Multvariate Extremes, Max-Stable Processes and the Analysis of Financial Risk

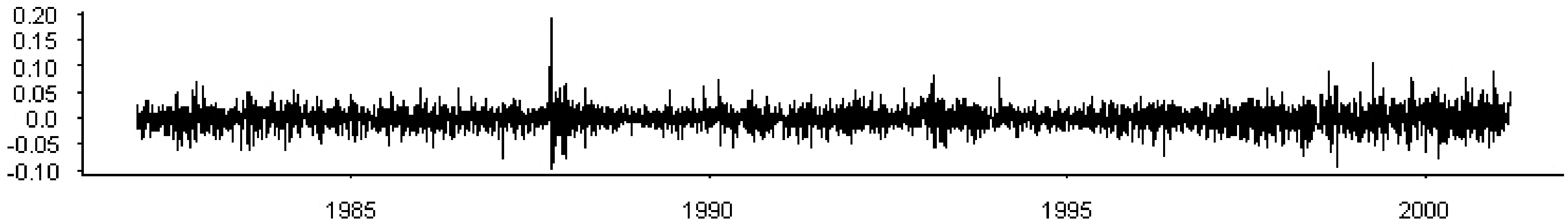
Richard L. Smith
Department of Statistics and Operations Research
University of North Carolina
Chapel Hill, N.C., U.S.A.

Oberwolfach Meeting "Statistics in Finance"
January 14 2004

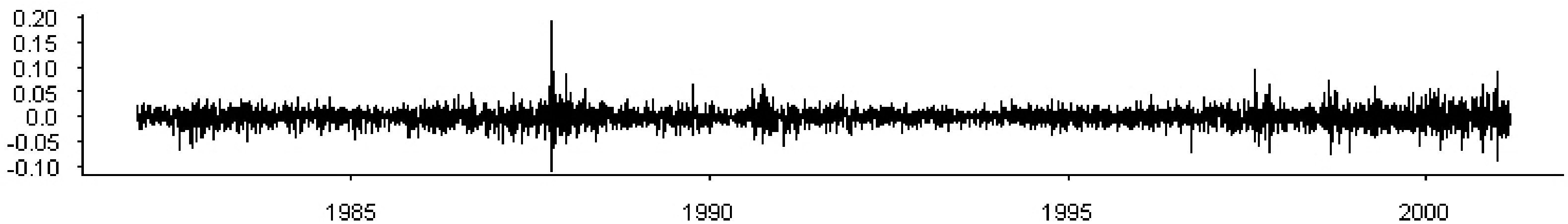
Co-authors:
Ishay Weissman
Zhengjun Zhang
Francescu Chamú

Neg Daily Returns 1982-2001

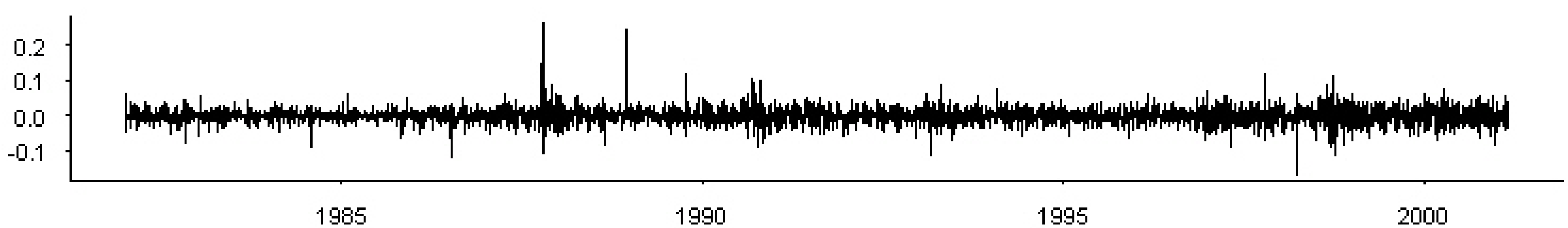
Pfizer



GE



Citibank



Motivation

We show negative daily returns ($R_t = log(X_{t-1}/X_t)$ where X_t is closing price on day t) of 1982–2001 stock prices in three companies, Pfizer, GE and Citibank. We are interested in assessing the risk of a portfolio that includes these stocks.

