High-frequency data modelling using Hawkes processes

Valérie Chavez-Demoulin[1]
joint work with A.C. Davison and A.J. McNeil (2005), J.A McGill (2012)

[1]Faculty of Business and Economics,
University of Lausanne, Switzerland

September 9, 2013
EVT2013, Vimeiro

High frequency financial data

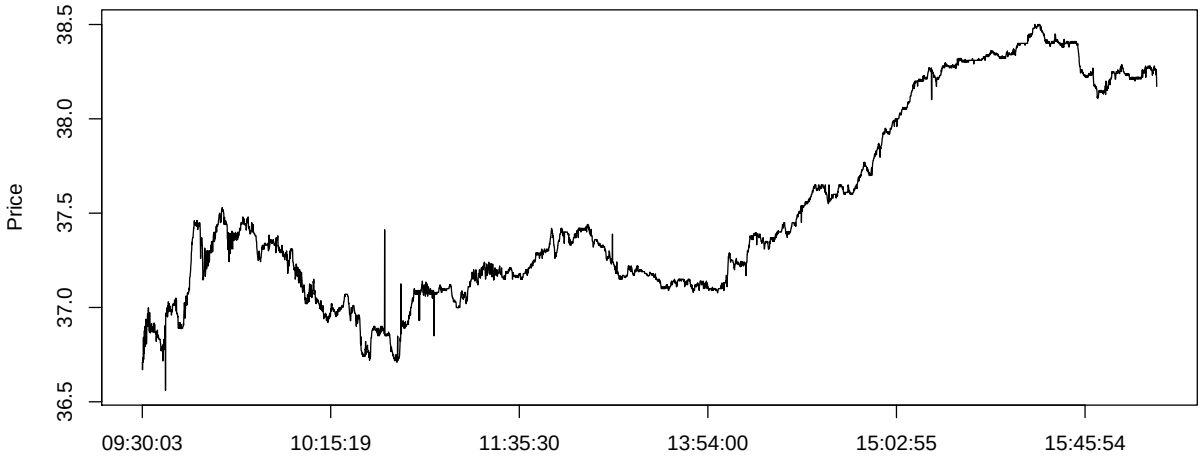
- The availability of high frequency data on transactions, quotes and order flow has revolutionized data processing and has led to statistical modeling challenges
- At the same time, the frequency of submission of orders has increased and the time to execution of market orders has dropped from more than 25 milliseconds in 2000 to less than a millisecond in 2010.

	Average number of orders in 10s	Number of price changes (June 26, 2008)
Citigroup	4469	12499
General Electric	2356	7862
General Motors	1275	9016

Table : Taken from R.Cont, 2011, http://ssrn.com/abstract=1748022

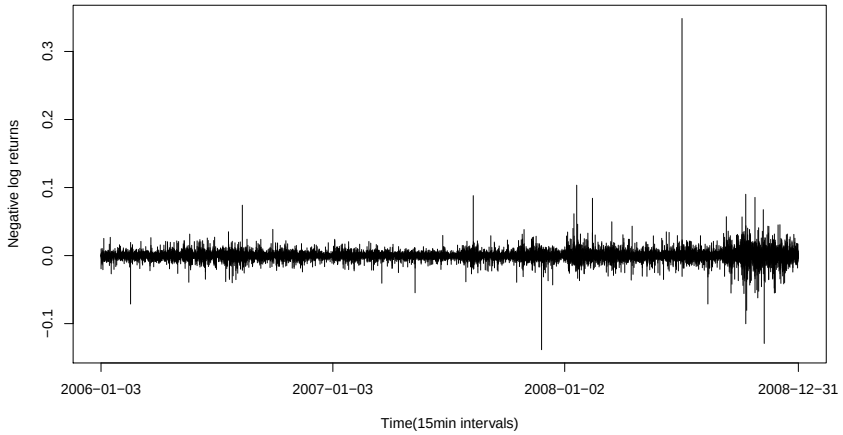
Trade database structure

SYMBOL	DATE	TIME	PRICE	SIZE	G127	CORR	COND
NVDA	20060901	9:30:04	28.67	264600	0	0	-
NVDA	20060901	9:30:04	28.67	264509	0	2	-
NVDA	20060901	9:30:05	28.65	100	0	0	-
NVDA	20060901	9:30:05	28.65	200	0	0	-
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

