CHALMERS | GÖTEBORG UNIVERSITY

PREPRINT 2005:22

Mixture Models for Extremes

A.-L. FOUGÈRES
J. P. NOLAN
H. ROOTZÉN

Department of Mathematical Sciences
Division of Mathematical Statistics
CHALMERS UNIVERSITY OF TECHNOLOGY
GÖTEBORG UNIVERSITY
Göteborg Sweden 2005

Preprint 2005:22

Mixture Models for Extremes

A.-L. Fougères, J. P. Nolan, H. Rootzén

CHALMERS | GÖTEBORG UNIVERSITY

Department of Mathematical Sciences
Division of Mathematical Statistics
Chalmers University of Technology and Göteborg University
SE-412 96 Göteborg, Sweden

Göteborg, June 2005

Preprint 2005:22
ISSN 1652-9715

Matematiskt centrum
Göteborg 2005

MIXTURE MODELS FOR EXTREMES

A.-L. FOUGÈRES, J. P. NOLAN, AND H. ROOTZÉN

Abstract. This paper develops "logistic" multivariate extreme value distributions obtained by mixing with positive stable distributions. The mixing variables are used as a modelling tool, to introduce new classes of mixture models for extreme value data, and as a road to better understanding and use of the models. A distinguishing feature is that the models lead to extreme value distributions both conditionally on the mixing variables and unconditionally. In many situations this is important for easy interpretation and extrapolation and it is natural for data obtained as maxima of other underlying variables. One way of understanding the models is as block-size mixtures of extreme value distributions, where the mixing is by positive stable distributions. For Gumbel distributions a second interpretation is as exponential-stable location mixtures of independent Gumbel distributions with the same scale parameter. The corresponding interpretation for non-Gumbel EV distributions is as power-stable scale mixtures of independent EV distributions. A third interpretation is through a Peaks over Thresholds model with random intensity. We develop analogues of the components of variance models in ANOVA, and new time series, spatial, and continuous parameter models for extreme value data. The results are applied to data from a pitting corrosion investigation and to an interest rate time series. The models present many challenging problems of interpretation and use, on numerical methods for estimation, and on asymptotic analysis.

1. Introduction

Multivariate models for extreme value data is attracting substantial interest, see e.g. Kotz and Nadarajah (2000) and Fougères (2004). However, with the exception of Smith (2004) and Heffernan and Tawn (2004), few applications involving more than two or three dimensions have been reported. One main application area is environmental extremes. Dependence between extreme wind speeds and rain fall can be important for reservoir safety (Anderson and Nadarajah (1993), Ledford and Tawn (1996)), high mean water levels occurring together with extreme waves may cause flooding (Bruun and Tawn (1998), de Haan and de Ronde (1998)), and simultaneous high water levels at different spatial locations pose risks for large floods (Coles and Tawn (1991)). Another set of applications is in economics where multivariate extreme value theory has been used to model the risk that extreme fluctuations of several exchange rates or of prices of several assets, such as stocks, occur together (Mikosch (2004), Smith (2004), Stărică (1999)). A third use, perhaps somewhat unlikely, is in the theory of rational choice (McFadden (1978)). Below we will also consider a fourth problem, analysis of pitting corrosion measurements.

1991 *Mathematics Subject Classification.* Primary 62P30, secondary 62G32.

Key words and phrases. logistic distribution, multivariate extreme value distribution, random effect, max-stable, pitting corrosion.

Research supported by VCC/Ford and by the Knut and Alice Wallerberg foundation.

The most popular choice of multivariate extreme value distributions are various families of "logistic" distributions. The first such family was described early by Gumbel (1960). The most general versions, termed the asymmetric logistic distribution and the nested logistic distribution were introduced by Tawn (1990) and McFadden (1978) and further studied in Coles and Tawn (1991). Tawn built on earlier results in survival analysis (Hougaard (1986), Crowder (1989)), and obtained the distributions assuming independence conditionally on a positive stable variable. He also gave a physical motivation for the models (see below).

In this paper we attempt to use the stable mixing variables as a modelling tool. This allows new understanding, physical motivation, and use of the models. In particular, the starting point for our work was a need for analogues of the random effects model in analysis of variance, for use in analysis of pit corrosion measurements. One of our results is that the symmetric logistic distribution indeed provides such a model. Further insights obtained from this result are better understanding of identifiability of parameters, and new model checking tools.

By suitable choices of the mixing variables we also obtain natural time series models, spatial models, and continuous parameter models for extreme value data. The first two of these are subfamilies of the general asymmetric logistic distribution, but their special significance doesn't seem to have been realized before. In particular, these models provide conceptually, analytically and computationally tractable models for extreme value data which go beyond dimensions two and three.

It is not immediately obvious from the form of the multivariate distributions how to simulate values from them, see e.g. Kotz and Nadarajah (1999, Section 3.7). However the representation as stable mixtures makes simulation straightforward. According to it, one can first simulate the stable variables, using the method of Chambers *et al.* (1976), and then simulate independent variables from the conditional distribution given the stable variables, cf. Stephenson (2003). This adds substantially to the usefulness of the models.

The results can be presented in two very closely related ways, as mixture models for Gumbel distributions, and as mixture models for the generalized Extreme Value (EV) distribution. The Gumbel models are more parsimonious, and we first present the results in this setting. An additional reason is that we think that the Gumbel distribution has a special importance in extreme value theory, for several reasons. One is that it occurs as the limit of maxima of most standard distributions, specifically so for the normal distribution. In fact, it is the only possible limit for the entire range of tail behavior between polynomial decrease and (essentially) a finite endpoint. Another reason is the (approximate) lack of memory property of the locally exponential tails of the underlying variables which goes together with the Gumbel distribution for maxima. Finally, from experience, the Gumbel distribution is known to fit well in many situations. In particular this is the case for most pit corrosion measurements, see Kowaka (1994).

However, we also develop the models in the EV setting. In it, two out of three physical motivations for the model, as "block size mixtures" and as maxima in a Peaks over Thresholds (PoT) model with a doubly stochastic Poisson number of large values are the same as for the Gumbel model. The counterpart to the remaining Gumbel interpretation, as a location parameter mixture, is that the multivariate EV distributions are obtained as scale mixtures with an accompanying location change which keeps the endpoints of the distributions fixed.

In the models both conditional and unconditional distributions are Gumbel (or EV) and maxima of all kinds, e.g. over a number of "groups" with differing numbers of elements, are also Gumbel (or EV). This is natural in settings where all observations are obtained as maxima of some underlying variables and it also leads to substantial economies of understanding, analysis and prediction.

The basic motivation and explanation of the models for the Gumbel case is given in Section 2 below. In Section 3 we rederive and remotivate the asymmetric and nested logistic multivariate Gumbel distributions and introduce new classes of multivariate Gumbel models for time series, spatial, and continuous parameter applications. Properties of the exponential-stable mixing distributions are given in Section 4. In Section 5 we first discuss estimation in the random effects model and in a hidden MA(1) model. These models are then used to analyze two data sets. The first one is from an investigation of risks of penetration by pitting corrosion on the lower hemflange of a car door. The second one concerns a time series of extreme interest rate fluctuations. The section also uses new model checking tools.

Section 6 translates the Gumbel results and models from Sections 2, 3, and 5 to the general EV family. Section 7 contains a small concluding discussion.

2. Mixtures of Gumbel distributions

We will motivate our models by two physical situations. The first one is a standard type of pitting corrosion measurement. In it a number of test specimens of a metal are divided up into subareas, here called test areas, and the deepest corrosion pit in each of the test areas is measured. The presumption is that there may be an extra variation between specimens which is not present between test areas from the same specimen. In Section 5 below we analyze such an experiment. In this experiment the extra variation was caused by randomness in the proportion of the surface which was covered by corrosion-preventing glue or coating.

The second situation is from Tawn (1990) and concerns maximum windspeeds at different locations. Tawn assumes that the yearly maximum at a location is obtained as the largest of the maximum wind speeds in the individual storms which affected the location that year. The highest wind speed in individual storms are assumed to be i.i.d, but the number of storms varies randomly between years. There is an unavoidable Poisson variation in this number. Tawn assumes there is an additional variation between "stormy" and "stormfree" years, and that this may affect nearby locations in a similar way.

In the present section we introduce the ideas in the one-dimensional case. The physical motivations, however, extend directly to the multivariate models which are the main interest of this paper, and are treated in the subsequent sections.