Typical questions are

1. How to determine *Value at Risk*, i.e. the amount which might be lost in a portfolio of assets over a specified time period with a specified small probability,

2. Dependence among the extremes of different series, and application to the portfolio management problem,

3. Modeling extremes in the presence of volatility.

Univariate Methods

Univariate methods of extreme value analysis broadly fall into three categories,

1. *sample maximum methods*, exploiting the "three types" of limiting distributions of extreme value theory,

2. *threshold approaches*, using all exceedances over a high threshold, often employing the generalized Pareto distribution,

3. the *point process approach* (Smith 1989, Coles 2001) in which all values and time points above a threshold are plotted as a 2-dimensional point process, and various models fitted.

All of these methods have been applied to financial data to assess, e.g. the drop in price that a stock may suffer with specified probability (e.g. 0.01) over a specified time period (e.g. 1 month) — the *value at risk* problem. A practical difficulty is how to deal with *volatility*. One solution is to apply extreme value analysis to the residuals from a GARCH model instead of directly to returns (McNeil and Frey 1999). We will present a variant of that idea here. The main topic of this talk, however, is to extend these methods to consider multivariate aspects of the problem, including short-term dependence between the extremes of a time series that may not be captured by GARCH-type behavior.

Multivariate Extremes

Basic definitions: Suppose $\mathbf{X}_i = (X_{i1}, ..., X_{iD})$, $i = 1, 2, ...$ is an i.i.d. sequence of D-dimensional random vectors. For each $d \in \{1, ..., D\}$, let $M_{nd} = \max\{X_{id},\ 1 \le i \le n\}$.

If normalizing constants a_{nd}, b_{nd} and a D-dimensional distribution function G exist such that as $n \to \infty$,

$$\Pr\left\{\frac{M_{nd} - b_{nd}}{a_{nd}} \le x_d,\ 1 \le d \le D\right\}$$
$$\to G(x_1, ..., x_D)$$

then G is called a *multivariate extreme value distribution*.

Representations: Pickands, de Haan and Resnick, Deheuvels (1970s) gave general representation formulae for MEVDs (see Resnick's (1987) book for full description). However these formulae are too general to be useful for statistics.

Statistics: Much work on parametric subfamilies (Tawn, Coles, ...) and on nonparametric estimation methods but these work well only for small D.

Problem 1: What to do about large D? (e.g. $D \approx 100$ for a typical portfolio)

Problem 2: How to extend these methods to take into account also time-series dependence within each series?

Max-Stable Processes

Suppose $\{Y_{id},\ i = 0, \pm 1, \pm 2,\ d = 1, ..., D\}$ is a D-dimensional time series with discrete time index i.

W.l.o.g. we may assume $\Pr\{Y_{id} \le y\} = e^{-1/y}$ for $0 < y < \infty$ (unit Fréchet assumption). In practice, this would be achieved only above a given threshold, by first fitting a GPD to the marginal distribution above a threshold, but the properties we are interested in can all be expressed in terms of (multivariate) exceedances of high thresholds.

The process is *max-stable* if for any $n \ge 1,\ N \ge 1,\ y_{id} \ge 0$ for $i = 1, ..., n,\ d = 1, ..., D$,

$$\begin{aligned} &\Pr{}^N \{Y_{id} \le N y_{id},\ 1 \le i \le n,\ 1 \le d \le D\} \\ &= \Pr\{Y_{id} \le y_{id},\ 1 \le i \le n,\ 1 \le d \le D\}. \end{aligned}$$

Why consider max-stable processes?

1. Natural generalization of multivariate extremes to infinite dimensions

2. (Smith and Weissman). Suppose we are interested in calculating the *multivariate extremal index* — a natural measure of the clustering and dependency of extremes in a multivariate time series. Under some mixing conditions, if the finite-dimensional distributions of our time series converge to those of a max-stable process, then the multivariate extremal index of our time series is the same as that of the limiting max-stable process. Therefore, for statistical modeling based on exceedances over a high threshold, we may as well assume we are observing the max-stable process directly.