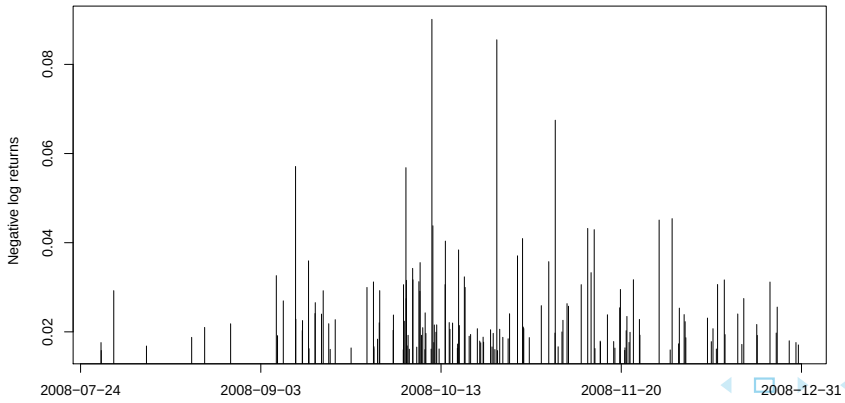


- Presence of erroneous trades: moving average filtering
- Irregularly spaced in time: linear interpolation

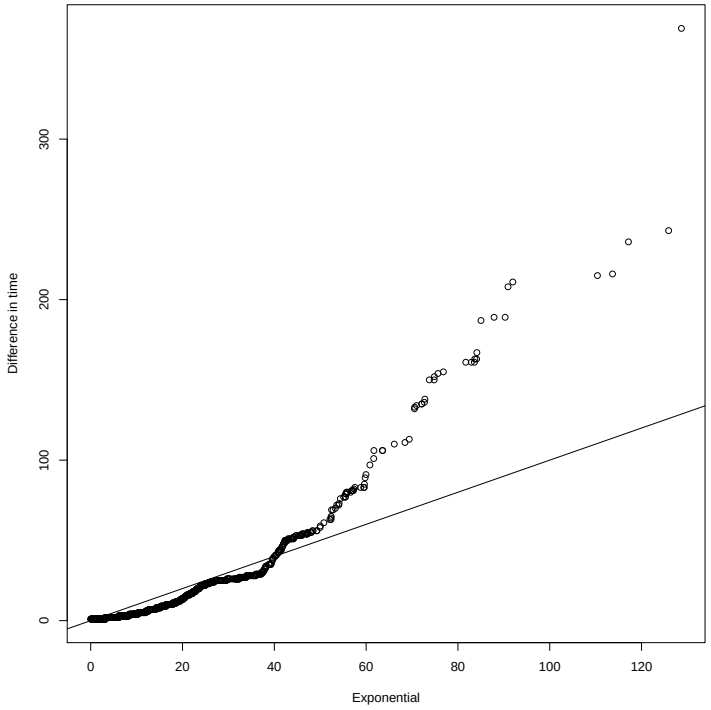
NVDA negative log-returns from 2006 to 2008 (15-minutes linearly interpolated)



NVDA, July to December 2008 negative returns exceedances :



NVDA log-returns from 2006 to 2008 (15-minutes interpolated) time differences between exceedances over high threshold

