Modeling financial time series data as moving maxima processes

Zhengjun Zhang and Richard L. Smith
Department of Statistics
University of North Carolina
Chapel Hill, NC 27599-3260
2001 NBER/NSF time series conference
Raleigh, NC. USA

September 21, 2001

Abstract

Studies have shown time series data from finance, insurance and environment etc. are fat tailed and clustered when extremal events occur. In order to characterize such extremal processes, max-stable processes or min-stable processes have been proposed since the 1980s and some probabilistic properties have been obtained, but applications are very limited due to the lack of efficient statistical estimation methods. Zhang (2001) have shown some probabilistic properties of the processes and proposed a series of estimation procedures to estimate the underlying max-stable processes, i.e. multivariate maxima of moving maxima processes. This work illustrates how to model financial data as moving maxima processes. Examples will be illustrated with GE, Citibank, Pfizer stock index data.

1 Introduction

There have been many attempts to characterize the possible limits for multivariate extreme value distributions with unit Fréchet margins. Examples include the point process approach of de Haan and Resnick (1977), de Haan (1985) and Resnick(1987), and Pickands's (1981) representation theorem. Although there are well-developed approaches to model univariate extremal processes, problems concerning the environment, finance and insurance etc. are multivariate in character: for example, floods may occur at different sites along a coastline; the failure of a portfolio management may be caused by a single extreme price movement or multiple movements. Here multivariate extreme modeling is essential for risk management.

Davis and Resnick (1989) studied what they called the max-autoregressive moving average (MARMA(p,q)) process of a stationary process $\{X_n\}$ which satisfy the MARMA recursion,

$$X_n = \phi_1 X_{n-1} \vee \cdots \vee \phi_p X_{n-p} \vee Z_n \vee \theta_1 Z_{n-1} \vee \cdots \vee \theta_q Z_{n-q}$$

for all n where $\phi_i, \theta_j \geq 0, 1 \leq i \leq p, 1 \leq j \leq q$ and $\{Z_n\}$ is *i.i.d.* with common distribution function $F(x) = \exp\{-\sigma x^{-1}\}$. The case $\sigma = 1$ is unit Fréchet. For any given $\{\phi_i\}, \{\theta_j\}$, the corresponding process is a max-stable process. A naive estimation procedure for ϕ_i, θ_j's when $q = 1$ is given in Davis and Resnick(1989).

Deheuvels (1983) defined what he called the moving minimum (MM) corresponding process as

$$T_n = \min\{\delta_k Z_{n-k}, -\infty < k < \infty\}, -\infty < n < \infty,$$

where $\delta_k > 0$, and $\{Z_k\}$ is *i.i.d.* exponential 1. If T_n is MM, then $X_n = 1/T_n$ is a MARMA process with $p = 0$ and $q = \infty$. If a MARMA(p,q) X_n is causal, a process is called causal if there exist non-negative constants a_i such that $\max\limits_{i\geq 0}(a_i Z_{n-i}) < \infty$ a.s. and $X_n = \max\limits_{i\geq 0}(a_i Z_{n-i})$, then $T_n = 1/X_n$ is MM.

Smith and Weissman (1996) extended this definition to a more general framework which is more realistic and is called multivariate maxima of moving maxima (henceforth M4) process. The definition is

$$Y_{id} = \max_l \max_k a_{lkd} Z_{l,i-k}, \quad d = 1, \cdots, D, \tag{1}$$

where $\{Z_{li}, l \geq 1, -\infty < i < \infty\}$ are an array of independent unit Fréchet random variables. The constants $\{a_{lkd}, l \geq 1, -\infty < k < \infty, 1 \leq d \leq D\}$ are nonnegative constants satisfying

$$\sum_{l=1}^{\infty} \sum_{k=-\infty}^{\infty} a_{lkd} = 1 \text{ for } d = 1, \ldots, D \tag{2}$$

Notice that $M4$ processes deal with D dimensional random processes whereas MM processes deal with univariate processes ($D = 1$). Although MM processes are only specified over one index there are possibilities to easily extend to over two indexes. The main theorem of Smith and Weissman (1996) shows that any max-stable process without deterministic component may be approximated arbitrarily closely by an $M4$ process. Following de Haan(1984), (1) defines max-stable processes because for any finite number r and positive constants $\{y_{id}\}$ we have

$$\begin{aligned}
&\Pr\{Y_{id} \leq y_{id}, 1 \leq i \leq r, 1 \leq d \leq D\} \\
&\quad = \Pr\{Z_{l,i-k} \leq \tfrac{y_{id}}{a_{lkd}} \text{ for } l \geq 1, -\infty < k < \infty, 1 \leq i \leq r, 1 \leq d \leq D\} \\
&\quad = \Pr\{Z_{l,m} \leq \min_{1-m\leq k\leq r-m} \min_{1\leq d\leq D} \tfrac{y_{m+k,d}}{a_{lkd}}, l \geq 1, -\infty < m < \infty\} \\
&\quad = \exp[-\textstyle\sum_{l=1}^{\infty} \sum_{m=-\infty}^{\infty} \max\limits_{1-m\leq k\leq r-m} \max\limits_{1\leq d\leq D} \tfrac{a_{lkd}}{y_{m+k,d}}]
\end{aligned} \tag{3}$$

This is (2.5) of Smith and Weissman (1996) and we have

$$\Pr^n\{Y_{id} \le ny_{id}, 1 \le i \le r, 1 \le d \le D\} = \Pr\{Y_{id} \le y_{id}, 1 \le i \le r, 1 \le d \le D\}$$

which is independent of n, so the finite-dimensional distributions of $\{\mathbf{Y}_i\}$ are max-stable (also called multivariate extreme value).

Recently Hall, Peng and Yao (2001) discussed moving maximum models

$$Y_i = \sup\{a_{j-i}Z_i,\ -\infty < i < \infty\}$$

where the distribution of Z_i is assumed either $F(z|\theta) = \exp(-z^{-\theta})$ or the generalized Pareto distribution $F(z|\theta) = 1-(1+z)^{-\theta}$. Then for a finite number of parameters, they choose $(\theta, a_{(m)})$ to minimize

$$\begin{aligned} D_m(\theta, a_{(m)}) = \textstyle\int (\widehat{G}(y) - \prod_{i=2-m}^{k} & F[\min\{a_{j-i}^{-1}y_j, \\ & \max(i,1) \le j \le \min(i+m,k)\}|\theta])^2 w(y)dy, \end{aligned} \tag{4}$$

where the integral is over $y = (y_1, \dots, y_k) \in \mathbb{R}_+^k$ and

$$\widehat{G}(y) = (n-k)^{-1}\sum_{i=1}^{n-k} I_{(Y_{i+j-1} \le y_j \text{ for } 1 \le j \le k)}, \tag{5}$$

and w is a nonnegative weight function. We state their main theorem as follows.

Theorem 1.1 *Under conditions*

- *F has support on the positive half-line, and is in the domain of attraction of a Type III extreme value distribution.*
- *each a_i is nonnegative and, for some $\epsilon \in (0, r)$, $0 < \sum_i a_i^{r-\epsilon} < \infty$.*

then

$$\sup_{-\infty < y_1, \dots, y_k < \infty} |\Pr(Y_1^* \le y_1, \dots, Y_k^* \le y_k | Y_1, \dots, Y_n) - \Pr(Y_1 \le y_1, \dots, Y_k \le y_k)| \to 0 \tag{6}$$

where Y_j^* is defined by

$$Y_j^* = \sup\{\widehat{a}_{j-i}Z_i^*,\ -\infty < i < \infty\}$$

$\widehat{a}_{j-i}$ and $\widehat{\theta}$ are solutions of (4) and Z_i^* has distribution function $F(.|\widehat{\theta})$. Moreover, if $m \ge C_4(\log n)^2$ for C_4 sufficiently large, the rate of convergence in (6) is $O_p(n^{-(1/2)+\delta})$ for all $\delta > 0$.