The mathematical basis is the following observation. Let S be a standard positive α-stable variable, specified by its Laplace transform

$$E(e^{-tS}) = e^{-t^{\alpha}}, \quad t \geq 0, \tag{2.1}$$

where necessarily $\alpha \in (0, 1]$. (When $\alpha = 1$, S is taken to be identically 1, see the discussion in Section 4.) Further, let the random variable X be Gumbel distributed conditionally on S,

$$P(X \le x|S) = \exp(-Se^{-\frac{x-\mu}{\sigma}}) = \exp(-e^{-\frac{x-(\mu+\sigma\log(S))}{\sigma}}). \tag{2.2}$$

Then by (2.1),

$$P(X \le x) = \exp(-(e^{-\frac{x-\mu}{\sigma}})^{\alpha}) = \exp(-e^{-\frac{x-\mu}{\sigma/\alpha}}). \tag{2.3}$$

Hence unconditionally X also has a Gumbel distribution, but the mixing increases the scale parameter σ of the Gumbel distribution by $100(1/\alpha-1)\%$.

We will sometimes use the terminology that the distribution of X is *directed* by the stable variable S. Let $G \sim \text{Gumbel}(\mu,\sigma)$ mean that the random variable G has the distribution function (d.f.) $\exp(-e^{-\frac{x-\mu}{\sigma}})$. If S has the distribution specified by (2.1), the variable $M=\mu+\sigma\log(S)$ will be called exponential-stable with parameters α,μ, and σ. The symbols $M \sim \text{ExpS}(\alpha,\mu,\sigma)$ will be used to denote such a distribution. Now, equation (2.3) has the following three interpretations:

(i) *Gumbel distribution as a location mixture of Gumbel distributions:* Replacing μ in (2.2) and (2.3) by $\mu_1+\mu_2$, it follows that if G and M are independent and $G \sim \text{Gumbel}(\mu_1,\sigma)$ and $M \sim \text{ExpS}(\alpha,\mu_2,\sigma)$ then $G+M \sim \text{Gumbel}(\mu_1+\mu_2,\sigma/\alpha)$.

For the pitting corrosion measurements, the interpretation would be that the maximal pit depth in a test area had a Gumbel distribution with a random location parameter μ_1+M. The value of M would depend on the proportion of the specimen which was exposed to corrosion. The form of this dependence is made more clear in the next interpretation.

Briefly going beyond the one-dimensional model, it would be natural to assume that different test areas would have different G-s but that the variable M would be the same for all test areas on the same specimen, and different for different test specimens. A further remark is that in this model it is not possible to separate μ_1 and μ_2. However, the parameters can be made identifiable by assuming that either μ_1 or μ_2 is zero.

(ii) *Gumbel distribution as a size mixture of Gumbel distributions:* If the maximum over a unit block has the Gumbel d.f. $\exp(-e^{-\frac{x-\mu_1}{\sigma}})$ and blocks are independent then the maximum over n blocks, or equivalently over one block of size n, has the d.f.

$$(\exp(-e^{-\frac{x-\mu_1}{\sigma}}))^n = \exp(-ne^{-\frac{x-\mu_1}{\sigma}}). \tag{2.4}$$

In this equation it also can make sense to think of non-integer block sizes and random block sizes. In particular, it makes sense to replace n by $Se^{\mu_2/\sigma}$ in (2.4) to obtain the d.f. $\exp(-Se^{\mu_2/\sigma}e^{-\frac{x-\mu_1}{\sigma}})$. It then again follows from (2.1) that the unconditional distribution is $\text{Gumbel}(\mu_1+\mu_2,\sigma/\alpha)$. Thus the $\text{Gumbel}(\mu_1+\mu_2,\sigma/\alpha)$ distribution is obtained as a "size mixture" of $\text{Gumbel}(\mu_1,\sigma)$ distributions, by using the stable size distribution $Se^{\mu_2/\sigma}$. As before, to make the model identifiable, one should assume that either μ_1 or μ_2 is zero.

The interpretation in the corrosion example is simply that $Se^{\mu_2/\sigma}$ is the size of the area which is exposed to corrosion. This size of course cannot be negative. Further it could reasonably be expected to be determined as the sum of many individually negligible contributions. Suitably interpreted, these two properties together characterize the positive stable distributions.

Next, a variant of the Tawn (1990) physical motivation. It is well known that maxima of underlying i.i.d. variables asymptotically has a Gumbel distribution if the point process of large values asymptotically is a Poisson process. Slightly more precisely, if $\{Y_{n,i}\}$ are suitably linearly renormalized values of an i.i.d. sequence $\{Y_i\}$ and $t_i = i/n$, then the point process $\sum_i \epsilon_{(t_i, Y_{n,i})}$ tends to a Poisson process in the plane with intensity $d\Lambda = dt \times d(e^{-(x-\mu)/\sigma})$ if and only if the probability that $\max_{1 \leq i \leq n} Y_{n,i} \leq x$ tends to $\exp(-e^{-\frac{x-\mu}{\sigma}})$, see e.g. Leadbetter et al. (1983). One direction of this result is immediate. Specifically, it follows since $\max_{1 \leq i \leq n} Y_{n,i}$ is less than x precisely if the point process has no points in $(0,1] \times (x, \infty)$ and since the latter event has probability $\exp(-\Lambda\{(0,1] \times (x,\infty)\}) = \exp(-e^{-\frac{x-\mu}{\sigma}})$. Our third interpretation of the Gumbel mixture model is obtained by replacing the constant intensity in the point process by a stable one.

(iii) Gumbel distribution as the maximum of a conditionally Poisson point process: Suppose X is the maximum y-coordinate of a point process in $(0,1] \times R$ such that conditionally on the stable variable S the point process is Poisson with intensity $d\Lambda = Se^{\mu_2/\sigma} dt \times d(e^{-(x-\mu_1)/\sigma})$. Then, by the same argument as above, conditionally on S the variable X has d.f. $\exp(-Se^{\mu_2/\sigma} e^{-\frac{x-\mu_1}{\sigma}})$, and as for (2.3), it follows that the unconditional distribution of X is Gumbel$(\mu_1 + \mu_2, \sigma/\alpha)$.

Tawn's interpretation is that the points in the point process correspond to the maximum wind speeds in the storms that occur during a year. The random intensity $Se^{\mu_2/\sigma}$ then describes an extra stochastic variation from year to year (which may be similar for nearby locations). Again this has to be positive and perhaps obtained as the sum of many individually negligible influences, and hence perhaps positive stable. The interpretation in the corrosion experiment is the same as for (ii) above. As above, one of μ_1 or μ_2 should be assumed to be zero for identifiability.

It may also be noted that in some situations it may be possible to use PoT observations, i.e. to actually observe the underlying large values, say all large storms during a year or all deep corrosion pits in each square. Such measurements could also be handled within the present framework, by substituting the likelihoods in this paper with the corresponding point process (or PoT) likelihoods. However, we will not pursue this further in the present paper.

3. New classes of Gumbel processes

In this section we introduce a number of concrete Gumbel models directed by linear stable processes: a random effects model, time series models with directing stable linear processes, and a spatial model with a stable moving average as directing process. We also consider a hierarchical setup and continuous parameter models.

We first state a slight generalization (a restriction on the size of the set A is removed) of the main result of Tawn (1990). The result is given in three variations which correspond to the three interpretations in Section 2. Let T and A be discrete index sets, where in addition T is assumed to be finite. Further let $\{c_{t,a}\}$ be non-negative constants and let $\{S_a, a \in A\}$ be independent positive α-stable variables with distribution specified by (2.1). We assume without further comment that $\sum_{a\in A} c_{t,a}S_a$ converges almost surely for each t.

Proposition 1. *Conditions (i) and (ii) as stated below are equivalent, and Condition (iii) implies the other two:*

(i)

$$X_t = G_t + \sigma_t \log(\sum_{a\in A} c_{t,a}S_a), \qquad t \in T,$$

where $G_t \sim Gumbel(\mu_t, \sigma_t)$, and all variables are mutually independent,

(ii) the random variables $X_t, t \in T$ are conditionally independent given $S_a, a \in A$, with marginal distributions

$$P(X_t \le x_t | S_a, a \in A) = \exp\left(-(\sum_{a\in A} c_{t,a}S_a)e^{-\frac{x_t-\mu_t}{\sigma_t}}\right), \quad t \in T, \tag{3.1}$$

and

(iii) for $t \in T$, X_t is the maximum y-coordinate of a point process in $(0,1] \times R$ such that conditionally on $S_a, a \in A$ the point processes are independent and Poisson with intensities $\left(\sum_{a\in A} c_{t,a}S_a\right) dt \times d(e^{-(x-\mu_t)/\sigma_t})$.

