Records in a changing world

Joachim Krug
Institute of Theoretical Physics, University of Cologne

- What are records, and why do we care?
- Record-breaking temperatures and global warming
- Records in stock prices and random walks

with Gregor Wergen, Jasper Franke and Miro Bogner

The fascination of records

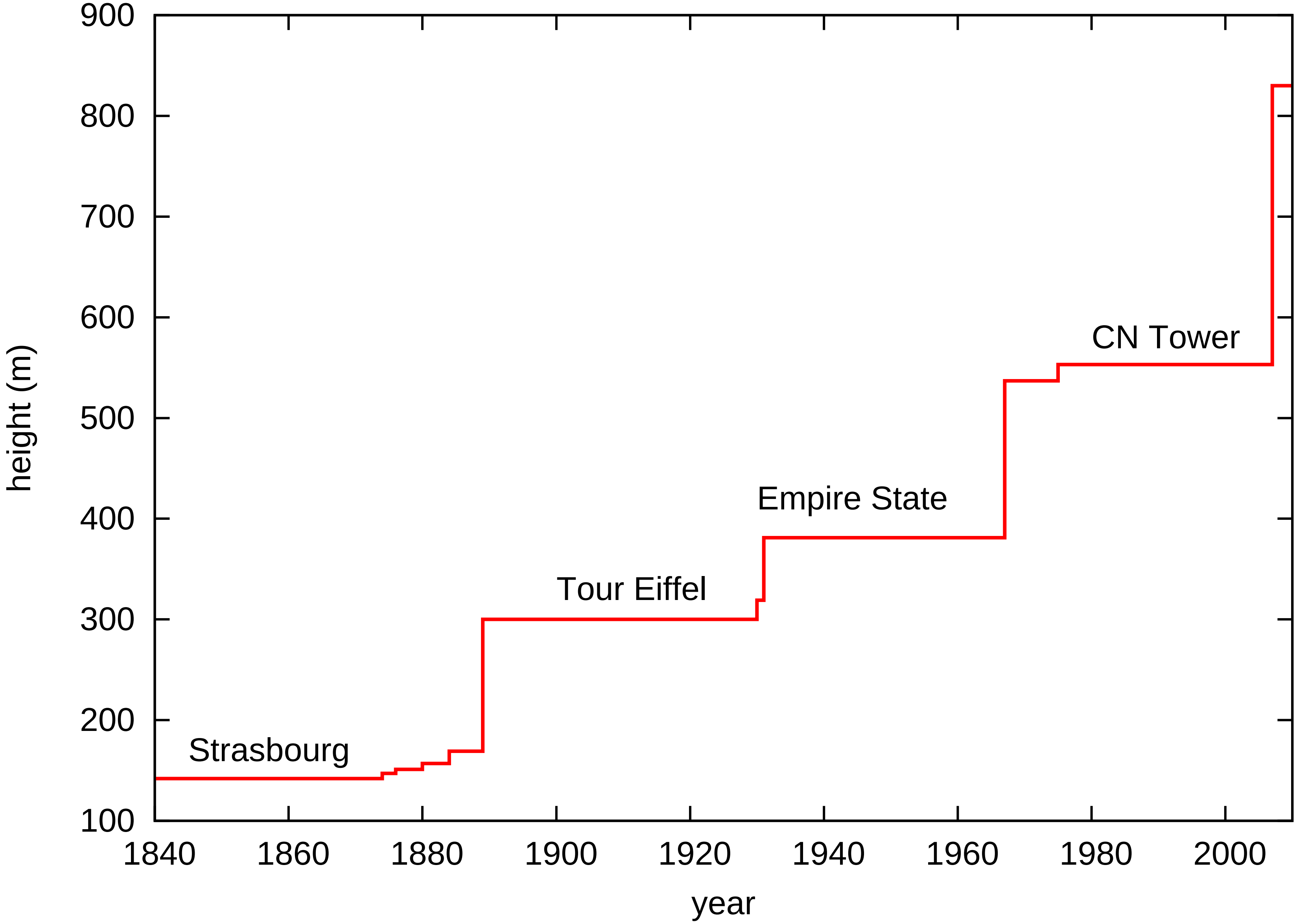
The fascination of records

World’s largest ancient castle

The fascination of records

World’s tallest building from 1880-1884

The world’s tallest buildings over time



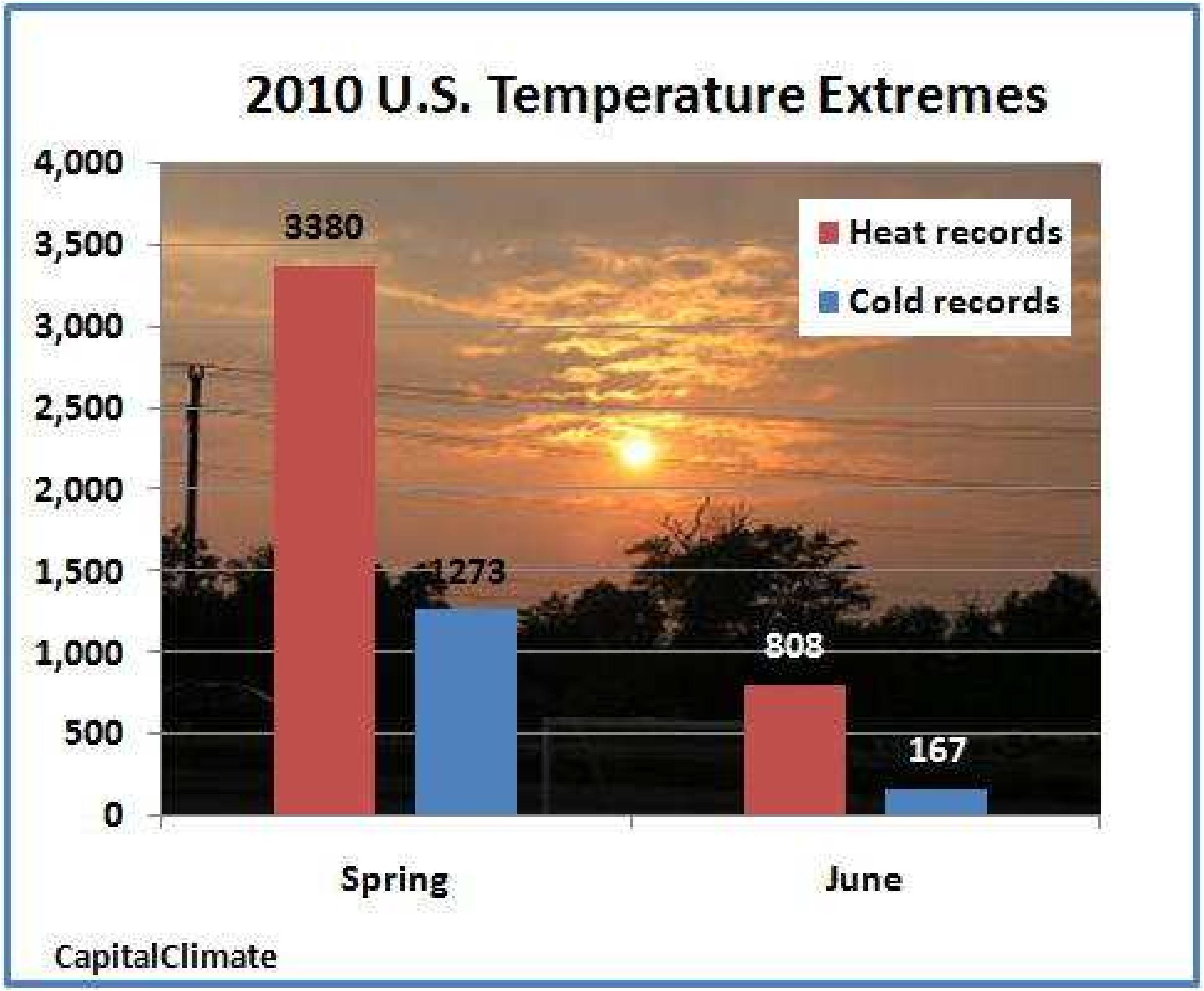
Temperature records

G. Wergen, JK, EPL **92** 30008 (2010)

