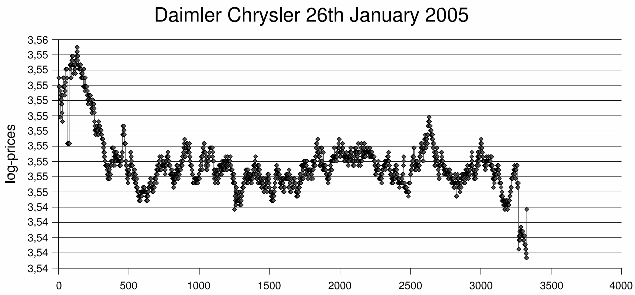
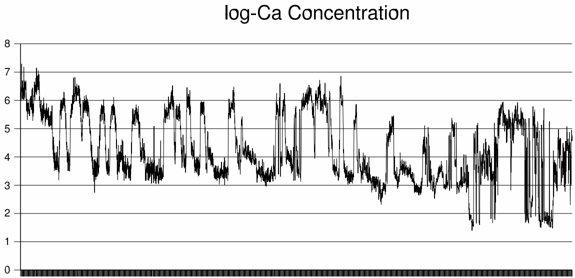
Brownian motion based versus fractional Brownian motion based models

Jeannette Woerner
University of Göttingen

- Comparison of models based on
  - -Brownian motion
  - -Brownian motion with iid noise
  - -fractional Brownian motion
- Identification of jump components
- Applications to financial and climate data

Motivation


Daimler Chrysler 26th January 2005
log-prices
3,56
3,55
3,54
0
500
1000
1500
2000
2500
3000
3500
4000


log-Ca Concentration
8
7
6
5
4
3
2
1
0

Motivation

discrete data $X_{t_{n,0}}, \cdots, X_{t_{n,n}}$
$t_{n,n} = t$ =fixed, $\Delta_{n,i} \to 0$ as $n \to \infty$
Assume stochastic volatility model

$$
\begin{aligned}
X_t &= Y_t + \int_0^t \sigma_s dB_s + \delta Z_t \\
X_t &= Y_t + \int_0^t \sigma_s dL_s + \delta Z_t \\
X_t &= Y_t + \int_0^t \sigma_s dB_s^H + \delta Z_t .
\end{aligned}
$$

Aim:

Determine which **model is suitable** for a specific data set.
Estimate $\int_0^t \sigma_s^2 ds$**.**

How can we infer $\int_0^t \sigma_s^2 ds$

First we consider **Brownian motion based models**.
Use the concept of **quadratic variation**, i.e. realized volatility

$$\sum_i |Y_{t_i} - Y_{t_{i-1}} + \int_{t_{i-1}}^{t_i} \sigma_s dB_s|^2 \xrightarrow{p} \int_0^t \sigma_s^2 ds$$

Advantages:

- almost model free, only need some Brownian motion based model
- very simple to compute
- distributional theory is known and Gaussian

Empirical studies versus theoretical results

Statistical principle:

Use all available data.

Problem:
For tick-by-tick data realized volatility increases.
Possible Explanation:

Market microstructure or **market friction**

i.e. effects due to bid-ask bounces, discreteness of prices, liquidity problems, asymmetric information,...

Model with iid noise

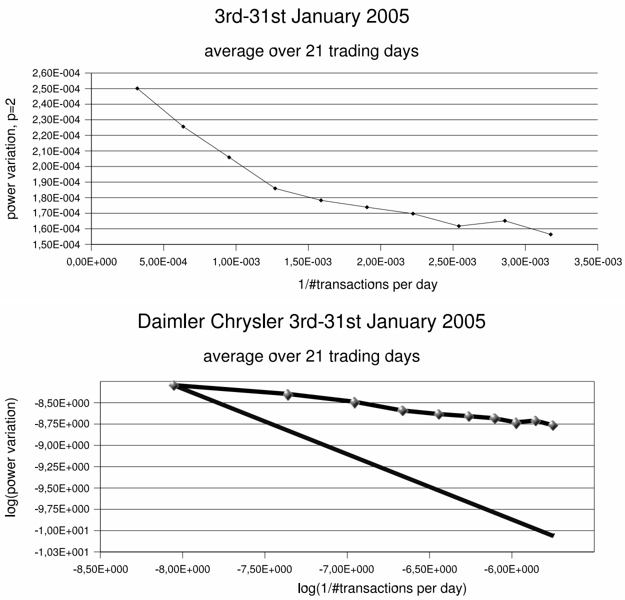
(cf. Ait-Sahalia, Mykland and Zhang (2006))

