Detecting Long-Range Dependence

A Comparative Approach

Henning Rust

hrust@uni-potsdam.de

E2C2 Summer School, Comorova, September 2007

F^3 E^2C^2 Summer Cup

Prices

- one answer – price

Questions

- What is the difference between the spectral density and the periodogram?
- What is the difference between SRD and LRD?
- Which theorem links the ACF and the spectral density?

F^3 E^2C^2 Summer Cup

Prices

- one answer – price

Questions

- What is the difference between the spectral density and the periodogram?
- What is the difference between SRD and LRD?
- Which theorem links the ACF and the spectral density?

F^3 E^2C^2 Summer Cup

Prices

- one answer – price

Questions

- What is the difference between the spectral density and the periodogram?
- What is the difference between SRD and LRD?
- Which theorem links the ACF and the spectral density?

F^3 E^2C^2 Summer Cup

Prices

- one answer – price

Questions

- What is the difference between the spectral density and the periodogram?
- What is the difference between SRD and LRD?
- Which theorem links the ACF and the spectral density?

LRD Parameters

Time Domain

$$\tau \to \infty : \quad \rho(\tau) \propto \tau^{-\gamma} \quad ; \quad 0 < \gamma < 1 \tag{1}$$

Spectral Domain

$$\omega \to 0 : \quad S(\omega) \propto |\omega|^{-\beta} \quad ; \quad 0 < \beta < 1 \tag{2}$$

Relation

$$1 - \gamma = \beta = 2d \tag{3}$$

LRD Parameters

Time Domain

$$\tau \to \infty : \quad \rho(\tau) \propto \tau^{-\gamma} \quad ; \quad 0 < \gamma < 1 \tag{1}$$

Spectral Domain

$$\omega \to 0 : \quad S(\omega) \propto |\omega|^{-\beta} \quad ; \quad 0 < \beta < 1 \tag{2}$$

Relation

$$1 - \gamma = \beta = 2d \tag{3}$$

LRD Parameters

Time Domain

$$\tau \to \infty : \quad \rho(\tau) \propto \tau^{-\gamma} \quad ; \quad 0 < \gamma < 1 \tag{1}$$

Spectral Domain

$$\omega \to 0 : \quad S(\omega) \propto |\omega|^{-\beta} \quad ; \quad 0 < \beta < 1 \tag{2}$$

Relation

$$1 - \gamma = \beta = 2d \tag{3}$$

Outline

Effects of Long-Range Dependence

Suppose x_t realisation of LRD process:

- reduced rate of convergence, e.g.
  - of estimators $N^{-1/2}$, e.g. mean, quantiles, regression parameters, etc
  - of extremes towards the GEV
- increased uncertainty of estimation
- altered levels of significance in tests (compared to white noise/SRD hypothesis), e.g.
  - change point detection
  - trend detection
- reduced gain in information due to dependence

Effects of Long-Range Dependence

Suppose x_t realisation of LRD process:

- reduced rate of convergence, e.g.
  - of estimators $N^{-1/2}$, e.g. mean, quantiles, regression parameters, etc
  - of extremes towards the GEV
- increased uncertainty of estimation
- altered levels of significance in tests (compared to white noise/SRD hypothesis), e.g.
  - change point detection
  - trend detection
- reduced gain in information due to dependence

Effects of Long-Range Dependence

Suppose x_t realisation of LRD process:

- reduced rate of convergence, e.g.
  - of estimators $N^{-1/2}$, e.g. mean, quantiles, regression parameters, etc
  - of extremes towards the GEV
- increased uncertainty of estimation
- altered levels of significance in tests (compared to white noise/SRD hypothesis), e.g.
  - change point detection
  - trend detection
- reduced gain in information due to dependence

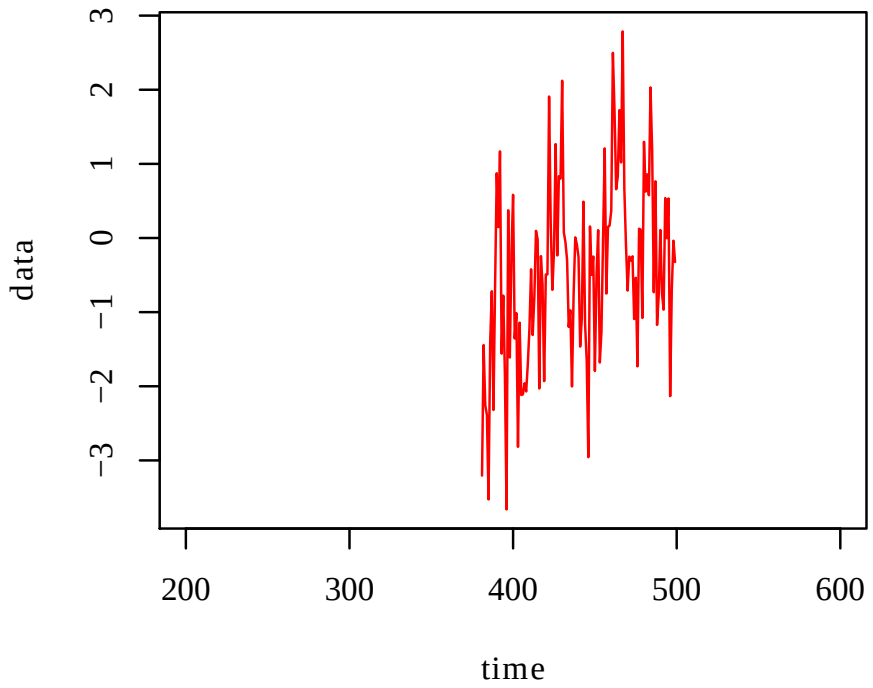
Effects of Long-Range Dependence

Suppose x_t realisation of LRD process:

- reduced rate of convergence, e.g.
  - of estimators $N^{-1/2}$, e.g. mean, quantiles, regression parameters, etc
  - of extremes towards the GEV
- increased uncertainty of estimation
- altered levels of significance in tests (compared to white noise/SRD hypothesis), e.g.
  - change point detection
  - trend detection
- reduced gain in information due to dependence