Further, if either one of the conditions hold then

$$P(X_t \le x_t, t \in T) = \prod_{a\in A} \exp\left(-(\sum_{t\in T} c_{t,a}e^{-\frac{x_t-\mu_t}{\sigma_t}})^\alpha\right). \tag{3.2}$$

Clearly (i) and (ii) of the proposition are just different ways of saying the same thing, and that (iii) implies the others follows as in the discussion of the interpretation (iii) in Section 2. Further, that (ii) implies (3.2) follows immediately from (2.1) since, by conditional independence,

$$P(X_t \le x_t, t \in T) = E\left(\exp(-\sum_{t\in T}\sum_{a\in A} c_{t,a}S_a e^{-\frac{x_t-\mu_t}{\sigma_t}})\right) = \prod_{a\in A} E\left(\exp[-S_a(\sum_{t\in T} c_{t,a}e^{-\frac{x_t-\mu_t}{\sigma_t}})]\right).$$

As discussed in the introduction, for modelling the most useful case is when all kinds of maxima have distribution of the same type as the marginal distributions. This holds when all the scale parameters have the same value, i.e. when $\sigma_t = \sigma$, for $t \in T$, and then, if $T_0 \subset T$,

$$P(\max_{t\in T_0} X_t \le x) = \prod_{a\in A} \exp\left(-(\sum_{t\in T_0} c_{t,a}e^{\frac{\mu_t}{\sigma}})^\alpha e^{-\frac{x}{\sigma/\alpha}}\right), \tag{3.3}$$

or equivalently

$$\max_{t\in T_0} X_t \sim \text{Gumbel}\left((\sigma/\alpha)\log(\sum_{a\in A}(\sum_{t\in T_0} c_{t,a}e^{\mu_t/\sigma})^\alpha), \sigma/\alpha\right).$$

In particular, by letting T_0 be a one point set we see that in this case marginals are Gumbel distributed,

$$X_t \sim \text{Gumbel}\left((\sigma/\alpha)\log(\sum_{a\in A}(c_{t,a}e^{\mu_t/\sigma})^\alpha), \sigma/\alpha\right).$$

Conditions (i) - (iii) in Proposition 1 correspond to the three "physical" interpretations in Section 2. We now turn to a number of specific models. Which interpretation is most relevant of course varies from model to model. E.g the first model below is the standard logistic model for extreme value data, but with a new interpretation as a random effects model. We will use it on a pit corrosion example, where perhaps interpretation (ii) is most compelling. However, to streamline presentation, we will for the rest of this section formulate the models as in (i), but of course could equally well have used (ii) or (iii).

Example: A one-way random effects model. This is the model

$$X_{i,j} = \mu + \tau_i + G_{i,j}, \qquad 1 \le i \le m,\ 1 \le j \le n_i \tag{3.4}$$

with μ a constant, $\tau_i \sim$ExpS$(\alpha, 0, \sigma)$, $G_{i,j} \sim$Gumbel$(0, \sigma)$ and all variables independent.

Setting $T = \{(i,j);\ 1 \le i \le m,\ 1 \le j \le n_i\}$, $A = \{1, 2, \ldots m\}$ and $c_{(i,j),k} = 1_{\{i=k\}}$, this is a special case of the situation in Proposition 1 and we directly get the distribution function

$$P(X_{i,j} \le x_{i,j},\ 1 \le i \le m,\ 1 \le j \le n_i) = \prod_{i=1}^{m} \exp(-(\sum_{j=1}^{n_i} e^{-\frac{x_{i,j}-\mu}{\sigma}})^\alpha). \tag{3.5}$$

According to Proposition 1 this model is max-stable, and explicit formulas are directly available for the distribution of all kinds of unconditional and conditional maxima. In particular the marginal distributions are Gumbel(μ, σ^*) for $\sigma^* = \sigma/\alpha$. □

This model can be extended to higher order random effects models which are "linear on an exponential scale". We next turn to time series models. A linear stationary positive stable process may be obtained as $H_t = \sum_{i=-\infty}^{\infty} b_i S_{t-i}$, where the b_i are nonnegative constants, and the sum converges if $\sum b_i^\alpha < \infty$. Defining

$$X_t = \mu_t + \sigma \log(H_t) + G_t, \tag{3.6}$$

for some constants μ_t, then gives a Gumbel time series model. In particular (3.6) includes hidden ARMA models. We will look closer at the two simplest cases of this.

Example: A hidden MA-process model. Suppose $H_t = b_0 S_t + b_1 S_{t-1} + \ldots b_q S_{t-q}$ and X_t is defined by (3.6), where the S_i have distribution (2.1), $G_t \sim$Gumbel$(0, \sigma)$ and all variables are mutually independent. Then, by Proposition 1 with $T = \{1, \ldots n\}$ and $A = \{0, \pm 1, \ldots\}$,

$$P(X_t \le x_t,\ 1 \le t \le n) = \prod_{k=1-q}^{n} \exp(-(\sum_{t=1\vee k}^{n\wedge(k+q)} b_{t-k} e^{-\frac{x_t-\mu_t}{\sigma}})^\alpha). \tag{3.7}$$

□

Example: A hidden AR-process model. For $0 < \rho < 1$ define the positive stable AR-process H_t by $H_t = \sum_{i=0}^{\infty} \rho^i S_{t-i}$, and let X_t be given by (3.6), with the S_i and G_t as before. From the definition of H_t,

$$\begin{aligned} H_0 &= \sum_{i=0}^{\infty} \rho^i S_{-i} \\ H_1 &= \rho H_0 + S_1 \\ &\vdots \\ H_n &= \rho^n H_0 + \rho^{n-1} S_1 + \cdots + \rho S_{n-1} + S_n, \end{aligned} \tag{3.8}$$

and in addition, by (2.1) H_0 has the same distribution as

$$(\sum_{i=0}^{\infty} \rho^{i\alpha})^{1/\alpha} S_0 = (1-\rho^\alpha)^{-1/\alpha} S_0,$$

and is independent of $S_1, \dots S_n$. Thus, the model is again of the form considered in Proposition 1, with $T = \{0, \dots n\}$, $A = \{0, \pm 1, \dots\}$ and $c_{t,0} = \rho^t (1-\rho^\alpha)^{-1/\alpha}$, $c_{t,a} = \rho^{t-a}$ for $a = 1, \dots t$ and $c_{t,a} = 0$ otherwise. Thus by Proposition 1 the distribution function is

$$P(X_t \leq x_t,\ 0 \leq t \leq n) = \exp[-(1-\rho^\alpha)^{-1} (\sum_{t=0}^{n} \rho^t e^{-\frac{x_t - \mu_t}{\sigma}})^\alpha] \prod_{i=1}^{n} \exp(-(\sum_{t=i}^{n} \rho^{t-i} e^{-\frac{x_t-\mu_t}{\sigma}})^\alpha).$$

□

In the next example we consider models on the integer lattice in the plane. Let $n_{(i,j)}$ be a system of neighborhoods with the standard properties $(i,j) \in n_{(i,j)}$ and $(k,l) \in n_{(i,j)} \Leftrightarrow (i,j) \in n_{(k,l)}$. A simple example is when the neighbors are the four closest points and the point itself, i.e. when $n_{(i,j)} = \{(i,j), (i-1,j), (i+1,j), (i,j-1), (i,j+1)\}$.

Example: A spatial hidden MA-process model. Let $\{S_{i,j}; -\infty < i,j < \infty\}$ be independent standard positive α-stable variables and set $H_{i,j} = \sum_{(k,l) \in n_{(i,j)}} \delta S_{k,l}$ where δ is a positive constant. Put

$$X_{i,j} = \mu_{i,j} + \sigma \log(H_{i,j}) + G_{i,j}, \quad 1 \leq i,j \leq n,$$

where the $G_{i,j}$ are mutually independent and independent of the $S_{i,j}$, and $G_{i,j} \sim$ Gumbel$(0,\sigma)$. Again this is of the form considered in Proposition 1, now with $c_{(i,j),(k,l)} = \delta$ if $(i,j) \in n_{(k,l)}$ and zero otherwise. To write down the joint distribution function it is convenient to use the notation $\bar{n}_{(k,l)} = n_{(k,l)} \cap \{(i,j);\ \ 1 \leq i,j \leq n\}$. We then get that

$$P(X_{i,j} \leq x_{i,j};\ 1 \leq i,j \leq n) = \prod_{(k,l)} \exp(-\delta^\alpha (\sum_{(i,j) \in \bar{n}_{(k,l)}} e^{-\frac{x_{i,j} - \mu_{i,j}}{\sigma}})^\alpha).$$

□

We now turn to a situation not covered by Proposition 1, the so-called nested logistic model of McFadden (see Tawn (1990)).