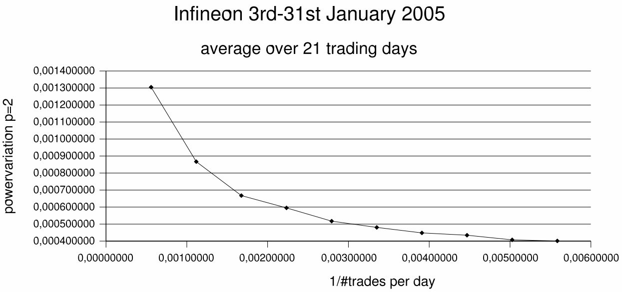
$$X_t = \int_0^t \sigma_s dB_s + \epsilon_t,$$

where ϵ denotes iid noise.
Then the realized volatility is of the order

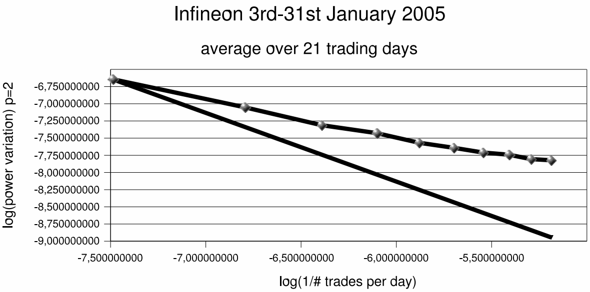
$$2\Delta^{-1} E(\epsilon^2),$$

hence the noise term leads to a bias with dominates the quadratic variation estimate for small Δ.


3rd-31st January 2005
average over 21 trading days
power variation, p=2
2,60E-004
2,50E-004
2,40E-004
2,30E-004
2,20E-004
2,10E-004
2,00E-004
1,90E-004
1,80E-004
1,70E-004
1,60E-004
1,50E-004
0,00E+000
5,00E-004
1,00E-003
1,50E-003
2,00E-003
2,50E-003
3,00E-003
3,50E-003
1/#transactions per day
Daimler Chrysler 3rd-31st January 2005
average over 21 trading days
log(power variation)
-8,50E+000
-8,75E+000
-9,00E+000
-9,25E+000
-9,50E+000
-9,75E+000
-1,00E+001
-1,03E+001
-8,50E+000
-8,00E+000
-7,50E+000
-7,00E+000
-6,50E+000
-6,00E+000
log(1/#transactions per day)



Infineon 3rd-31st January 2005
average over 21 trading days
powervariation p=2
0,001400000
0,001300000
0,001200000
0,001100000
0,001000000
0,000900000
0,000800000
0,000700000
0,000600000
0,000500000
0,000400000
0,00000000
0,00100000
0,00200000
0,00300000
0,00400000
0,00500000
0,00600000
1/#trades per day



Infineon 3rd-31st January 2005
average over 21 trading days
log(power variation) p=2
-6,750000000
-7,000000000
-7,250000000
-7,500000000
-7,750000000
-8,000000000
-8,250000000
-8,500000000
-8,750000000
-9,000000000
-7,500000000
-7,000000000
-6,500000000
-6,000000000
-5,500000000
log(1/# trades per day)

Non-normed and normed power variation

$$\begin{aligned} V_p^n(\int_0^t \sigma_s dB_s) &= \\ \sum_i |\int_{t_{i-1}}^{t_i} \sigma_s dB_s|^p &\xrightarrow{p} \begin{cases} 0 & : \quad p > 2 \\ \int_0^t \sigma_s^2 ds & : \quad p = 2 \\ \infty & : \quad p < 2 \end{cases} \end{aligned}$$

$$\Delta^{1-p/2} V_p^n(\int_0^t \sigma_s dB_s) \xrightarrow{p} \mu_p \int_0^t \sigma_s^p ds,$$

as $n \to \infty$, where $\mu_p = E(|u|^p)$ with $u \sim N(0,1)$ and $\Delta = t_i - t_{i-1}$. (cf. Barndorff-Nielsen and Shephard (2003))

Power variation for the model with iid noise

Using Minkowski's inequality with $p > 1$ we obtain

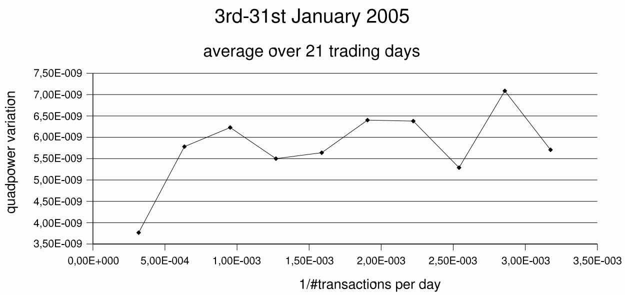
$$(\sum |\epsilon_{t_i} - \epsilon_{t_{i-1}}|^p)^{1/p} - (\sum |\int_{t_{i-1}}^{t_i} \sigma_s dB_s|^p)^{1/p} \leq (\sum |X_{t_i} - X_{t_{i-1}}|^p)^{1/p}$$