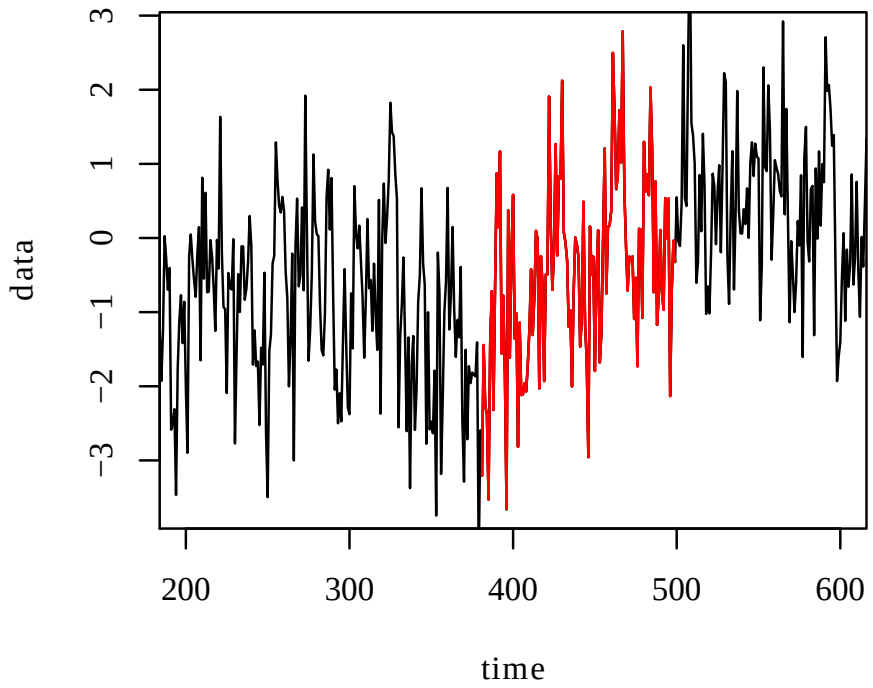
Spurious Trends

Standard Example



Spurious Trends

Standard Example



- mission imposible for statistics
- recall Michael: "Consider the physics behind it!"
- your choice: physics or noise

Why is a reliable detection important?

If choice is made, than

Ignoring LRD although present

- false detection of trends
- underestimation of confidence intervals, e.g. mean or maxima, return levels, regression parameters

Falsely detecting LRD

- ignoring trends/signals although present
- overestimation of uncertainty
- put off from paths of physical explanation

Why is a reliable detection important?

If choice is made, than

Ignoring LRD although present

- false detection of trends
- underestimation of confidence intervals, e.g. mean or maxima, return levels, regression parameters

Falsely detecting LRD

- ignoring trends/signals although present
- overestimation of uncertainty
- put off from paths of physical explanation

Detecting Long-Range Dependence

. . . is in principle impossible

- asymptotic property of a stochastic process
- for an observed series exists always a SRD model
- LRD can never be identified unambiguously
- but

. . . one can find decent arguments

- study largest time scales available
- use SRD and LRD models

- principle: exclude complex SRD models in favour of simpler LRD models, if fit equally well

Detecting Long-Range Dependence

. . . is in principle impossible

- asymptotic property of a stochastic process
- for an observed series exists always a SRD model
- LRD can never be identified unambiguously
- but

. . . one can find decent arguments

- study largest time scales available
- use SRD and LRD models
- principle: exclude complex SRD models in favour of simpler LRD models, if fit equally well

Detecting Long-Range Dependence

. . . is in principle impossible

- asymptotic property of a stochastic process
- for an observed series exists always a SRD model
- LRD can never be identified unambiguously
- but

. . . one can find decent arguments

- study largest time scales available
- use SRD and LRD models
  - LRD: if all SRD alternatives can be excluded (would be perfect, but impossible)
  - but: try to exclude most likely SRD models
- principle: exclude complex SRD models in favour of simpler LRD models, if fit equally well

Detecting Long-Range Dependence

. . . is in principle impossible

- asymptotic property of a stochastic process
- for an observed series exists always a SRD model
- LRD can never be identified unambiguously
- but

. . . one can find decent arguments

- study largest time scales available
- use SRD and LRD models
  - LRD: if all SRD alternatives can be excluded (would be perfect, but impossible)
  - but: try to exclude most likely SRD models
- principle: exclude complex SRD models in favour of simpler LRD models, if fit equally well

Detecting Long-Range Dependence

. . . is in principle impossible

- asymptotic property of a stochastic process
- for an observed series exists always a SRD model
- LRD can never be identified unambiguously
- but

. . . one can find decent arguments

- study largest time scales available
- use SRD and LRD models
  - LRD: if all SRD alternatives can be excluded (would be perfect, but impossible)
  - but: try to exclude most likely SRD models
- principle: exclude complex SRD models in favour of simpler LRD models, if fit equally well

Various Stratgegies

Semi-parametric

- parametrise asympt. relation $I(\omega_j) \propto \omega_j^{-\beta}$
- choose bandwidth $\omega_{\min} 0 < \omega_j < \omega_{\max}$

Full-parametric

- parametrise spectral density FARIMA$[p, d, q]$
- model selection choose p and q

Heuristic Approaches

- e.g. Detrended Fluctuation Analysis (DFA)

Various Stratgegies

Semi-parametric

- parametrise asympt. relation $I(\omega_j) \propto \omega_j^{-\beta}$
- choose bandwidth $\omega_{\min} 0 < \omega_j < \omega_{\max}$

Full-parametric

- parametrise spectral density FARIMA$[p, d, q]$
- model selection choose p and q

Heuristic Approaches

- e.g. Detrended Fluctuation Analysis (DFA)

Various Stratgegies

Semi-parametric

- parametrise asympt. relation $I(\omega_j) \propto \omega_j^{-\beta}$
- choose bandwidth $\omega_{\min} 0 < \omega_j < \omega_{\max}$

Full-parametric

- parametrise spectral density FARIMA$[p, d, q]$
- model selection choose p and q

Heuristic Approaches

- e.g. Detrended Fluctuation Analysis (DFA)

Various Stratgegies

Semi-parametric

- parametrise asympt. relation $I(\omega_j) \propto \omega_j^{-\beta}$
- choose bandwidth $\omega_{\min} 0 < \omega_j < \omega_{\max}$

Full-parametric

- parametrise spectral density FARIMA$[p, d, q]$
- model selection choose p and q

Heuristic Approaches

- e.g. Detrended Fluctuation Analysis (DFA)