Example: A two-layer hierarchical model. Consider the model

$$X_{i,j,k} = \mu + \tau_i + \eta_{i,j} + G_{i,j,k}, \qquad 1 \le i \le m,\ 1 \le j \le n_i,\ 1 \le k \le r_{i,j},$$

with μ a constant, $\tau_i \sim$ExpS$(\beta, 0, \sigma/\alpha)^{1/\alpha}$, $\eta_{i,j} \sim$ExpS$(\beta, 0, \sigma)$, $G_{i,j} \sim$Gumbel$(0, \sigma)$, and all variables independent. By repeated conditioning we obtain, after some calculations similar to the proof of Proposition 1,

$$\begin{aligned} &P(X_{i,j,k} \le x_{i,j,k},\ 1 \le i \le m,\ 1 \le j \le n_i,\ 1 \le k \le r_{i,j}) \\ &\qquad = \prod_{i=1}^{m} \exp[-\{\sum_{j=1}^{n_i}(\sum_{k=1}^{r_{i,j}} e^{-\frac{x_{i,j,k}-\mu}{\sigma}})^{\alpha}\}^{\beta}]. \end{aligned}$$

□

There also are continuous parameter versions of Proposition 1. Let $\{S_j(\mathbf{s}); \mathbf{s} \in R^k\}$ be independently scattered positive stable noise (see Samorodnitsky and Taqqu (1994, Chapter 3)). We assume that the noise is standardized, so that for nonnegative functions $f \in L_\alpha$,

$$E[\exp\{-\int_{-\infty}^{\infty} f(\mathbf{s}) S_j(d\mathbf{s})\}] = \exp(-\int_{-\infty}^{\infty} f(\mathbf{s})^{\alpha} d\mathbf{s}). \tag{3.9}$$

In the sequel we will without comment assume that functions are such that integrals converge, and integrals are taken to be over R^k.

Proposition 2. *Suppose that there are non negative functions $f_j(\mathbf{t}, \mathbf{s})$ with $\mathbf{t} \in R^\ell$, $\mathbf{s} \in R^k$ such that*

$$X_{\mathbf{t}} = G_{\mathbf{t}} + \sigma_{\mathbf{t}} \log(\sum_{j=1}^{m} \int f_j(\mathbf{t}, \mathbf{s}) S_j(d\mathbf{s})), \quad \mathbf{t} = \mathbf{t}_1, \dots \mathbf{t}_n,$$

where $G_t \sim$Gumbel(μ, σ_t), and all variables are mutually independent. Then

$$P(X_{\mathbf{t}_i} \le x_{\mathbf{t}_i};\ i = 1, \dots n) = \prod_{j=1}^{m} \exp(-\int (\sum_{i=1}^{n} f_j(\mathbf{t}_i, \mathbf{s}) e^{-\frac{x_{\mathbf{t}_i} - \mu_{\mathbf{t}_i}}{\sigma_{\mathbf{t}_i}}})^{\alpha} d\mathbf{s}). \tag{3.10}$$

The proof follows from (3.9) in the same way as Proposition 1 follows from (2.1). The interpretations (ii), as size mixtures, and (iii) as a random Poisson intensity could equally well have been used as assumptions. However, this we leave to the reader.

Proposition 2 gives a natural model for environmental extremes, such as yearly maximum wind speeds or water levels, at irregularly located measuring stations. E.g. one could assume years to be independent and obtain a simple isotropic model for one year by choosing $k = \ell = 2$, $m = 1$ and $f_1(\mathbf{t}, \mathbf{s}) = \exp(-d|\mathbf{t} - \mathbf{s}|^{\beta})$, for some constants $d, \beta > 0$. One extension to non-isotropic situations is by letting D be a diagonal matrix with positive diagonal elements and taking $f_1(\mathbf{t}, \mathbf{s}) = \exp(-((\mathbf{t} - \mathbf{s})^t D(\mathbf{t} - \mathbf{s})\beta)$. (Formally the entire distribution function for n years is also of the form (3.10), as can be seen by taking $\ell = 3$, $m = n$ and letting the different S_j correspond to different years.) It is possible to derive recursion formulas for the densities of these models in a similar but more complicated way as for the random

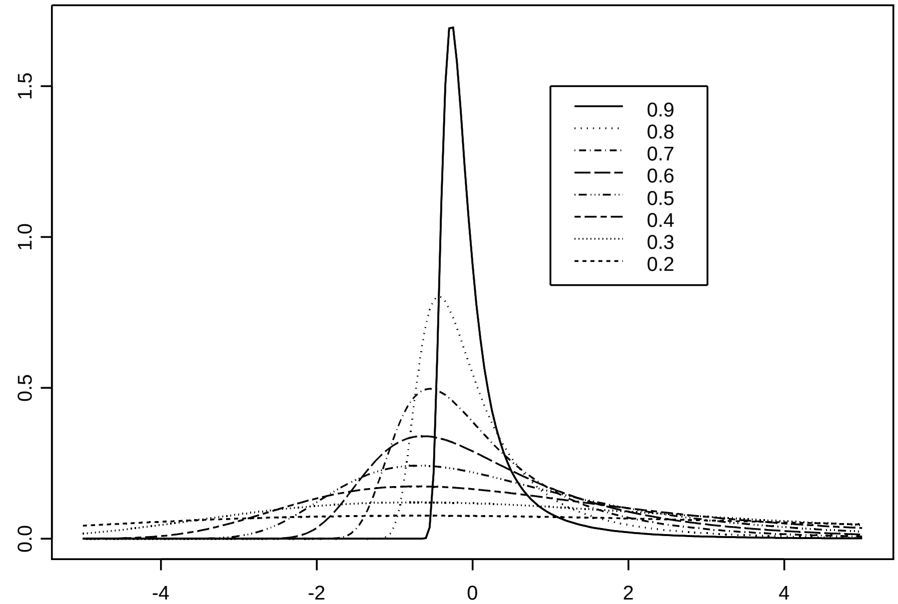

FIGURE 4.1. Plot of densities of standardized exponential-stable distributions ExpS$(\alpha, 0, 1)$, with varying α.

effects model. If the number of measuring stations is not too large, these expressions may be numerically tractable. However, we will not investigate this further in this paper.

4. Some properties of the mixing distribution

This section discusses some of the basic facts about the models. In the notation of Samorodnitsky and Taqqu (1994), the r.v. S in (2.1) is $S_\alpha((\cos \pi\alpha/2)^{1/\alpha}, 1, 0)$; in the notation of Zolotarev (1986), $S \sim S_C(\alpha, 1, 1)$. It has characteristic function

$$E \exp(itS) = \exp\left\{-\cos(\pi\alpha/2)|t|^\alpha \left[1 - i\tan(\pi\alpha/2)(\operatorname{sign} t)\right]\right\}.$$

Let $F_S(s)$ be the d.f. and $f_S(s)$ be the density of S. If $M \sim$ExpS(α, μ, σ), then the d.f. and density of M are $F_M(x) = F_S[\exp\{(x-\mu)/\sigma\}]$ and $f_M(x) = \exp\{(x-\mu)/\sigma\} f_S[\exp\{(x-\mu)/\sigma\}]/\sigma$. Using the programs for computing with stable distributions described in Nolan (1997), it is possible to compute densities, d.f., quantiles and simulate values for M. Figure 4.1 shows the density of some log-stable distributions. The densities all have support $(-\infty, \infty)$ and appear to be unimodal. Note that as $\alpha \uparrow 1$, S converges in distribution to 1 and hence $M = \log S$ converges in distribution to 0.