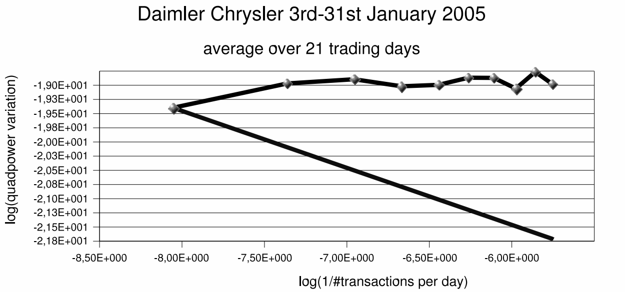
$$\leq (\sum |\epsilon_{t_i} - \epsilon_{t_{i-1}}|^p)^{1/p} + (\sum |\int_{t_{i-1}}^{t_i} \sigma_s dB_s|^p)^{1/p}$$

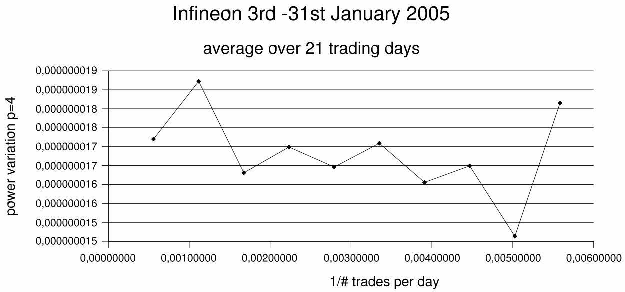
Hence if $E|\epsilon|^p < \infty$ for some $p > 2$, then

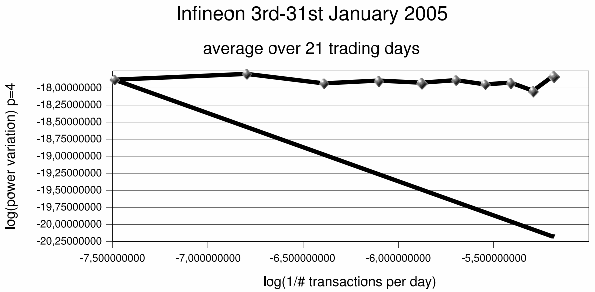
$$\sum |X_{t_i} - X_{t_{i-1}}|^p \to \infty,$$

which does not coincide with empirical findings.


3rd-31st January 2005
average over 21 trading days
quadpower variation
7,50E-009
7,00E-009
6,50E-009
6,00E-009
5,50E-009
5,00E-009
4,50E-009
4,00E-009
3,50E-009
0,00E+000
5,00E-004
1,00E-003
1,50E-003
2,00E-003
2,50E-003
3,00E-003
3,50E-003
1/#transactions per day


Daimler Chrysler 3rd-31st January 2005
average over 21 trading days
log(quadpower variation)
-1,90E+001
-1,93E+001
-1,95E+001
-1,98E+001
-2,00E+001
-2,03E+001
-2,05E+001
-2,08E+001
-2,10E+001
-2,13E+001
-2,15E+001
-2,18E+001
-8,50E+000
-8,00E+000
-7,50E+000
-7,00E+000
-6,50E+000
-6,00E+000
log(1/#transactions per day)


Infineon 3rd -31st January 2005
average over 21 trading days
power variation p=4
0,000000019
0,000000019
0,000000018
0,000000018
0,000000017
0,000000017
0,000000016
0,000000016
0,000000015
0,000000015
0,00000000
0,00100000
0,00200000
0,00300000
0,00400000
0,00500000
0,00600000
1/# trades per day


Infineon 3rd-31st January 2005
average over 21 trading days
log(power variation) p=4
-18,00000000
-18,25000000
-18,50000000
-18,75000000
-19,00000000
-19,25000000
-19,50000000
-19,75000000
-20,00000000
-20,25000000
-7,500000000
-7,000000000
-6,500000000
-6,000000000
-5,500000000
log(1/# transactions per day)

Fractional Brownian Motion

A fractional Brownian motion (fBm) with **Hurst parameter** $H \in (0,1)$, $B^H = \{B_t^H, t \geq 0\}$ is a zero mean Gaussian process with the covariance function

$$E(B_t^H B_s^H) = \frac{1}{2}(t^{2H} + s^{2H} - |t-s|^{2H}), \quad s,t \geq 0.$$

The fBm is a **self-similar** process, that is, for any constant $a > 0$, the processes
$\{a^{-H} B_{at}^H, t \geq 0\}$ and $\{B_t^H, t \geq 0\}$ have the same distribution.
For $H = \frac{1}{2}$, B^H coincides with the classical Brownian motion. For $H \in (\frac{1}{2}, 1)$ the process possesses **long memory** and for $H \in (0, \frac{1}{2})$ the behaviour is chaotic.