Our present work on the estimation of $M4$ processes is somewhat parallel to Hall et al. (2001)'s work. In contrast to the bootstrapped processes which Hall et al. (2001) used to construct confidence intervals and prediction intervals, we directly construct parameter estimators and prove their asymptotic properties.

In practice we usually have $l = 1, \cdots, L$ and $-K_1 \le k \le K_2$ for some finite numbers L, K_1 and K_2. When an extreme event occurs or when a large Z_{li} occurs, $Y_{id} \propto a_{l,i-k,d}$ for $i \approx k$, i.e. if some Z_{lk} is much larger than all neighboring Z values, we will have $Y_{id} = a_{l,i-k,d} Z_{lk}$ for i near k. This indicates a moving pattern of the time series, call it signature pattern. Hence L corresponds to the maximum number of signature patterns. And K_1 and K_2 characterize the range of the sequence dependence. $K_1 + K_2 + 1$ is the order of maxima moving processes. We will focus on the finite dimensional $M4$ process,

$$Y_{id} = \max_{1 \le l \le L} \max_{-K_1 \le k \le K_2} a_{lkd} Z_{l,i-k}, \quad d = 1, \cdots, D, \tag{7}$$

where$\sum_{l=1}^{L} \sum_{k=-K_1}^{K_2} a_{lkd} = 1$ for $d = 1, \cdots, D$.

Under model (7), it is easy to obtain the finite distribution of $\{Y_{id},\ 1 \le i \le r,\ 1 \le d \le D\}$ from (3). The goal is to estimate all parameters $\{a_{lkd}\}$ under the constraints that the parameters are nonnegative and the summation is equal to 1 for each $d = 1, \ldots, D$.

Due to the degeneracy of the multivariate joint distribution function of the $M4$ processes, the method of maximum likelihood is not applicable in this instance. We are going to develop estimators based on the joint empirical distribution function.

The model can be rewritten as

$$\begin{aligned} Y_i &= \max_{1 \le l \le L} \max_{-K_1 \le k \le K_2} a_{lk} Z_{l,i-k} \\ &= \max_{1 \le l \le L} b_l \max_{-K_1 \le k \le K_2} c_{lk} Z_{l,i-k} \end{aligned} \tag{8}$$

where b_l is the weight of l's signature pattern and such that $\sum_l b_l = 1$ and $\sum_k c_{lk} = 1$ for each l. In section 2 we illustrate procedures to estimate c_{lk}s first and then to estimate b_ls after c_{lk}s are obtained. In section 3 we illustrate how to estimate a_{lk}s directly. In section 4 we give two examples. The first example is to show the effectiveness of the estimating procedures when applying to an $M4$ process and a mixed process of $M4$ and Gaussian noise. The second example is to demonstrate how we model multivariate financial time series as $M4$ processes.

2 Estimation based on two steps

When we plot the time series of an $M4$ process, we expect to have L signature patterns on the plot. Zhang (2001) shows that $\frac{Y_{t+m}}{Y_t + Y_{t+1} + \cdots + Y_{t+K_2+K_1}} = c_{l,-K_1+m}$ infinitely many times in $-\infty < t < \infty$ for each $m \in \{0, 1, \ldots, K_1 + K_2\}$, if Y_t follows an $M4$ process. These values $c_{l,m}$ lead to estimates of $\frac{a_{l,-K_1+m}}{a_{l,-K_1} + a_{l,-K_1+1} + \cdots + a_{l,K_2}}$, $1 \le l \le L$. But these estimates are only defined if $M4$ model really is an exact model for the process. However, if the

$M4$ process is viewed as an approximation to a max-stable process, rather than an exact model, such estimators do not make sense since we would not expect to have any two values of $\frac{Y_{t+m}}{Y_t+Y_{t+1}+\cdots+Y_{t+K_2+K_1}}$ are the same. Example 1 in section 4 shows the existence of such phenomenon. As an example, we consider an $M4$ process with 3 signature patterns here. Figure 1 shows three different significant patterns (points fall onto 3 horizontal lines) which correspond to $L = 3$. As we have already seen, the plots give accurate

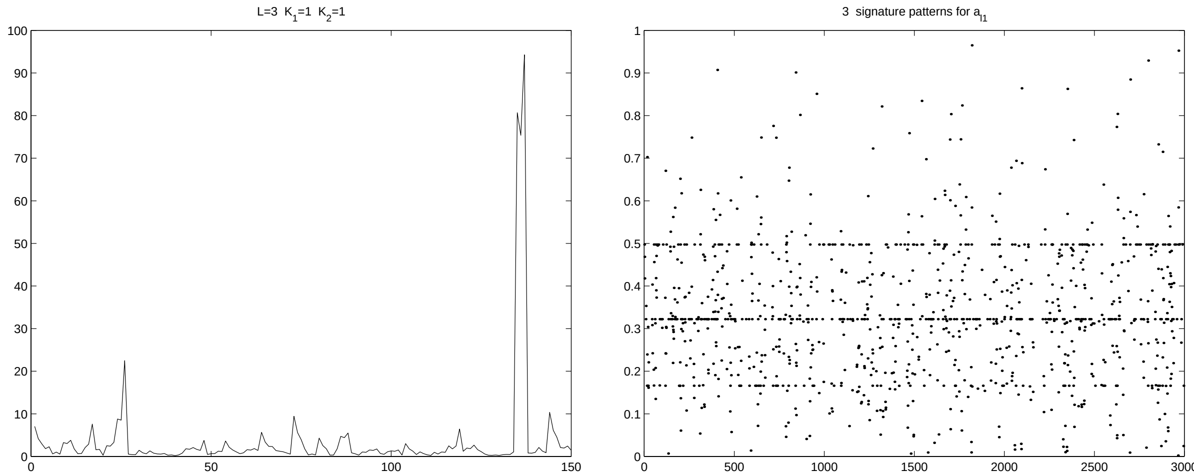

Figure 1: A demo of multiple signature patterns.

estimates of the ratios. When $L = 1$, we can get all the exact values of a_k. But for $L > 1$ we cannot. Even for $L = 1$, we have assumed that the model assumptions are exactly satisfied, not something we would expect to use in practice. Also, the whole method presupposes that the margins are transformed into unit Fréchet margin and this wouldn't be exact in practice, either. We need to develop estimation procedures to obtain estimates of a_{lk} in a more practical way. We will study this in section 3.

We now assume that the model assumptions are exactly satisfied but L is greater than 1.

It follows immediately from (3) that $P(Y_1 \le y_1) = e^{-\frac{1}{y_1}}$ and

$$P(Y_1 \le y_1, Y_2 \le y_2) = \exp[-\sum_{l=1}^{L} b_l \sum_{m=1-K_2}^{2+K_1} \max(\frac{c_{l,1-m}}{y_1}, \frac{c_{l,2-m}}{y_2})] \tag{9}$$

where $c_{l,K_2+1} = 0, c_{l,-K_1-1} = 0$

These results tell that if we only consider the marginal distribution function it makes no sense to our estimation problem since it has no parameter at all. Through the whole work, we now project each original observed process into a bivariate process. And all the theories and methods are developed under the bivariate process.