Representations of Max-Stable Processes

The process Y_{id} is said to be a *multivariate maxima of moving maxima* (M4) process if

$$Y_{id} = \max_{\ell=1}^{\infty} \max_{k=-\infty}^{\infty} a_{\ell,k,d} Z_{\ell,i-k},$$

where $Z_{\ell,i}$ are independent unit Fréchet for all ℓ, i; $a_{\ell,k,d} \geq 0$; and

$$\sum_{\ell=1}^{\infty} \sum_{k=-\infty}^{\infty} a_{\ell,k,d} = 1, \quad d = 1, ..., D.$$

For this process,

$$\begin{aligned} &\Pr\{Y_{id} \leq y_{id},\ i = 1, ..., n,\ d = 1, ..., D\} \\ &= \exp\left(- \sum_{\ell=1}^{\infty} \sum_{m=-\infty}^{\infty} \max_{k=1-m}^{n-m} \max_{d=1}^{D} \frac{a_{\ell,k,d}}{y_{m+k,d}} \right). \end{aligned}$$

Corollary. The M4 process is max-stable.

The converse property: Can all max-stable processes be approximated by M4 processes?

If we exclude certain degenerate cases,the answer is yes. This directly generalizes a result of Deheuvels (1983) for one-dimensional max-stable processes, which in turn gneeralizes the representation of multivariate extreme value distributions due to Deheuvels (1978).

Relationship to Multivariate Regular Variation

Consider a multivariate vector $\mathbf{Y}$. Thomas Mikosch defined MRV by the property

$$\lim_{N\to\infty} \frac{\Pr\left\{\frac{\mathbf{Y}}{N} \in A\right\}}{\Pr\left\{|\mathbf{Y}| \geq N\right\}} = \mu(A)$$

where μ is some measure on $\mathcal{R}^D - \{\mathbf{0}\}$ satisfying a homogeneity property.

Let us apply this when the components of $\mathbf{Y}$ are $Y_{id},\ i = 1, ..., n,\ d = 1, ..., D$ and $|\mathbf{Y}| = \max_{i=1}^{n} \max_{d=1}^{D} Y_{id}$.

Suppose $Y_{id} = \max_\ell \max_k a_{\ell,k,d} Z_{\ell,i-k} > N$ for some $i \in \{1, ..., n\}$, $d \in \{1, ..., D\}$.

Fact: If $N \to \infty$, then with probability tending to 1, there will be a *single* index pair (ℓ^*, m^*) such that

$$Y_{id} = a_{\ell^*, i-m^*, d} Z_{\ell^*, m^*}, \ \ i = 1, ..., n, \ \ d = 1, ..., D,$$

where

$$\left[\max_{i=1}^{n} \max_{d=1}^{D} a_{\ell^*, i-m^*, d} \right] Z_{\ell^*, m^*} > N.$$

If $\max_{i=1}^{n} \max_{d=1}^{D} a_{\ell^*, i-m^*, d} = a_{\ell^*, i^*-m^*, d^*}$, then for $1 \leq i \leq n, \ 1 \leq d \leq D$,

$$\frac{Y_{id}}{Y_{i^* d^*}} = \frac{a_{\ell^*, i-m^*, d}}{a_{\ell^*, i^*-m^*, d^*}}. \tag{1}$$

The measure μ is degenerate, concentrating all its mass on vectors $\mathbf{Y}/|\mathbf{Y}|$ that satisfy (1) for some ℓ^*, m^*. We call such relationships *signature patterns*.

Estimation

The existence of signature patterns essentially means that standard methods such as maximum likelihood are not applicable.

Zhang and Smith (2004, to appear) showed that for a long enough series, all the parameters are exactly identifiable from the signature patterns, but this is obviously a pathological result, and could not apply to real time series that arise in fields such as finance.

Earlier methods were based on simpler forms of model,but ran into essentially the same difficulty. For example, the one-dimensional model $Y_i = \max_k a_{i-k} Z_k$ is call the *moving maximum* process. The *max ARMA* processes of Davis and Resnick (1989, 1993) are special cases of this.

Hall, Peng and Yao (2002) estimated moving maxima processes through the empirical distribution function, thus avoiding the issue of degeneracy.

Zhang (2002, PhD thesis) generalized this to M4 processes.

Another idea is to write the model with noise, e.g.

$$X_i = Y_i + \epsilon_i$$

where Y_i is MM and ϵ_i are normal noise, and solve by MCMC techniques. In PhD work in progress, F. Chamú has considered the use of *particle filters* to solve this problem.