The 2010 summer heat wave

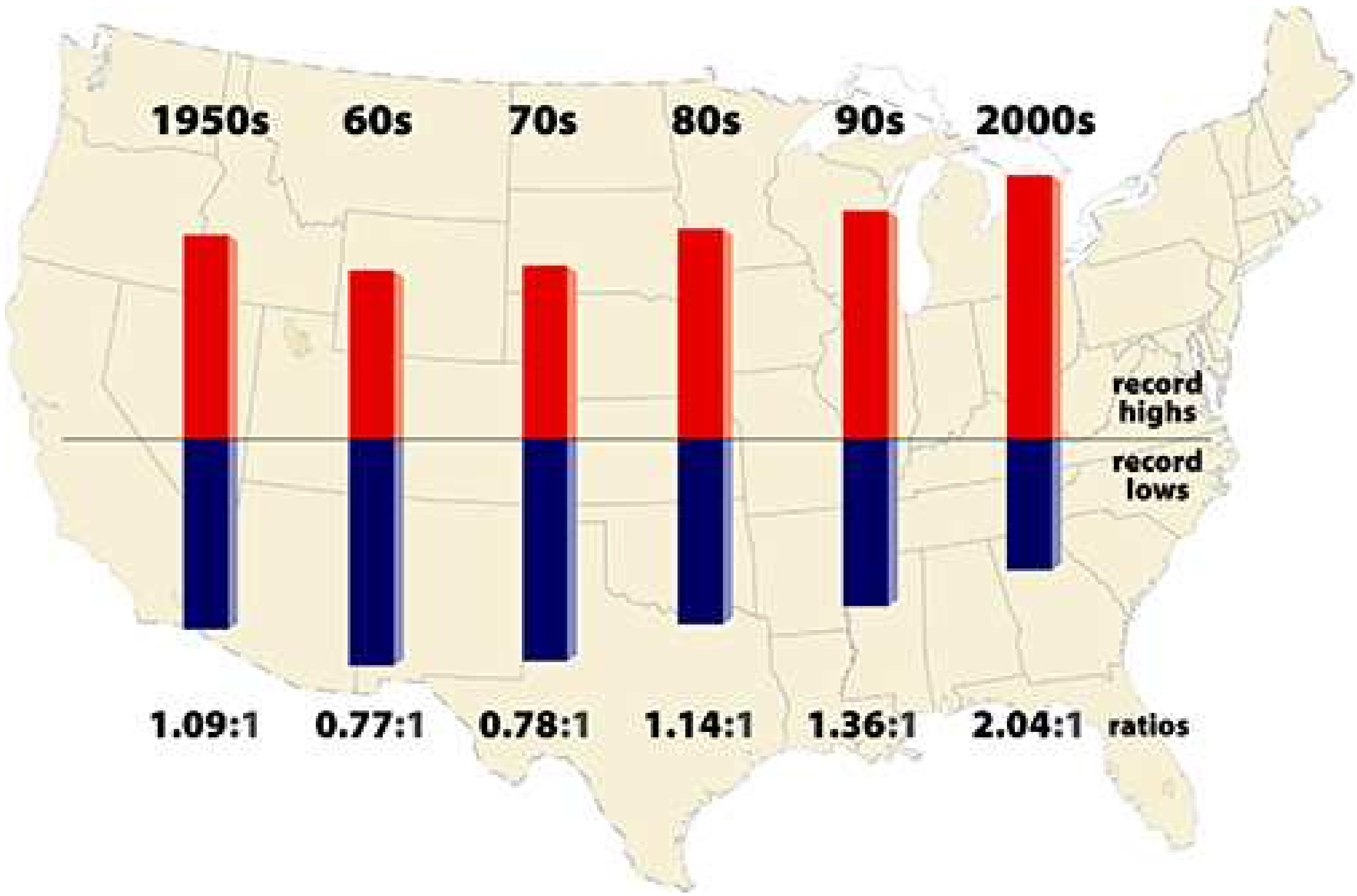
http://www.spiegel.de/

The 2010 summer heat wave



http://climateprogress.org/2010/07/05/heat-wave-global-warming/

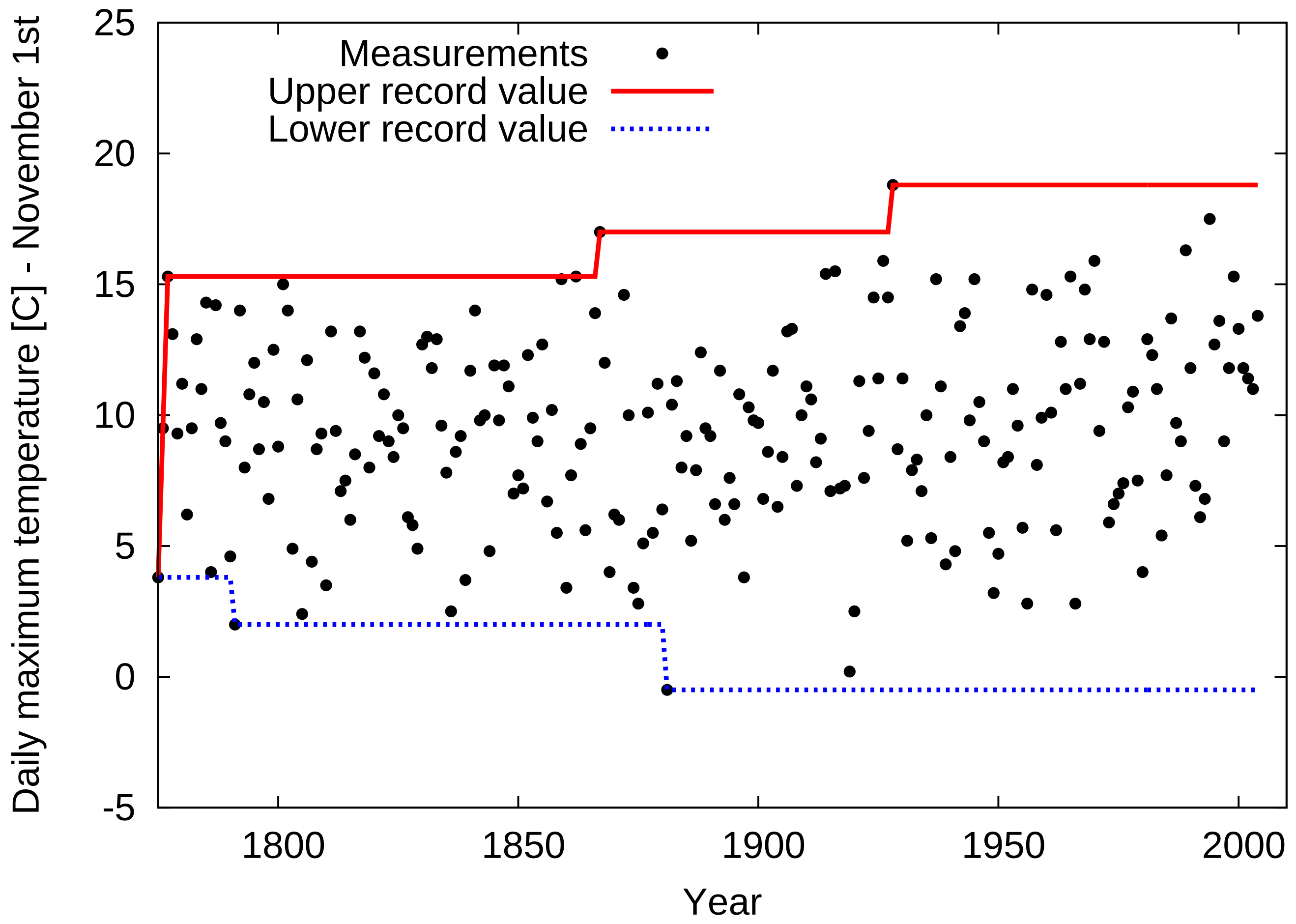
Temperature records in the USA



http://www.ucar.edu/news/releases/2009/maxmin.jsp

based on G.A. Meehl et al., Geophys. Res. Lett. **36** (2009) L23701

Daily temperature at Klementinum, Prague, on November 1

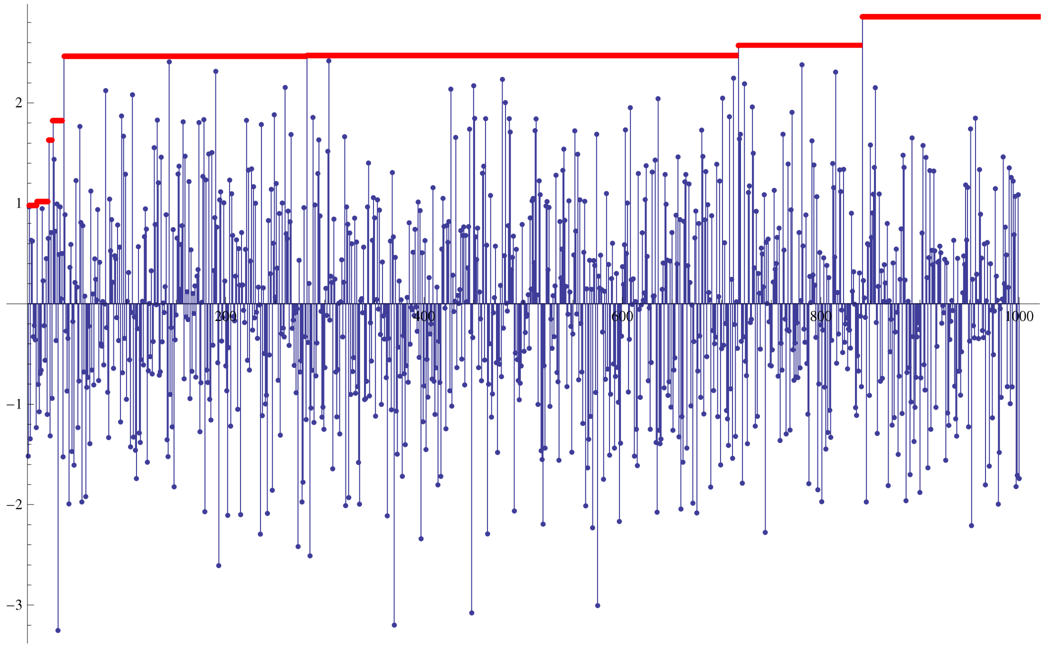


- 6 upper records and 3 lower records in 235 years
- How many records should we expect if the climate did not change?