Non-normed Power Variation for fractional Brownian motion

$$\sum_i |\int_{t_{i-1}}^{t_i} \sigma_s dB_s^H|^p \xrightarrow{p} \begin{cases} 0 & : \quad p > 1/H \\ \mu_{1/H} \int_0^t \sigma_s^{1/H} ds & : \quad p = 1/H \\ \infty & : \quad p < 1/H \end{cases},$$

The integral is a **pathwise Riemann-Stieltjes integral** and we need that σ is a stochastic process with paths of finite q-variation, $q < \frac{1}{1-H}$.
Idea: Empirical behaviour of tick-by-tick data may also be explained by fBB with $H < 0.5$.

Consistency

joint with J.M. Corcuera and D. Nualart (2006)

Theorem

Suppose that σ_t is a stochastic process with finite q-variation, where $q < \frac{1}{1-H}$. Set

$$Z_t = \int_0^t \sigma_s dB_s^H.$$

Then,

$$\Delta^{1-pH} V_p^n(Z) \xrightarrow{P} \mu_p \int_0^T |\sigma_s|^p ds,$$

as $n \to \infty$.

Estimate for quadratic variation

We can explain the empirical findings by considering

$$\sum_i |X_{t_i} - X_{t_{i-1}}|^2 = \Delta^{2H-1}(\Delta^{1-2H} \sum_i |X_{t_i} - X_{t_{i-1}}|^2),$$

where

$$\Delta^{1-2H} \sum_i |X_{t_i} - X_{t_{i-1}}|^2 \xrightarrow{p} \int_0^t \sigma_s^2 ds$$

and $\Delta^{2H-1} \to \infty$ for $H < 0.5$.

More details:

We look at the test statistics:

$$S = \frac{\sum_{i=1}^{[nT]-1}(X_{\frac{i+1}{n}} - X_{\frac{i}{n}})(X_{\frac{i}{n}} - X_{\frac{i-1}{n}})}{\sum_{i=1}^{[nT]}(X_{\frac{i}{n}} - X_{\frac{i-1}{n}})^2}$$

model based on Brownian motion: 0
model based on Brownian motion with iid noise: -1/2
model based on fractional Brownian motion: $\frac{1}{2}(2^{2H} - 2)$
confidence interval:

$$\left[-c_\gamma\sqrt{\frac{\sum_{i=1}^{[nt]}|X_{\frac{i}{n}} - X_{\frac{i-1}{n}}|^4}{3\left(\sum_{i=1}^{[nt]}|X_{\frac{i}{n}} - X_{\frac{i-1}{n}}|^2\right)^2}}, c_\gamma\sqrt{\frac{\sum_{i=1}^{[nt]}|X_{\frac{i}{n}} - X_{\frac{i-1}{n}}|^4}{3\left(\sum_{i=1}^{[nt]}|X_{\frac{i}{n}} - X_{\frac{i-1}{n}}|^2\right)^2}}\right],$$

where c_γ denotes the γ-quantile of a $N(0,1)$-distributed random variable.

Daimler Chrysler, January 3rd-31st 2005, 1% level

# transactions	mean distance	S	l. bound BM
66140	7s	-0.1061	-0.0796
33070	14s	-0.1606	-0.1094
22046	21s	-0.1574	-0.1244
16535	28s	-0.1192	**-0.1295**
13228	35s	-0.1156	**-0.1367**

# transactions	mean distance	R	l. bound BM
66140	7s	-0.424	-0.0357
33070	14s	-0.4411	-0.0405
22046	21s	-0.3837	-0.0622
16535	28s	-0.2744	-0.0467
13228	35s	-0.2532	-0.0518
11023	42s	-0.2202	-0.0597
9448	49s	-0.1749	-0.0601
8267	56s	-0.1117	-0.0675
7348	63s	-0.1336	-0.0623
6614	70s	-0.0972	-0.0702

Model with market microstructure

Daimler Chrysler, January 3rd-31st 2005, 1% level

# transactions	mean distance	S	u. bound iid
66140	7s	-0.1061	-0.3621
33070	14s	-0.1606	-0.3106
22046	21s	-0.1574	-0.2845
16535	28s	-0.1192	-0.2758
13228	35s	-0.1156	-0.2632
11023	42s	-0.0994	-0.2413
9448	49s	-0.0827	-0.2354
8267	56s	-0.0542	-0.2472
7348	63s	-0.0608	-0.2134
6614	70s	-0.0487	-0.2285

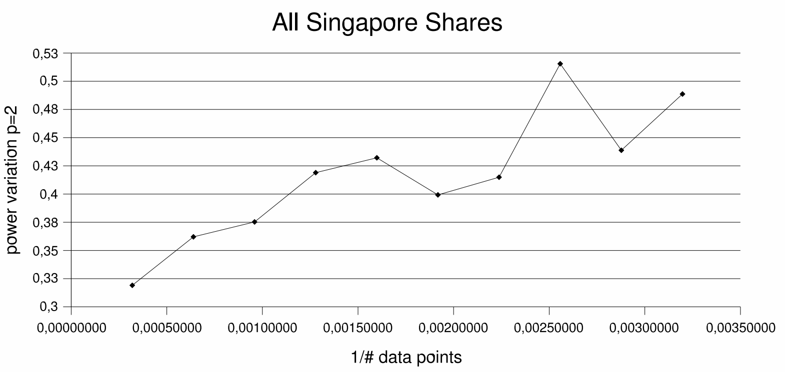
What are the effects of these results?