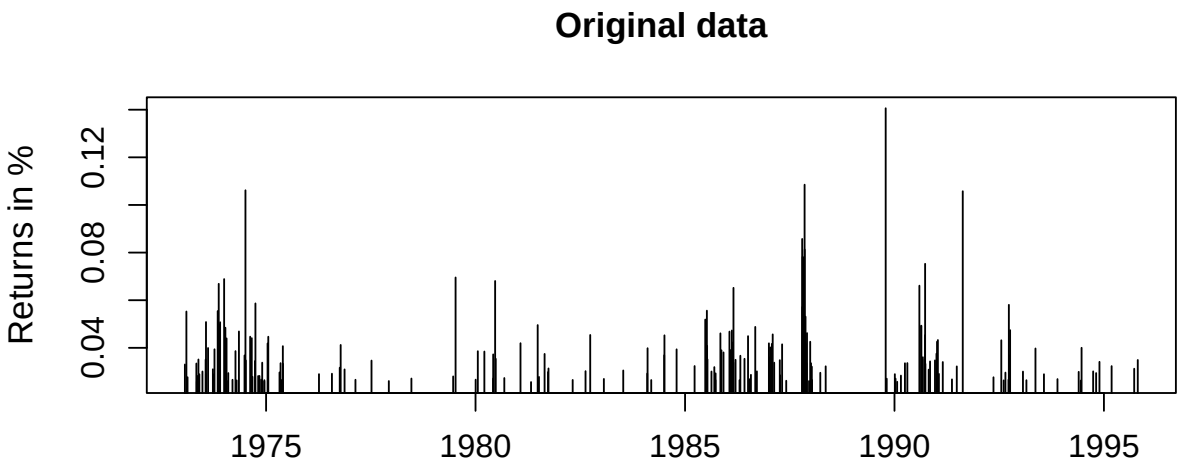


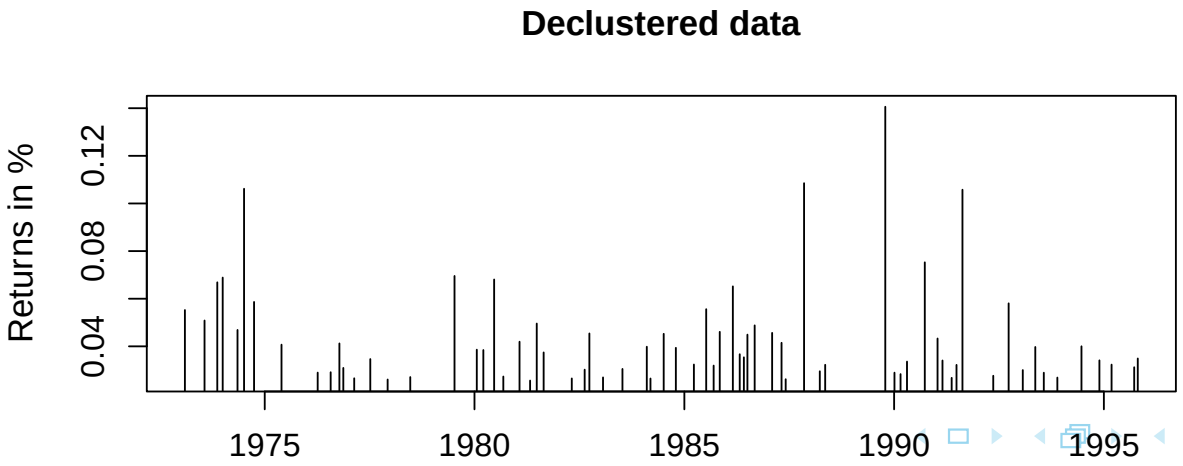
- does not seem exponentially distributed
- higher than expected frequencies for short distance between extremes

High-frequency financial data features

- Returns not iid
- Absolute returns highly correlated
- *Volatility* appears to change randomly with time
- Returns are *leptokurtic* or *heavy–tailed*
- *Extremes* appear in clusters

Declustering methods exist: An example for daily returns





Idea

- the classical Peaks-Over-Thresholds (POT) model of EVT assumes an homogeneous Poisson process
- the inter-event times would be independent exponential random variables
- the Figures above contradict the classical model

Aim: Treats the negative log-returns above a high threshold u as a realization of a marked point process:

- Using a first model that combines a self-exciting process for the threshold exceedance **times** with Generalized Pareto model for the threshold excess sizes (**mark sizes**)
- Using a second model that considers a self-exciting **POT model**

Classical Peaks-Over-Thresholds method

- The number of exceedances over u has a Poisson distribution with mean λ
- Conditional on n exceedances, their sizes $W_j = Z_j - u$ are a random sample of size n from the generalized Pareto distribution (GPD)

$$G_{\xi,\sigma}(w) = \begin{cases} 1 - (1 + \xi w/\sigma)_+^{-1/\xi}, & \xi \neq 0, \\ 1 - \exp(-w/\sigma), & \xi = 0. \end{cases}$$

Overall loglikelihood

Since number and mark size for the threshold exceedances are assumed independent

$$
\begin{aligned}
l(\lambda, \sigma, \xi) &= \log \left\{ P(N_u = n) \prod_{j=1}^{n} g_{\xi,\sigma}(w_j) \right\} \\
&\doteq n \log \lambda - \lambda - n \log \sigma - (1 + 1/\xi) \sum_{j=1}^{n} \log(1 + \xi w_j/\sigma)_+
\end{aligned}
$$

so inference can be performed separately for the frequency of exceedances and their sizes.

Marked Point Processes (1)

- consider an event process $(T_i, W_i), i \in \mathbb{Z}$
- observed over the period $(0; t_0]$ gives data $(t_1, w_1), \ldots, (t_n, w_n)$
- let $\mathcal{H}_t$, entire history of process up to time t
- the joint density of the data seen over $(0, t_0]$ can be written

$$\prod_{i=1}^{n} f_{T_i, W_i | \mathcal{H}_{t_{i-1}}}(t_i, w_i | \mathcal{H}_{t_{i-1}}) P(T_{n+1} > t_0 | \mathcal{H}_{t_n})$$

Marked Point Processes (2)

Assuming independence of times T_i and marks W_i the conditional density function is

$$f_{T_i,W_i|\mathcal{H}_{t_{i-1}}}(t_i, w_i|\mathcal{H}_{t_{i-1}}) \quad = \quad f_{T_i|\mathcal{H}_{t_{i-1}}}(t_i|\mathcal{H}_{t_{i-1}}) \cdot f_{W_i|\mathcal{H}_{t_{i-1}}}(w_i|\mathcal{H}_{t_{i-1}})$$

leading to the log-likelihood

$$\begin{aligned} \ell \quad = \quad & \left\{ \sum_{i=1}^{n} \log f_{T_i|\mathcal{H}_{i-1}}(t_i) + \log P(T_{n+1} > t_0|\mathcal{H}_{t_n}) \right\}^{A} \\ & + \left\{ \sum_{i=1}^{n} \log f_{W_i|\mathcal{H}_{t_{i-1}}}(w_i|\mathcal{H}_{t_{i-1}}) \right\}^{B} \end{aligned}$$