A surprising fact: the MM(1) process is second-order Markov.

$$Y_t = \max\{\alpha Z_t, (1-\alpha)Z_{t-1}\}$$

I. $Y_t = \alpha Z_t > (1-\alpha)Z_{t-1},\ Y_{t-1} = \alpha Z_{t-1};\ \frac{Y_t}{Y_{t-1}} > \frac{\alpha}{1-\alpha}$.

II. $Y_t = \alpha Z_t,\ Y_{t-1} = (1-\alpha)Z_{t-2}$; no restriction on $\frac{Y_t}{Y_{t-1}}$.

III. $Y_t = (1-\alpha)Z_{t-1},\ Y_{t-1} = \alpha Z_{t-1};\ \frac{Y_t}{Y_{t-1}} = \frac{\alpha}{1-\alpha}$.

IV. $Y_t = (1-\alpha)Z_{t-1},\ Y_{t-1} = (1-\alpha)Z_{t-2} > \alpha Z_{t-1};\ \frac{Y_t}{Y_{t-1}} < \frac{\alpha}{1-\alpha}$.

$\frac{Y_t}{Y_{t-1}} > \frac{\alpha}{1-\alpha}$ implies **I** or **II**; $Z_t = \frac{Y_t}{\alpha}$.

$\frac{Y_t}{Y_{t-1}} = \frac{\alpha}{1-\alpha}$ implies **III**; $Z_t < \frac{Y_t}{\alpha}$.

$\frac{Y_t}{Y_{t-1}} < \frac{\alpha}{1-\alpha}$ implies **IV** or **II**; $Z_{t-1} < \frac{Y_{t-1}}{\alpha}$.

In all cases the future prediction of $Y_s,\ s > t$ depends on $Y_s,\ s \le t$ only though Y_{t-1} and Y_t.

In very recent work, Chamú has shown that the MM(2) process is third-order Markov.

A plausible conjecture?? Every M4 process with finitely many non-zero coefficients is Kth order Markov for some finite K.

If true, this could be very helpful in generating approximate conditional distributions for the particle filter.

Heuristic Method

1. For each univariate series, fit the standard EV model to exceedances above a threshold and transform the margins to unit Fréchet.

2. Fix a finite range for k (here, -2 to 2) and an upper bound for ℓ (here, 25)

3. For each local maximum above the threshold, define a signature pattern from the ratio of neighbouring Y values to the maximum. In our example, $D = 3$ and there are 607 local maxima, so we end up with 607 candidate signature patterns in 15 dimensions.

