-2})$, $\hat{\beta} = -1.7 \times 10^{-2}(1.1 \times 10^{-2})$, $\hat{\gamma} = 2.4 \times 10^{-2}(0.8 \times 10^{-2})$; the values are not at all comparable with those in the Gaussian case but this is presumably explained by the fact that the t density with 2.5 degrees of freedom is quite different from a Gaussian density. A plot of fitted extremal coefficient against distance is in Figure 4, and of residuals against distance is in Figure 6. It would appear that the t model is a significant improvement in fit over the Gaussian model, and that the trend apparent in Figure 5 has largely disappeared in Figure 6.

We now consider the higher-order extremal coefficients. The fitted values here used the simulation formula (3.4). In the case of the t model, this required sumulating from the density (3.5) - this is discussed in Appendix 4. Three subsets of the sites were considered:

(a) Sites 0,1,2,3,9, raw extremal coefficient 2.1 (standard error 0.3),

(b) Sites 4,5,6,7,8, raw extremal coefficient 1.7 (standard error 0.2),

(c) All 10 sites, raw extremal coefficient 2.7 (standard error 0.5).

The fitted values, estimated using the simulation technique based on

(3.4), were respectively 1.77, 1.62, 2.54 under the Gaussian model, 2.15, 1.93, 3.48 under the t model. In this case the comparison between the two models is not clear-cut, though either seems to be performing adequately when the standard errors are taken into account.

We also consider the second subset of 10 sites depicted in Figure 1. In this case the region is much larger (200 by 140 km.) and we might expect the extremal coefficients to be larger, and this is in fact the case as shown in Figure 7. In fact we now have several raw extremal coefficients greater than 2 though their individual standard errors are such that this does not by itself contradict the model described in this paper.

Again the Gaussian and t models were both fitted to the data, the resulting values of $\sum r_{ij}^2$ being respectively 22.1 and 16.4, with $\hat{v} = 2.62$ under the t model. The parameter estimates are $\hat{\alpha} = 1.4 \times 10^{-4}(0.9 \times 10^{-4})$, $\hat{\beta} = -0.7 \times 10^{-4}(0.5 \times 10^{-4})$, $\hat{\gamma} = 0.5 \times 10^{-4}(0.2 \times 10^{-4})$ under the Gaussian model, and $\hat{\alpha} = 1.3 \times 10^{-2}(1.1 \times 10^{-2})$, $\hat{\beta} = -1.1 \times 10^{-2}(1.1 \times 10^{-2})$, $\hat{\gamma} = 1.3 \times 10^{-2}(1.2 \times 10^{-2})$ under the t model. In this case the standard errors, which were large enough to cause concern in the earlier example, create serious doubts over the credibilty of the parameter estimates. No single explanation for this phenomenon stands out, though it seems likely that the indirect method of estimation (essentially, an attempt to estimate storm profiles from annual maxima data) is responsible, and that a direct method of fitting through daily rainfall data would be superior. Note that, in both the cases we have considered, under the t model the estimated $\hat{v}$ is far from ∞ and in fact quite close to its limiting value 2, when the density ceases to be defined. Plots of residuals against distance show similar features to those of Figure 4 and 6 (less markedly), but the most interesting plots in this case are those of fitted versus raw values of extremal coefficient (Figures 8 and 9). Under the Gaussian model, many fitted values are almost exactly 2, corresponding to independence, whereas this is not the case under the t model. The fact that the t model retains some dependence over the larger

distances in the sample might be thought an additional reason for preferring the t model, though the argument for this is not clear-cut.

Once again, the models were tested by assessing their ability to reproduce higher-order extremal coefficients:

(a) Sites 1,2,4,5,6, raw extremal coefficient 3.3 (standard error 0.6),

(b) Sites 0,3,7,8,9, raw extremal coefficient 3.5 (standard error 0.8),

(c) All 10 sites, raw extremal coefficient 5.7 (standard error 1.3).

The fitted values are 3.40, 4.78 and 7.09 under the Gaussian model, 3.35, 4.05 and 6.69 under the t model. In this case the t model seems to be performing the better of the two, though only slightly, and both models are underestimating the true dependence to some extent.

It would of course be unwise to draw firm conclusions on the basis of just two comparatively small examples, but the two examples do seem to indicate that both models can be adequately fitted to the data, and give sensible conclusions, with a preference for the t model though it is not completely clear-cut.