Hawkes process

Combining separately a Hawkes process for the counting process of the times (part A) and a GPD model for the contributions from the marks (part B), the likelihood becomes

$$\begin{aligned}\ell &= \log\left\{\prod_{i=1}^{n}\lambda(t_i|\mathcal{H}_{t_i})\exp\left(-\int_0^{t_0}\lambda(u|\mathcal{H}_u)du\right)\right\}\\ &\quad -n\log\sigma-(1+1/\xi)\sum_{i=1}^{n}\log(1+\xi w_i/\sigma)\end{aligned}$$

with $\lambda(t|\mathcal{H}_t)$ being the conditional intensity

Self-exciting process

We chose the following general form with a constant loading μ_0

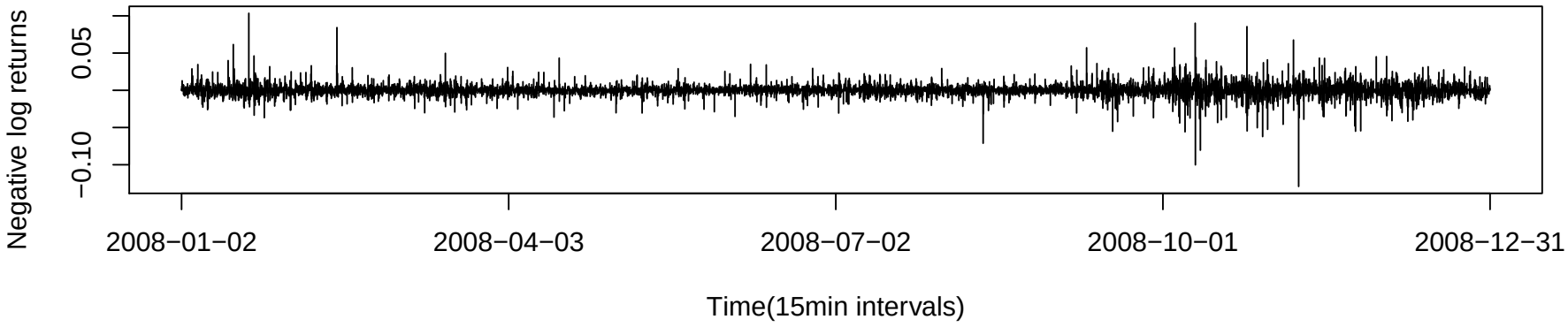
$$\lambda(t|\mathcal{H}_t) \quad = \quad \mu_0 + \sum_{j:t_j<t} \omega(t - t_j; w_j),$$

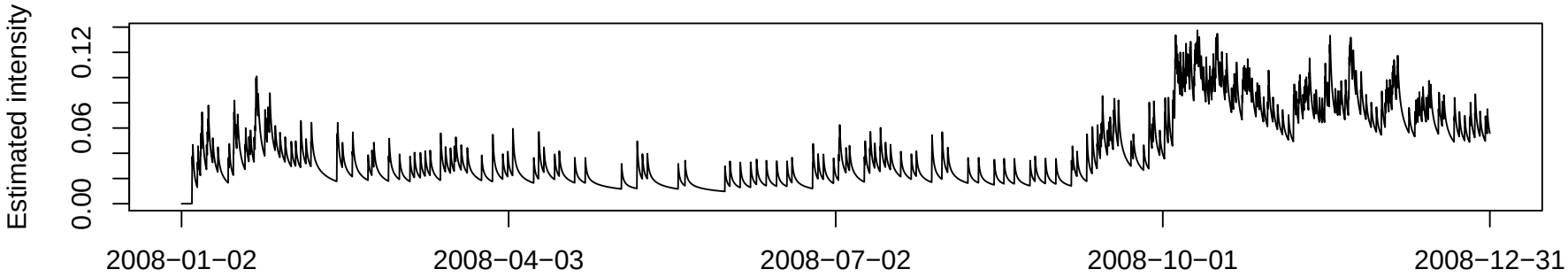
where $\mu_0 > 0$ and the self-exciting function $\omega(s)$ has to be positive when $s \geq 0$ and 0 elsewhere. For example, the conditional intensity model proposed by Ogata (1988) is

$$\lambda(t|\mathcal{H}_t) \quad = \quad \mu_0 + \psi \sum_{j:t_j<t} \frac{e^{\delta w_j}}{(t - t_j + \gamma)^{1+\rho}} \qquad \textit{ETAS}$$

where $\mu_0, \psi, \delta, \gamma, \rho > 0$.