Semi-Parametric Estimation

Log-Periodogram Regression

$$\log I(\omega_j) = c - \beta \log \omega_j + u_j \tag{4}$$

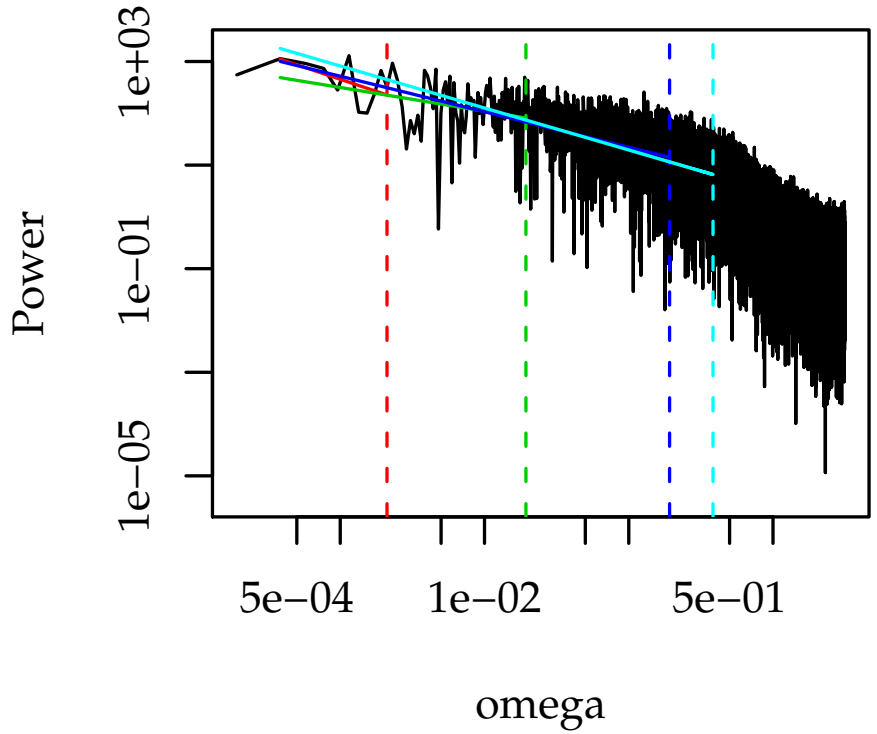
Properties

- slope as an estimator for β
- bandwidth: $\omega_{\min} \leq \omega_j \leq \omega_{\max}$, m number of ω_j included

$$m^{1/2}(\hat{\beta} - \beta) \to \xi \quad ; \qquad \xi \sim N\left(0, \frac{\pi^2}{24}\right) \tag{5}$$

- confidence intervals and testing

Semi-Parametric Estimation



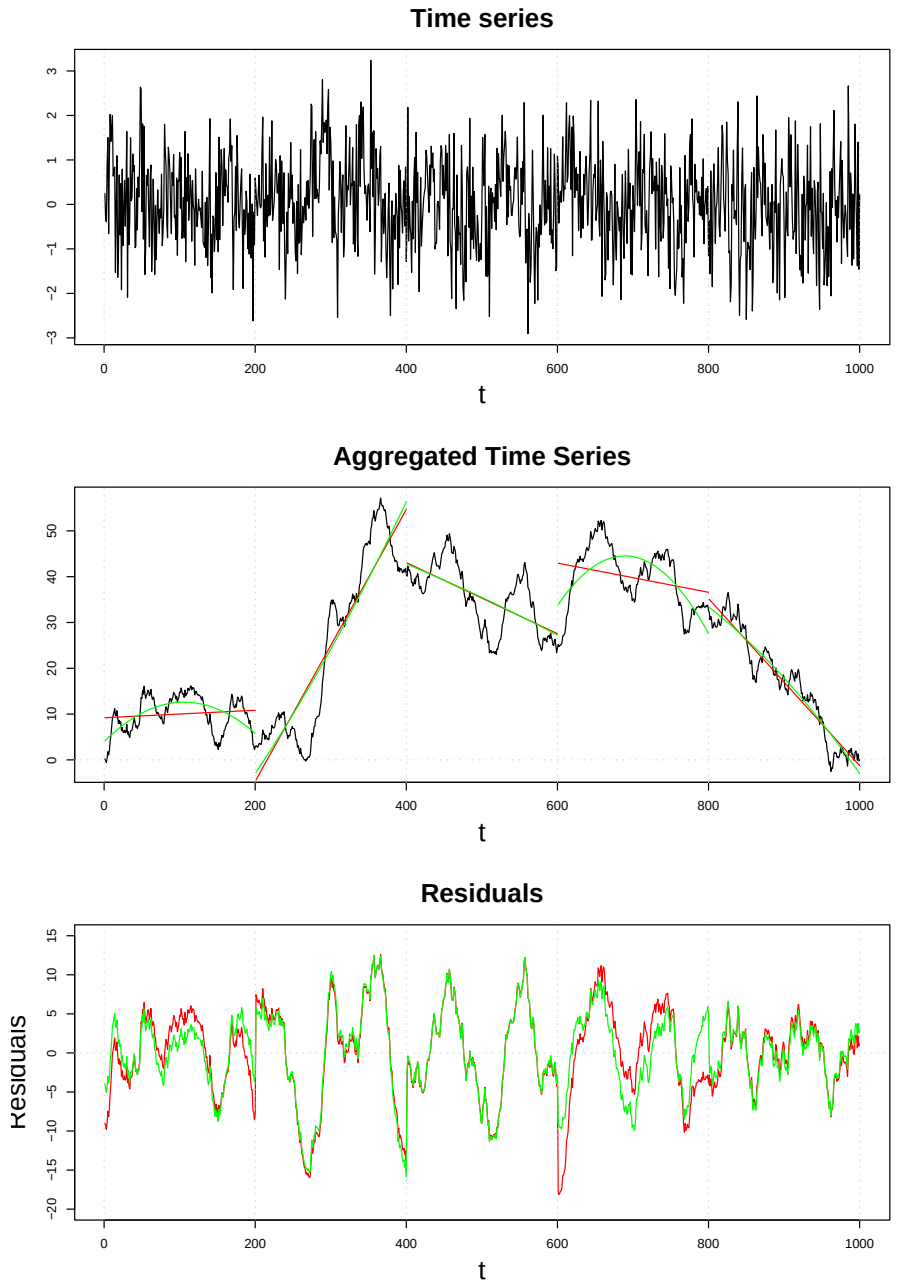
Properties

- slope as an estimator for β
- bandwidth: $\omega_{\min} \leq \omega_j \leq \omega_{\max}$, m number of ω_j included