Let $A_1 = (0, x_1) \times (0, x_1'), \cdots, A_{L-1} = (0, x_{L-1}) \times (0, x_{L-1}')$ be different and define

$$\bar{X}_{A_j} = \frac{1}{n} \sum_{i=1}^{n} I_{A_j}(Y_i, Y_{i+1}) \tag{10}$$

Since we consider finite order moving process, we can choose pairs (Y_i', Y_{i+1}') from the observed process such that those pairs form an $i.i.d$ time series sequence. For example, let $Y_1' = Y_1, Y_2' = Y_2, Y_3' = Y_{K_1+K_2+3}, Y_4' = Y_{K_1+K_2+4}, \dots$, and so on. Then by the strong law of large numbers, we have

$$\bar{X}_{A_j} \xrightarrow{a.s.} P(A_j) = P(Y_1' \le x_j, Y_2' \le x_j'). \tag{11}$$

In fact, it is not necessary to insist on the pairs (Y_i', Y_{i+1}') being independent if we use all pairs (Y_i, Y_{i+1}), by strong law for M-dependent sequences we also have

$$\bar{X}_{A_j} \xrightarrow{a.s.} P(A_j) = P(Y_1 \le x_j, Y_2 \le x_j'). \tag{12}$$

Now let

$$\exp[-\sum_{l=1}^{L} \widehat{b}_l \sum_{m=1-K_2}^{2+K_1} \max(\frac{c_{l,1-m}}{x_j}, \frac{c_{l,2-m}}{x_j'})] = \bar{X}_{A_j},\ j = 1, \cdots, L-1 \tag{13}$$

then we can construct parameter estimators by solving a system of linear equations

$$\begin{cases} \sum_{l=1}^{L} \widehat{b}_l \sum_{m=1-K_2}^{2+K_1} \max(\frac{c_{l,1-m}}{x_1}, \frac{c_{l,2-m}}{x_1'}) & = & -\log(\bar{X}_{A_1}) \\ & \vdots & \\ \sum_{l=1}^{L} \widehat{b}_l \sum_{m=1-K_2}^{2+K_1} \max(\frac{c_{l,1-m}}{x_{L-1}}, \frac{c_{l,2-m}}{x_{L-1}'}) & = & -\log(\bar{X}_{A_{L-1}}) \\ \sum_{l=1}^{L} \widehat{b}_l = 1 & & \end{cases} \tag{14}$$

We have

Theorem 2.1

$$\sqrt{n}(\widehat{\mathbf{b}} - \mathbf{b}) \xrightarrow{d} N(0, D(\Sigma + \sum_{k=1}^{K_1+K_2+1} (W_k + W_k'))D'),$$

where

$$\widehat{\mathbf{b}} = \begin{bmatrix} \widehat{b}_1 \\ \widehat{b}_2 \\ \vdots \\ \widehat{b}_L \end{bmatrix}, \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_L \end{bmatrix}, D = \begin{bmatrix} d_{11}/\mu_1 & d_{12}/\mu_2 & \cdots & d_{1,L-1}/\mu_{L-1} \\ d_{21}/\mu_1 & d_{22}/\mu_2 & \cdots & d_{2,L-1}/\mu_{L-1} \\ \vdots & \vdots & \ddots & \vdots \\ d_{L,1}/\mu_1 & d_{L,2}/\mu_2 & \cdots & d_{L,L-1}/\mu_{L-1} \end{bmatrix}.$$

$\mu_i = \Pr(Y_1 \le 1, Y_2 \le x_i)$, $\mu_{ij} = \Pr(Y_1 \le 1, Y_2 \le \min(x_i, x_j))$, *the elements of* Σ*:* $\sigma_{ij} = \mu_{ij} - \mu_i\mu_j$,
$w_k^{ij} = \Pr(Y_1 \le 1, Y_2 \le x_i, Y_{1+k} \le 1, Y_{2+k} \le x_j) - \mu_i\mu_j$, $\mu_{ii} = \mu_i$. *and the* d_{lk}*s are elements of the inverse of the matrix constructed from (14).*

Note: the limiting covariance matrix is singular because $\sum_{l=1}^{L} \widehat{b}_l = 1$.

3 Direct Estimation

The methods developed in previous sections may be just the ideal. In practice we may not be able to estimate the ratios accurately as stated in the previous section, especially when the data are with error or the process is a mix of $M4$ and noise. In this section, we will develop methods which can be applied to estimate a_{lk} directly. The estimators are the solution of the system of nonlinear equations.

$$\begin{cases} \sum_{l=1}^{L}\sum_{m=1-K_2}^{2+K_1} \max(\frac{\widehat{a}_{l,1-m}}{x_1}, \frac{\widehat{a}_{l,2-m}}{x'_1}) & = & -\log(\bar{X}_{A_1}) \\ \sum_{l=1}^{L}\sum_{m=1-K_2}^{2+K_1} \max(\frac{\widehat{a}_{l,1-m}}{x_2}, \frac{\widehat{a}_{l,2-m}}{x'_2}) & = & -\log(\bar{X}_{A_2}) \\ & \vdots & \\ \sum_{l=1}^{L}\sum_{m=1-K_2}^{2+K_1} \max(\frac{\widehat{a}_{l,1-m}}{x_{L\times(K_1+K_2+1)}}, \frac{\widehat{a}_{l,2-m}}{x'_{L\times(K_1+K_2+1)}}) & = & -\log(\bar{X}_{A_{L\times(K_1+K_2+1)}}) \end{cases}$$

For the simplicity, consider now the case of $L = 1$ and study the structure of the bivariate distribution function

$$P(Y_1 \le y_1, Y_2 \le y_2) = \exp[-\sum_{m=1-K_2}^{2+K_1} \max(\frac{a_{1-m}}{y_1}, \frac{a_{2-m}}{y_2})] \tag{15}$$

where $a_{K_2+1} = 0, a_{-K_1-1} = 0$. Now define

$$q(x) = a_{-K_1} + \sum_{j=-K_1}^{K_2-1} \max(xa_j, a_{j+1}) + xa_{K_2}, \tag{16}$$

then $P(Y_1 \le 1, Y_2 \le x) = \exp[-q(x)/x]$.
As an example, consider the process

$$Y_i = \max(\frac{1}{3}Z_{i-2}, \frac{1}{6}Z_{i-1}, \frac{1}{6}Z_i, \frac{1}{3}Z_{i+1})$$

In this case, $q(x)$ is as in Figure 2. The plot tells at each change point the slope of the $q(x)$ changes to a larger value and is eventually 1. Each change point corresponds to a ratio of two adjacent parameters. The plot also enable us to choose a finite number of points such that we can draw piecewise straight line and hence determine the parameter ratios and solve for their values.

The following proposition gives a sufficient condition that we can determine all parameter values from the $q(x)$ function.

Proposition 3.1 *If all $\binom{K_1+K_2+1}{2}$ ratios $\frac{a_j}{a_{j'}}$ are distinct, the model is uniquely identified by $q(x)$.*

Zhang (2001) has studied general cases on $q(x)$ and some complicated models. The following results are restrict to the case of $L = 1$. The proofs and more general cases can be found in Zhang (2001).