4. Use K-means clustering to group the 607 signature patterns into 25 clusters (less than a minute in S-PLUS)

5. Estimate $a_{\ell,k,d}$ coefficients by aggregating signature patterns across clusters.

Extremes in Financial Time Series

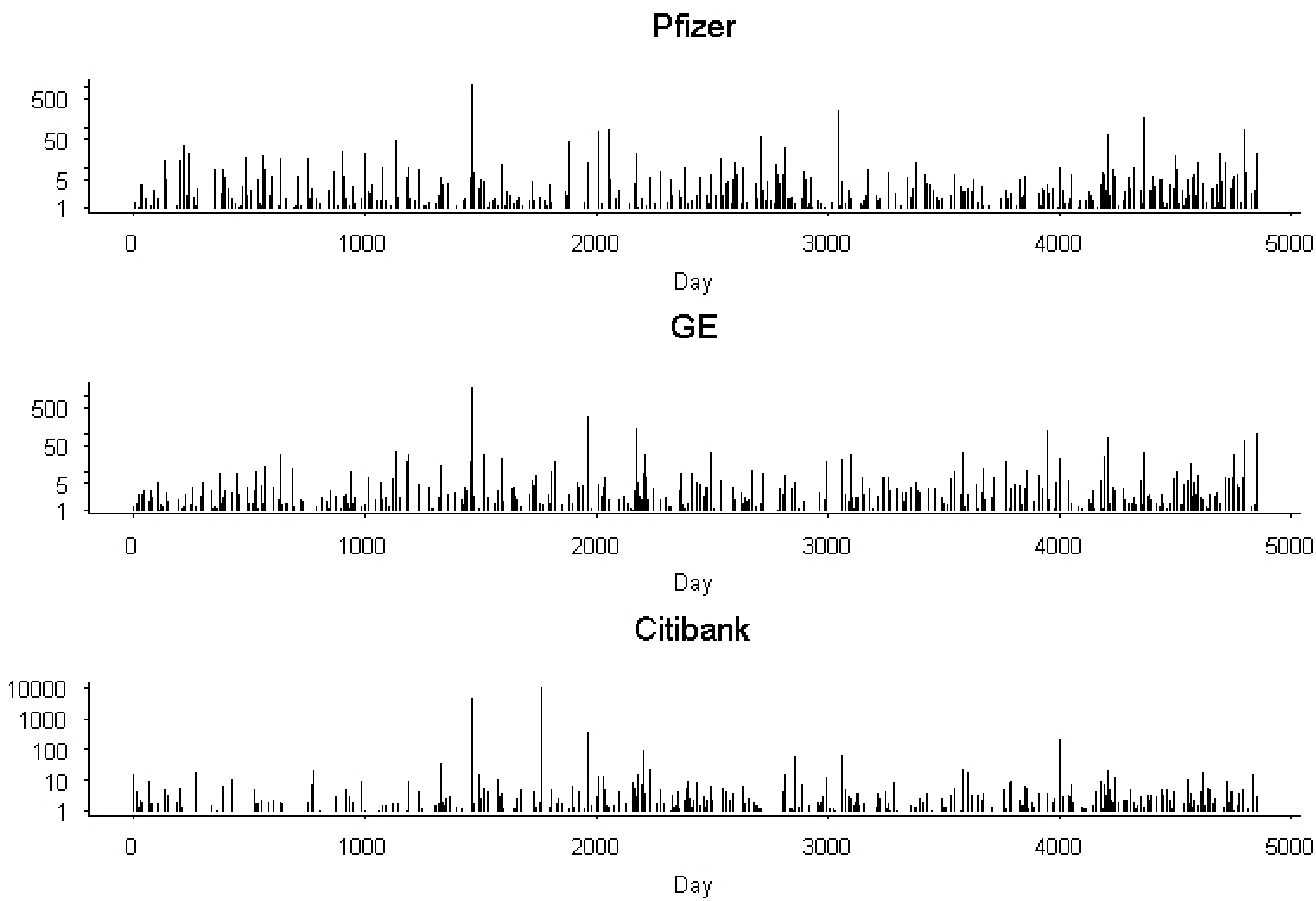
We return to the three series of daily returns for Pfizer, GE and Citibank, considered in the introduction.