Risk induced by model misspecification
We look at Daimler Chrysler data of 12.1.2005:
Assuming a model based on **Brownian motion**:

$$\int_0^T \sigma_s^2 ds = 0.000309$$

Assuming a model based on **fractional Brownian motion** with $H = 0.4$:

$$\int_0^T \sigma_s^2 ds = 0.000059$$


All Singapore Shares
power variation p=2
0,53
0,5
0,48
0,45
0,43
0,4
0,38
0,35
0,33
0,3
0,00000000
0,00050000
0,00100000
0,00150000
0,00200000
0,00250000
0,00300000
0,00350000
1/# data points

What are the effects of these results?

Risk induced by model misspecification
We look at index data from Singapore:
Assuming a model based on **Brownian motion**:

$$\int_0^T \sigma_s^2 ds = 0.319$$

Assuming a model based on **fractional Brownian motion** with $H = 0.6$:

$$\int_0^T \sigma_s^2 ds = 1.595$$

Do we have an additional jump component?

A measure for the **activity** of the jump component of a Lévy process is the **Blumenthal-Getoor** index β,

$$\beta = \inf\{\delta > 0 : \int (1 \wedge |x|^{\delta})\nu(dx) < \infty\}.$$

This index ensures, that for $p > \beta$ the sum of the p-th power of jumps will be finite.

Comparison of non-normed power variation

$$\sum_i \left| \int_{t_{i-1}}^{t_i} \sigma_s dB_s \right|^p \xrightarrow{p} \begin{cases} 0 & : \quad p > 2 \\ \int_0^t \sigma_s^2 ds & : \quad p = 2 \\ \infty & : \quad p < 2 \end{cases}$$

and the case for the Lévy model

$$\sum_i \left| \int_{t_{i-1}}^{t_i} \sigma_s dL_s \right|^p$$
$$\xrightarrow{p} \begin{cases} \sum (\left| \int_{u-}^{u} \sigma_s dL_s \right|^p : 0 < u \leq t) & : \quad p > \beta \\ \infty & : \quad p < \beta \end{cases}$$

under appropriate regularity conditions, where β denotes the Blumenthal-Getoor index of L.

Non-normed Power Variation for fractional Brownian motion

$$\sum_i \left| \int_{t_{i-1}}^{t_i} \sigma_s dB_s^H \right|^p \xrightarrow{p} \begin{cases} 0 & : \quad p > 1/H \\ \mu_{1/H} \int_0^t \sigma_s^{1/H} ds & : \quad p = 1/H \\ \infty & : \quad p < 1/H \end{cases},$$

where $H > 1/2$. The integral is a **pathwise Riemann-Stieltjes integral** and we need that σ is a stochastic process with paths of finite q-variation, $q < \frac{1}{1-H}$.
Hence one over the **Hurst exponent** plays a similar role as the **Blumenthal-Getoor index**.

Log-Power Variation Estimators

Theorem

Assume that for some $k \in \mathbb{R}$ *and* $p \in (a,b)$*, s.t.* $1 - pk \neq 0$

$$\Delta^{1-pk} V_p^n(X) \stackrel{p}{\to} C, \tag{1}$$

with $0 < C < \infty$*, then*

$$\frac{\ln(\Delta V_p^n(X))}{p \ln \Delta} \stackrel{p}{\to} k \tag{2}$$

holds as $n \to \infty$*, if on the other hand*

$$V_p^n(X) \stackrel{p}{\to} C, \tag{3}$$

with $0 < C < \infty$*, then as* $n \to \infty$

$$\frac{\ln(\Delta V_p^n(X))}{p \ln \Delta} \stackrel{p}{\to} \frac{1}{p}. \tag{4}$$

Question:

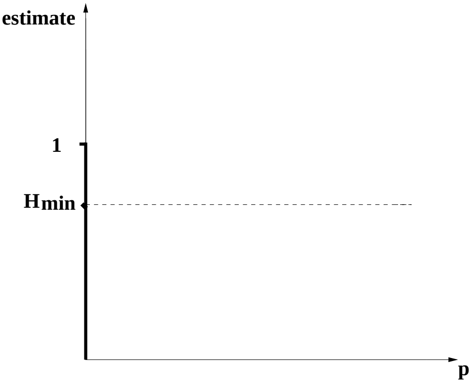
When is condition (1) or (3) satisfied

The definition of the Blumenthal-Getoor index for $p > \beta$ yields (3).