6. CONCLUSIONS.

The main purpose of this study has been to present some of the properties of max-stable models. There remain many questions about their statistical application, including the choice of an appropriate model class, the possibility of improved estimation algorithms or diagnostic tools, and (perhaps most importantly) the extension from an annual maximum to a threshold approach.

We conclude by mentioning another aspect of the models which points to a potentially much greater range of applicability. The Gaussian model (or the t model for fixed v) is defined by a covariance matrix Σ so that the dependence between two sites t_i and t_j depends on $(t_i - t_j)^T \Sigma^{-1} (t_i - t_j)$. However, there is another way to think about this as a standardised model in which

Σ is the 2×2 identity matrix, applied to transformed sites $t_i^* = \Sigma^{-1/2} t_i$. In other words, by applying a *linear* transformation to the positions of the measurement sites, we reduce the max-stable process to a single completely defined process. This is reminiscent of some recent work on estimating spatial correlations. For example, two discussion contributions by Lewis (1989) and by Guttorp and Sampson (1989) focussed on the possibility of improving a spatial model by "moving" one of the sites, and in further as yet unpublished work, P. Guttorp and P. Sampson have proposed a systematic scheme of this form, a form of multidimensional scaling, in which they used bivariate splines to transform the site positions. Clearly there are possibilities of this nature in the context of the work described here - instead of assuming a single linear transformation to all the sites, we could apply different linear transformations to different parts of space. This idea could provide a basis for extending the scheme from small numbers of sites, as have been considered here, to a much more general scheme involving large number of sites.

ACKNOWLEDGEMENTS

Additional statistical results on max-stable processes, some of which motivated some of the ideas developed here, were obtained by Mr. Stuart Neil as part of an MPhil thesis at the University of Surrey. Some further computational work was done by Mr. Philip Summers as part of a final year project at Surrey. I am grateful to both for their contributions. The data were supplied by the Institute of Hydrology. The work was supported by the Wolfson Foundation in the form of a personal research fellowship, and used equipment obtained as part of an SERC grant.

REFERENCES.

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

Coles, S.G. and Tawn, J.A. (1990), Statistics of coastal flood prevention. *Phil. Trans. R. Soc. Lond. A* **332**, 457-476.

Coles, S.G. and Tawn, J.A. (1991), Modelling multivariate extreme events. *J.R. Statist. Soc. B* **53**, No. 1.

Cox, D.R. and Isham, V. (1988), A simple spatial-temporal model of rainfall. *Proc. R. Soc. Lond. A* **415**, 317-328.

Davison, A.C. and Smith, R.L. (1990), Models for exceedances over high thresholds (with discussion). *J.R. Statist. Soc. B* **52**, 393- 442.

Efron, B. (1982), *The Jackknife, the Bootstrap and Other Resampling Plans.* CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia.

Galambos, J. (1987), *The Asymptotic Theory of Extreme Order Statistics* (2nd. edn.). Krieger, Melbourne, Fl. (First edn. published 1978 by John Wiley, New York.)

Gumbel, E.J. (1958), *Statistics of Extremes.* Columbia University Press, New York.

Gumbel, E.J. (1960), Distributions des valeurs extremes en plusiers dimensions. *Publ. Inst. Statist. Univ. Paris* **9**, 171-173.

Guttorp, P. and Sampson, P. (1989), Discussion of paper by Haslett and Raftery. *Applied Statistics* **38**, 32-33.

Haan, L. de (1984), A spectral representation for max-stable processes. *Ann. Probab.* **12**, 1194-1204.

Haan, L. de (1985), Extremes in higher dimensions: the model and some statistics. *Proc. 45th Session I.S.I., Amsterdam*, Paper 26.3.

Haan, L. de and Pickands, J. (1986), Stationary min-stable stochastic processes. *Prob. Th. Rel. Fields* **72**, 477-492.

Hosking, J.R.M., and Wallis, J.R. (1988), The effect of intersite dependence on regional flood frequency analysis. *Wat. Res. Res.* **24**, 588-600.

Hosking, J.R.M., Wallis, J.R. and Wood, E.F. (1985), Estimation of the generalised extreme-value distribution by the method of probability-weighted moments. *Technometrics* **27**, 251-261.