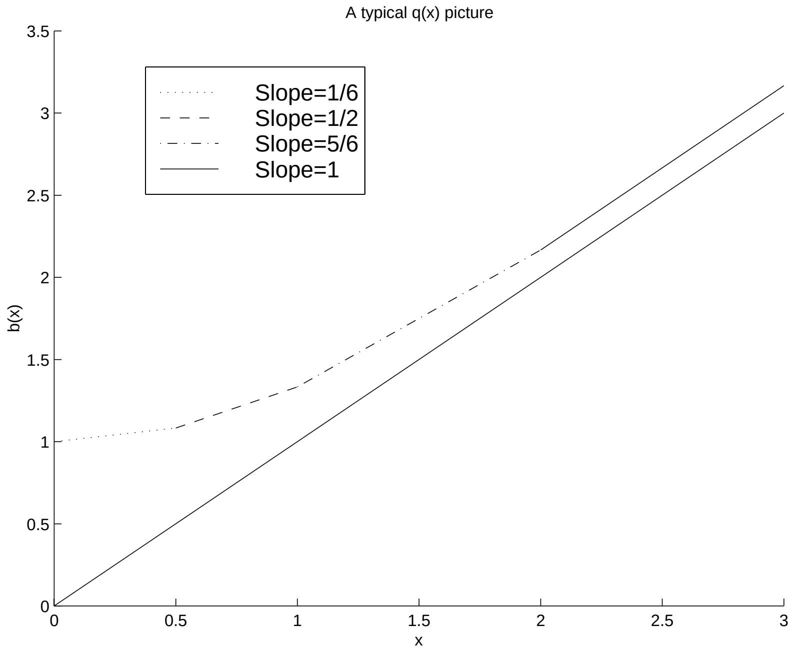

Figure 2: A demo of $q(x)$ and its slope $q'(x)$ and ratio change points.

We define

$$\begin{aligned} a_{K_2} &+ \max(a_{K_2-1}, \frac{a_{K_2}}{x}) + \max(a_{K_2-2}, \frac{a_{K_2-1}}{x}) \\ &+ \max(a_{K_2-3}, \frac{a_{K_2-2}}{x}) + \cdots + \max(a_{-K_1}, \frac{a_{-K_1+1}}{x}) + \frac{a_{-K_1}}{x} = b(x) \end{aligned} \tag{17}$$

then we have $q(x) = xb(x)$. Let $Y_1, Y_2, \cdots, Y_n$ be observed values and

$$\widehat{b}(x) = -\log(\frac{1}{n}\sum_{i=1}^{n} I_{(Y_i \le 1, Y_{i+1} \le x)}), \tag{18}$$

then $\widehat{q}(x) = x\widehat{b}(x)$.

Suppose the solutions of

$$\left\{ \begin{array}{lll} a_{K_2} & + & \max(a_{K_2-1}, \frac{a_{K_2}}{x_1}) + \max(a_{K_2-2}, \frac{a_{K_2-1}}{x_1}) \\ & + & \max(a_{K_2-3}, \frac{a_{K_2-2}}{x_1}) + \cdots + \max(a_{-K_1}, \frac{a_{-K_1+1}}{x_1}) + \frac{a_{-K_1}}{x_1} = \widehat{b}_\omega(x_1) \\ \cdots & & \cdots \\ a_{K_2} & + & \max(a_{K_2-1}, \frac{a_{K_2}}{x_m}) + \max(a_{K_2-2}, \frac{a_{K_2-1}}{x_m}) \\ & + & \max(a_{K_2-3}, \frac{a_{K_2-2}}{x_m}) + \cdots + \max(a_{-K_1}, \frac{a_{-K_1+1}}{x_m}) + \frac{a_{-K_1}}{x_m} = \widehat{b}_\omega(x_m) \end{array} \right. \tag{19}$$

are $\widehat{\mathbf{a}}_\omega$. This is equivalent to $C_{n\omega}\widehat{\mathbf{a}}_\omega = \widehat{\mathbf{b}}_\omega$ in matrix notations where $C_{n\omega}$ is uniquely determined by $\widehat{\mathbf{a}}_\omega$. The elements of $C_{n\omega}$ are either 1, $\frac{1}{x_i}$ or $1+\frac{1}{x_i}$. And the solutions of

$$\begin{cases} a_{K_2} & + \max(a_{K_2-1}, \frac{a_{K_2}}{x_1}) + \max(a_{K_2-2}, \frac{a_{K_2-1}}{x_1}) \\ & + \max(a_{K_2-3}, \frac{a_{K_2-2}}{x_1}) + \cdots + \max(a_{-K_1}, \frac{a_{-K_1+1}}{x_1}) + \frac{a_{-K_1}}{x_1} = b(x_1) \\ \cdots & \cdots \\ a_{K_2} & + \max(a_{K_2-1}, \frac{a_{K_2}}{x_m}) + \max(a_{K_2-2}, \frac{a_{K_2-1}}{x_m}) \\ & + \max(a_{K_2-3}, \frac{a_{K_2-2}}{x_m}) + \cdots + \max(a_{-K_1}, \frac{a_{-K_1+1}}{x_m}) + \frac{a_{-K_1}}{x_m} = b(x_m) \end{cases} \tag{20}$$

are $\mathbf{a}$.

Theorem 3.2 *Suppose the model is identifiable from $x_1, \ldots, x_m$, where these values are different from the ratios of true parameters, then the solutions of (19) converge to the solutions of (20) almost surely. i.e.* $\widehat{\mathbf{a}} \xrightarrow{a.s.} \mathbf{a}$ *and* $C_n \xrightarrow{a.s.} C$.

Theorem 3.3 *Suppose the model is identifiable from $x_1, \ldots, x_m$, where these values are different from the ratios of true parameters, then*

$$\sqrt{n}(\widehat{\mathbf{a}} - \mathbf{a}) \xrightarrow{d} N(0, BD(\Sigma + \sum_{k=1}^{K_1+K_2+1} (W_k + W_k'))D'B')$$

where $B = (C'C)^{-1}C$, $\mu_i = \Pr(Y_1 \le 1, Y_2 \le x_i)$, $\mu_{ij} = \Pr(Y_1 \le 1, Y_2 \le \min(x_i, x_j))$, $\sigma_{ij} = \mu_{ij} - \mu_i\mu_j$, $w_k^{ij} = \Pr(Y_1 \le 1, Y_2 \le x_i, Y_{1+k} \le 1, Y_{2+k} \le x_j) - \mu_i\mu_j$, $\mu_{ii} = \mu_i$, *and* $D = diag\{\frac{1}{\mu_1}, \cdots, \frac{1}{\mu_m}\}$.

4 Examples

In this section we give two examples. The first example illustrates a simulated $M4$ process with two signature patterns where each pattern has order of 2 and we also add Gaussian noise. In the second one, we study negative stock price returns of three stock products, i.e. GE, Citibank and Pfizer.

Example 1. We perform two simulation experiments with the following two processes.

$$Y_i = \max(.1Z_{1,i-1}, .4Z_{1,i}, .35Z_{2,i-1}, .15Z_{2,i}) \tag{21}$$

and

$$Y_i = \max(.1Z_{1,i-1}, .4Z_{1,i}, .35Z_{2,i-1}, .15Z_{2,i}) + N_i \tag{22}$$

where $N_i \sim N(0, .01)$ are *i.i.d.*

We plot the ratios $\frac{Y_i}{Y_i+Y_{i+1}}$ for both models. Plots in Figure 3 look almost exactly the same. However, when a portion of the plot is magnified, as in Figure 4, we can see the difference.

We now apply estimating methods developed in previous sections and list all results in the following tables.