(1) has been considered in the framework of estimating the **integrated volatility** for many models :

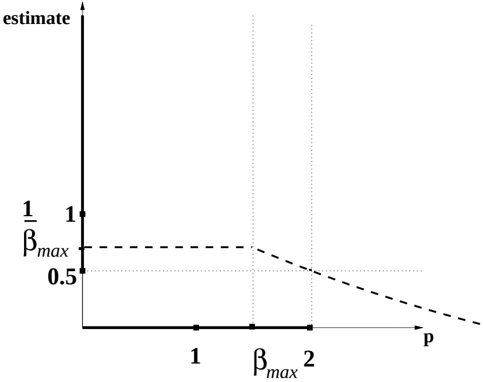
- classical stochastic volatility models based on Brownian motion with general mean process and additional jump component.
- models based on fractional Brownian motion.
- models based on Lévy processes.

Purely continuous model



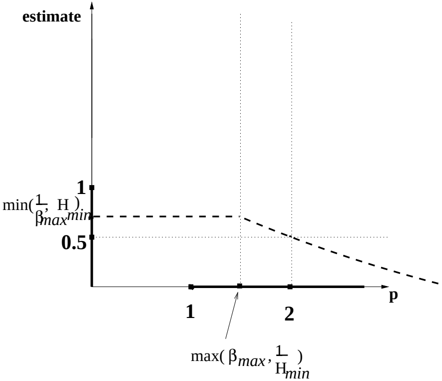
$$0 < H_{min} < 1$$

Pure jump model



$$\frac{1}{2} < \frac{1}{\beta_{max}} < \infty$$

Mixed model



$$0 < \beta_{max} < 2, \quad 0 < H_{min} < 1$$

$$0 < \min(\frac{1}{\beta_{max}}, H_{min}) < 1, \quad 1 < \max(\beta_{max}, \frac{1}{H_{min}}) < \infty$$

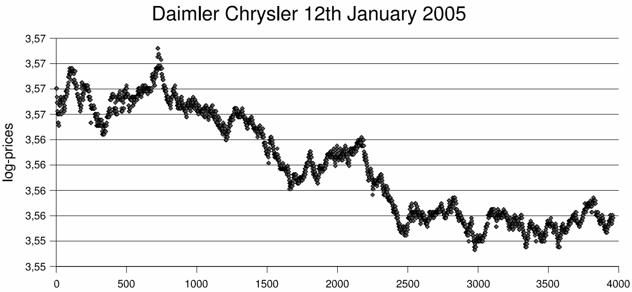
How to determine a jump component

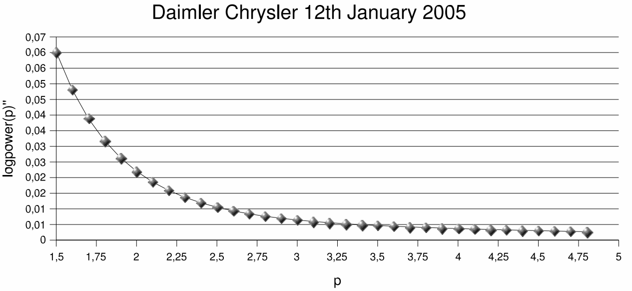
Look at the behaviour of the **second derivative** of the log-power variation in p.

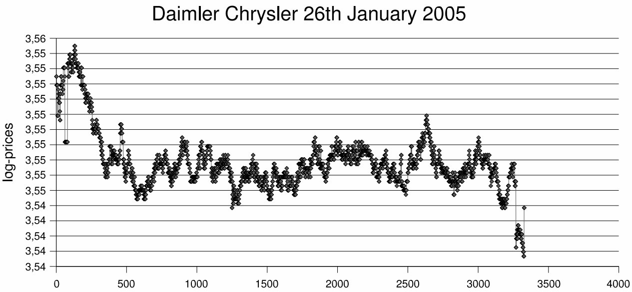
Example 1: Daimler Chrysler data
12th January 2005: 3960 transactions
26th January 2005: 3328 transactions

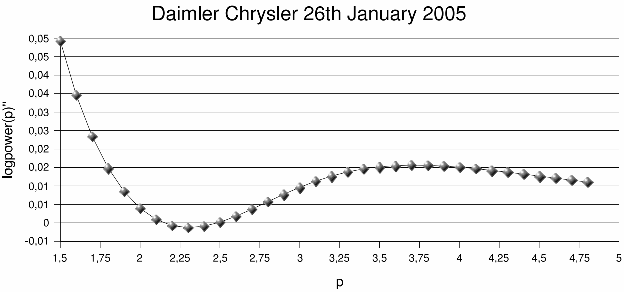
Example 2: Infineon data
12th January 2005: 2806 transactions
26th January 2005: 1977 transactions