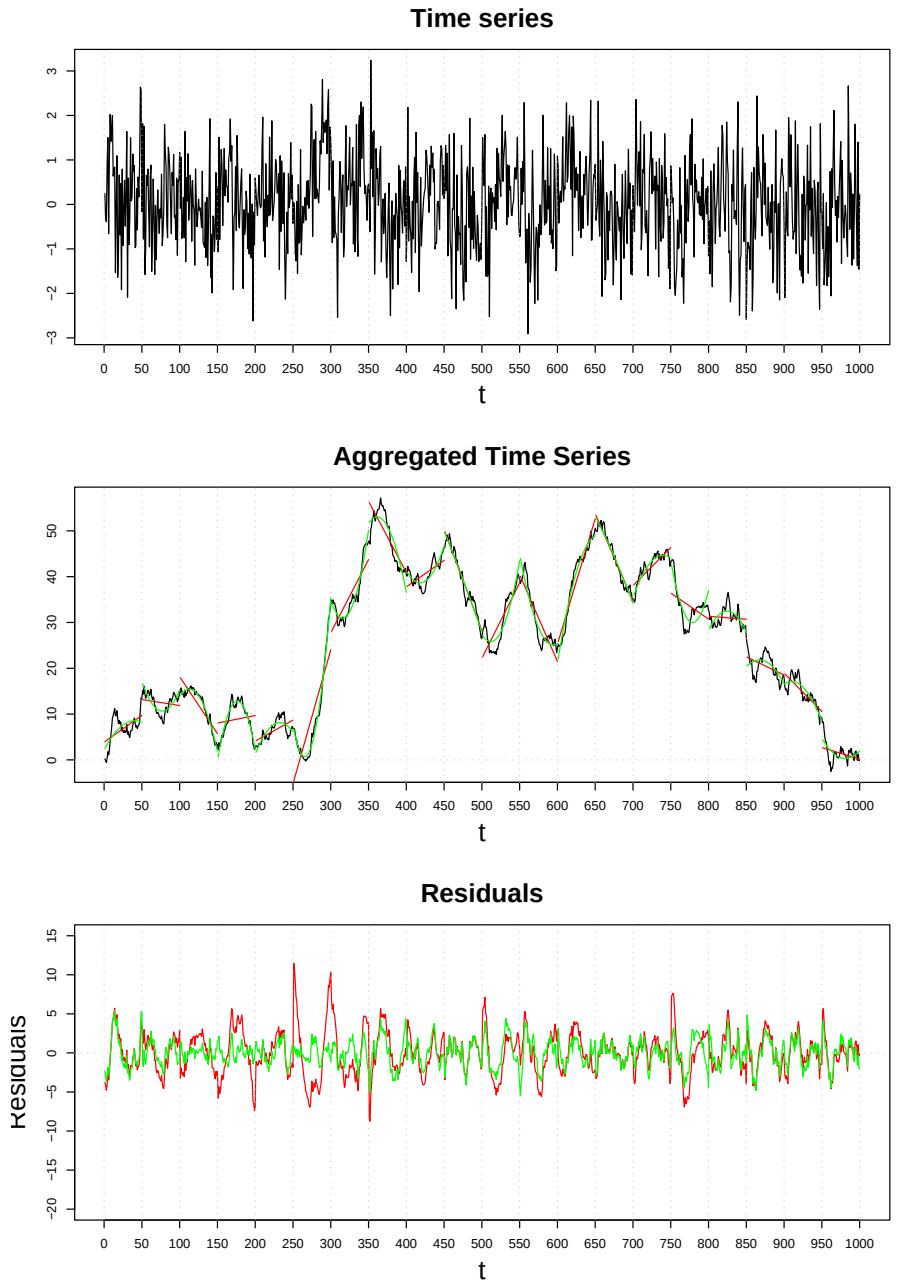
$$m^{1/2}(\hat{\beta} - \beta) \to \xi \quad ; \qquad \xi \sim N\left(0, \frac{\pi^2}{24}\right) \tag{4}$$

- confidence intervals and testing

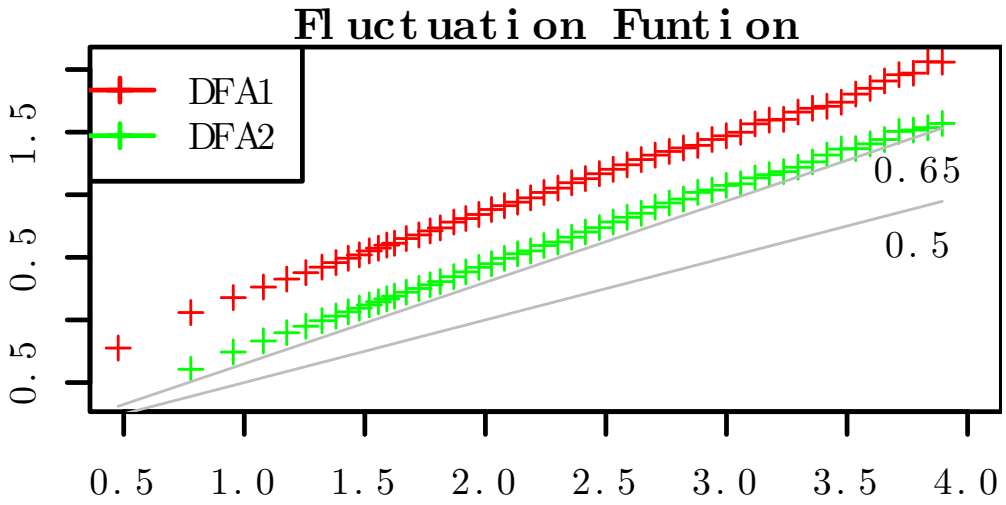
Detrended Fluctuation Analysis (DFA)



Detrended Fluctuation Analysis (DFA)



Detrended Fluctuation Analysis (DFA)



log-log plot of mean fluctuation at scale s over s

Interpretation

- white noise: slope $H = 0.5$ independent of s
- SRD: slope $H = 0.5$, asymptotically
- LRD: slope $0.5 < H < 1$, asymptotically

Detrended Fluctuation Analysis

LRD Parameter

$$\beta = 2H - 1 \tag{5}$$

Problems

- many plots look reasonably smooth in log-log
- slope constant? $\longrightarrow$ local slopes
- no limiting distribution, no CI, no testing
- but: easy to use, good for first impression

Detrended Fluctuation Analysis

LRD Parameter

$$\beta = 2H - 1 \tag{5}$$

Problems

- many plots look reasonably smooth in log-log
- slope constant? $\longrightarrow$ local slopes
- no limiting distribution, no CI, no testing
- but: easy to use, good for first impression

Full-Parametric Estimation

Recall from Yesterday

- specify a suitable family of models
- e.g. FARIMA$[p, d, q]$
- estimate parameters
- e.g. Whittle estimator
- select model
- e.g. AIC/HIC, LRT, simulation based non-nested model selection
- test for $d > 0$ in best model