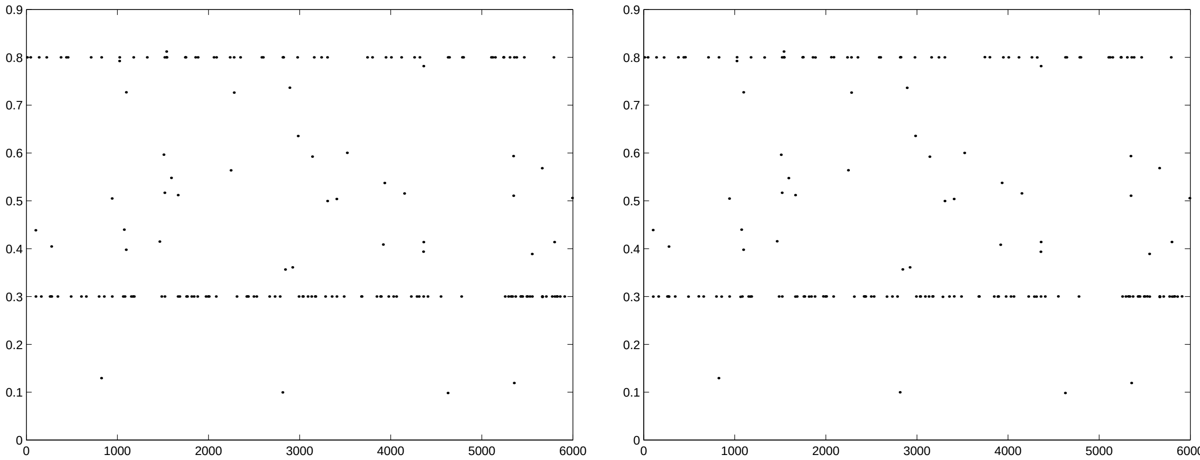

Figure 3: The left plot is the ratios of $\frac{Y_i}{Y_i+Y_{i+1}}$ at the threshold level 10 under the model (21). The right plot is the ratios of $\frac{Y_i}{Y_i+Y_{i+1}}$ at the threshold level 10 under the model (22).

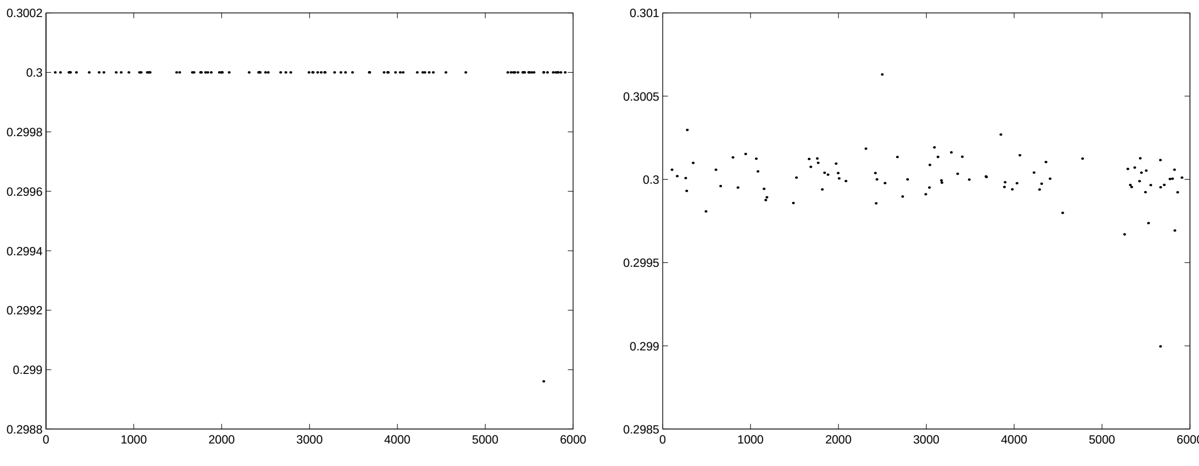

Figure 4: The left plot is the ratios around .3 with distance .01 at the threshold level 10 under the model (21). The right plot is the ratios .3 with distance .01 at the threshold level 10 under the model (22).

Table 1. Simulation results for model (21)

Parameter	$a_{1,-1}$	$a_{1,0}$	$a_{2,-1}$	$a_{2,0}$
True value	.1	.4	.35	.15
Estimated value	.0912	.3247	.3747	.2170
Variance	(1.8365)	(26.5205)	(32.4965)	(7.0202)

Table 2. Simulation results for model (22)

Parameter	$a_{1,-1}$	$a_{1,0}$	$a_{2,-1}$	$a_{2,0}$
True value	.1	.4	.35	.15
Estimated value	.0958	.3437	.3477	.2258
Variance	(1.8592)	(27.2590)	(33.7253)	(7.0585)

The estimated values are based on a sample of size 10000. The variances are obtained by evaluating the formula in Theorem 3.3 with the true values approximated by the empirical values. These simulation experiments show that the effectiveness of the estimating procedures proposed. Notice that the variance values are large. One main reason is that we use the empirical distribution function as the first step to estimate probability. Then we transform them into a logarithmic scale which results in a large variance since all probability are less than 1, especially for those small probability values which correspond to extreme events.

Example 2. In this example we model stock prices of GE, CITIBANK and Pfizer as an $M4$ process. Parameter estimates are based on a multivariate time series of approximately 7000 days. We first plot the three time series and transform the negative returns into Fréchet scale by first fitting all values above certain threshold(.02 is used in this study). Further diagnostics are applied in order to apply $M4$ process modeling.

The underlying idea behind these analysis of Figure 8 is the point-process approach to univariate extreme value modeling due to Smith (1989). Smith and Shively (1995) introduced a number of diagnostic devices to examine the fit of the generalized extreme value distributions. One idea is based on what we shall call Z-statistics

$$Z_k = \int_{T_{k-1}}^{T_k} \Lambda_s(u) ds$$

where T_k denotes the time of the k'th exceedance of u. $\Lambda_t(x)$ is given by

$$\Lambda_t(x) = (1 + \xi_t \frac{x - \mu_t}{\psi_t})_+^{-1/\xi_t}$$

the intensity of a nonhomogeneous Poisson process of exceedances of a level x. If the model is correct, then $Z_1, Z_2, \ldots,$ will be independent exponentially distributed random variables with mean 1. The Z-statistics are an indication of how closely the exceedances of a fixed level u are represented by a nonhomogeneous Poisson process, but they do not