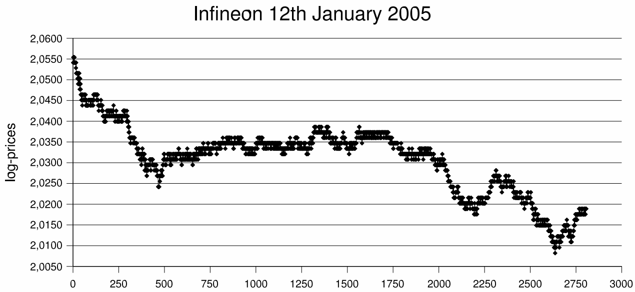
Example 3: daily index data of Singapore All Shares
6.1.1986-31.12.1997: 3128 transactions

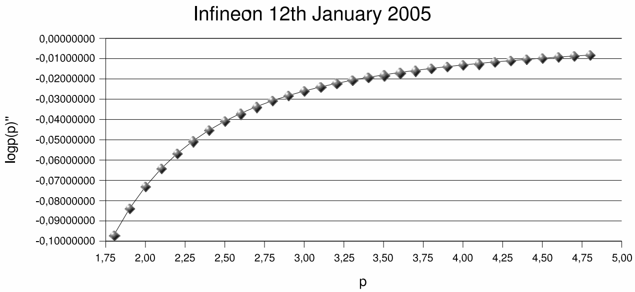

Daimler Chrysler 12th January 2005
log-prices
3,57
3,57
3,57
3,57
3,56
3,56
3,56
3,56
3,55
3,55
0
500
1000
1500
2000
2500
3000
3500
4000

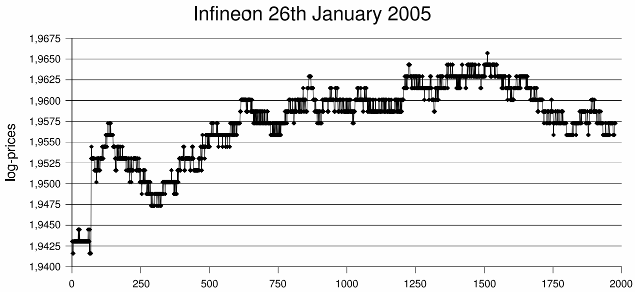

Daimler Chrysler 12th January 2005
logpower(p)"
0,07
0,06
0,06
0,05
0,05
0,04
0,04
0,03
0,03
0,02
0,02
0,01
0,01
0
1,5
1,75
2
2,25
2,5
2,75
3
3,25
3,5
3,75
4
4,25
4,5
4,75
5
p

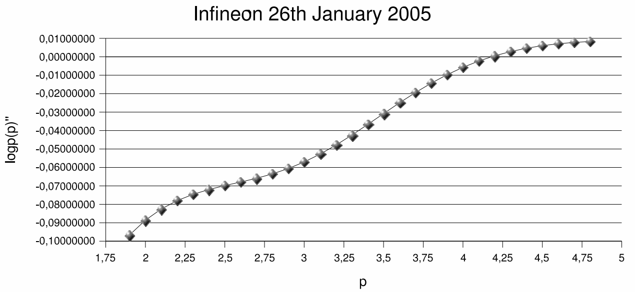

Daimler Chrysler 26th January 2005
log-prices
3,56
3,55
3,54
0
500
1000
1500
2000
2500
3000
3500
4000

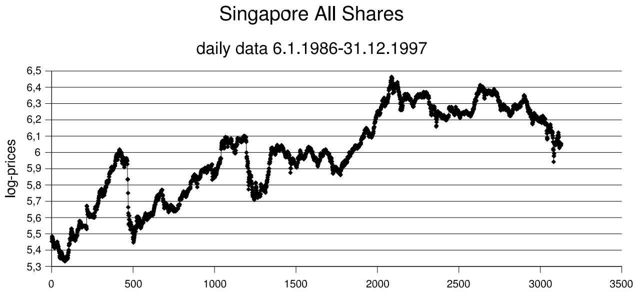

Daimler Chrysler 26th January 2005
logpower(p)"
0,05
0,04
0,03
0,02
0,01
0
-0,01
1,5
1,75
2
2,25
2,5
2,75
3
3,25
3,5
3,75
4
4,25
4,5
4,75
5
p


Infineon 12th January 2005
log-prices
2,0600
2,0550
2,0500
2,0450
2,0400
2,0350
2,0300
2,0250
2,0200
2,0150
2,0100
2,0050
0
250
500
750
1000
1250
1500
1750
2000
2250
2500
2750
3000