Full-Parametric Estimation

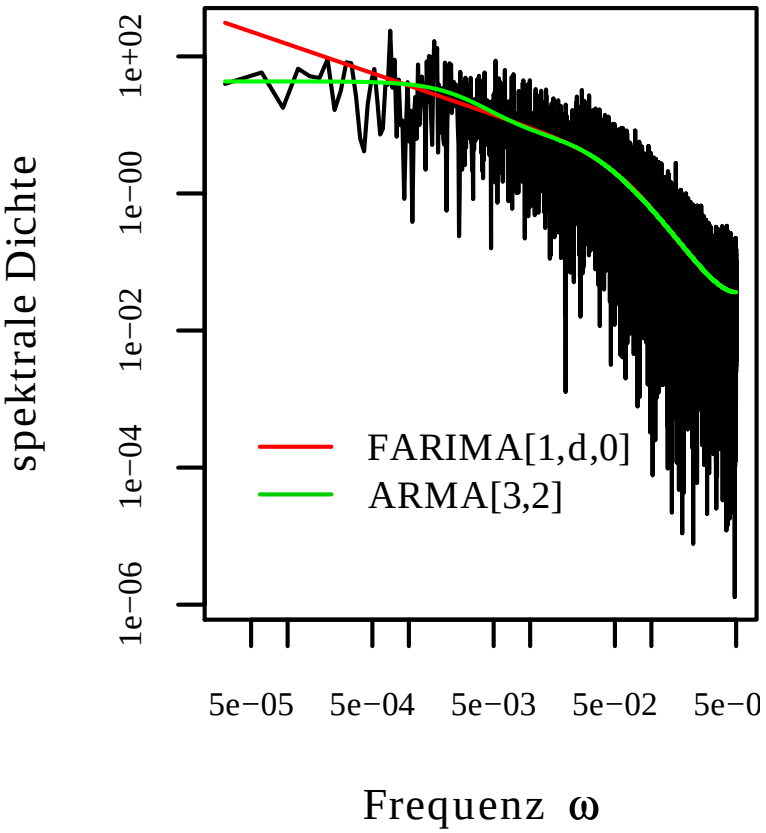
Recall from Yesterday

- specify a suitable family of models
- e.g. FARIMA$[p, d, q]$
- estimate parameters
- e.g. Whittle estimator
- select model
- e.g. AIC/HIC, LRT, simulation based non-nested model selection
- test for $d > 0$ in best model

Problematic Situations

Example: Realisation of an ARMA[3, 2]

1. FARIMA$[p, d, q]$
2. model selection (LRT)
   $\rightarrow$ FARIMA$[1, d, 0]$
3. $\hat{d} = 0.30(0.02)$
   $\rightarrow$ LRD



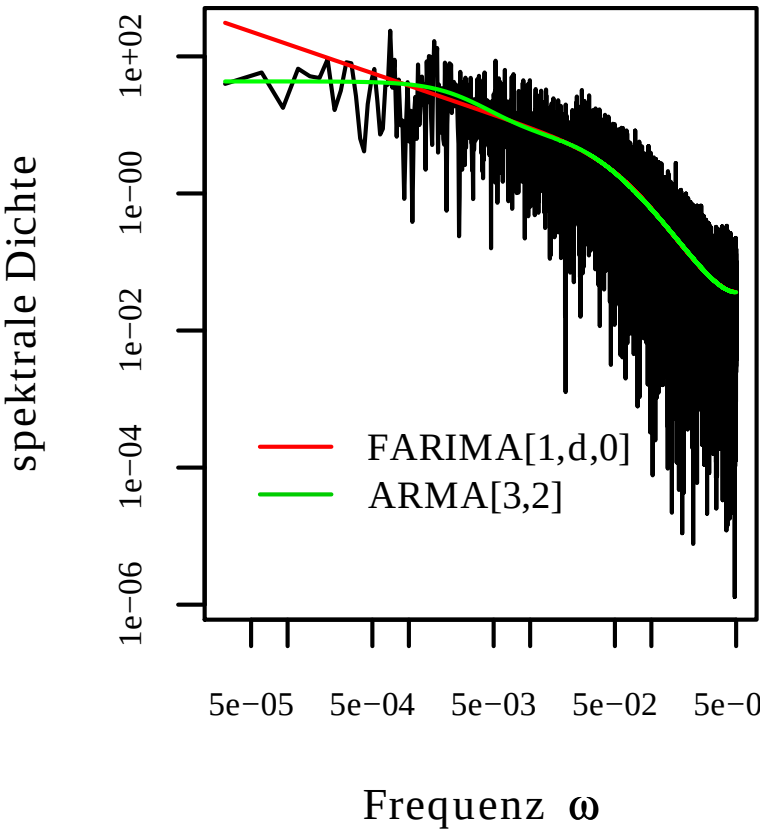
Problem

- why not FARIMA$[3, d, 2]$ with d=0?
- probably N too small
- way out?

Problematic Situations

Example: Realisation of an ARMA[3, 2]

1. FARIMA$[p, d, q]$
2. model selection (LRT)
   $\longrightarrow$ FARIMA$[1, d, 0]$
3. $\hat{d} = 0.30(0.02)$
   $\longrightarrow$ LRD



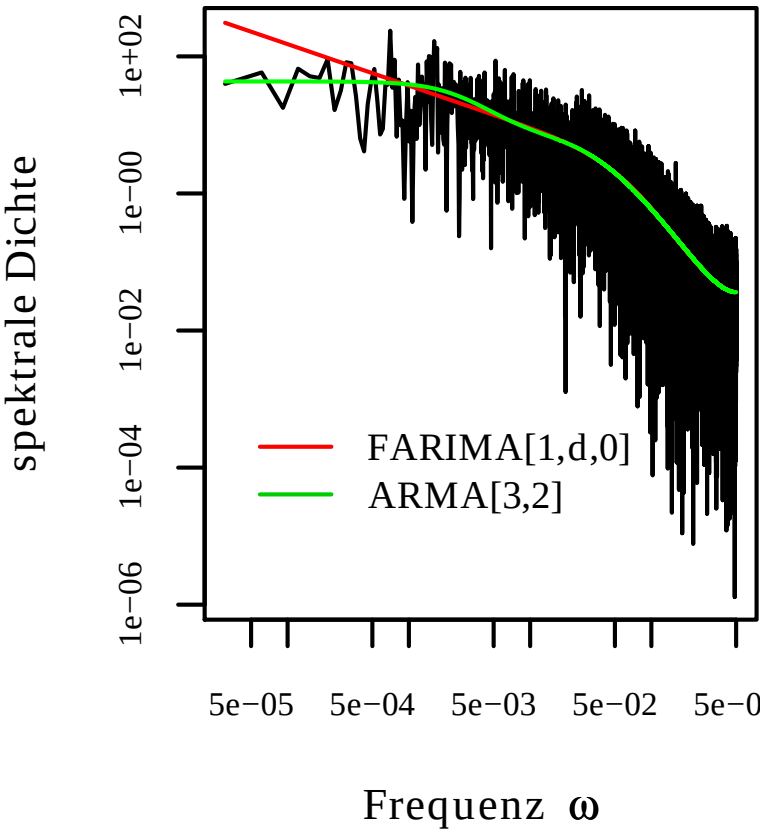
Problem

- why not FARIMA$[3, d, 2]$ with d=0?
- probably N too small
- way out?