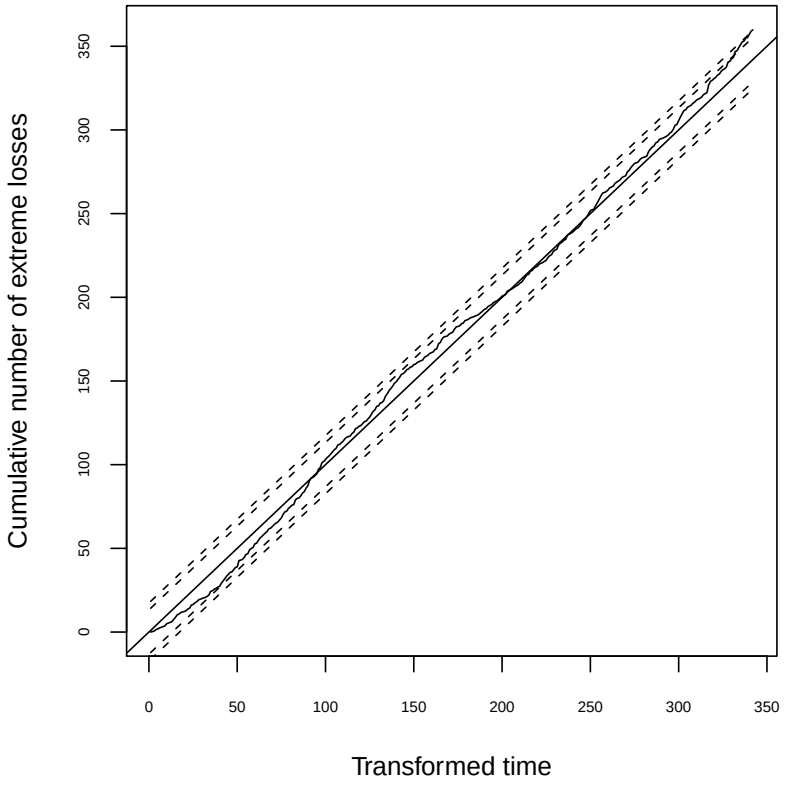
Applications to high-frequency data: 15-minutes linearly interpolated NVDA 2008 negative log-returns





Estimated cumulative events number

$\hat{\Lambda}_{\mathcal{H}_t}(t) = \int_0^t \hat{\lambda}(u|\mathcal{H}_u)du$ and two-sided 95% and 99% confidence limits based on the Kolmogorov-Smirnov statistic



Alternative models (1)

- Hawkes process with exponential decay

$$\lambda(t|\mathcal{H}_t) \quad = \quad \mu_0 + \psi \sum_{j:t_j<t} \exp\{\delta w_j - \gamma(t - t_j)\}.$$

- First order Markov chain for the marks

$$W_i \mid W_{i-1} = w_i \sim GPD(\xi, \sigma_i = \exp(a + bw_{i-1}))$$

whose parameters a, b and ξ can be estimated by maximising

$$\prod_{i=2}^{n} f_{W_i|W_{i-1}}(w_i \mid w_{i-1}) f_{w_1}(w_1)$$

Diagnostic for the marginal distribution based on the residuals

$$R_j = \hat{\xi}^{-1} \log\left(1 + \hat{\xi} w_j / \hat{\sigma}_j\right), \qquad j = 1, \ldots, n$$