Hüsler, J. and Reiss, R.-D. (1989), Maxima of normal random vectors: Between independence and complete dependence. *Statistics and Probability Letters* **7**, 283-286.

Joe, H., Smith, R.L. and Weissman, I. (1991), Bivariate threshold models for extremes. *J.R. Statist. Soc. B*, to appear.

Leadbetter, M.R. (1983), Extremes and local dependence in stationary sequences. *Z. Wahrsch. v. Geb.* **65**, 291-306.

Lewis, T. (1989), Discussion of paper by Haslett and Raftery. *Applied Statistics* **38**, 29.

N.E.R.C. (1975), *The Flood Studies Report.* The Natural Environment Research Council, London.

Reed, D.W. and Dales, M.Y. (1988), Regional rainfall risk: a study of spatial dependence. Proceedings of IAHR Symposium.

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

Sibuya, M. (1960), Bivariate extreme statistics. *Ann. Inst. Stat. Math.* **11**, 195-210.

Smith, R.L. (1989), Extreme value analysis of environmental time series: an example based on ozone data (with discussion). *Statistical Science* **4**, 367-393.

Smith, R.L. (1990), Extreme value theory. In *Handbook of Applicable Mathematics* **7**, ed. W. Ledermann. John Wiley, Chichester, 437-471.

Smith, R.L., Tawn, J.A. and Yuen, H.-K. (1990), Statistics of multivariate extremes. *Int. Statist. Rev.* **58**, 47-58.

Stedinger, J.R. (1983), Estimating a regional flood frequency distribu-

tion. *Wat. Res. Res.* **19**, 503-510.

Tawn, J.A. (1988), Bivariate extreme value theory - models and estimation. *Biometrika* **75**, 397-415.

Tawn, J.A. (1990a), Modelling multivariate extreme value distributions. *Biometrika* **77**, 245-253.

Tawn, J.A. (1990b), Discussion of paper by Davison and Smith. *J.R. Statist. Soc. B* **52**, 428-429.

Tawn, J.A. (1990c), Estimating probabilities of extreme sea levels. Submitted.

Vatan, Pirooz (1985), Max-infinite divisibility and max-stability in infinite dimensions. In *Probability in Banach Spaces V*, ed. A. Beck et al. *Lecture Notes in Mathematics* **1153**, Springer, Berlin, 400-425.

APPENDIX 1: PROOF OF (2.7).

We require to show

$$\int_{S_p} \max_i \frac{s_i^{-\alpha}}{y_i} ds = \prod_{i=1}^{p-1} (i-\alpha)^{-1} \left(\sum_{i=1}^{p} y_i^{-1/\alpha} \right)^{\alpha} \qquad (A1.1)$$

where S_p is the unit simplex in $\Re^p$, ds is Lebesgue measure on S_p, $y_i > 0$ for all i and $0 < \alpha < 1$.

Let μ be a measure on $[0,\infty]^p$ such that the set $\{(x_1,...,x_p) \in \Re^p : \ 0 \leq x_i \leq z_i$ for at least one $i\}$ has μ-measure $(\sum z_i^{1/\alpha})^{\alpha} = H(z_1,...,z_p)$ say. The corresponding density is

$$\begin{aligned} h(z_1,...,z_p) &= (-1)^{p-1} \frac{\partial^p}{\partial z_1 ... \partial z_p} H(z_1,...,z_p) \\ &= \prod_{i=1}^{p-1} \frac{i-\alpha}{\alpha} \prod_{i=1}^{p} z_i^{1/\alpha - 1} \left(\sum_{i=1}^{p} z_i^{1/\alpha} \right)^{\alpha - p} . \end{aligned} \qquad (A1.2)$$

Now make the transformation

$$z_i = r s_i^{\alpha} \ (1 \leq i \leq p-1), \quad z_p = r(1 - s_1 - ... - s_{p-1})^{\alpha}$$

where $r = (\sum z_i^{1/\alpha})^{\alpha}$. For convenience of notation, let $s_p = 1 - s_1 - ... - s_{p-1}$. Let $J_1(r, s_1, ..., s_p)$ denote the Jacobian of the transformation from $(z_1, ..., z_p)$ to $(r, s_1, ..., s_{p-1})$. By expanding the first row or column of the determinant it can be seen that J_p satisfies the recurrence relation

$$J_p(r, s_1, ..., s_p) = \alpha r s_1^{\alpha-1} J_{p-1}(r, s_2, ..., s_p) + s_1^{\alpha} (\alpha r)^{p-1} \prod_{i=2}^{p} s_i^{\alpha-1}$$

from which we deduce that

$$J_p(r, s_1, ..., s_p) = (\alpha r)^{p-1} \left(\prod_{i=1}^{p} s_i^{\alpha-1} \right) \left(\sum_{i=1}^{p} s_i \right). \qquad (A1.3)$$