It is well-known that the upper tail of S is asymptotically Pareto: as $x \to \infty$, $P(S > x) \sim c_\alpha x^{-\alpha}$ where $c_\alpha = \Gamma(\alpha)\sin(\pi\alpha)/\pi$. This implies that the right tail of

$M \sim$ExpS(α, μ, σ) is asymptotically exponential: as $t \to \infty$,

$$P(M > t) = P\left(S > \exp\left(\frac{t-\mu}{\sigma}\right)\right) \sim c_\alpha \exp\left(-\frac{t-\mu}{\sigma/\alpha}\right).$$

The left tail of S is light, see e.g. Section 2.5 of Zolotarev (1986), so the left tail of M is even lighter. Thus all moments of M exist; in particular, using the results of Section 3.6 of Zolotarev (1986),

$$\mathrm{E}(M) = \mu + \sigma\gamma_{Euler}\left(\frac{1}{\alpha} - 1\right), \quad \mathrm{Var}(M) = \frac{\pi^2\sigma^2}{6}\left(\frac{1}{\alpha^2} - 1\right),$$

where $\gamma_{Euler} \approx 0.57721$ is Euler's constant.

As a simple consequence we derive the correlation between two variables in the same group in the random effects model (3.4). Thus, suppose $X_i = \mu + \tau + G_i,\ i = 1,2$ with $\tau \sim$ExpS$(\alpha, 0, \sigma)$, $G_i \sim$Gumbel$(0, \sigma)$ and the three variables independent. Then $\mathrm{Cov}(X_1, X_2) = \mathrm{Var}(\tau)$ and $\mathrm{Var}(X_i) = \mathrm{Var}(\tau) + \mathrm{Var}(G_i)$. Since $\mathrm{Var}(G_i) = \frac{\pi^2\sigma^2}{6}$ we obtain that $\mathrm{Cor}(X_1, X_2) = 1 - \alpha^2$, which varies from 0 in the independent case $\alpha = 1$ to 1 as $\alpha \to 0$, as it should since the limit corresponds to full dependence.

5. Data analysis

In this section we illustrate the random effects model and the hidden MA(1) model from Section 3 by using them to analyze a set of pit corrosion measurements and an interest rate data set. As preliminaries we first discuss maximum likelihood estimation in the two models.

5.1. **Estimation in the random effects model.** Let $0 < \sigma < \sigma^*$, $-\infty < \mu^* < \infty$, so $\alpha := \sigma/\sigma^* \in (0,1)$. Assume a data set $\mathbf{X}$ that comes from m groups,

$$\begin{aligned} &\text{group 1}: \quad X_{1,1}, X_{1,2}, \ldots, X_{1,n_1} \\ &\text{group 2}: \quad X_{2,1}, X_{2,2}, \ldots, X_{2,n_2} \\ &\qquad\qquad\vdots \\ &\text{group m}: \quad X_{m,1}, X_{m,2}, \ldots, X_{m,n_m}. \end{aligned} \tag{5.1}$$

The groups are assumed to be independent and the i^{th} group comes from a Gumbel$(0, \sigma)$ distribution, where the location parameter μ_i for group i is drawn from an ExpS$(\alpha = \sigma/\sigma^*, \mu^*, \sigma)$ distribution. The goal is to estimate the three parameters $\theta = (\sigma, \sigma^*, \mu^*)$ from the data by maximum likelihood.

The likelihood $L(\theta|\mathbf{X}) = \prod_{i=1}^m L_i(\theta|X_{i,1}, \ldots, X_{i,n_i})$ is the product of the group likelihoods. Each of these terms can be derived by differentiating (3.5) with respect to $x_1, \ldots, x_n$. The direct calculations are complicated, but Property (1) of Shi (1995) gives the likelihood function for the group. To simplify formulas we suppress the group index i and get

$$\frac{\partial^n}{\partial x_1 \ldots \partial x_n} P(X_1 \le x_1, \ldots, X_n \le x_n) = \frac{\alpha^n}{\sigma^n} z^{1-n/\alpha} e^{-z} Q_n(z, \alpha) \prod_{j=1}^n e^{-(x_j - \mu)/\sigma}, \tag{5.2}$$

where $z = \left(\sum_{j=1}^{n} e^{-(x_j-\mu)/\sigma}\right)^{\alpha}$ and $Q_n(z,\alpha)$ is a polynomial in z defined by

$$Q_n(z,\alpha) = \left(\frac{n-1}{\alpha} - 1 + z\right) Q_{n-1}(z,\alpha) - z\frac{\partial Q_{n-1}(z,\alpha)}{\partial z}, \qquad Q_1(z,\alpha) = 1. \tag{5.3}$$

The conditions show that $Q_n(z,\alpha)$ is a polynomial of degree $n-1$, say $Q_n(z,\alpha) = \sum_{j=0}^{n-1} q_{n,j} z^j$. Using this in (5.3) gives a recursive equation for the coefficients:

$$q_{n,j} = \begin{cases} \left(\frac{n-1}{\alpha} - 1\right) q_{n-1,0} & j = 0 \\ \left(\frac{n-1}{\alpha} - j\right) q_{n-1,j} + q_{n-1,j-1} & j = 1, \ldots, n-2 \\ 1 & j = n-1. \end{cases}$$

Usually this makes it straightforward to find maximum likelihood estimates by numerical optimization. However, if a group is large or α is small, the coefficients of $Q_n(z,\alpha)$ can be very large. E.g. the constant term is

$$q_{n,0} = \left(\frac{n-1}{\alpha} - 1\right)\left(\frac{n-2}{\alpha} - 1\right)\left(\frac{n-3}{\alpha} - 1\right)\cdots\left(\frac{1}{\alpha} - 1\right).$$

This can in bad cases cause numerical overflow in the optimization routines. Further, if all groups only have one value or if there is only one group then parameters are not identifiable. Presumably parameter estimates will be bad also if data is close to these situations.

An alternative way to derive the likelihood, which in addition indicates a possibility to compute it by simulation, is as follows: A group likelihood, conditional on τ, is

$$\prod_{j=1}^{n} \frac{1}{\sigma} e^{-\frac{x_j-\mu-\tau}{\sigma}} \exp\left\{-e^{-\frac{x_j-\mu-\tau}{\sigma}}\right\} = \frac{1}{\sigma^n} S^n e^{-\sum_{j=1}^{n} \frac{x_j-\mu}{\sigma}} \exp\left\{-S \sum_{j=1}^{n} e^{-\frac{x_j-\mu}{\sigma}}\right\},$$

where $\tau = \sigma \log S$ and S is a standard $\alpha-$stable variable, as previously. Hence, a group likelihood is

$$\frac{1}{\sigma^n} e^{-\sum_{j=1}^{n} \frac{x_j-\mu}{\sigma}} E\left[S^n \exp\left\{-S \sum_{j=1}^{n} e^{-\frac{x_j-\mu}{\sigma}}\right\}\right].$$

Let $\Delta = z^{1/\alpha} = \sum_{j=1}^{n} e^{-(x_j-\mu)/\sigma}$. Then, the expectation in the last expression reduces to

$$E\left[S^n e^{-S\Delta}\right] = E\left[\frac{d^n}{d\Delta^n}\left\{e^{-S\Delta}\right\}\right] = (-1)^n \frac{d^n}{d\Delta^n}\left\{e^{-\Delta^\alpha}\right\},$$

where the second equality makes one more use of the stable distribution of S. The last expression of course is the same as (5.2).

The maximum likelihood algorithm has been implemented in S-Plus. The estimation procedure numerically evaluates $\ell(\theta|\mathbf{X}) = \log L(\theta|\mathbf{X})$ and numerically maximizes it to find the estimate of θ. The search is initialized at $\theta_0 := (\sigma_0/2, \sigma_0, \mu_0)$, where μ_0 and σ_0 are estimates of the Gumbel parameters for the (ungrouped) data set $\mathbf{X}$. This estimate is found by using the probability-weighted moment estimator, see e.g. Section 1.7.6 of Kotz and Nadarajah (2000).