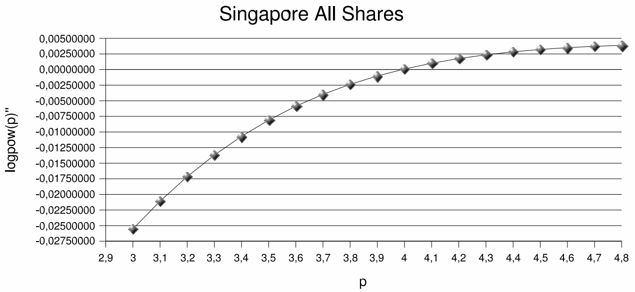

Infineon 12th January 2005
logp(p)"
0,00000000
-0,01000000
-0,02000000
-0,03000000
-0,04000000
-0,05000000
-0,06000000
-0,07000000
-0,08000000
-0,09000000
-0,10000000
1,75
2,00
2,25
2,50
2,75
3,00
3,25
3,50
3,75
4,00
4,25
4,50
4,75
5,00
p


Infineon 26th January 2005
log-prices
1,9675
1,9650
1,9625
1,9600
1,9575
1,9550
1,9525
1,9500
1,9475
1,9450
1,9425
1,9400
0
250
500
750
1000
1250
1500
1750
2000


Infineon 26th January 2005
logp(p)"
0,01000000
0,00000000
-0,01000000
-0,02000000
-0,03000000
-0,04000000
-0,05000000
-0,06000000
-0,07000000
-0,08000000
-0,09000000
-0,10000000
1,75
2
2,25
2,5
2,75
3
3,25
3,5
3,75
4
4,25
4,5
4,75
5
p



Singapore All Shares
daily data 6.1.1986-31.12.1997
log-prices
6,5
6,4
6,3
6,2
6,1
6
5,9
5,8
5,7
5,6
5,5
5,4
5,3
0
500
1000
1500
2000
2500
3000
3500



Singapore All Shares
logpow(p)''
0,00500000
0,00250000
0,00000000
-0,00250000
-0,00500000
-0,00750000
-0,01000000
-0,01250000
-0,01500000
-0,01750000
-0,02000000
-0,02250000
-0,02500000
-0,02750000
2,9
3
3,1
3,2
3,3
3,4
3,5
3,6
3,7
3,8
3,9
4
4,1
4,2
4,3
4,4
4,5
4,6
4,7
4,8
p

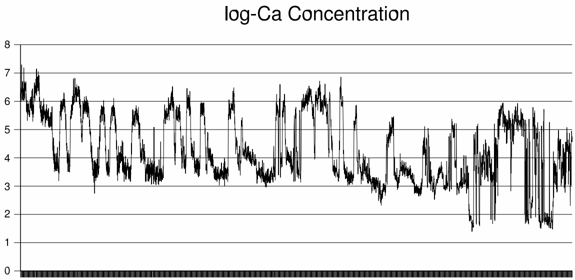
Modelling of transitions between climate states

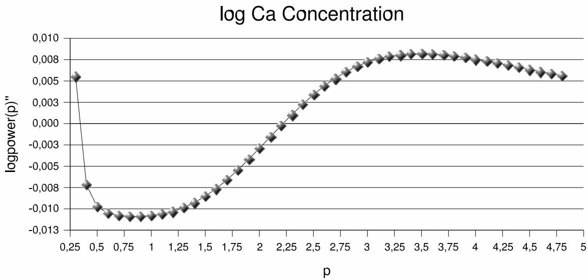
Calcium concentration in ice cores is proportional to one over the temperature.

suggested model: dynamical system with a Lévy component

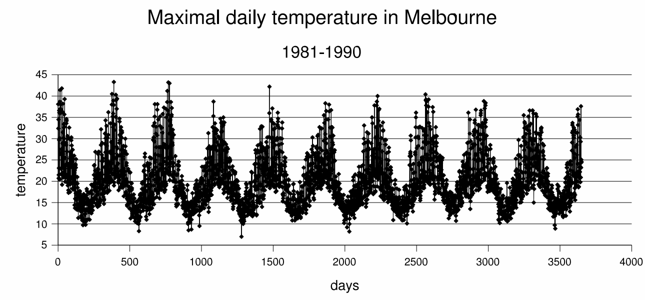
$$dX_t^\epsilon = -U'(X_t^\epsilon)dt + \epsilon dL_t.$$

(cf. Ditlevsen (1999), Imkeller and Pavlyukevich (2006))
However we get $H = 0.36$.


log-Ca Concentration
8
7
6
5
4
3
2
1
0


log Ca Concentration
0,010
0,008
0,005
0,003
0,000
-0,003
-0,005
-0,008
-0,010
-0,013
logpower(p)"
0,25 0,5 0,75 1 1,25 1,5 1,75 2 2,25 2,5 2,75 3 3,25 3,5 3,75 4 4,25 4,5 4,75 5
p

Temperature data



- no jumps
- H=0.35

Conclusion:

- Increasing limits in quadratic variation for data may be explained by fractional Brownian motion with $H < 0.5$.
- This approach may be applied to financial and climate data.
- log-power variation may be used to detect jump components in both Brownian and fractional Brownian motion based models.