Combining $(A1.2)$ and $(A1.3)$, the density of μ in $(r, s_1, ..., s_{p-1})$ co-ordinates turns out to be the constant $\prod_{i=1}^{p-1}(i - \alpha)$.

To complete the calculation, we write

$$\begin{aligned} (\sum z_i^{1/\alpha})^{\alpha} &= \mu\{r \leq \max_i z_i s_i^{-\alpha}\} \\ &= \prod_{i=1}^{p-1}(i-\alpha) \int_{S_p} \int_0^{\infty} I(r \leq \max_i \ z_i s_i^{-\alpha}) dr ds \\ &= \prod_{i=1}^{p-1}(i-\alpha) \int_{S_p} \max_i (z_i s_i^{-\alpha}) ds \end{aligned}$$

which is equivalent to $(A1.1)$, as required.

APPENDIX 2: PROOF OF (3.1).

We use (3.4). It can be seen that $f_0(X)/y_1 > f_0(X - t_1 + t_1)/y_2$ if and only if

$$X^T \Sigma^{-1} X + 2 \log y_1 < (X - t_2 + t_1)^T \Sigma^{-1} (X - t_2 + t_1) + 2 \log y_2$$

and this occurs if and only if

$$X^T\Sigma^{-1}(t_1 - t_2) > \log(y_1/y_2) - (t_1 - t_2)^T\Sigma^{-1}(t_1 - t_2)/2. \qquad (A2.1)$$

However, $X^T\Sigma^{-1}(t_1 - t_2)$ is a random variable with mean 0, variance $(t_1 - t_2)^T\Sigma^{-1}(t_1 - t_2) = a^2$ and hence the probability of the event $(A2.1)$ is $\Phi(a/2 + \log(y_1/y_2)/a)$. Combining this with the similar expression with the indices 1,2 interchanged, leads to the result.

APPENDIX 3: PROOF OF (3.6).

Again using (3.4), the task is to show that

$$\Pr\{f_0(X) > f_0(X - t_2 + t_1)\} = \frac{1}{2}\left\{1 + B\left(\frac{a^2}{a^2 + 4v^2}; \frac{1}{2}, \frac{v-d}{2}\right)\right\} \quad (A3.1)$$

when X has density f_0 defined by (3.5).

First, note that $f_0(X) > f_0(X - t_2 + t_1)$ if and only if $X^T\Sigma^{-1}(t_1 - t_2) > -a^2/2$. Writing $X = \Sigma^{1/2}Y$, the density of Y, say g_0, is given by

$$g_0(y) = (\pi v)^{-d/2}\frac{\Gamma(v/2)}{\Gamma((v-d)/2)}\left(1 + \frac{y^T y}{v}\right)^{-v/2}. \qquad (A3.2)$$

By making a further orthogonal transformation of Y, it can be seen that the desired probability is the same as

$$\Pr\{Y_1 > -a/2\} \qquad (A3.3)$$

when $Y = (Y_1, ..., Y_p)$ has the density g_0. Thus we concentrate on $(A3.3)$.

Consider the integral

$$J = \int_{y:|y_1|>b}\left(1 + \frac{y^T y}{v}\right)^{-v/2} dy.$$

We may write

$$\begin{aligned} J &= \lim_{\sigma\to\infty} \int_{y:|y_1|>b} \exp\left(-\frac{y^T y}{2\sigma^2}\right)\left(1+\frac{y^T y}{v}\right)^{-v/2} dy \\ &= \lim_{\sigma\to\infty} (2\pi\sigma^2)^{d/2} E\left\{\left(1+\frac{Y^T Y}{v}\right)^{-v/2} I(|Y_1|>b)\right\} \end{aligned}$$

where Y is d-variate normal with mean 0 and covariance σ^2 times the identity matrix.