5.2. **Estimation in the hidden MA(1) model.** By (3.7) the hidden MA(1) model with constant location parameter, $\mu_t = \mu$ and, for identifiability, $b_0 = 1, b_1 = b$ has distribution function

$$F = P(X_t \le x_t,\ 1 \le t \le n) = \exp\left(-\left\{(bz_1)^\alpha + \sum_{t=1}^{n-1}(z_t + bz_{t+1})^\alpha + z_n^\alpha\right\}\right), \tag{5.4}$$

where $z_t = \exp(-(x_t - \mu)/\sigma)$. The parameters of the model are $\theta = (\mu, b, \sigma, \alpha)$. By differentiation with respect to $x_1, \ldots, x_n$ the likelihood function can be seen to be of the form

$$L(\theta|\mathbf{X}) = Q_n F \prod_{t=1}^{n} \frac{z_t}{\sigma},$$

with F from (5.4) and Q_n defined recursively as follows. Set $u_1 = bz_1$, $u_t = z_{t-1} + bz_t$ for $t = 2, \ldots, n$, $u_{n+1} = z_n$. Then $F = \exp(-\sum_{t=1}^{n+1} u_t^\alpha)$ and

$$\begin{aligned} Q_0 &= 1, \qquad Q_1 = \alpha\left(bu_1^{\alpha-1} + u_2^{\alpha-1}\right), \\ Q_i &= -Q_{i-2}\alpha(\alpha-1)bu_i^{\alpha-2} + Q_{i-1}\alpha\left(bu_i^{\alpha-1} + u_{i+1}^{\alpha-1}\right), \quad i = 2, \ldots, n. \end{aligned}$$

When $b = 0$, the Q_1 term above should be interpreted as $Q_1 = \alpha u_2^{\alpha-1}$, which makes the likelihood formula valid in the case where the x_t are independent. This has been implemented in S-Plus, where

$$\log\{L(\theta|\mathbf{X})\} = \log Q_n - \sum_{t=1}^{n+1} u_t^\alpha - \sum_{t=1}^{n}\left(\frac{x_t - \mu}{\sigma}\right) - n\log\sigma$$

is computed and numerically maximized. As default the search is started at $(\mu = \mu_0, b = 0, \sigma = \sigma_0/0.5, \alpha = 0.5)$, where (μ_0, σ_0) are the Gumbel power weighted moment estimators for the data set. However, for small sample sizes results were sensitive to the choice of starting values. In such cases we started the search at many different randomly chosen points and chose as estimator the final values which gave the highest likelihood.

5.3. **Pitting corrosion data analysis.** The pitting corrosion investigation which generated this data set was briefly mentioned in the beginning of Section 2. Specifically, pieces (or "test specimens") were cut out from different parts of the bottom hemflange of the aluminum back door of a twelve year old station wagon. The corrosion products were dissolved from the pieces, and the deepest corrosion pit was measured in a number of one centimetre long test areas on each specimen. The hemflange had been glued together and had also been treated with a corrosion preventing coating. Surface areas where the glue or coating was intact showed no corrosion. However, in some places the glue and coating had not penetrated well or had fallen of, leaving the surface exposed to corrosion. The proportion of the area which could corrode varied between specimens, and this was a potential cause of extra variation in the corrosion measurements.

Interest was centered on the risk of penetration by deepest corrosion pit on the outer surface of the hemflange. The data set for this surface consisted of microscope measurements (in microns) of the maximum pit depth in 11 to 15 test areas on each of 12 specimens. There was no corrosion on 5 of the test specimens, and on one only two test areas showed any corrosion. These 6 specimens were excluded from our

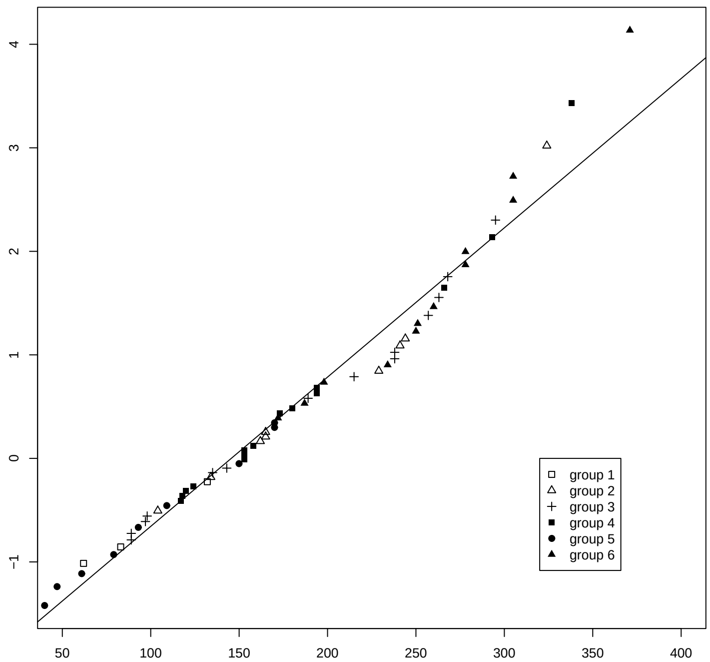

FIGURE 5.1. Gumbel plot for the pooled corrosion measurements. A different symbol is used for each group.

analysis. Also in the remaining specimens there were some corrosion free test areas, and the data we used for analysis hence consisted of 6 groups (=test specimens) with varying numbers (ranging from 4 to 14) of measured maximum pit depths.

The engineers who performed the experiment disregarded the group structure and considered the pooled data set as an i.i.d Gumbel sample. The maximum likelihood parameter estimates under this model were $(\mu_{\text{pool}}, \sigma_{\text{pool}}) = (145.6, 69.4)$. It was remarked by the engineers that there seemed to be some deviation from a straight line in Gumbel plot, see Figure 5.1.

We instead analysed the data with the random effects Gumbel model from Subsection 5.1. The aim was both to see if this model fitted better and to check wether it lead to a substantially different risk estimate. In addition to the extra variation between test specimens there might also be a short range dependence between neighboring test areas. We tried to judge the size of this effect by fitting a hidden MA(1) model.

The maximum likelihood estimates in the random effects Gumbel model were $(\mu, \sigma, \alpha) = (140.9, 54.1, 0.716)$ with standard deviations $(21.75, 5.71, 0.118)$ estimated from the inverse of the empirical information matrix. A very rough calculation of the risk of perforation can then be made as follows. There are about 15 test specimens on a hemflange. Let us assume, as was the case with the present data, that typically about 6 of the test specimens will show corrosion and that on average about 11 test areas on each specimen will be corroded. Then, by (3.3) the estimated distribution function of the maximum pit depth for one car would be

$$\hat{F}(x) = \exp(-6(11e^{-\frac{x-140.9}{54.1}})^{54.1/75.6}).$$

The thickness of the aluminum was 1.1 mm = 1100 microns and hence we estimate that there on the average will be perforation in one out of $1/(1-\hat{F}(1100)) = 9671$ cars. A delta method 95% confidence interval for this estimate is $(8392, 10950)$. If we instead, following the engineering analysis, use the pooled Gumbel model with the assumption that typically there are $6 \times 11 = 66$ corroded test areas on a hemflange, the risk estimate is that on the average there is penetration in one out of $(1-\exp(-66e^{-\frac{1100-145.6}{69.4}})^{-1} \approx 14374$ cars. A delta method 95% confidence interval is $(13115, 15632)$.

The formulation as a random effects model gives a number of possibilities for model checking. From Figure 5.2 can be seen that the Gumbel distribution fits reasonably well to the separate groups, that there indeed seems to be an extra variation between groups, and that the fitted lines are approximately parallel. As a further formal check of the assumption that the σ- s in the group were equal, we made a conditional analysis, fitting separate Gumbel distributions to the groups by maximum likelihood. In this we considered two different models, one with separate μ-s and σ-s for the groups and one where all groups were assumed to have the same σ. A likelihood ratio test between the models gave $p = .53$. The σ estimate from the latter model was 47.6, which is reasonably close to the σ estimate 54.1 in the random effects model.

Similarly, $\sigma* = 75.6$ and $\sigma_{\text{pool}} = 69.4$ are rather close, as they should be. A further comparison is that the correlation coefficient estimated nonparametrically from the data was 0.44. This can be compared with the correlation coefficient $1-\hat{\alpha}^2 = 0.49$ computed from the fitted model. Finally, from looking at simulations of the fitted random effects model, we thought the apparent deviations from the marginal Gumbel distribution in the pooled Gumbel plot in Figure 5.1 seemed well within the range of what could be expected.

As a check on the fit of the mixing distribution, Figure 5.3 shows the quantiles of the estimated μ-s against the quantiles of the fitted exponential-stable distribution. According to the model, the μ-s are exponential-stable, and hence, apart from estimation error, the estimated μ-s are expected to be exponential-stable. The plot also shows a reasonable fit, and in fact looks much like the same qq-plots from simulated values from the model.