Initial analysis: Fit a GARCH(1,1) model to each series, divide the series by the estimated volatility to get an approximately volatility-standardized series.

Then perform a threshold analysis on each series, transform series above the threshold to unit Fréchet margins.

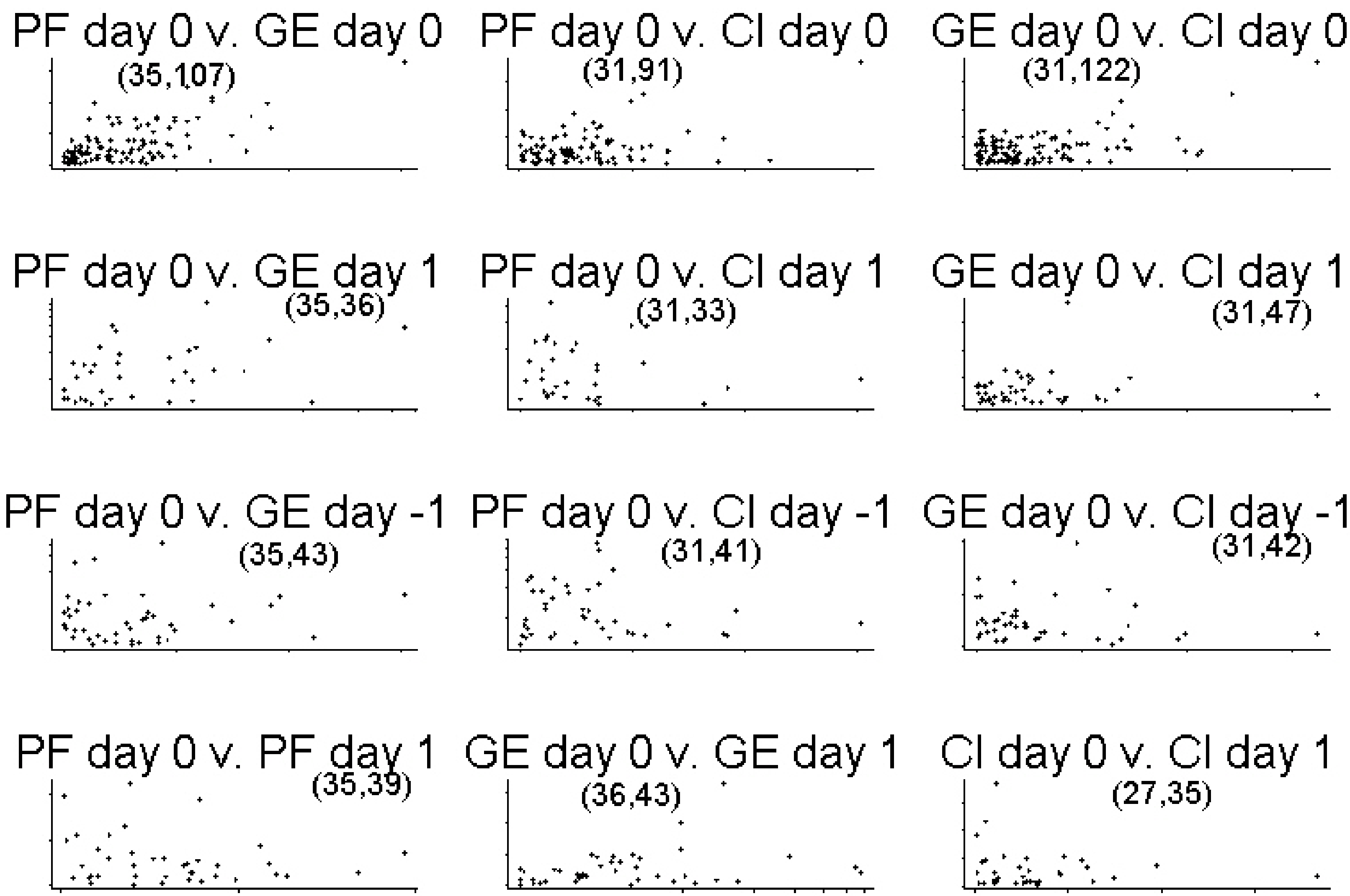
Our interest is in the dependence (both between and within series) among the resulting transformed threshold exceedances — specifically, what kind of clustering occurs among these exceedances?