approximately independent unit exponential variables

Second model: A self-exciting POT model

A model with predictable marks

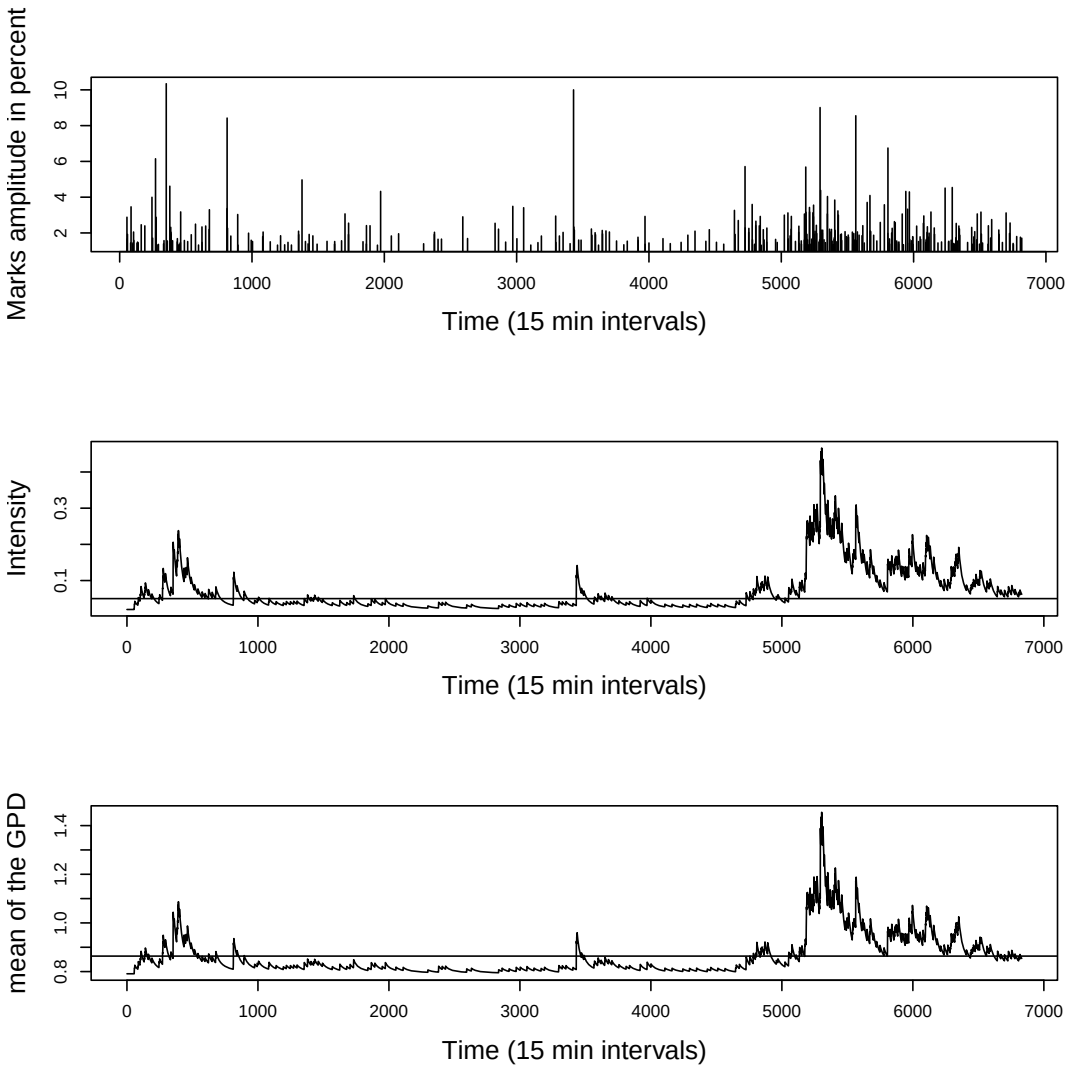
$$\lambda^\star(t,w) \quad = \quad \frac{\tau+\alpha_1\nu(t)}{\sigma+\alpha_2\nu(t)}\left(1+\xi\frac{w-u}{\sigma+\alpha_2\nu(t)}\right)^{-1/\xi-1}$$

Where $\nu(t)=\sum_{j:t_j<t}\omega(t-t_j)$ uses a self-exciting function ω. This uses the two-dimensional re-parametrisation of the POT:

- Homogeneous Poisson process for the exceedances of the level u with intensity $\tau=-\ln H(u;\mu,\beta,\xi)=\{1+\xi(u-\mu)/\beta\}_+^{-1/\xi}$, where $H(y;\mu,\beta,\xi)$ is the GEV distribution and $\sigma=\beta+\xi(u-\mu)$
- Inhomogeneous Poisson process for their sizes with intensity

$$\lambda(t,w)=\lambda(w)=\frac{1}{\sigma}\left(1+\xi\frac{w-u}{\sigma}\right)^{-1/\xi-1}$$

Estimated intensity of exceeding the threshold from self-exciting POT model for NVDA with Ogata function



Conditional risk measures

- The conditional approach models the intraday clustering of extremes and is used to calculate instantaneous conditional p =95% or 99% VaR over the next period

$$z_p^{t+1} = F^{-1}_{Z_{t+1|\mathcal{H}_t}}(p),$$

- $\bar{F}_{Z_{t+1|\mathcal{H}_t}}(z) = P(Z_{t+1} > z|\mathcal{H}_t)$ which is

$$P(Z_{t+1} - u > z|Z_{t+1} > u, \mathcal{H}_t) \times P(Z_{t+1} > u|\mathcal{H}_t).$$

- $P(Z_{t+1} > u|\mathcal{H}_t)$ represents the conditional probability of an event in $(t, t+1)$ that is:

$$1 - P\{N(t, t+1) = 0|\mathcal{H}_t\} = 1 - \exp\left(-\int_t^{t+1} \lambda(u|\mathcal{H}_u)du\right)$$

- $P(Z_{T+1} - u > z|Z_{t+1} > u, \mathcal{H}_t)$ is estimated using the GPD model with parameters σ and ξ.