Writing $Y^T Y/\sigma^2 = S + T$ with $S,\ T$ independent respectively χ^2_1 and χ^2_{d-1}, where $S = Y_1^T Y_1/\sigma^2$, we have

$$\begin{aligned} J &= \lim_{\sigma\to\infty} (2\pi\sigma^2)^{d/2} \int_{b^2/\sigma^2}^{\infty}\int_0^{\infty} \left\{1+\frac{(s+t)\sigma^2}{v}\right\}^{-v/2} \cdot \\ &\quad \cdot \frac{1}{2^{d/2}\Gamma((d-1)/2)\Gamma(1/2)} s^{-1/2} e^{-s/2} t^{(d-3)/2} e^{-t/2} dt ds \end{aligned}$$

which on substituting $s = uv/\sigma^2$, $t = wv/\sigma^2$ and taking the limit, reduces to

$$J = \frac{(\pi v)^{d/2}}{\Gamma((d-1)/2)\Gamma(1/2)} \int_{b^2/v}^{\infty}\int_0^{\infty} u^{-1/2} w^{(d-3)/2} (1+u+w)^{-v/2} dw du.$$

Writing this integral in the form

$$J = \frac{(\pi v)^{d/2}}{\Gamma((d-1)/2)\Gamma(1/2)} \int_c^{\infty} u^{-1/2} \int_u^{\infty} (w-u)^{(d-3)/2} (1+w)^{-v/2} dw du$$

with $c = b^2/v$, and making the substitution $w = (1+u)/x - 1$, $(0 \le x \le 1)$ the inner integral reduces to a standard beta integral from which we deduce

$$J = \frac{(\pi v)^{d/2}\Gamma((v-d+1)/2)}{\Gamma(v/2)\Gamma(1/2)} \int_c^{\infty} u^{-1/2} (1+u)^{(d-v-1)/2} du.$$

The further substitution $u = 1/(1-y) - 1$ $(c/(1+c) \leq y \leq 1)$ then reduces this to

$$J = \frac{(\pi v)^{d/2}\Gamma((v-d)/2)}{\Gamma(v/2)} \left\{1 - B\left(\frac{c}{1+c}; \frac{1}{2}, \frac{v-d}{2}\right)\right\}.$$

Writing $b = a/2$, we now see that

$$\Pr\{|Y_1| > a/2\} = \left\{1 - B\left(\frac{a^2}{a^2+4v}; \frac{1}{2}, \frac{v-d}{2}\right)\right\}$$

and the symmetry of the distribution of Y_1 then allows us to deduce $(A3.1)$ from $(A3.3)$.

APPENDIX 4: SIMULATING FROM THE MULTIVARIATE t DISTRIBUTION.

The density from which we wish to simulate is (3.5), but any transformation of the form $x = Ay$, where $AA^T = \Sigma$, reduces this to the density g_0 given in (A3.2). Thus we concentrate on the latter. Note that, for this purpose, it is not necessary that A be symmetric; in the 2×2 case, with say $\Sigma = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$, the most convenient choice is $A = \begin{pmatrix} a^{1/2} & 0 \\ ba^{-1/2} & (c - b^2/a)^{1/2} \end{pmatrix}$.

Now suppose $Y = (Y_1, ..., Y_d)$ is desired from (A3.2). The density is radially symmetric, so we may write $Y = (RW_1, ..., RW_d)$ where $R^2 = Y_1^2 + ... + Y_d^2$ and $(W_1, ..., W_d)$ are uniformly distributed over the unit sphere. The easiest way to simulate the W's is to generate $Z_1, ..., Z_d$ standard normal, and then set $W_i = Z_i/(\sum Z_j^2)^{1/2}$. For R, note that the marginal density at $R = r$ is proportional to $(1 + r^2/v)^{-v/2} r^{d-1}$ $(0 < r < \infty)$ which is the same thing as saying $R = \{vS/(1-S)\}^{1/2}$ where S has a beta distribution on (0,1) with parameters $d/2$ and $(v-d)/2$. When $d = 2$ this simplifies further and we may write $S = 1 - U^{2/(v-2)}$ where U is uniform on (0,1).

Figure Captions

Fig. 1. Plot of study region (units of 0.1km). Points 0–9 are subset 1 and A–J are subset 2.