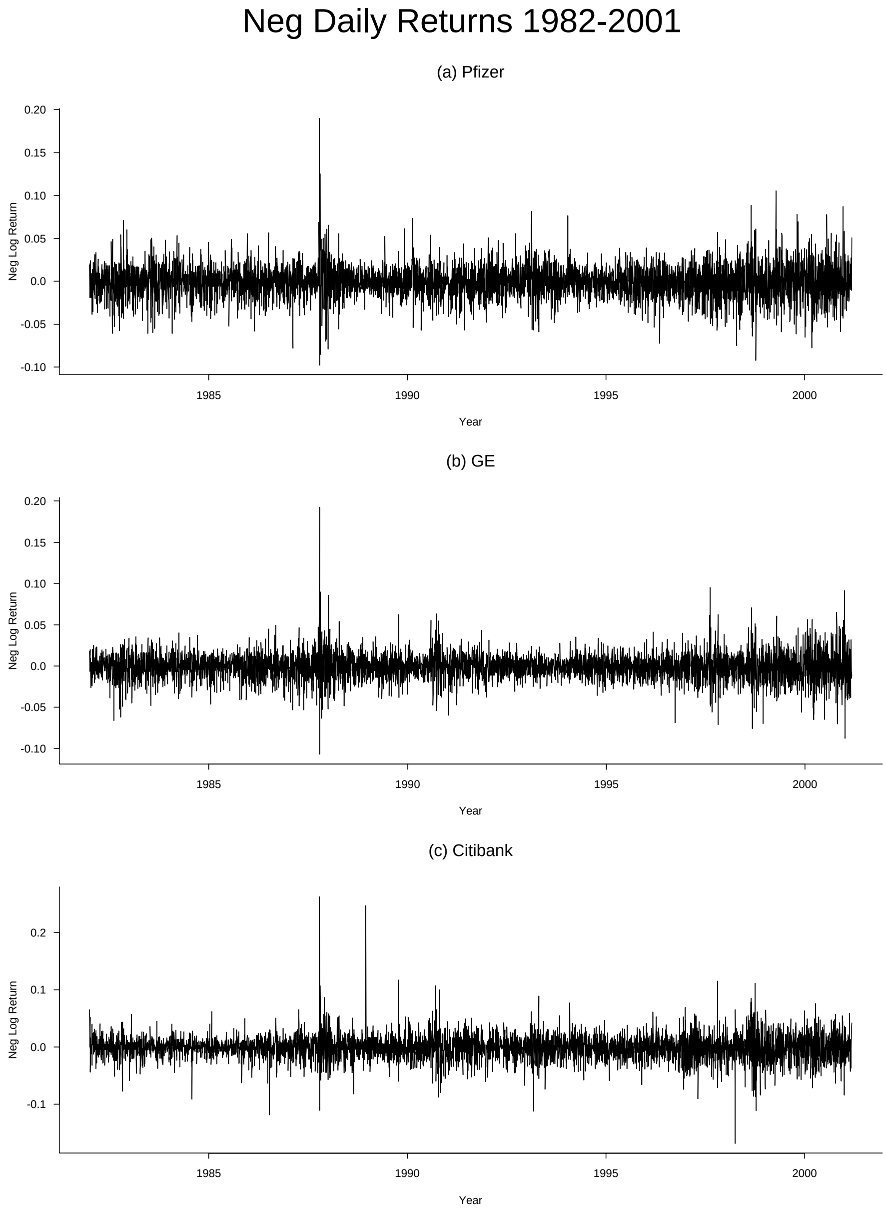

Figure 5: These figures show that there are extreme observations and the greatest drop happened in the same day in all three time series, i.e. October 19, 1987, the date of the Wall Street crash.

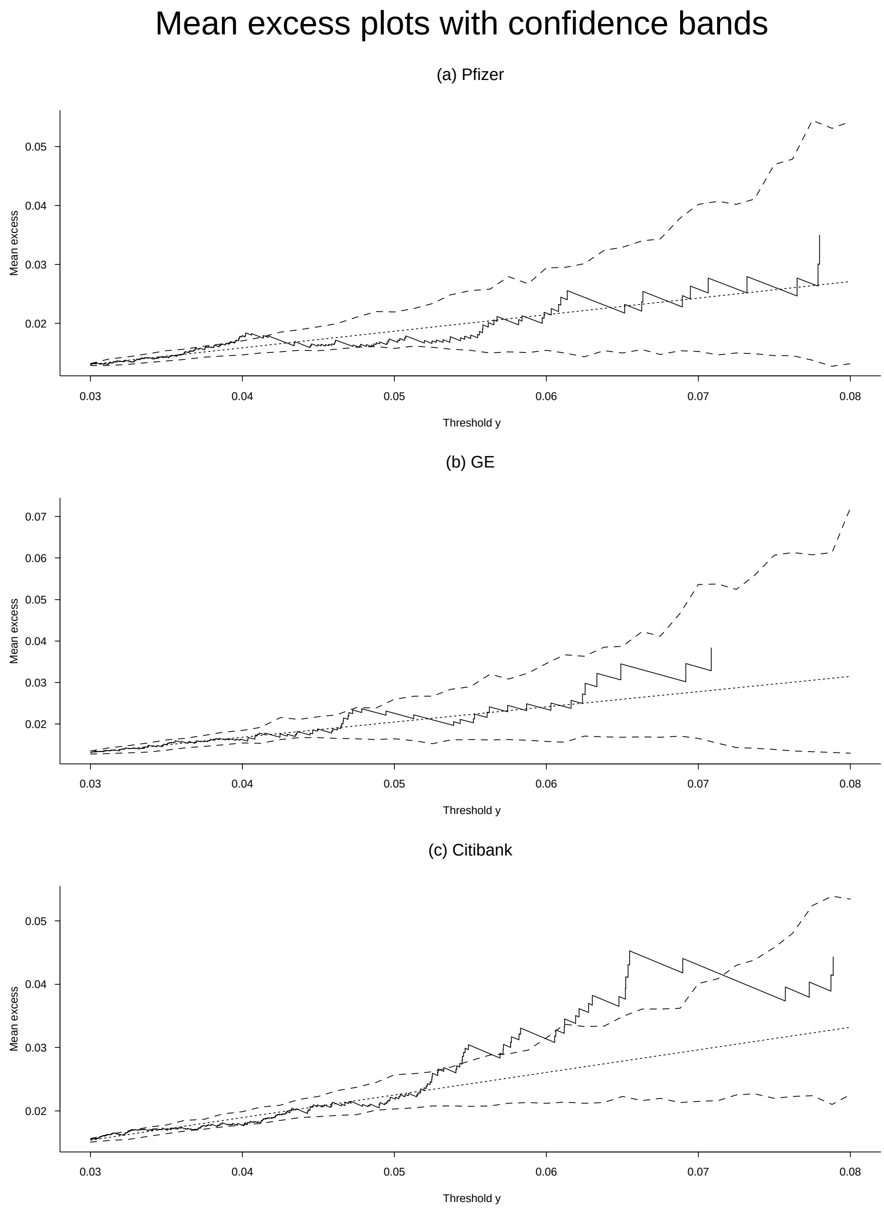

Figure 6: The mean excess plot usually suggests whether a extreme value distribution fitting is appropriate or not. The plot for the Pfizer data suggests extreme value distribution fitting since the plot is contained in its corresponding confidence interval. The other two are more doubtful since the plot goes outside the confidence bands, though further analysis shows that the extreme value approximation is reasonable in this case also.