Conditional VaR and Expected Shortfall

- The VaR z_p^{t+1} is then the solution of the equation

$$P(Z_{T+1} > z_p^{t+1} | \mathcal{H}_t) = 1 - p,$$

- Using the self-exciting POT model the condition VaR is

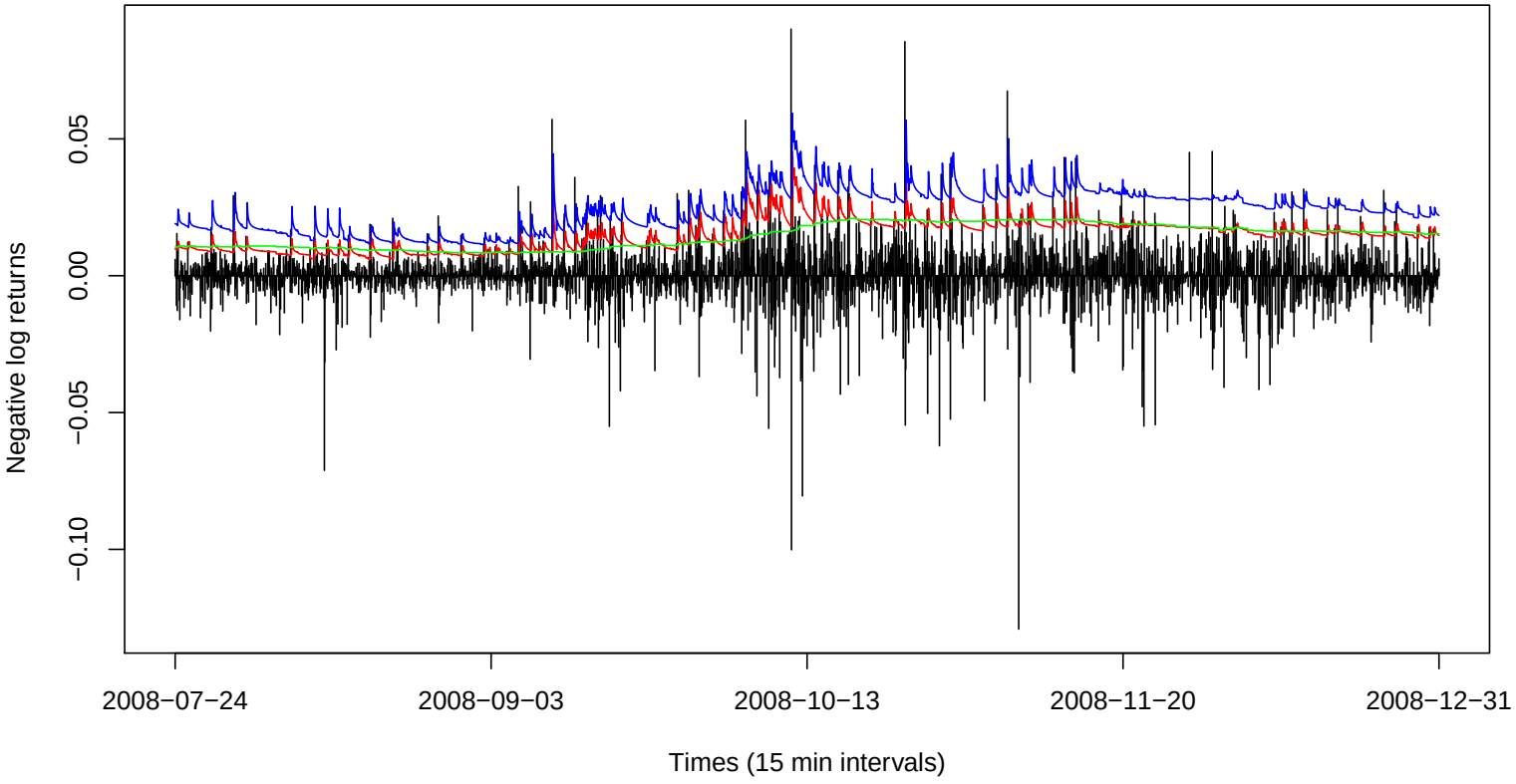
$$z_p^{t+1} = u + \frac{\sigma + \alpha_2 \nu(t+)}{\xi} \left\{ \left(\frac{1-p}{\tau + \alpha_1 \nu(t+)} \right)^{-\xi} - 1 \right\}$$

- From the estimated conditional VaR it is straightforward to get an estimation of the conditional Expected Shortfall (ES)