Mathematical theory of records I

N. Glick, Am. Math. Mon. **85**, 2 (1978)

- A record is an entry in a sequence of random variables (RV's) X_n which is larger (upper record) or smaller (lower records) than all previous entries
- Example: 1000 independent Gaussian random variables



$\Rightarrow$ 7 upper records

Mathematical theory of records II

- If the RV's are independent and identically distributed (i.i.d.), the probability for a record at time n is $P_n = 1/n$ by symmetry

- The expected number of records up to time n is therefore

$$\langle R_n \rangle = \sum_{k=1}^{n} \frac{1}{k} = \ln(n) + \gamma + \mathscr{O}(1/n)$$

where $\gamma \approx 0.5772156649....$ is the Euler-Mascheroni constant,

- Because record events are independent, the variance of the number of records is

$$\langle (R_n - \langle R_n \rangle)^2 \rangle = \sum_{k=1}^{n} \left(\frac{1}{k} - \frac{1}{k^2} \right) = \ln(n) + \gamma - \frac{\pi^2}{6} + \mathscr{O}(1/n)$$

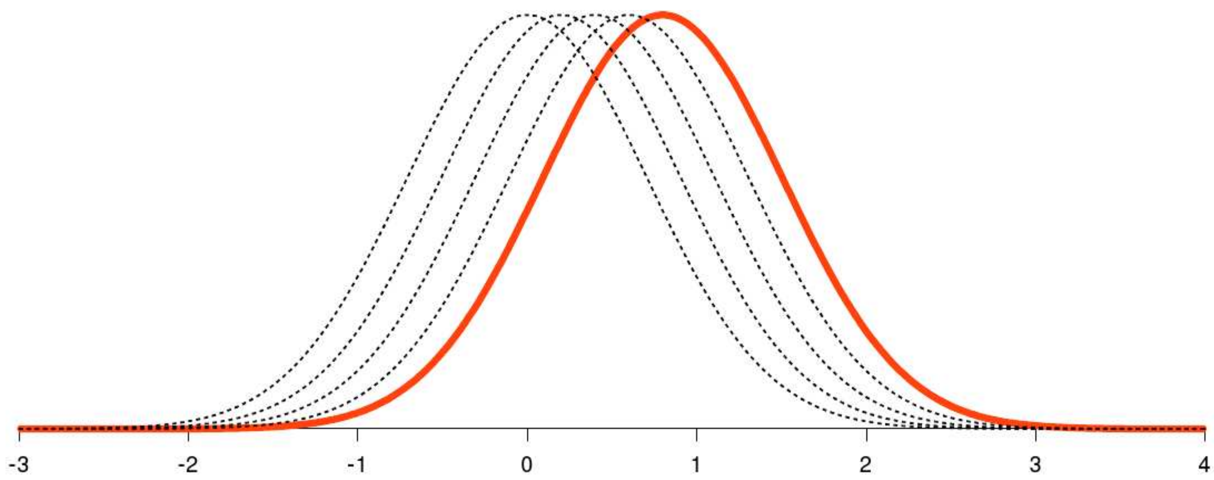
- In a constant climate we expect 6 records in 235 years and only 7.5 records in 1000 years!

Record-breaking temperatures and global warming

R.E. Benestad (2003); S. Redner & M.R. Petersen (2006)

- Question: Does global warming significantly increase the occurrence of record-breaking high daily temperatures?

- Model: The temperature on a given calendar day of the year is an independent Gaussian RV with constant standard deviation σ and a mean that increases at speed v

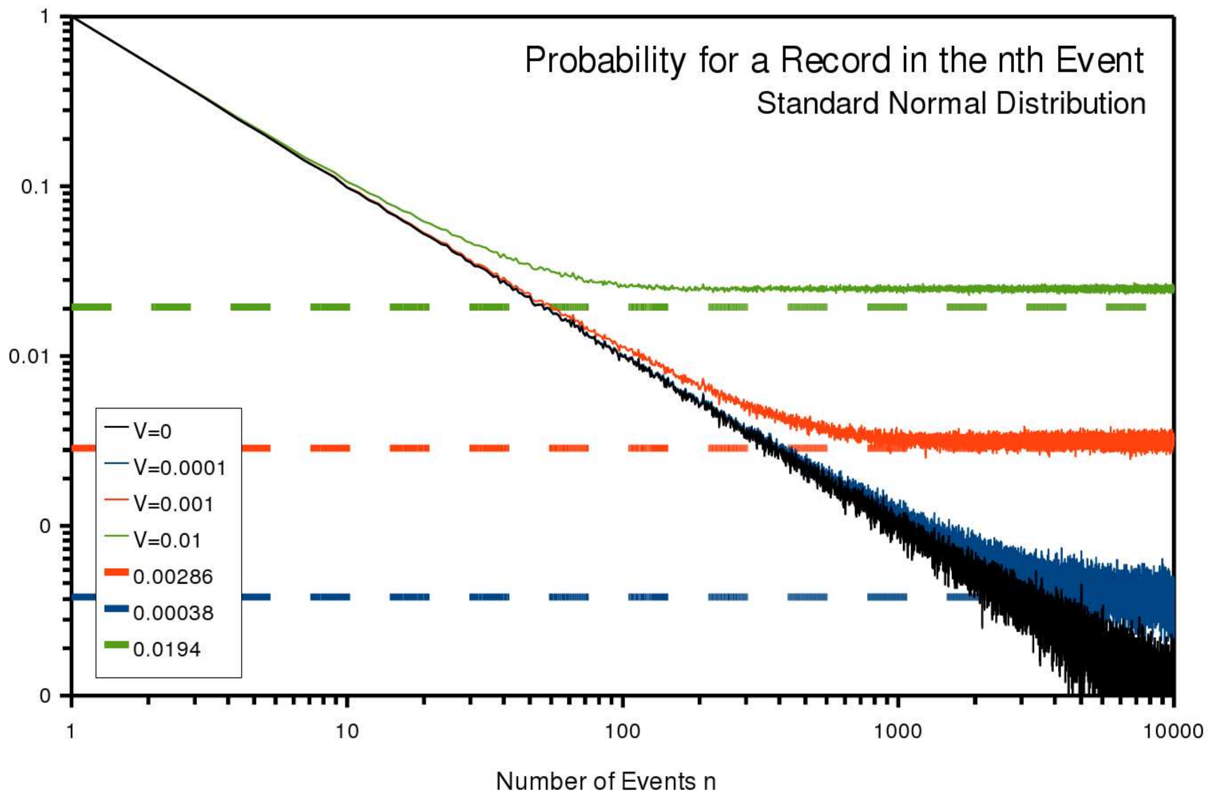


- Typical values: $v \approx 0.03^{\rm o}$C/yr, $\sigma \approx 3.5^{\rm o}$C $\Rightarrow v/\sigma \ll 1$

The linear drift model

R. Ballerini & S. Resnick (1985); J. Franke, G. Wergen, JK, JSTAT (2010) P10013

- General setting: Time series $X_n = Y_n + vn$ with i.i.d. RV's Y_n and $v > 0$
- For large n the record probability approaches a finite limit $\lim_{n\to\infty} P_n(v) > 0$



Approximate calculation of the record rate for small drift

- Let Y_n have probability density $p(y)$ and probability distribution function $q(x) = \int^x dy\ p(y)$. Then

$$P_n(v) = \int dx_n\ p(x_n - vn) \prod_{k=1}^{n-1} q(x_n - vk) = \int dx\ p(x) \prod_{k=1}^{n-1} q(x + vk)$$

- For small v we have $q(x+vk) \approx q(x) + vkp(x)$

$$\Rightarrow P_n \approx \int dx\ p(x) q(x)^{n-1} + \frac{vn(n-1)}{2} \int dx\ p(x)^2 q(x)^{n-2} = \frac{1}{n} + vI_n$$

with $I_n = \frac{n(n-1)}{2} \int dx\ p(x)^2 q(x)^{n-2}$

- For the Gaussian distributon a saddle point approximation for large n yields

$$P_n(v) \approx \frac{1}{n} + \frac{v}{\sigma} \frac{2\sqrt{\pi}}{e^2} \sqrt{\ln(n^2/8\pi)}$$

Comparison to temperature data

Data sets for daily temperatures

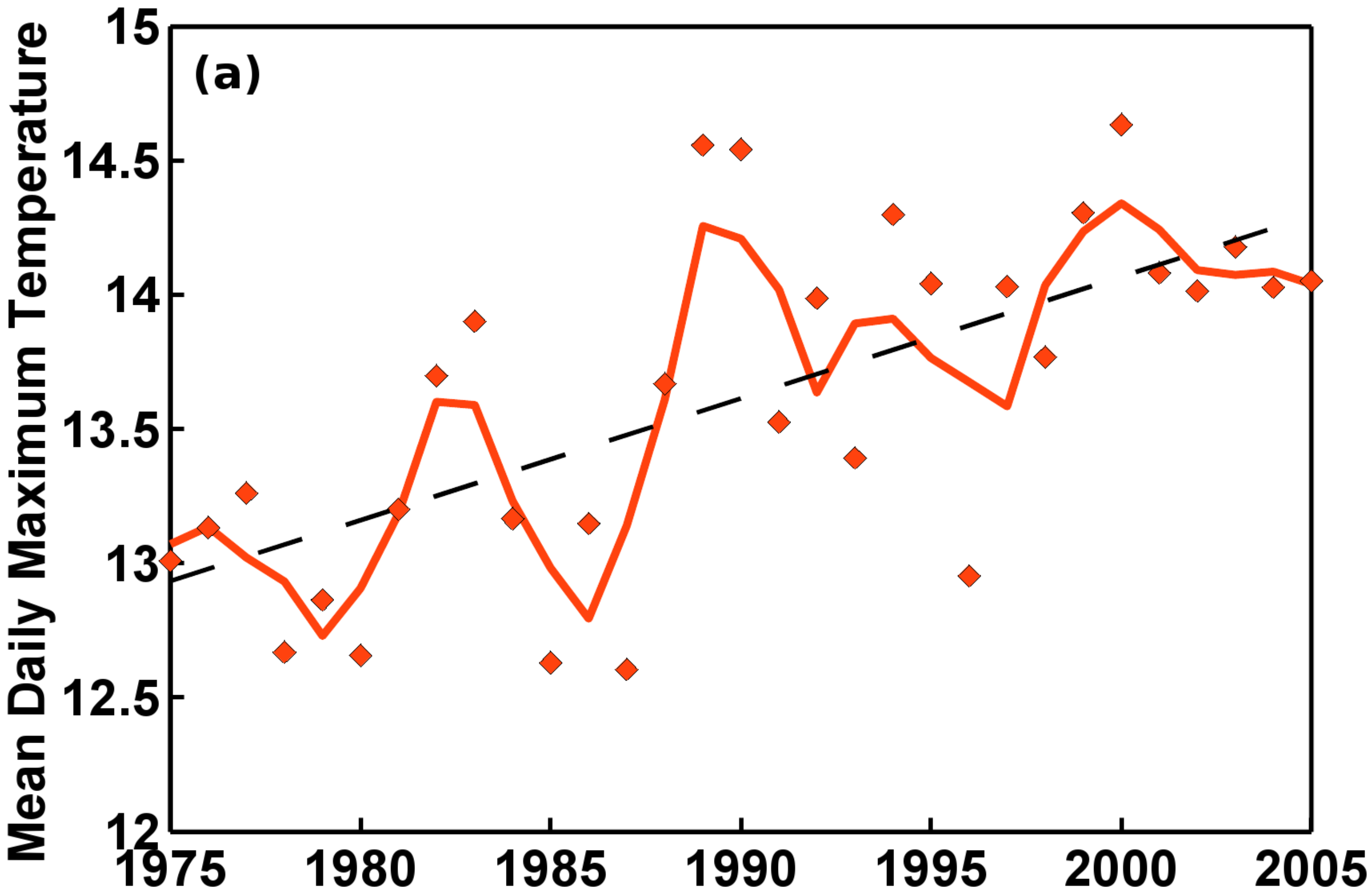
European data

- 43 stations over 100 year period 1906-2005
- 187 stations over 30 year period 1976-2005
- 30 year data: Constant warming rate $v \approx 0.047 \pm 0.003^{\mathrm{o}}$C/yr, standard deviation $\sigma \approx 3.4 \pm 0.3^{\mathrm{o}}$C $\Rightarrow v/\sigma \approx 0.014$