Fig. 2. Plot of raw extremal coefficient against distance, subset 1.

Fig. 3. Plot of fitted extremal coefficient against distance, Gaussian model, subset 1.

Fig. 4. Plot of fitted extremal coefficient against distance, t model, subset 1.

Fig. 5. Residuals against distance, Gaussian model, subset 1.

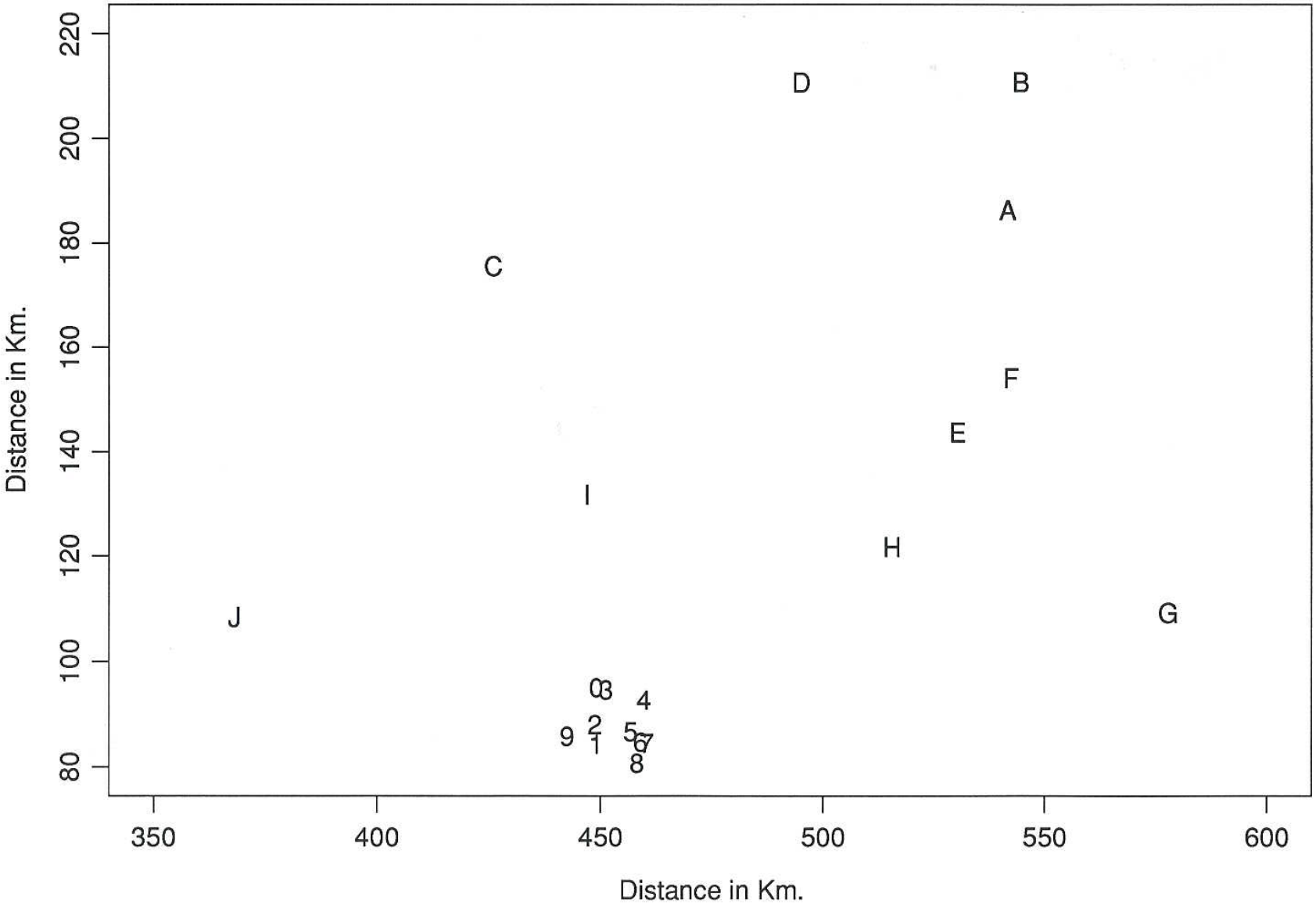
Fig. 6. Residuals against distance, t model, subset 1.