As a final control we fitted the hidden MA(1) model from Subsection 5.2 to the data. In this we assumed groups were independent and had their own μ-s, but that σ, α and b were the same in all 6 groups. Thus there were in all 9 parameters, the six group means $\mu_1, \mu_2, \mu_3, \mu_4, \mu_5, \mu_6$ and the parameters σ, α, b. Maximum likelihood estimation using the default initial values got stuck in a local maximum, and we hence did the optimization for 100 different starting values for σ, α, b, chosen at random from the cube $[7, 54] \times [0.1, 0.99] \times [0, 2]$. As estimates we took the final values which gave the highest likelihood. For the μ-s in the 6 groups these were $87.3, 142.0, 132.4, 140.0, 67.6, 214.8$ and the estimators for the remaining parameters were $\hat{\sigma} = 29.6, \hat{\alpha} = 0.58, \hat{b} = 0.13$.

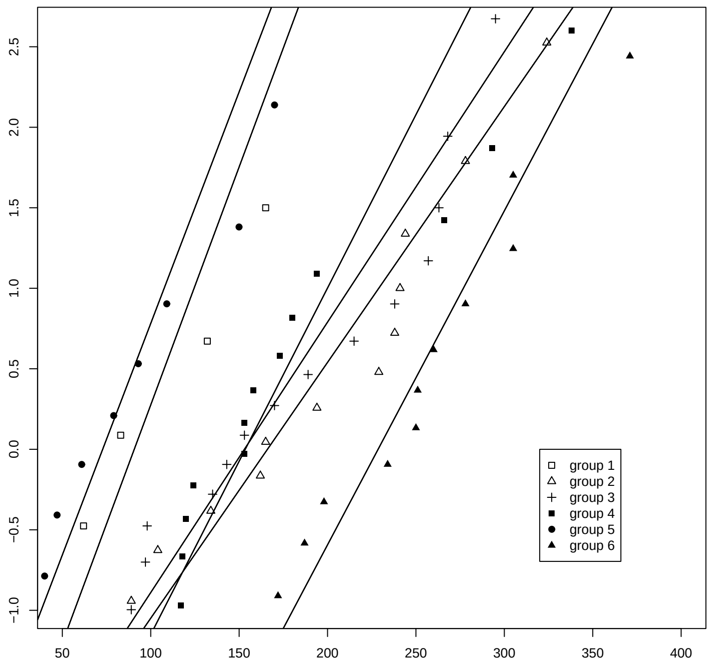

FIGURE 5.2. Gumbel plots made separately for the 6 groups. The solid lines are the different theoretical Gumbel fits for each group.

From the model, the marginal distributions in the groups are Gumbel with location parameter $\mu + \frac{\sigma}{\alpha}\log(1+b^{\alpha})$ and scale parameter σ/α. The estimates of these agreed to within 5% with their initial values, which indicated that these parameters were reasonably well determined by the data. The remaining two parameters, α and b, model the dependence structure. The smaller the α and the closer b is to one, the higher is the dependence. These parameters seemed harder to estimate. Their estimated values indicated a rather weak local dependence, and we stopped the analysis at this point. If this dependence had been judged important, we in principle could have fitted a model which included both random group means and a local MA(1) dependence. However, such an analysis would perhaps have been more than the present data set could bear.

There are weak points in this analysis. An obvious one is the assumption that a hemflange has 6 test specimens with 11 corroded test areas each. However the most important one may well be that the variation in pit depths from car to car is disregarded. If measurements on several cars had been available, it would have been natural to try to fit the hierarchical model from Section 3.

5.4. **Interest rate data analysis.** As a final example we considered a time series of extreme four week fluctuations of an interest rate, as shown in Figure 5.4, left plot. The data comes from the US Federal reserve system[1] and is the long-term or

[1]Available online at `http://www.federalreserve.gov/releases/h15/data.htm`

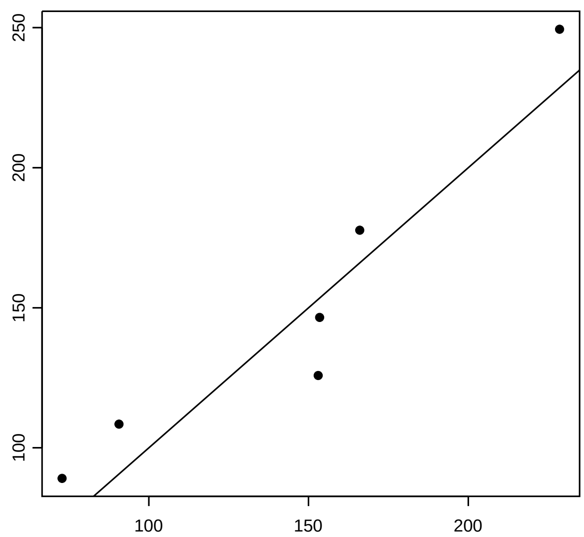

FIGURE 5.3. qq-plot of fitted exponential-stable distribution against estimated μ-s from the conditional analysis with the same σ in all groups.

capital market interest rate for "constant maturity, nominal 1 month". The unit is percentage points. The data are not seasonally adjusted, and concern business days (five days, Monday-Friday), for the time period Oct 9, 2001 to Jan 21, 2005. The

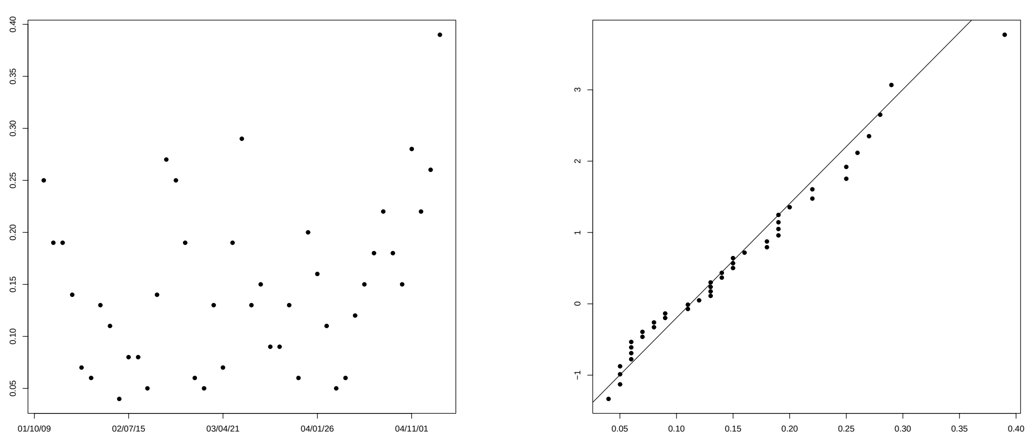

FIGURE 5.4. *Left: Four week fluctuations of an interest rate in terms of time. Right: Gumbel plot for the fluctuations data shown on the left.*

"extreme four week fluctuation" was computed as the maximum interest rate over the four week period minus the minimum interest rate over the four week period. There were in all 43 four week periods in the time span of the data, and hence the time series consists of 43 values.

The maximum likelihood estimates from fitting the hidden MA(1) model to the data set were $\hat{\mu} = 0.096$, $\hat{\sigma} = 0.017$, $\hat{\alpha} = 0.269$, $\hat{b} = 0.033$. Figure 5.4, right plot, shows a reasonable fit of the marginal Gumbel distribution. The first four partial autocorrelations were 0.476, -0.070, -0.004 and -0.261, which supports a 1-dependent model for the time series. To assess whether the fitted parameters capture the observed dependence, one thousand simulated time series of length 43 were generated with the parameters estimated above ($\hat{\sigma} = 0.017$, $\hat{\alpha} = 0.269$, $\hat{b} = 0.033$). In the simulations, the mean one-lag correlation was 0.314; this compares to the one-lag correlation coefficient of 0.476 found in the data. Neither of these checks contradicted the choice of model.

6. Mixtures of generalized extreme value distributions

The mixture models for the Gumbel distribution discussed so far in the paper carry over to the (generalized) Extreme Value distribution in a straightforward manner. However, the interpretation (i) is different.

The EV distribution has d.f. $\exp(-(1+\gamma\frac{x-\mu}{\sigma})^{-1/\gamma})$ with parameters $\mu, \gamma \in R$ and $\sigma > 0$. For positive γ this distribution has a finite left endpoint $\delta = \mu - \sigma/\gamma$ and for γ negative it has a finite right endpoint $\delta = \mu + \sigma/|\gamma|$. In analogy with (2.1) - (2.3) let S be positive stable with Laplace transform (2.1) and assume that

$$P(X \le x|S) = \exp[-S(1+\gamma\frac{x-\mu}{\sigma})^{-1/\gamma}] = \exp[-(\gamma\frac{x-\delta}{S^{\gamma}\sigma})^{-1/\gamma}]. \tag{6.1}$$

Then by (2.1),

$$P(X \le x) = \exp\left[-\left\{1+(\gamma/\alpha)\frac{x-\mu}{(\sigma/\alpha)}\right\}^{-1/(\gamma/\alpha)}\right]. \tag{6.2}$$

Thus, in the terminology of (ii) of Section 2, if X is a positive stable size mixture of an EV distribution with location μ, scale σ and shape parameter γ then also X itself has an EV distribution with the same location μ and the same right endpoint δ, but with a new scale parameter σ/α and new shape parameter γ/α. Hence in particular the unconditional distribution of X has heavier tails than the conditional one.