American data

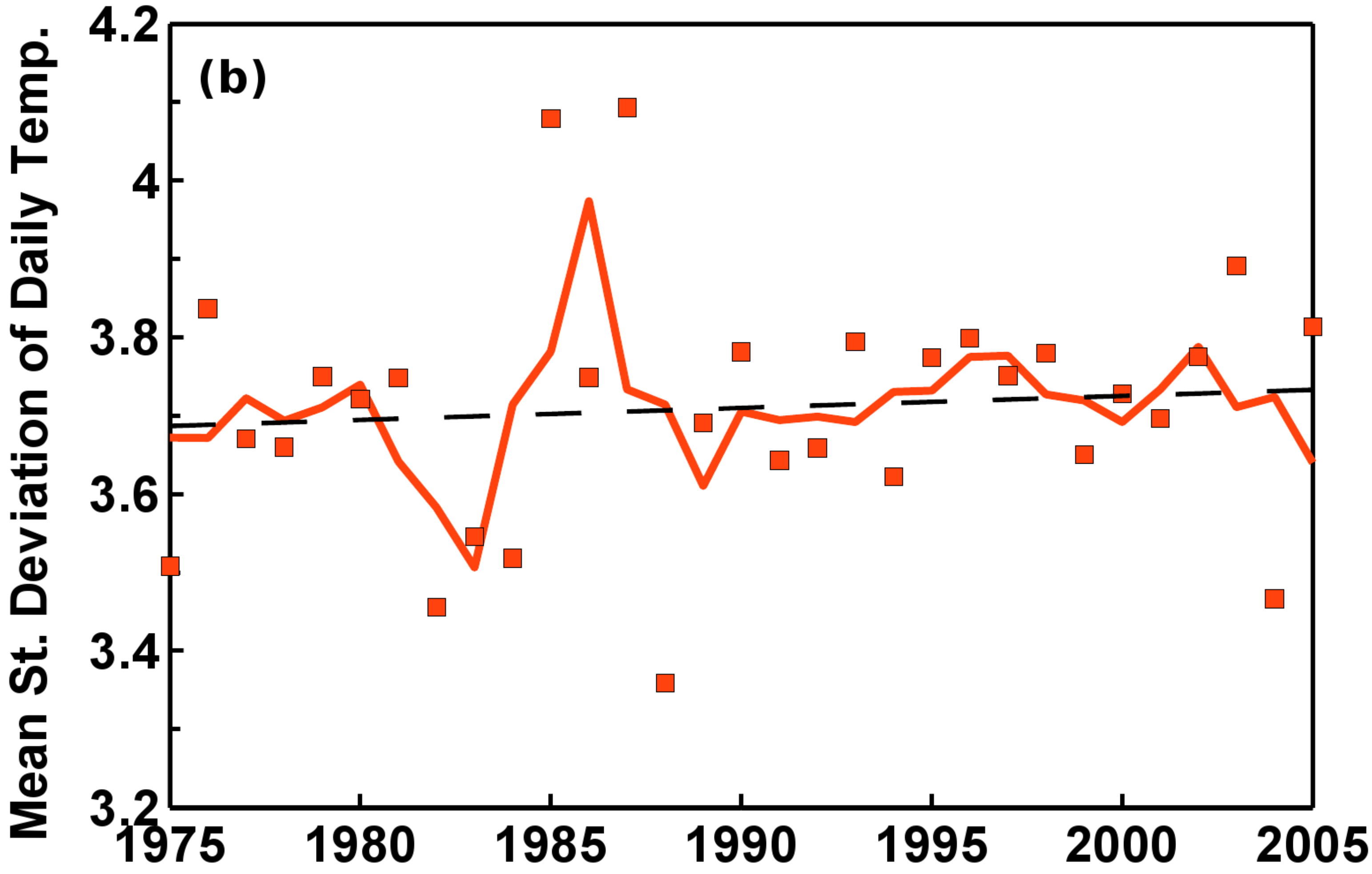
- 10 stations over 125 year period 1881-2005
- 207 stations over 30 year period 1976-2005
- Continental climate implies larger variability:
  $\sigma = 4.9 \pm 0.1^{\mathrm{o}}$C, $v = 0.025 \pm 0.002^{\mathrm{o}}$C/yr $\Rightarrow v/\sigma \approx 0.005$
- Significant effect of rounding to integer degrees Fahrenheit