Fig. 7. Plot of raw extremal coefficient against distance, subset 2.

Fig. 8. Fitted v. raw values of extremal coefficient, Gaussian model, subset 2.

Fig. 9. Fitted v. raw values of extremal coefficient, t model, subset 2.

Figure 1: Site Positions



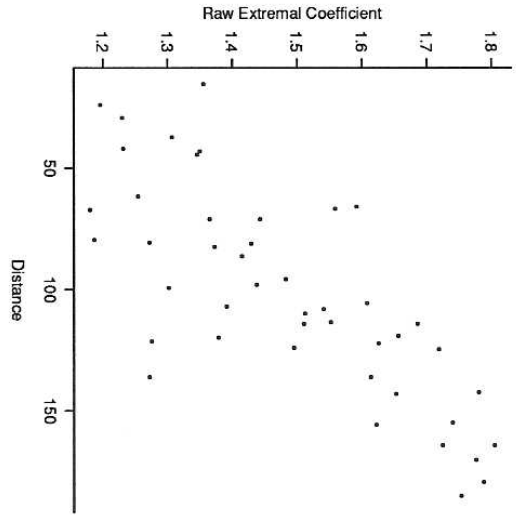


Figure 2

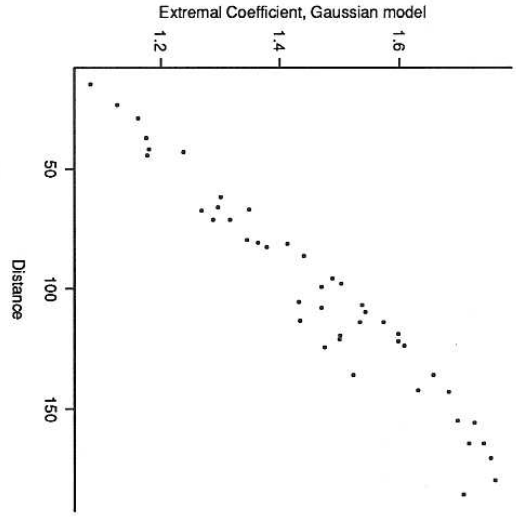