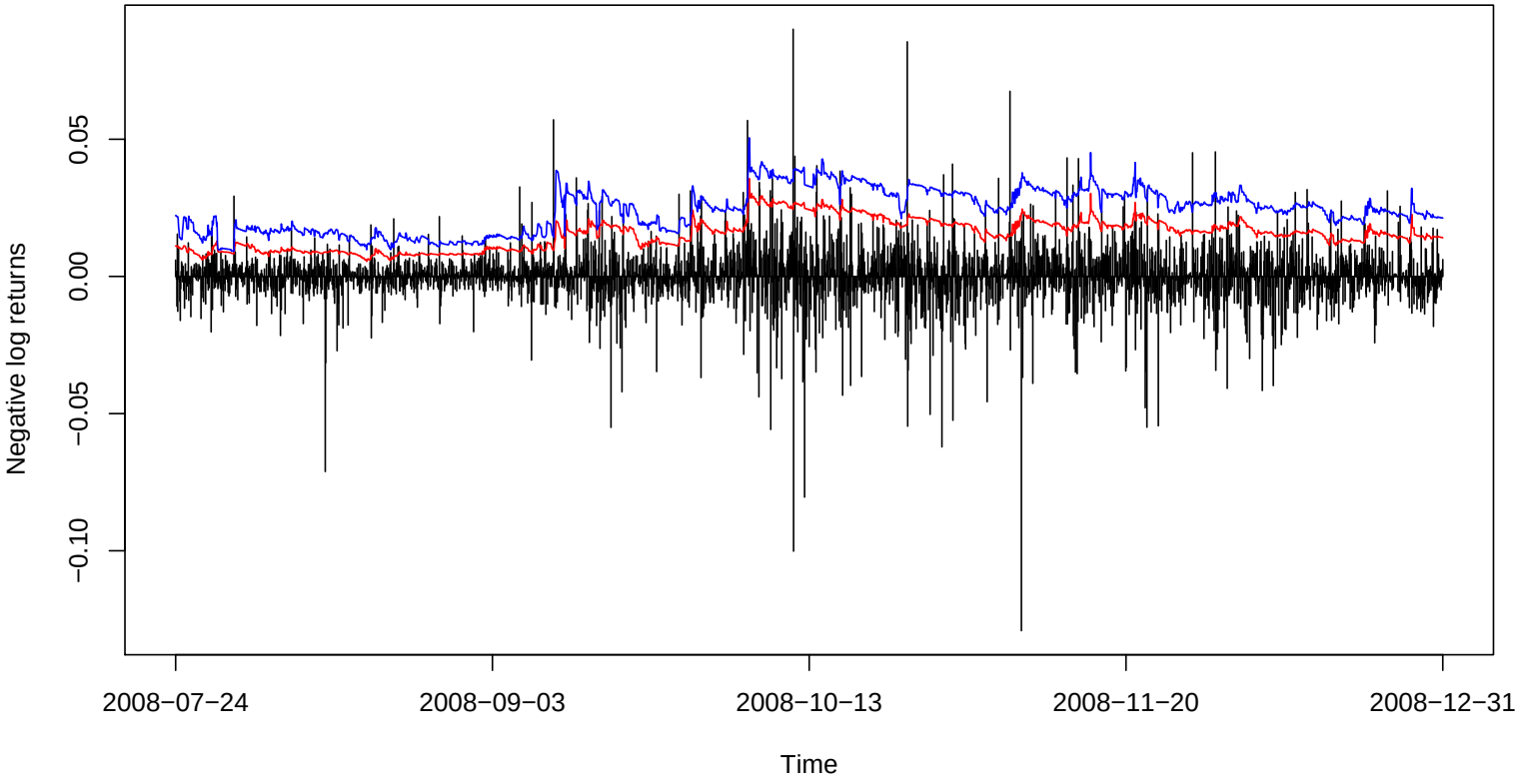
$$\mathrm{ES}_{\mathrm{p}}^{\mathrm{t+1}} = \frac{\mathrm{z}_{\mathrm{p}}^{\mathrm{t+1}}}{1-\xi} + \frac{\sigma - \xi \mathrm{u}}{1-\xi}.$$

with parameters replaced by their maximum likelihood estimates

15-minutes 95%VaR (red) and 95%ES (blue) estimates for NVDA negative log-returns using self-exciting POT models and 95%VaR using classical POT (green)



15-minutes 95%VaR and 95%ES estimates for NVDA negative log-returns using a competitor model (Chavez-Demoulin *et al.* (2013))



Backtesting results

0.95 Quantile	
Expected	25
POT	59 (0)
ETAS $\rho = 0$	21 (0.24)
ETAS $\rho \neq 0$	22 (0.31)
ETAS$\star$ $\rho = 0$	24 (0.47)
ETAS$\star$ $\rho \neq 0$	23 (0.39)
exp$\star$	25 (0.55)
NPOT	26 (0.39)

0.99 Quantile	
Expected	5
POT	14 (0)
ETAS $\rho = 0$	6 (0.39)
exp	6 (0.39)
ETAS$\star$ $\rho = 0$	6 (0.39)
exp$\star$	7 (0.24)
NPOT	9 (0.11)

References

- A.G. Hawkes, 1971, *Point spectra of some mutually exciting point processes*, **Biometrika**.
- Y. Ogata, 1988, *Statistical models for earthquake Occurrences and residuals analysis for point processes*, **JASA**.
- E. Errais, K. Giesecke and L.R. Goldberg, 2010, *Affine point processes and portfolio credit risk*, **SIAM J. Financial Math**.
- Y. Alt-Sahalia, J. Cacho-Diaz, and R. J. A. Laeven, 2010, *Modeling financial contagion using mutually exciting jump processes*, **The Review of Financial Studies**.
- V. Chavez-Demoulin, A.C. Davison, A.J. McNeil, 2005, *A point process approach to Value-at-Risk estimation*, **Quantitative Finance**.
- V. Chavez-Demoulin, P. Embrechts, S.Sardy, 2012 Extreme-quantile tracking for financial time series **Journal of Econometrics**, to appear.