European data: Mean daily maximum temperature

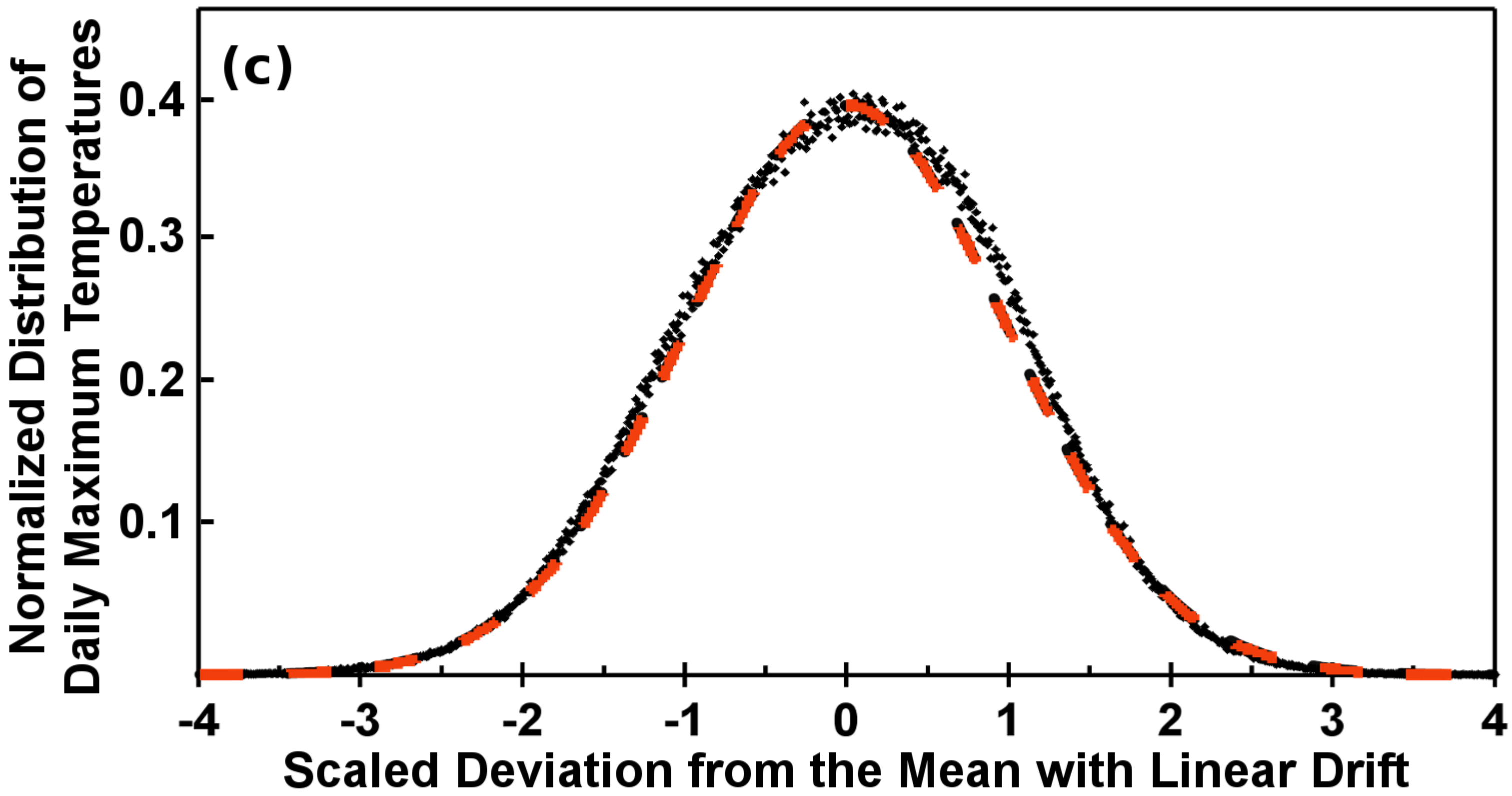


Full line: Sliding 3-year average

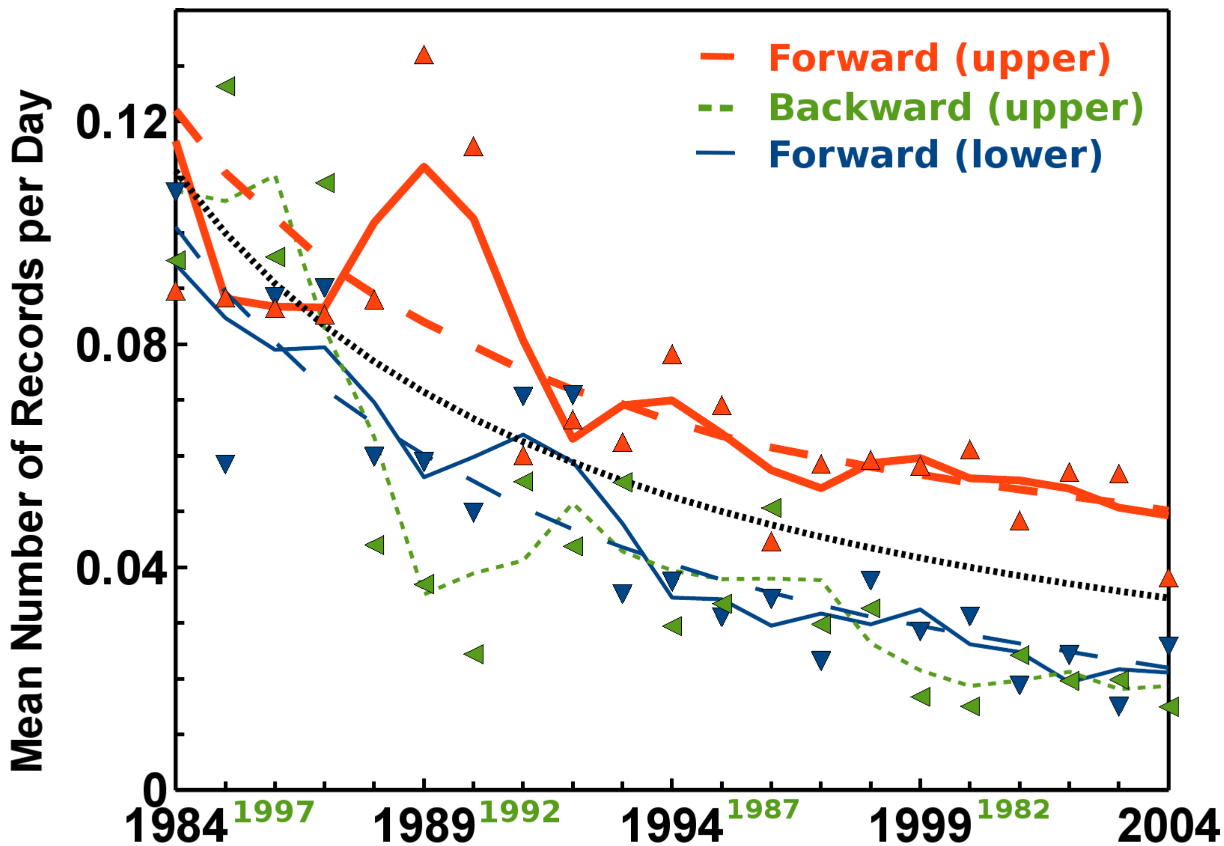
European data: No trend in the standard deviation



European data: Temperature fluctuations are Gaussian

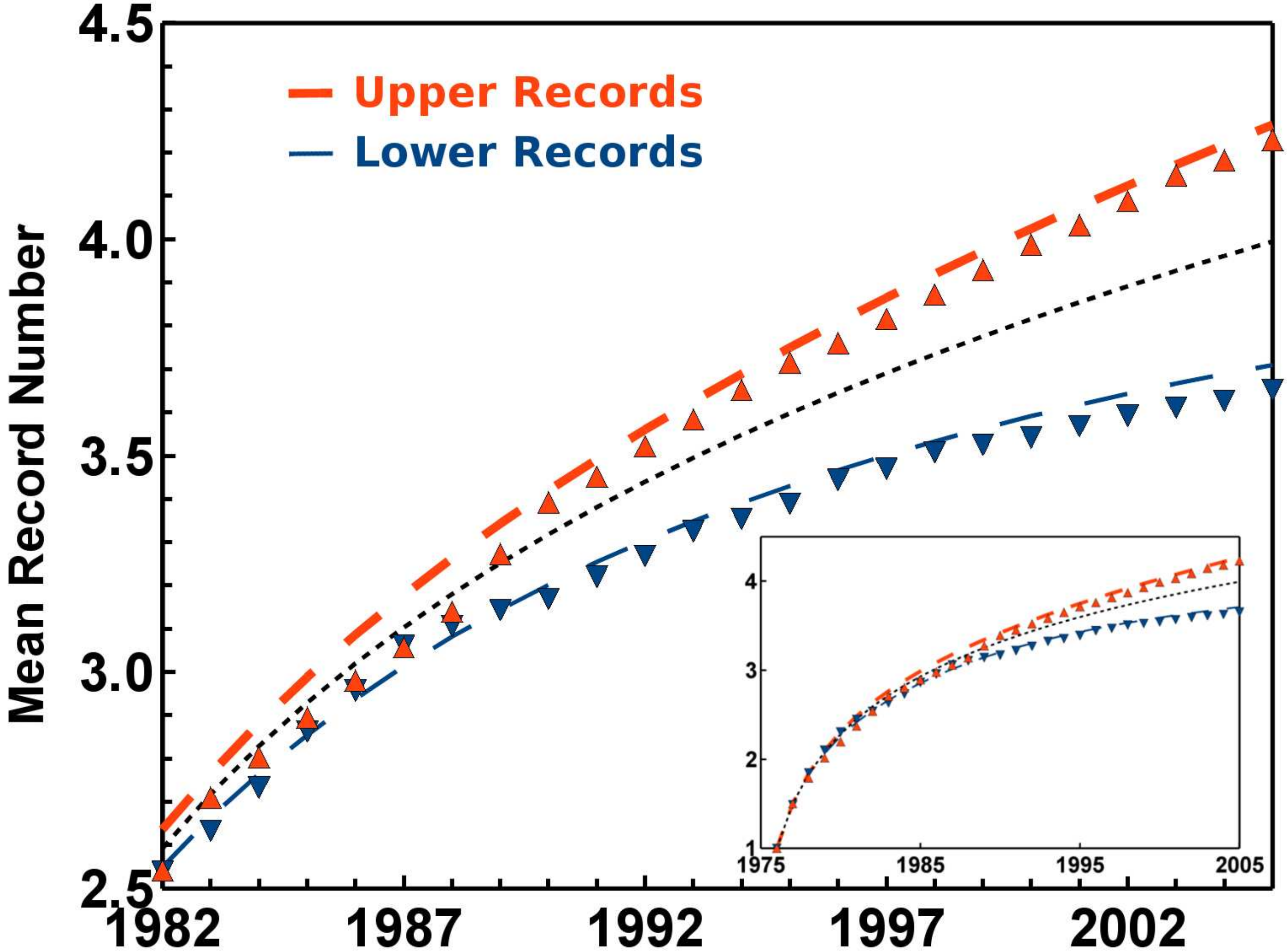


Record frequency in Europe: 1976-2005



- Expected number of records in stationary climate: $\frac{365}{30} \approx 12$
- Observed record rate is increased by about 40 % $\Rightarrow$ 5 additional records

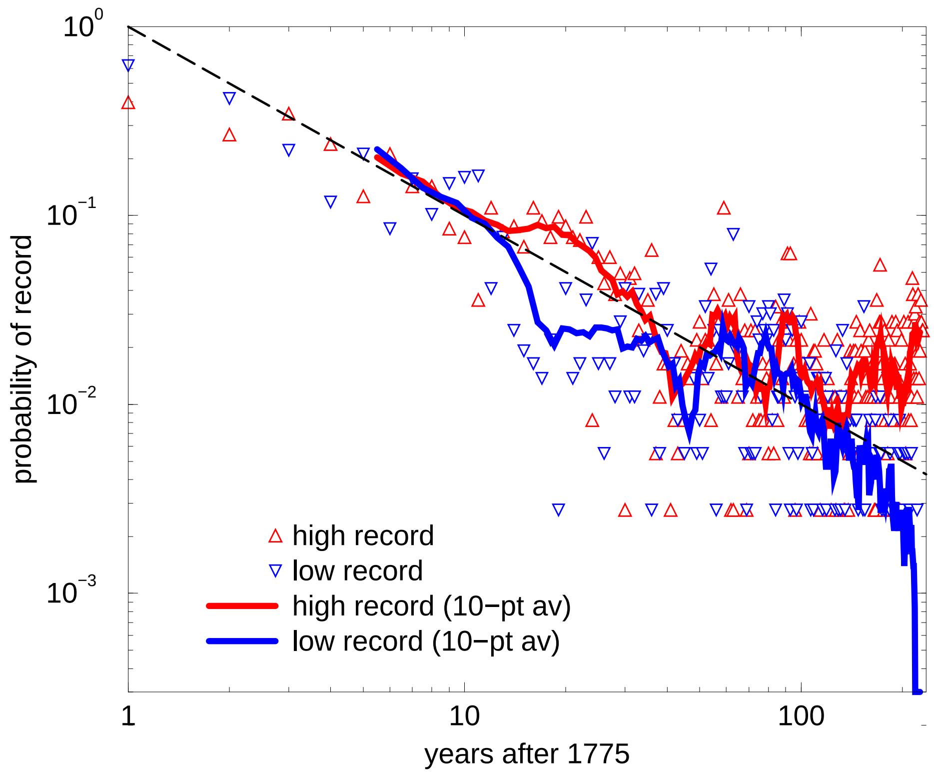
Mean record number: 1976-2005



Long term prospects

- If the current warming rate continues, the daily rate of upper records with respect to 1976 will saturate at $P^* \approx 1/30$ towards the end of this century

- Saturation is already visible in the 235 year data from Klementinum



Courtesy of Sid Redner

Correlations between record events

Record correlations in the linear drift model

G. Wergen, J. Franke, JK, J. Stat. Phys.**144** (2011) 1206

- Record events in series of i.i.d. random variables are independent
- To quantify dependence in the general case consider the normalized joint probability