Figure 3

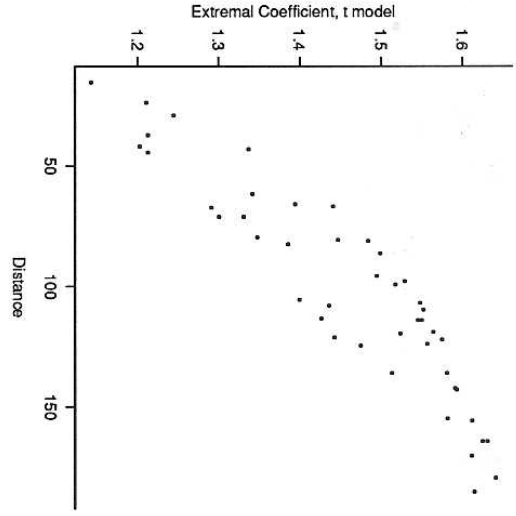


Figure 4

Figure 5

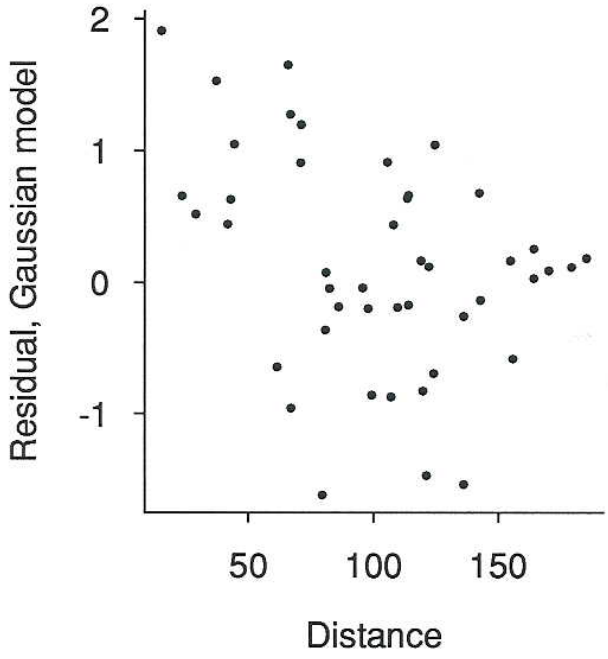
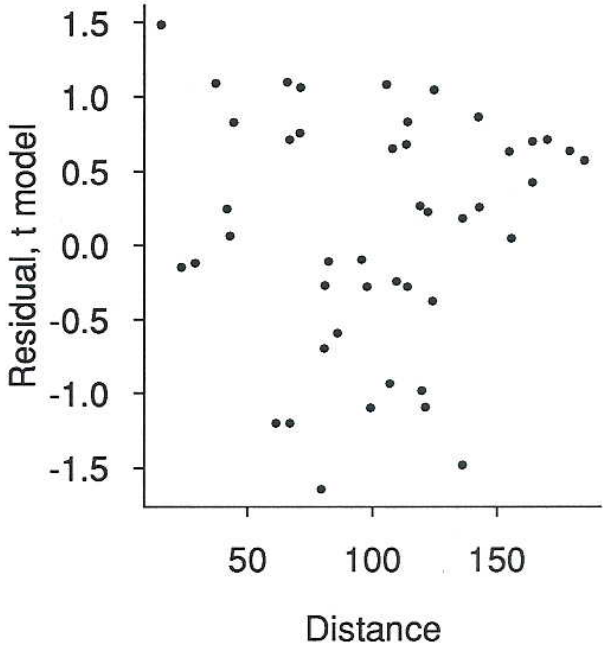


Figure 6



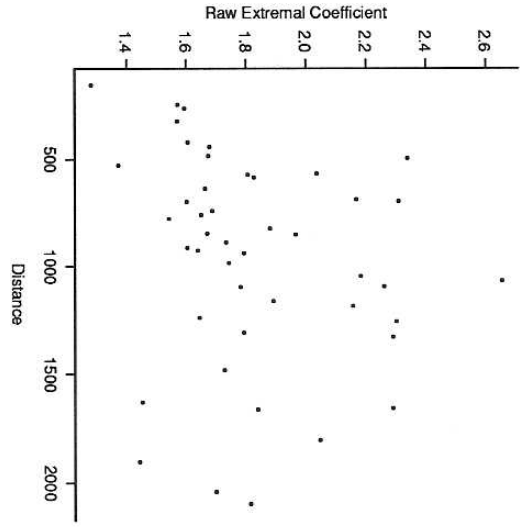


Figure 7

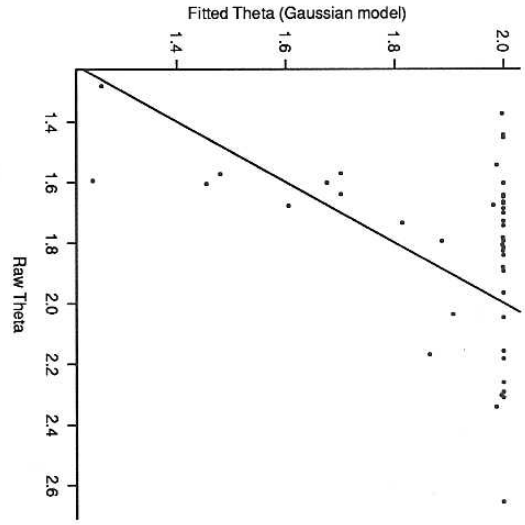


Figure 8

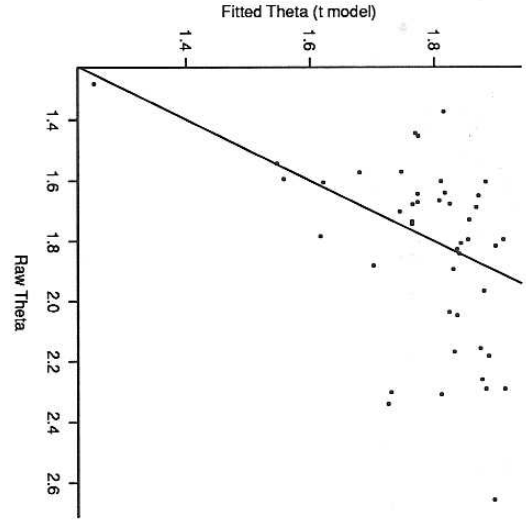


Figure 9