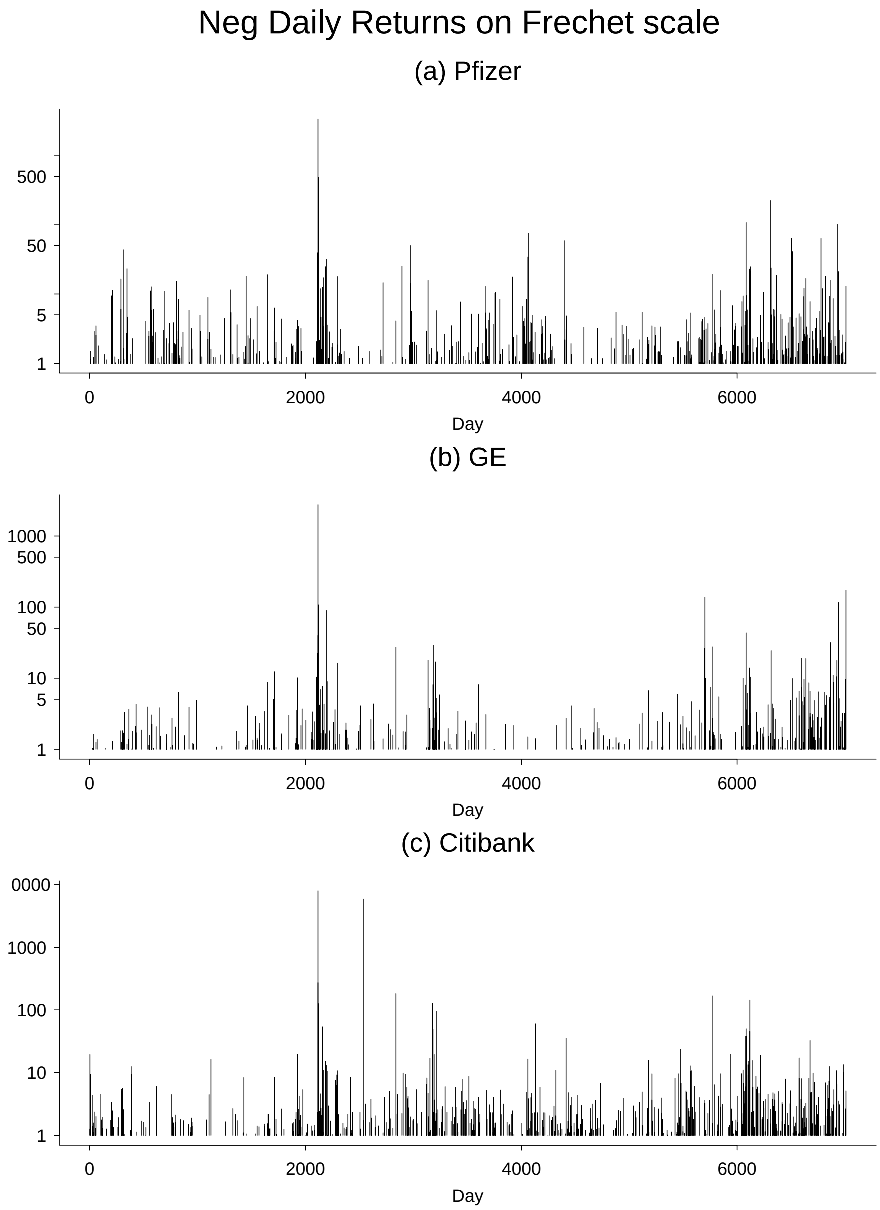

Figure 7: The negative returns after transformed into unit Fréchet scale.

test the generalized Pareto distribution assumption for the distribution of excesses over the threshold. This can be done via W-statistics:

$$W_k = \frac{1}{\xi_{T_k}} \log[1 + \xi_{T_k} \frac{Y_k - u}{\psi_{T_k} + \xi_{T_k}\{u - \mu_{T_k}\}}]$$

Then $W_1, W_2, \ldots$ are also independent exponential random variables with mean 1, if the model is correct. These techniques have been broadly used in model diagnostics, for example, Tsay (1999), Smith and Goodman (2000).

We observe throughout that the "observed" values are larger than the "expected", and therefore we conclude that the variables are dependent. We use the $M4$ process to model that dependence. Table 3 contains partial information in these plots and other diagnostic results.

Series 1	Series 2	Lag	Expected	Observed	z statistics
Pfizer	Pfizer	1	38.3	64	2
GE	GE	1	16.1	31	2
CITIBANK	CITIBANK	1	49.2	80	3
Pfizer	GE	-1	24.8	41	2
Pfizer	GE	0	24.8	127	14
Pfizer	CITIBANK	-1	43.4	74	3
Pfizer	CITIBANK	0	43.4	145	10
Pfizer	CITIBANK	1	43.4	63	2
GE	CITIBANK	-1	28.2	49	2
GE	CITIBANK	0	28.1	146	15
GE	CITIBANK	1	28.2	48	2

Table 3. Mutual exceedance statistics. Tabulated are the names of the two series, the lag by which series 1 leads series 2, and the expected and observed number of mutual exceedances.

The last column shows the integer value of a z statistic calculated under a Poisson assumption: only entries for which $z > 2$ are shown.

All the figures and the table suggest an $M4$ process fitting may be a good choice for financial time series data with multivariate temporal dependence.

Figure 7 and 9 suggest that a model of 3 signature patterns. Two of these have order of 2, corresponding to drops happened in two consecutive days, and the other one has order of 1, which corresponds to a single drop.

We now use the following model to fit the data.

$$\begin{aligned} Y_{id} = \max(&a_{1,-1}Z_{1,i-1}, \quad a_{1,0}Z_{1,i}, \\ &a_{2,-1}Z_{2,i-1}, \quad a_{2,0}Z_{2,i}, \\ &\qquad\qquad\quad a_{3,0}Z_{3,i}) \end{aligned} \tag{23}$$

We summarize all the estimated values in the following table.

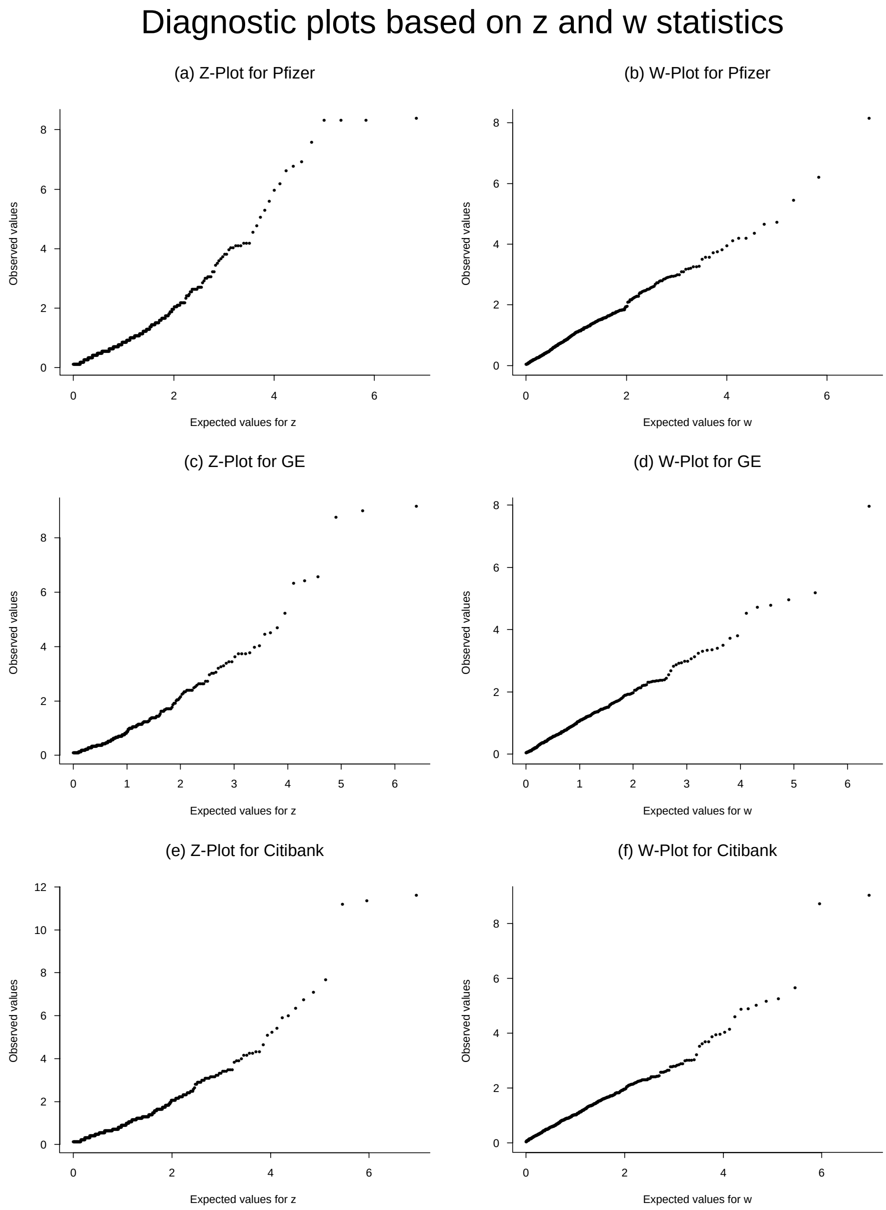

Figure 8: W-plots show a generalized extreme value distribution fitting is appropriate. Some caution should be given since a few points, partly the result of Oct87 crash, are away from the straight line.