Problematic Situations

Example: Realisation of an ARMA[3, 2]

1. FARIMA$[p, d, q]$
2. model selection (LRT)
   $\longrightarrow$ FARIMA$[1, d, 0]$
3. $\hat{d} = 0.30(0.02)$
   $\longrightarrow$ LRD



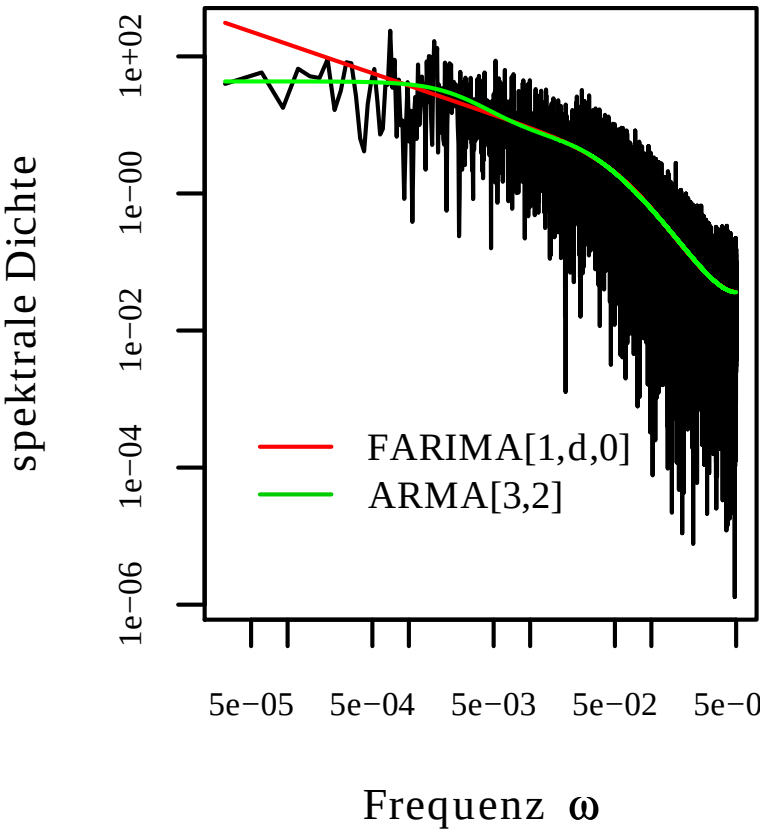
Problem

- why not FARIMA$[3, d, 2]$ with d=0?
- probably N too small
- way out?

Problematic Situations

Example: Realisation of an ARMA[3, 2]

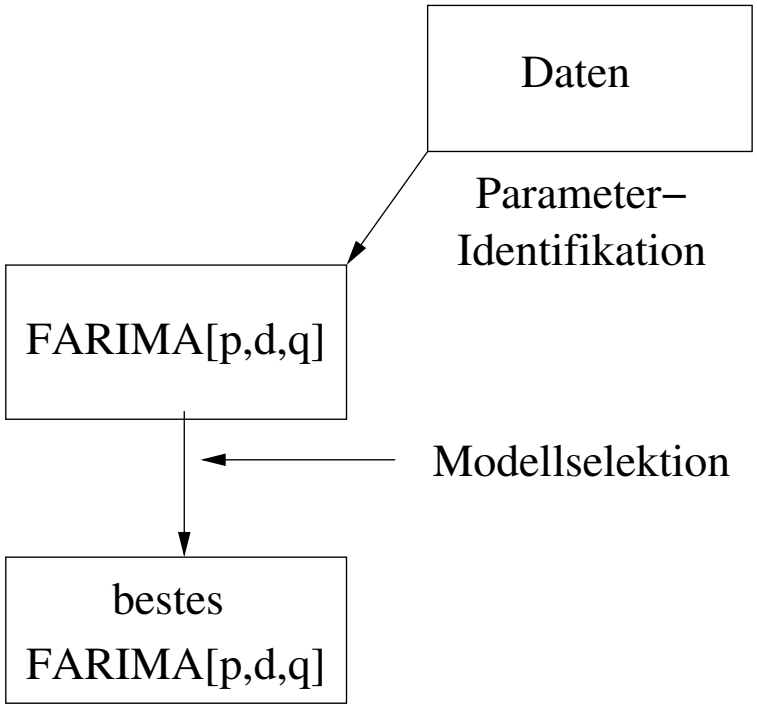
1. FARIMA$[p, d, q]$
2. model selection (LRT)
   $\longrightarrow$ FARIMA$[1, d, 0]$
3. $\hat{d} = 0.30(0.02)$
   $\longrightarrow$ LRD