The physical motivations (ii) and (iii) from Chapter 2 carry over to the present situation without change. Further, from (6.1) it can be seen that X may be obtained as a special random location-scale transformation of an EV distribution. Specifically, if E has an EV distribution with parameters μ, σ, γ and S is positive α-stable and independent of E, then X may be represented as

$$X = S^{\gamma}E + (1-S^{\gamma})\delta. \tag{6.3}$$

Thus X is obtained as a scale mixture with mixing distribution S^{γ}, but in addition there is an accompanying location change which is tailored to keep the endpoint of the distribution unchanged. This, of course, may be the most natural way to make scale mixtures of distributions with finite endpoints.

With this change, the motivations from Section 2 and the models from Section 3 carry over to the EV distribution. If the models in Section 3 are written as size mixtures, i.e. in the form (ii), the only changes needed to go from Gumbel to EV are to replace $e^{-\frac{x-\mu}{\sigma}}$ by $(1+\gamma\frac{x-\mu}{\sigma})^{-1/\gamma}$ in all expressions. The recursions for the likelihood functions from section 5 translate to the EV case similarly.

It is also straightforward to translate specifications using (i) to the EV case. E.g, in the formulation (i) the random effects model (3.4) becomes

$$X_{i,j} = S_i^{\gamma} E_{i,j} + (1 - S_i^{\gamma})\delta,$$

where $E_{i,j}$ has an EV distribution with parameters μ, σ, γ and S is positive α-stable, and all variables are mutually independent. In the same way, the hidden time series model (3.6) in EV form can be written as

$$X_t = H_t^{\gamma} E_t + (1 - H_t^{\gamma})\delta,$$

with H_t a linear stable process and E_t is EV distributed, and all variables are mutually independent.

Next,

$$\log(X - \delta) = \gamma \log S + \log(E - \delta),$$

and if X is of the form (6.3) with $\gamma > 0$ then $\log(E - \delta)$ has a Gumbel distribution with location parameter $\log(\sigma/\mu)$ and scale parameter γ. For $\gamma < 0$ we instead write

$$\log(\delta - X) = \gamma \log S + \log(\delta - E),$$

where $\log(\delta - E)$ has a Gumbel distribution with location parameter $\log(\sigma/\mu)$ and scale parameter γ. Thus the diagnostic plots for Gumbel mixtures could be used also for EV mixtures, except that δ isn't known. A pragmatic way to control the model assumptions then is to replace δ by some suitable estimate.

7. DISCUSSION

The pitting corrosion example discussed in Section 5 was the starting point for the present research. There it seemed important to use models where conditional and unconditional distributions and maxima over blocks of varying sizes all had Gumbel distributions, since this leads to simple and understandable results, and credible extrapolation into extreme tails. It seems important to stay within the extreme value framework throughout for many other applications too. This is the main reason for the present work.

The results include a large number of possibilities for new models. We have given some examples of this. They also throw new light on some much studied existing logistic models. In particular they point to possibilities for new kinds of model diagnostics.

However, we believe that the major part of the possibilities opened up still remains to be exploited. This will be examined in future research.

Acknowledgement: We want to thank Sture Holm for many stimulating discussions and ideas.

References

[1] Anderson, C. W. and Nadarajah, S. (1993). Environmental factors affecting reservoir safety. *In Statistics for the environment*, Barnett, V. and Turkman, K. F. (Ed.), 163-182, Wiley, New York.

[2] Bruun, J. T. and Tawn, J. A. (1998). Comparison of approches for estimating the probability of coastal flooding. *Appl. Statist.* **47**, 405-423.

[3] Chambers, J. L., Mallows, C.L. and Stuck, B.W. (1976). A method for simulating stable random variables. *JASA* **71**, 340-344. Correction in *JASA* **82** (1987), 704.

[4] Coles, S. G. (2001) *An Introduction to Statistical Modeling of Extreme Values.* Springer, London.

[5] Coles, S. G. and Tawn, J. A. (1991). Modelling multivariate extreme events. *J. Roy. Statist. Soc. B* **53**, 377-392.

[6] Coles, S. G. and Tawn, J. A. (1994). Statistical methods for multivariate extremes: an application to structural design (with discussion). *Appl. Statist.* **43**, 1-48.

[7] Crowder, M. J. (1989). A multivariate distribution with Weibull components. *J. Roy. Statist. Soc. B* **51**, 93-108.

[8] Fougères, A.-L.. (2004). Multivariate extremes. *In Extreme Values in Finance, Telecommunications and the Environment*, Finkenstädt, B. and Rootzén, H. (Ed.), 373-388, Chapman & Hall/CRC, Boca Raton.

[9] Gumbel, E. J. (1960). Distribution de valeurs extrêmes en plusieurs dimensions. *Publ. Inst. Statist. Paris* **9**, 171-173.

[10] de Haan, L. and de Ronde, J. (1998). Sea and wind: multivariate extremes at work. *Extremes* **1**, 7-45.

[11] Heffernan, J. E. and Tawn, J. T. (2004). A conditional approach for multivariate extreme values. *J. R. Statist. Soc. B* **66**, 497-546.

[12] Hougaard, P. (1986). A class of multivariate failure time distributions. *Biometrika* **73**, 671-678.

[13] Kotz, S. and Nadarajah, S. (2000). *Extreme Value Distributions : Theory and Applications.* Imperial College Press, London.

[14] Kowaka, M. (1994). *An Introduction to Life Prediction of Plant Materials. Application of Extreme Value Statistical Methods for Corrosion Analysis.* Allerton Press, New York.

[15] Leadbetter, M. R., Lindgren, G. & Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes.* Springer, New York.

[16] Ledford, A. W. and Tawn, J. A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika* **83**, 169-187.

[17] McFadden, D.(1978). Modelling the choice of residential location. In *Spatial Interaction Theory and Planning Models*, Karlquist, A. *et al.* (Ed.), 75-96, North-Holland, Amsterdam.

[18] Mikosch, T.. (2004). Modelling dependence and tails of financial time series. *In Extreme Values in Finance, Telecommunications and the Environment*, Finkenstädt, B. and Rootzén, H. (Ed.), 185-286, Chapman & Hall/CRC, Boca Raton.

[19] Nolan, J. P. (1997). Numerical calculation of stable densities and distribution functions, *Commun. Statist. Stochastic Models*, **13**, 759-774.

[20] Samorodnitsky, G. and Taqqu, M. S. (1994). *Stable Non-Gaussian Random Processes.* Chapman and Hall, New York.

[21] Shi, D. (1995). Fisher information for a multivariate extreme value distribution. *Biometrika* **82**, 644-649.

[22] Stărică, C. (1999). Multivariate extremes for models with constant conditional correlations. *J. Empirical Finance* **6**, 515-553.

[23] Smith, R. L. (2004). Statistics of extremes, with applications in environment, insurance, and finance. *In Extreme Values in Finance, Telecommunications and the Environment*, Finkenstädt, B. and Rootzén, H. (Ed.), 2-78, Chapman & Hall/CRC, Boca Raton.

[24] Stephenson, A. G. (2003). Simulating multivariate extreme value distributions of logistic type. *Extremes* **6**(1), 49-59.

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

[26] Zolotarev, V. M. (1986). *One-dimensional Stable Distributions*, Amer. Math. Soc. Transl. of Math. Monographs, Vol. **65**. Amer. Math. Soc., Providence, RI. (Transl. of the original 1983 Russian).

Département GMM, Institut National des Sciences Appliquées, 135 Avenue de Rangueil, 31077 Toulouse Cedex 04, France. Email: fougeres@insa-toulouse.fr

Department of Mathematics and Statistics, American University, 4400 Massachusetts Avenue, NW, Washington, DC 20016-8050 USA. Email: jpnolan@american.edu

Department of Mathematics, Chalmers University of Technology, S-412 96 Gothenburg, Sweden. Email: rootzen@math.chalmers.se