Neg Daily Returns on Frechet scale



On this transformed scale, pairwise scatterplots are shown of the three series against each other on the same day (top 3 plots), and against series on neighboring days. The two numbers on each plot show the expected number of joint exceedances based on an independence assumption, and the observed number of joint exceedances.

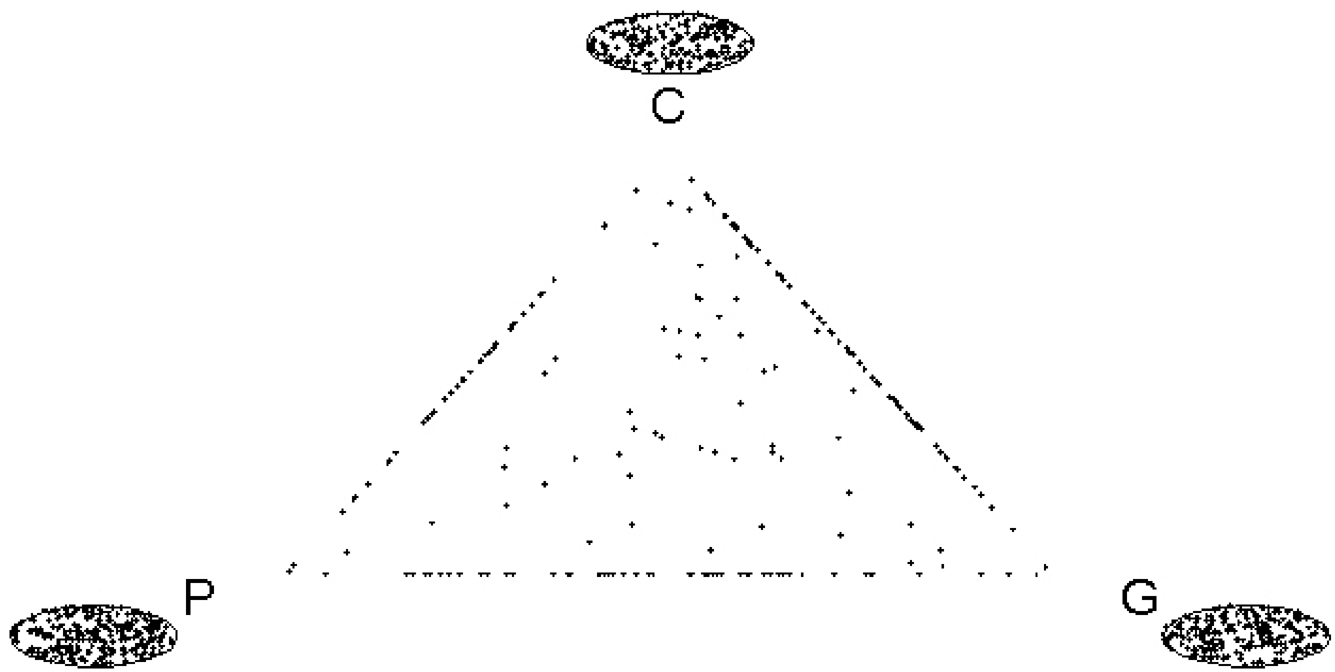
Exceedances on Frechet scale



Also shown is a plot of Fréchet exceedances for the three series on the same day, normalized to have total 1, plotted in barycentric coordinates. The three circles near the corner points P, G and C correspond to days for which that series along had an exceedance.

An M4 process was fitted to these three series by the technique described on an earlier slide.

Normalized Frechet exceedances



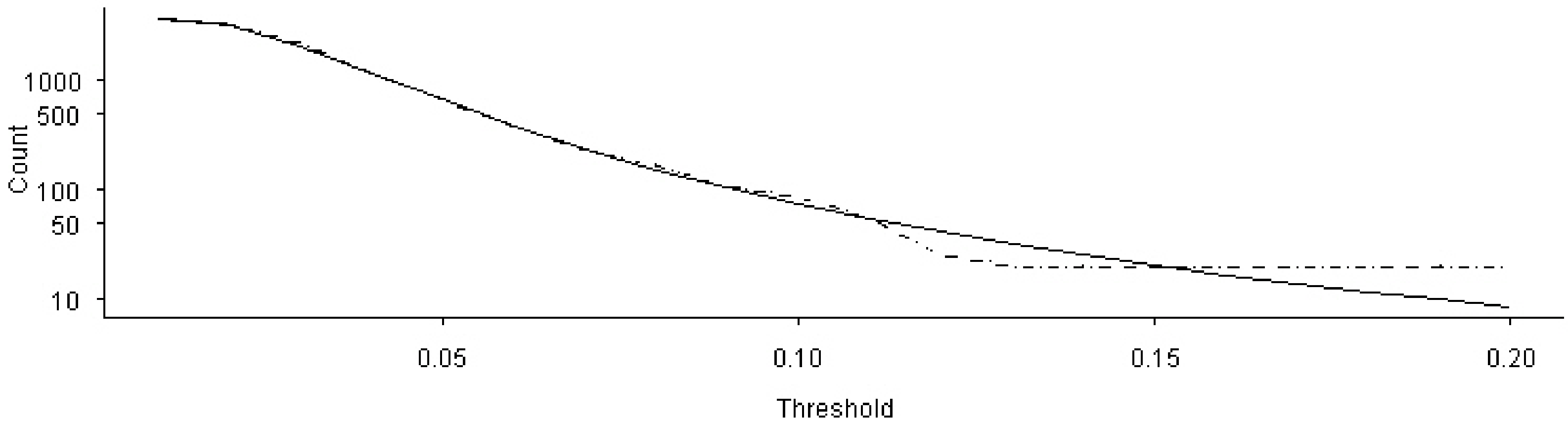
Finally, we attempt to validate the model by calibrating observed *vs.* expected probabilities of extreme events under the model.

The "extreme event" considered is that there is at least one exceedance of a specific threshold u by one of the three series in one of the next 10 days after a given day.

To make the comparison honest, the period of study is divided into four periods each of length just under 5 years. The univariate and multivariate EV model is fitted to each of the first three 5-year period, and used to predict extreme events in the following period.

The final plot shows observed (dashed lines) and expected (solid lines) counts for a sequence of thresholds u. There is excellent agreement between the two curves.