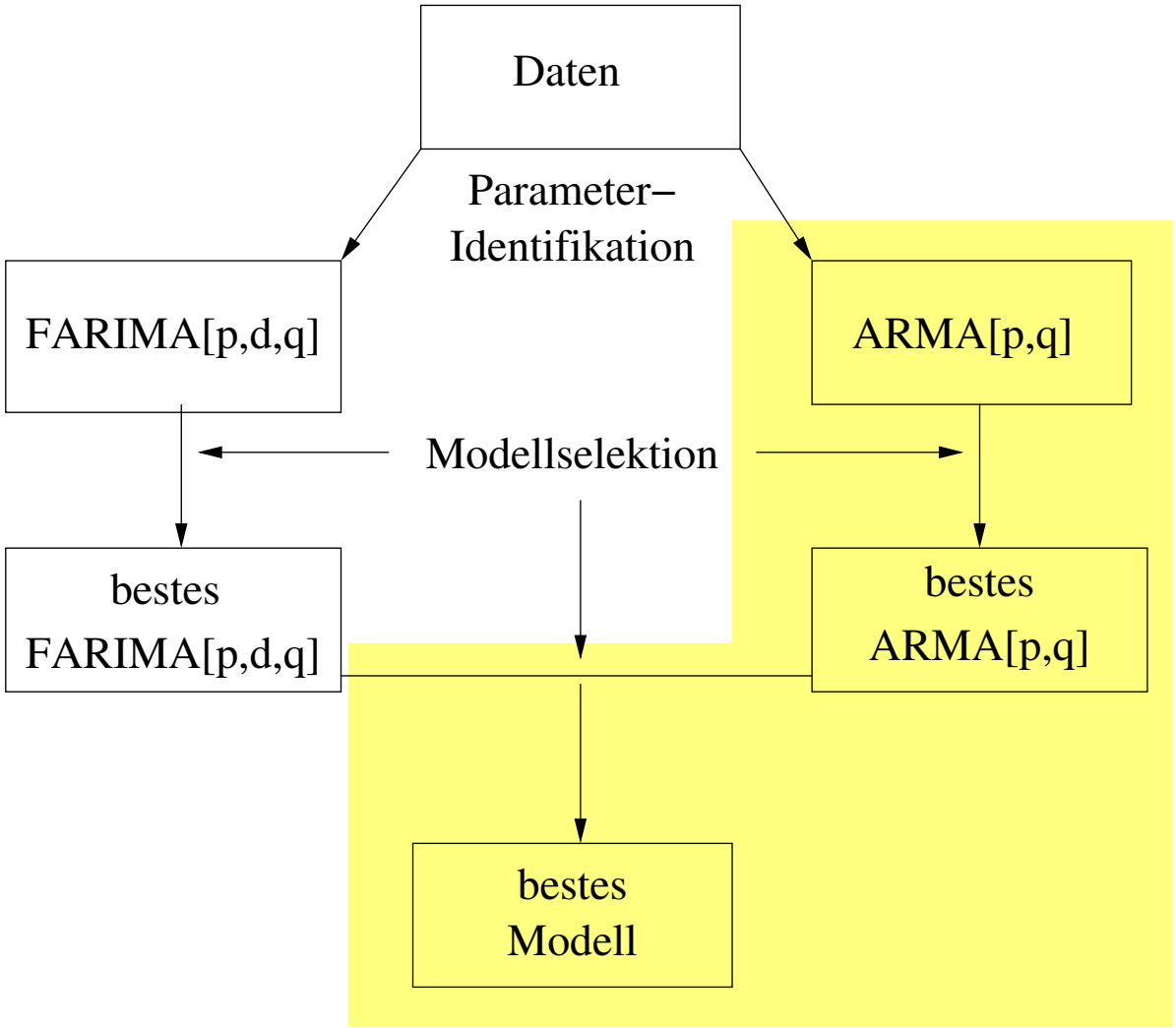
Problem

- why not FARIMA$[3, d, 2]$ with d=0?
- probably N too small
- way out? $\longrightarrow$ comparativ approach

Comparative Approach



Comparative Approach



Comparative Approach

Strategy

- walk through parametric approach, with
  1. ARMA$[p, q]$
  2. FARIMA$[p, d, q]$

- compare best ARMA vs. best FARIMA

Properties

- increases specificity of test for LRD
- not necessary for $N \to \infty$
- leads frequently to non-nested model selection problem, e.g. ARMA$[3, 2]$ vs. FARIMA$[1, d, 0]$

Comparative Approach

Strategy

- walk through parametric approach, with
  1. ARMA$[p, q]$
  2. FARIMA$[p, d, q]$
- compare best ARMA vs. best FARIMA

Properties

- increases specificity of test for LRD
- not necessary for $N \to \infty$
- leads frequently to non-nested model selection problem, e.g. ARMA$[3, 2]$ vs. FARIMA$[1, d, 0]$

Danube River

Achleiten/Danube River

- approx. 103 years daily discharge
- previously with DFA: LRD $d = 0.32$
- FARIMA$[p, d, q]$/ARMA$[p, q]$
- Result: ARMA$[7, 6]$ (no LRD!)

Map

DFA

Danube River

Achleiten/Danube River

- approx. 103 years daily discharge
- previously with DFA: LRD $d = 0.32$
- FARIMA$[p, d, q]$/ARMA$[p, q]$
- Result: ARMA$[7, 6]$ (no LRD!)

Map

DFA

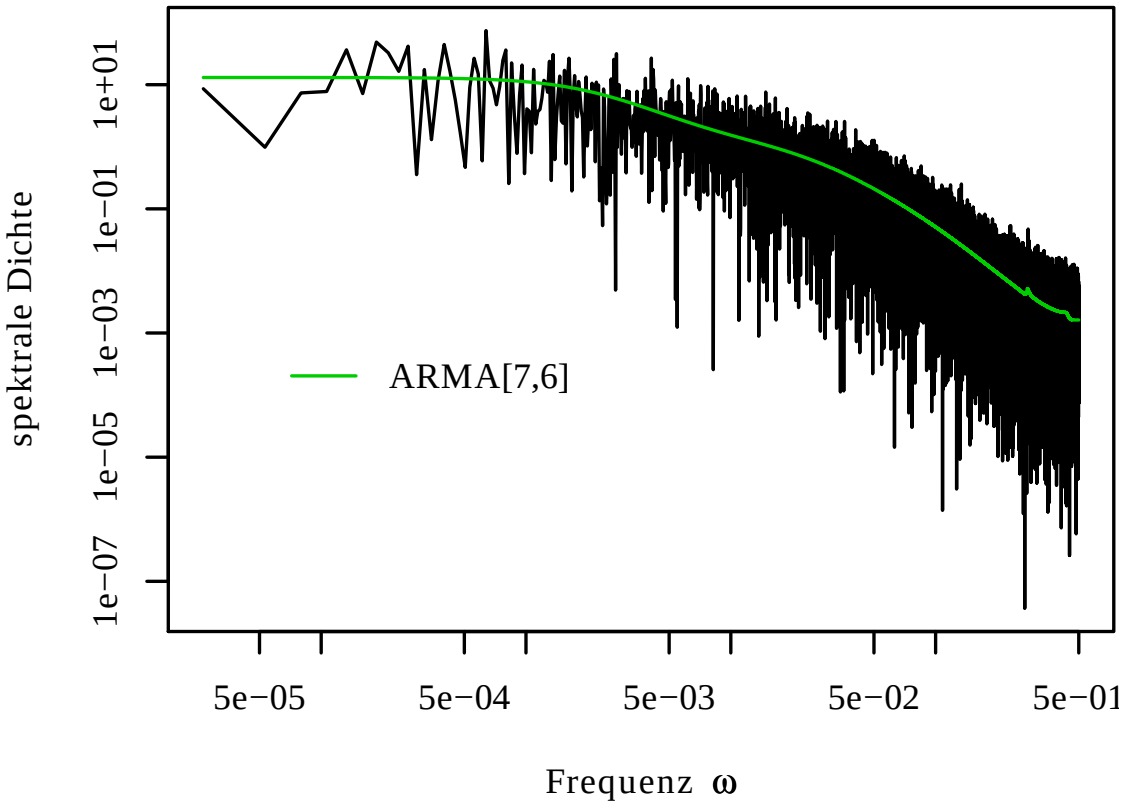
Danube River

Achleiten/Danube River

- approx. 103 years daily discharge
- previously with DFA: LRD $d = 0.32$
- FARIMA$[p, d, q]$/ARMA$[p, q]$
- Result: ARMA$[7, 6]$ (no LRD!)