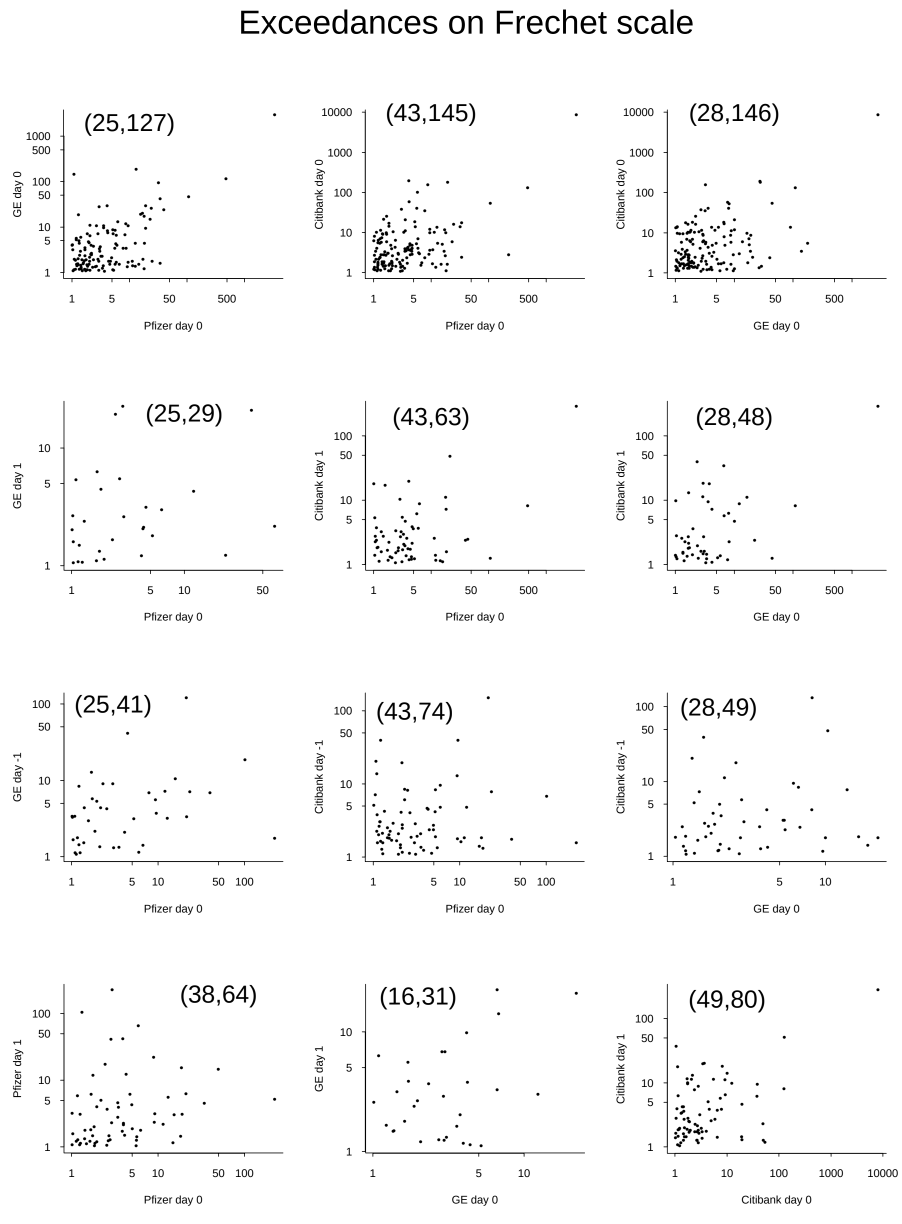

Figure 9: (i) All plots are based on Fréchet transformed exceedances of a high threshold based on negative log returns (so they represent price drops, not price rises); (ii) the purpose of the plots is to look for dependence among neighboring values; (iii) the numbers in parentheses show expected and observed numbers of simultaneous exceedances by the two variables, where "expected" is calculated on an independence assumption.

Table 4. Estimations for Moving Maxima Process Modeling

Parameter	$a_{1,-1}$	$a_{1,0}$	$a_{2,-1}$	$a_{2,0}$	$a_{3,0}$
GE	.0034	.0026	.0601	.0267	.9072
CITIBANK	.0017	.0017	.0692	.0641	.8632
Pfizer	.0042	.0016	.0805	.0377	.8760

This table is constructed based on the probability properties of $M4$ processes. Estimating variance may not be appropriate since there are few observed values in one pattern in every time series studied here. Threshold method or methods based on multivariate domain of attraction may be other choices to estimate all parameters. Zhang (2001) have developed related theory and methodology.

5 Discussion

The methods described in this paper represent a completely new approach to the modeling of financial time series data. The main goal here is to propose an approach which can efficiently model multivariate time series which are both spatial and temporal dependent. The results obtained can be used in many ways. For example, results from example 2 can be used to compute VaR or to optimize the portfolio under VaR constraints. Studies have shown financial data are fat tailed. They are not normally distributed. Compare with traditional assumption of normality of underlying distribution. These results provide more information to risk managers who may be most interested in the situation when an extreme price movement occurs what kind of risk the company is exposed to. The methods described can be used to other fields, such as modeling insurance data, environment data etc. Applications to VaR are being developed in future work which is ongoing.

Acknowledgment. Richard L. Smith's research was partly supported by NSF Grant DMS-9971980. The paper was partly written during a visit by both authors to the Isaac Newton Institute for Mathematical Sciences, Cambridge, U.K.

References

[1] Coles, S. G. and Tawn, J. A. (1991) Modeling Extreme Multivariate Events. *J. R. Statist. Soc. B,* **53**,pp. 377-392

[2] Coles, S. G. and Tawn, J. A. (1994) Statistical Methods for Multivariate Extrems: an Application to Structural Design(with discussion). *Appl. Statist.* **43**, No. 1, pp. 1-48

[3] Davis, R. A. and Resnick, S. I. (1989) Basic Properties and Prediction of Max-ARMA Processes. *Adv. Appl. Prob.* , **21**, 781-803

[4] Davis, R. A. and Resnick, S. I. (1993) Prediction of Stationary Max-Stable Processes. *Ann. Appl. Prob.* , **3**, no 2, 497-525

[5] Deheuvels, P. (1983) Point Processes and Multivariate Extreme Values. *Jour. of Multi. Anal.* **13**, pp. 257-272

[6] de Haan, L. (1985) Extremes in higher dimensions: the model and some statistics. In *Proc. 45th Sess. Int. Statist. Inst.,* paper 26.3. The Hague: International Statistical Institute.

[7] de Haan, L. and Resnick, S. I. (1977) Limit Theory for Multivariate Sample Extremes. *Z. Wahrscheinlichkeitstheorie verw. Gebiete,* **40**, pp.317-337

[8] Hall, P., Peng, L. and Yao. Q. (2001) Moving-maximum models for extrema of time series. Manuscript.

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

[10] Smith, R. L. (1989) Extreme Value Analysis of Environmental time Series: An Application to Trend Detection in Ground-Level Ozone. (with discussion) *Statistical Science.* **4**, No 4, 367-393

[11] Smith, R. L., and Shively, T. S. (1995) Point Process Approach to Modeling Trends in Troposphereic Ozone Based on Exceedances of A High Threshold.*Atmospheric Environment,* **29**, No. 23, pp.3489-3499

[12] Smith, R. L., and Weissman, I. (1996) Characterization and Estimation of the Multivariate Extremal Index. Manuscript, UNC.

[13] Tsay, R. S. (1999) Extreme Value Analysis of Financial Data. Manuscript, Univ. Of Chicago.

[14] Zhang, Z. (2001) *Multivariate extremes, max-stable process estimation and dynamic financial modeling*, PhD dissertation in preparation, Department of Statistics, University of North Carolina.