Theoretical (solid lines) and observed (dashed lines) counts of extreme events

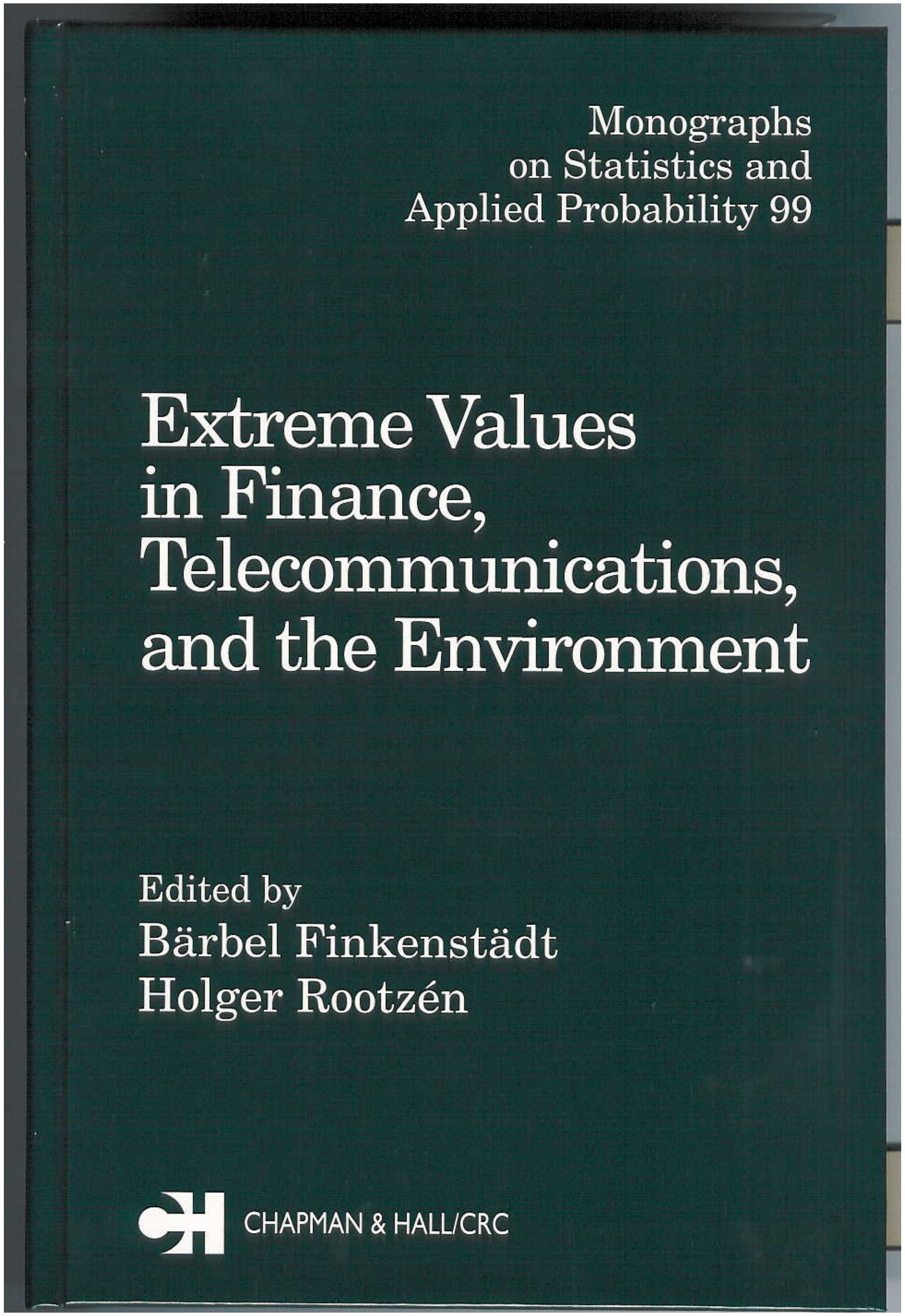


Conclusions

The representation in term of M4 processes contains the possibility of estimating both within-series and between-series dependence as part of the same model.

The key step in this method is the use of K-means clustering to identify a measure in a high-dimensional simplex of normalized exceedances. In contrast, existing methods of estimating multivariate extreme value distributions usually only work in low dimensions (up to 5 or so).

Ultimately the test of such methods will be whether they can be used for more reliable risk calculations than established methods such as RiskMetrics. The numerical example at the end shows that good progress has been made, but there are also many variations on the basic method which deserve to be explored.


Monographs
on Statistics and
Applied Probability 99
Extreme Values
in Finance,
Telecommunications,
and the Environment
Edited by
Bärbel Finkenstädt
Holger Rootzén
CHAPMAN & HALL/CRC