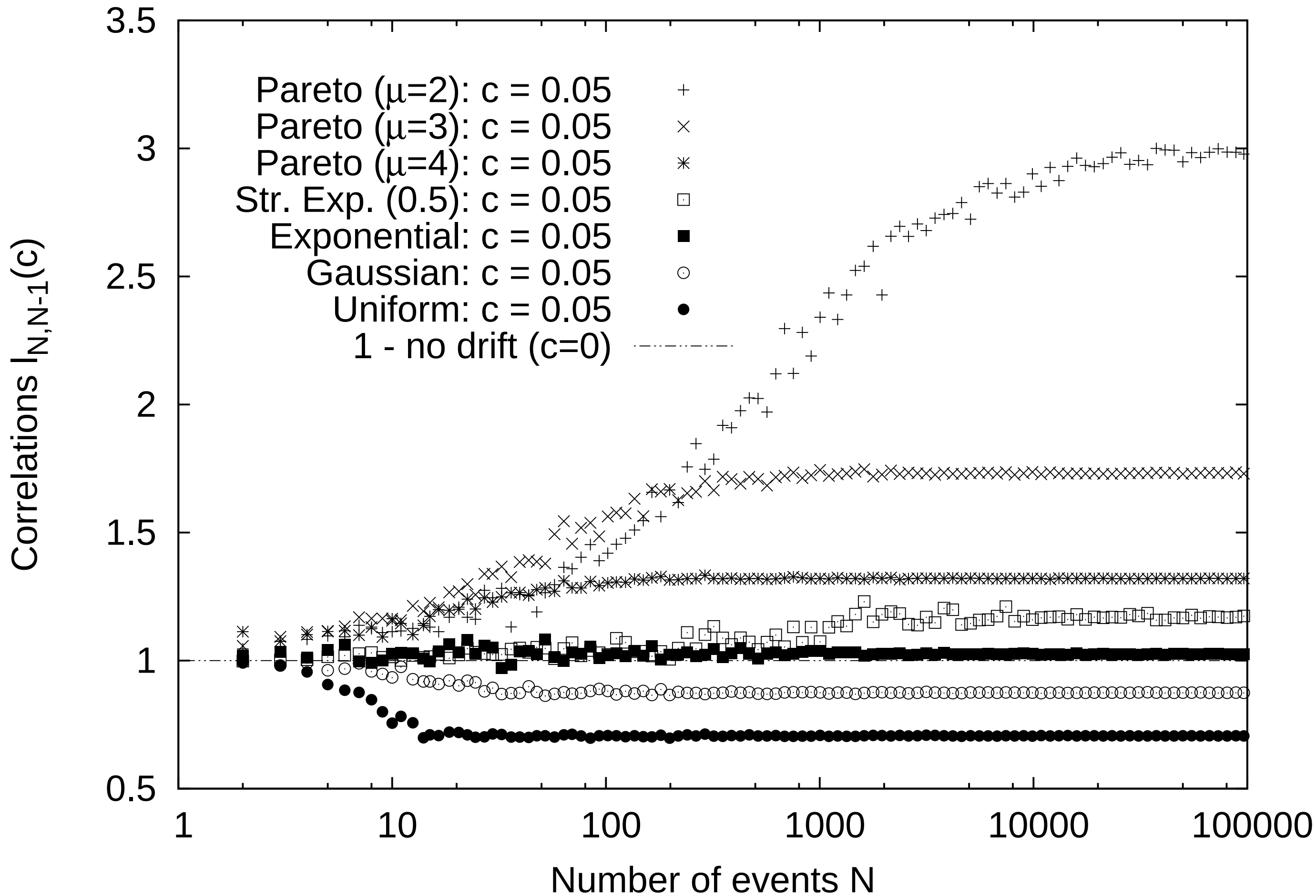
$$l_{N,N-1} = \frac{P_{N,N-1}}{P_N P_{N-1}} \quad \text{with} \quad P_{N,N-1} = \mathrm{Prob}[X_N \text{ record and } X_{N-1} \text{ record}]$$

- Small v expansion yields $l_{N,N-1}(v) \approx 1 + vJ_N(v)$ with

$$J_N \approx -\frac{1}{2}N^4 \frac{d}{dN}\left(\frac{2}{N^2}I_N\right) - 2NI_N \approx \kappa N I_N$$

where κ is the extreme value index of $p(x) \sim (1+\kappa x)^{-\frac{\kappa+1}{\kappa}}$

- Records cluster (repel) for distributions broader (more narrow) than an exponential:

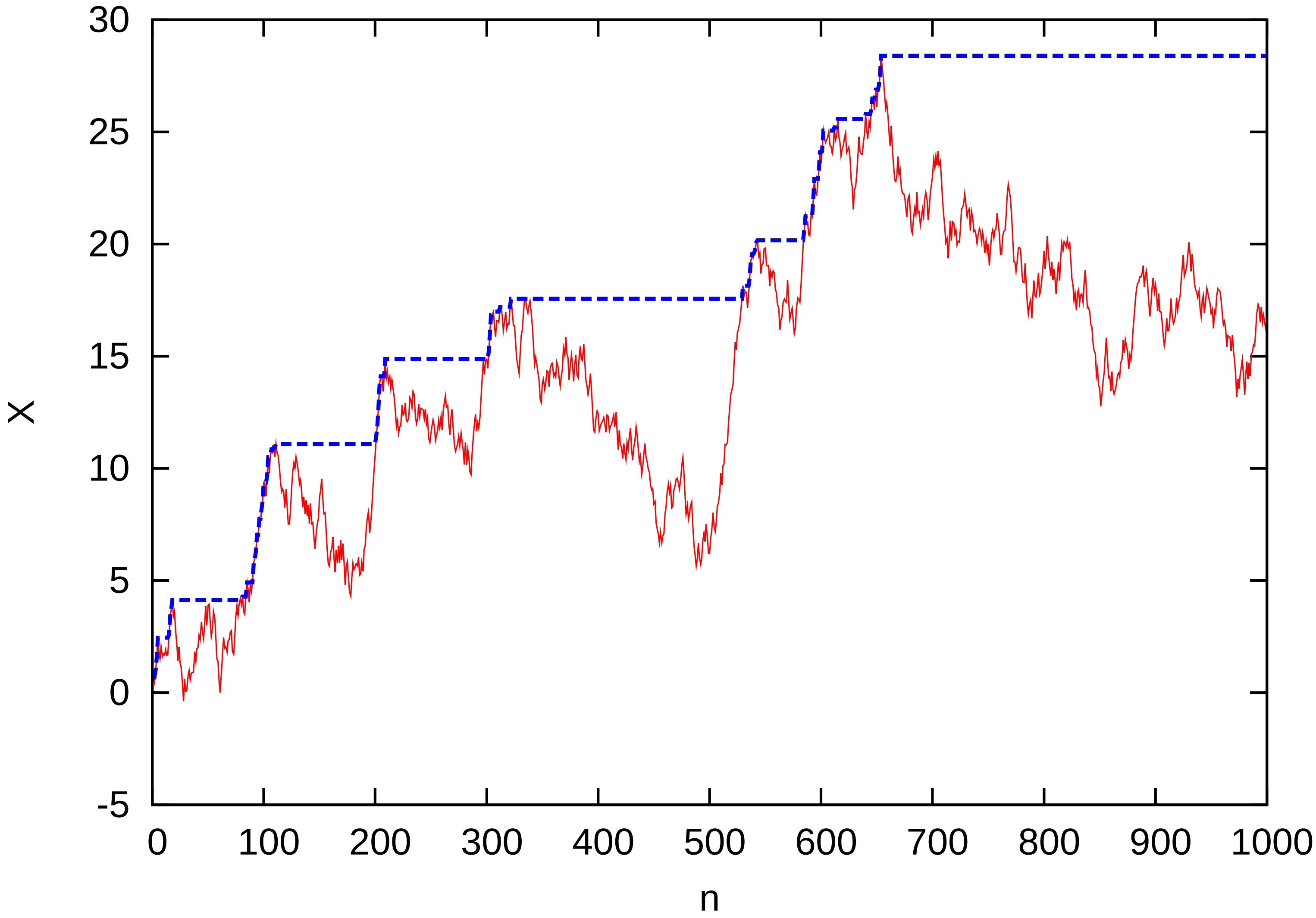


- This suggests a statistical test for fat-tailed distributions in small data sets

J. Franke, G. Wergen, JK, arXiv:1109.2061

Random walks & market fluctuations

Records in random walks



$\Longrightarrow$ 65 records in 1000 time steps

Records in random walks

S.N. Majumdar & R.M. Ziff, PRL **101**, 050601 (2008)

- Simple one-dimensional random walk is defined by

$$X_n = \sum_{k=1}^{n} \eta_k$$

with i.i.d. RV's η_k drawn from a symmetric, continuous distribution $\phi(\eta)$

- Based on a theorem of Sparre Andersen (1954), the probability of having m records in n steps is found to be

$$P(m,n) = \binom{2n-m+1}{n} 2^{-2n+m-1} \;\; \rightarrow \;\; \frac{1}{\sqrt{\pi n}} \exp[-m^2/4n]$$

- Mean number of records: $\langle R_n \rangle \approx \sqrt{4n/\pi} \gg \ln n + \gamma$

- This result does not require $\phi(\eta)$ to have finite variance
$\Rightarrow$ valid also for superdiffusive (Lévy) walks!

Biased random walks and stock market fluctuations

G. Wergen, M. Bogner, JK, PRE **83** 051109 (2011)

- Basic model of a fluctuating stock price S_n is the geometric random walk

$$S_n = e^{X_n} = \exp[\sum_{k=1}^{n} \eta_k]$$

which obviously has the same record statistics as the random walk itself.

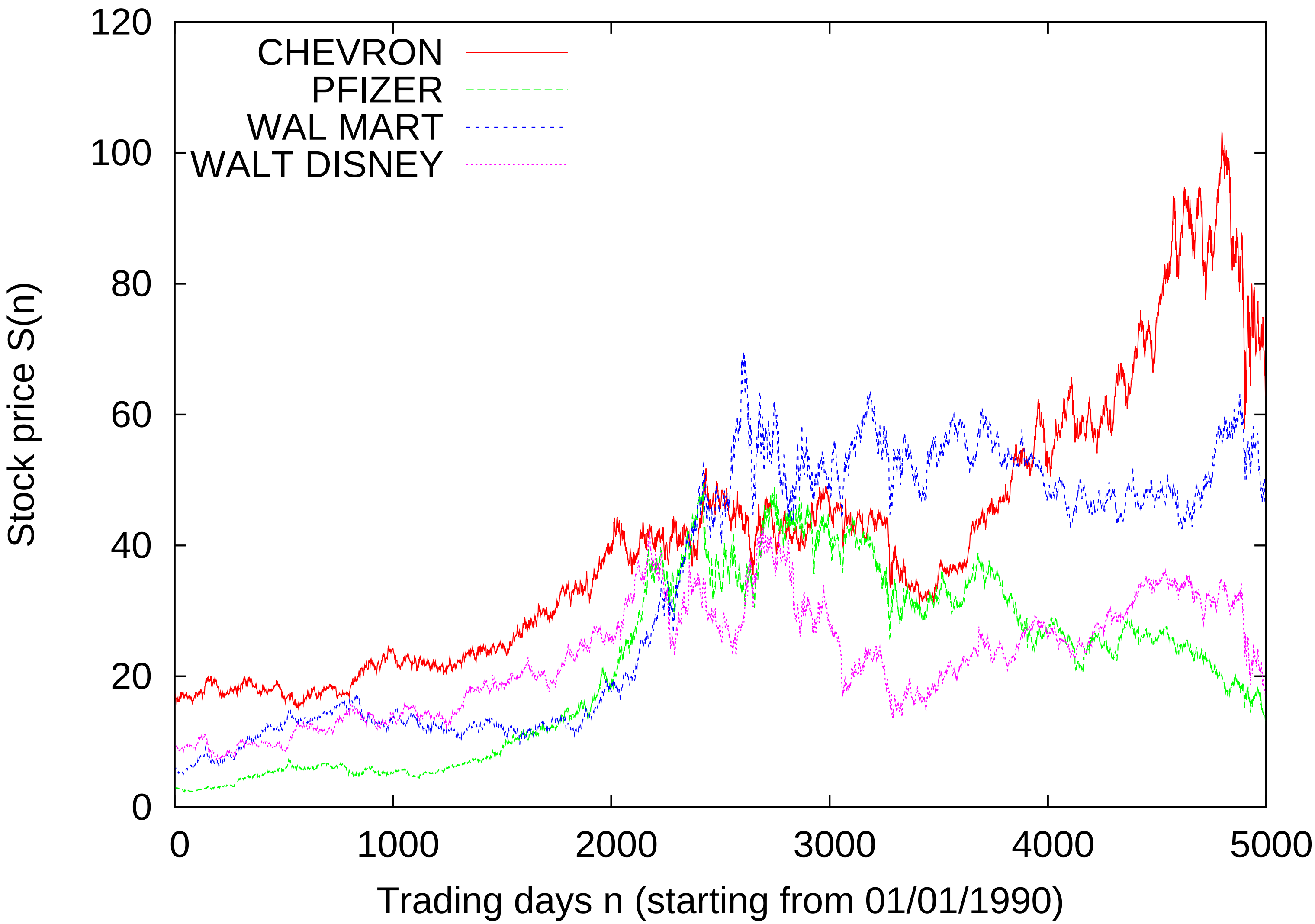
- Stock prices typically display an upward bias reflecting the long-term interest rate $\Rightarrow$ consider random walk with drift: $X_n \to X_n + vn$

- Leading order expansion in v yields

$$\langle R_n \rangle \approx \sqrt{\frac{4n}{\pi}} + \frac{v}{\sigma}\frac{\sqrt{2}}{\pi}\left[n \arctan(\sqrt{n}) - \sqrt{n}\right] \to \sqrt{\frac{4n}{\pi}} + \frac{vn}{\sqrt{2}\sigma}$$

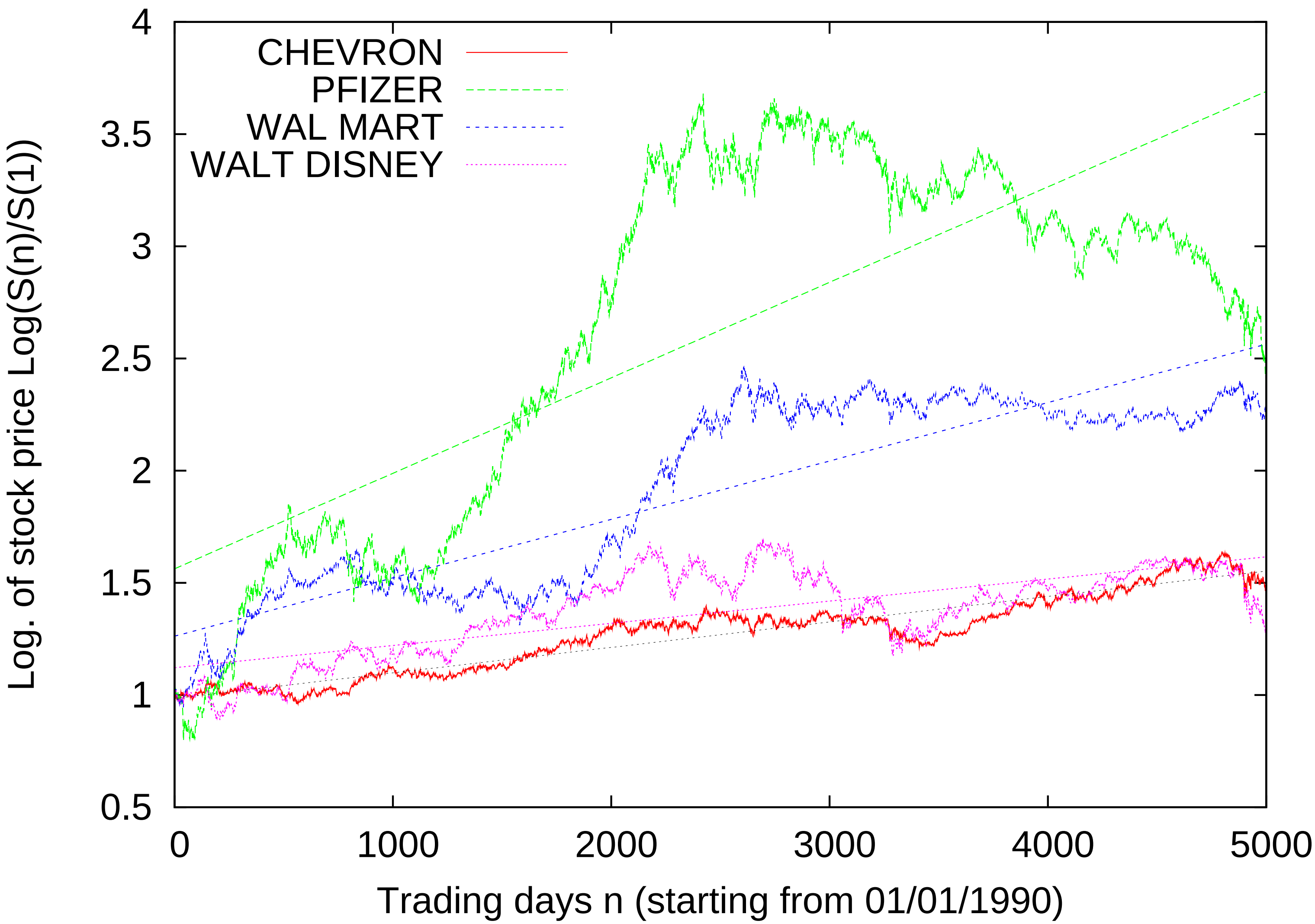
- For $n \to \infty$ the record probability P_n approaches a positive constant

The S&P 500 index 1.1.1990-31.3.2009



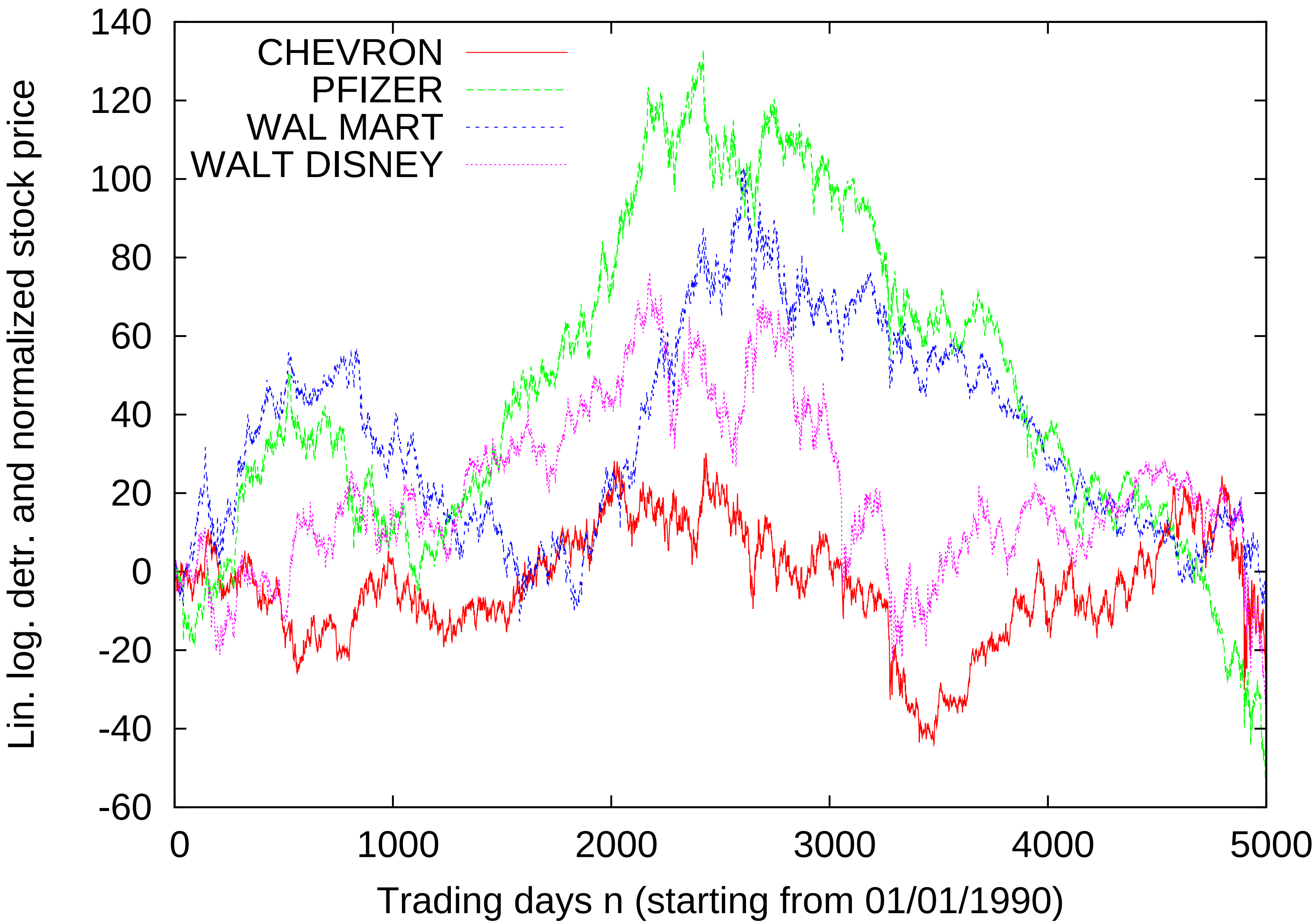
- raw data

The S&P 500 index 1.1.1990-31.3.2009



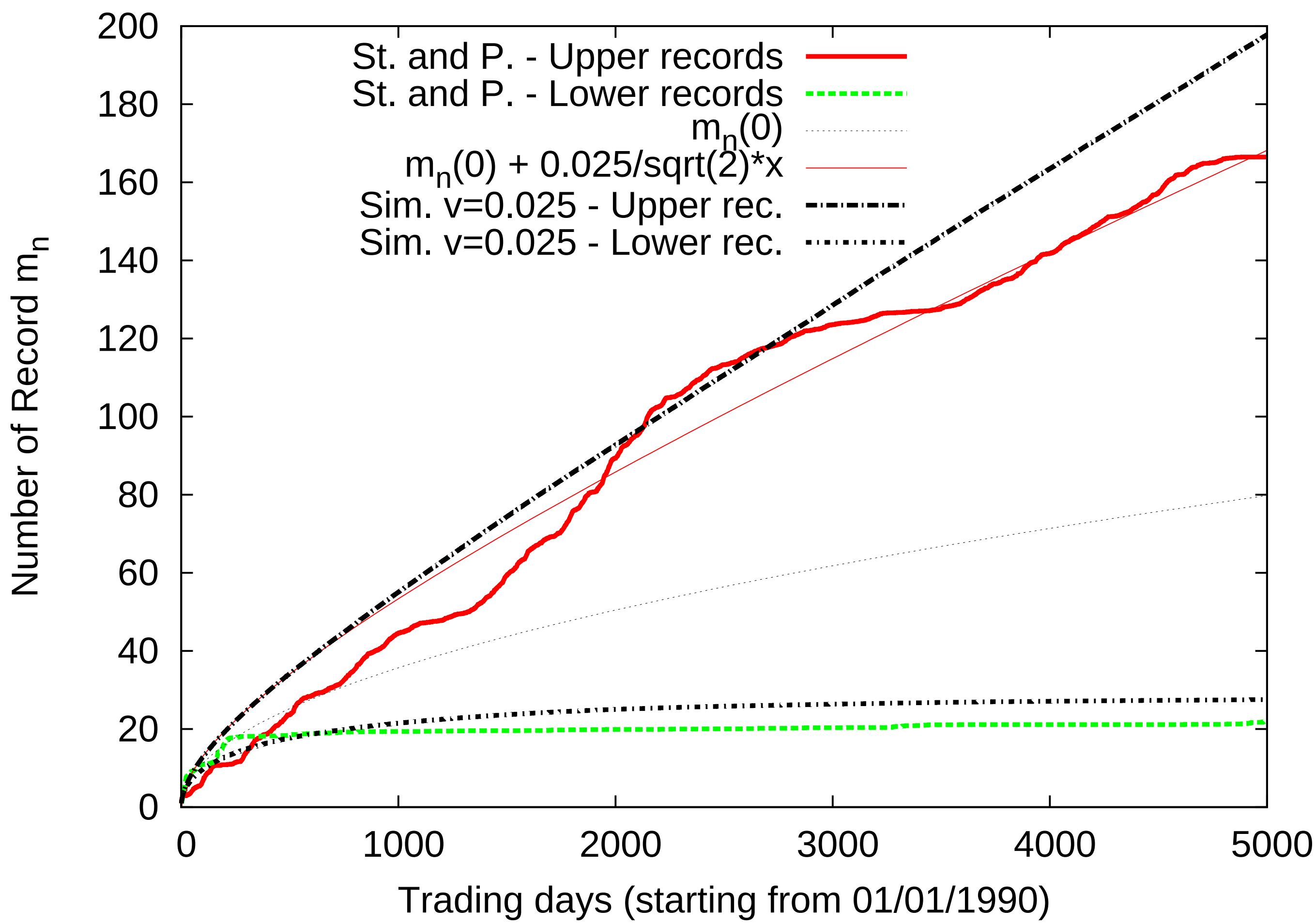
- logarithmic stock prices with linear fits

The S&P 500 index 1.1.1990-31.3.2009



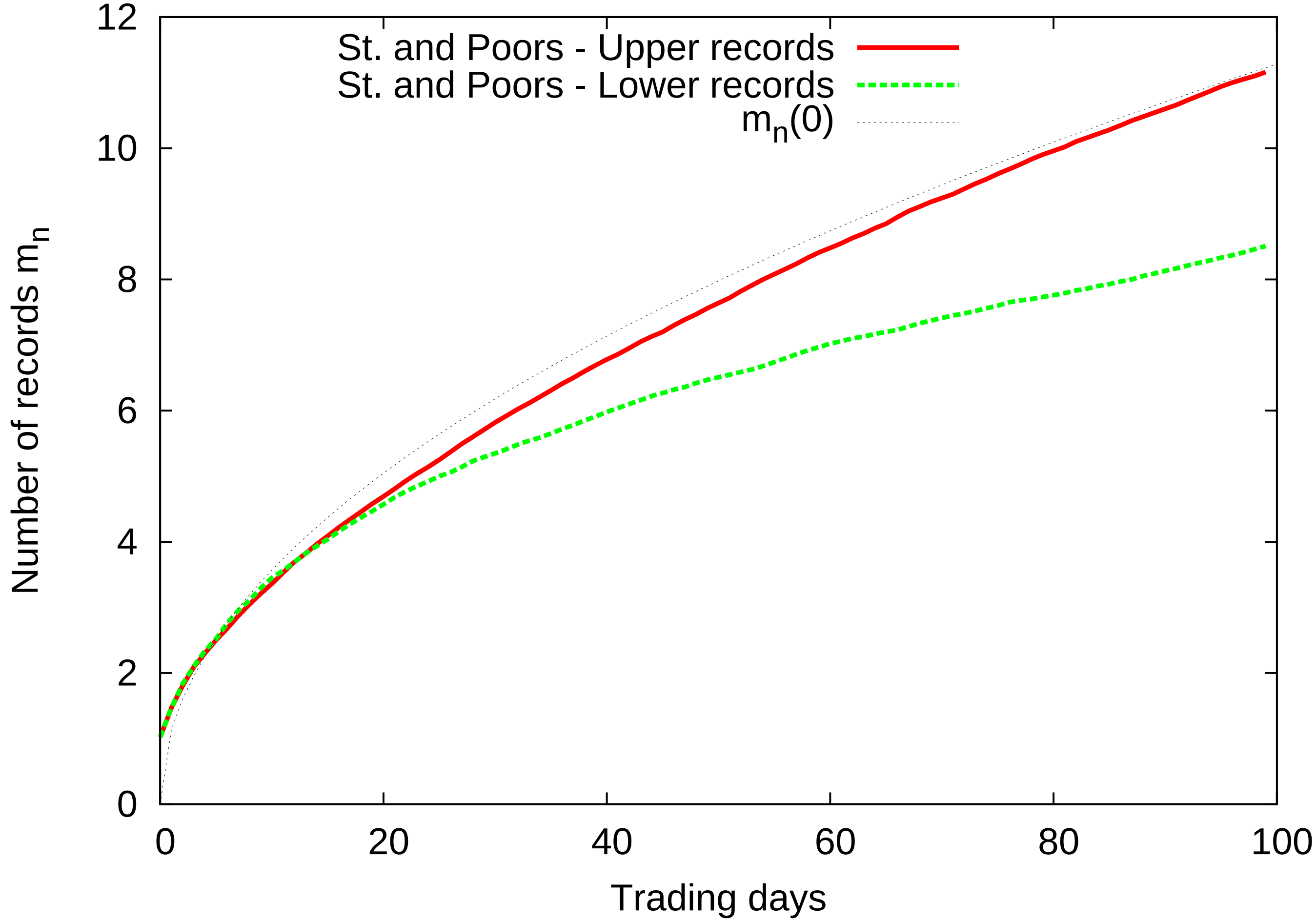
- logarithmic stock prices detrended and normalized

Upper and lower records in the S&P 500



- Record events averaged over 366 stocks
- Excess of upper records well predicted by analytic model with $v/\sigma = 0.025$

Upper and lower records in the S&P 500



- Time series were subdivided into pieces of length 100 and detrended
- Upper records conform to random walk prediction, but lower records do not

Conclusions

Record-breaking temperatures

- Global warming affects the rate of record-breaking temperatures in moderate but significant way
- Key predictor of excess record events is the ratio of warming rate to temperature variability v/σ
- If current trend persists, by the end of this century the rate of high temperature records relative to 1976 will become constant

Record-breaking stock prices

- Minimal model of biased random walk accounts quantitatively for the occurrence of upper records in the S&P 500
- Suppression of lower records remains to be explained

Thank you!