Map

DFA



Große Vils

Vilsbiburg/Große Vils

- approx. 62 years daily discharge
- FARIMA$[p, d, q]$/ARMA$[p, q]$
- LRT: FARIMA$[3, d, 0]$ and ARMA$[7, 6]$
- non-nested model selection: ARMA$[7, 6]$
- but: 5 parameters are compatible with 0!
- consider also FARIMA$[3, d, 0]$

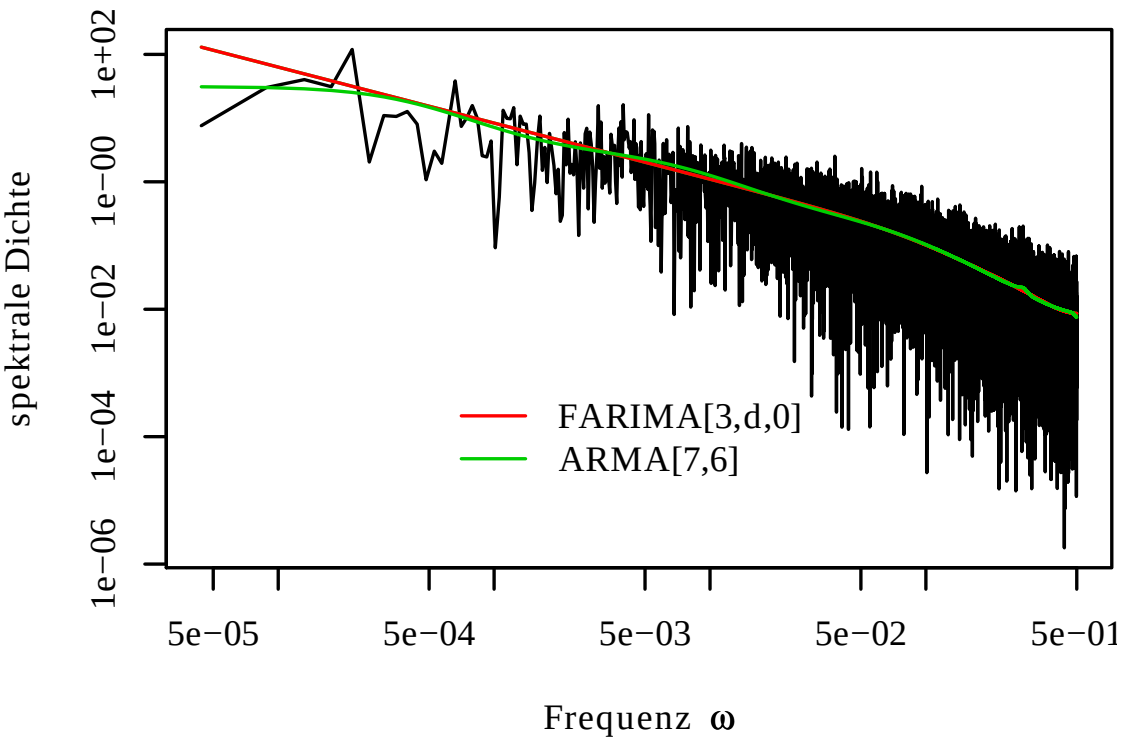
Map

Große Vils

Vilsbiburg/Große Vils

- approx. 62 years daily discharge
- FARIMA$[p, d, q]$/ARMA$[p, q]$
- LRT: FARIMA$[3, d, 0]$ and ARMA$[7, 6]$

Map

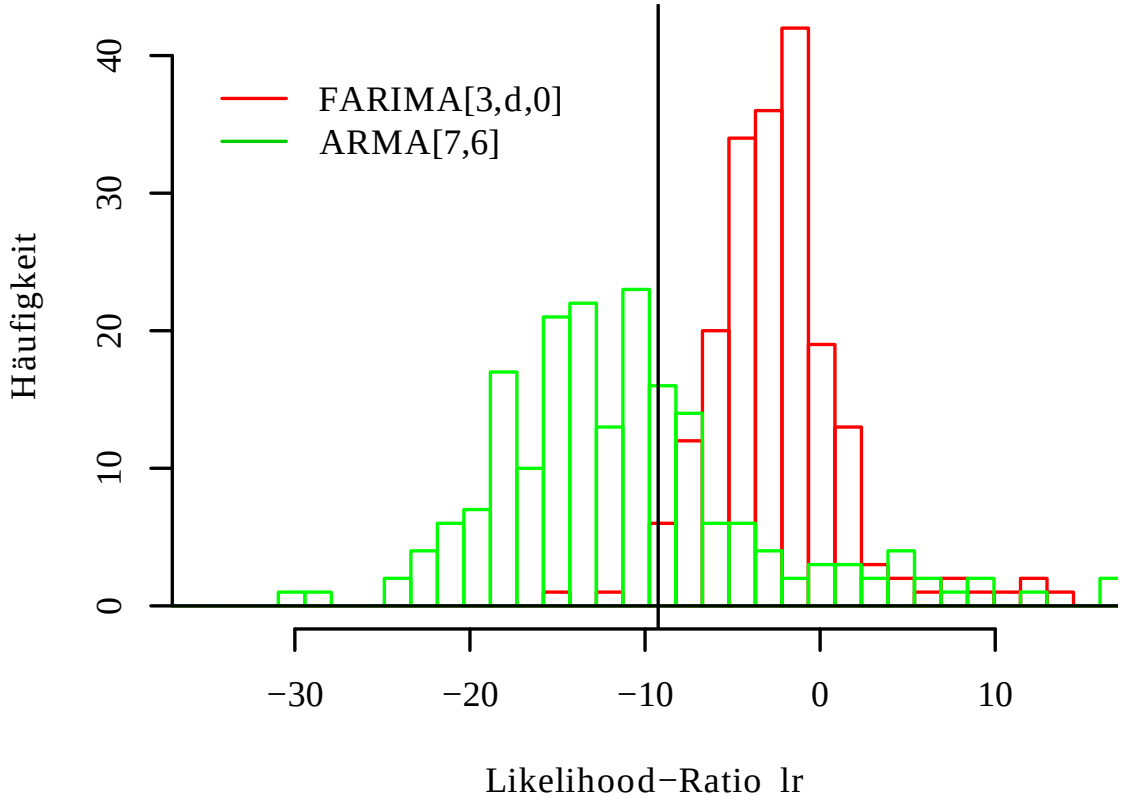


Große Vils

Vilsbiburg/Große Vils

Map

- approx. 62 years daily discharge
- FARIMA$[p, d, q]$/ARMA$[p, q]$
- LRT: FARIMA$[3, d, 0]$ and ARMA$[7, 6]$
- non-nested model selection: ARMA$[7, 6]$



Große Vils

Vilsbiburg/Große Vils

Map

- approx. 62 years daily discharge
- FARIMA$[p, d, q]$/ARMA$[p, q]$
- LRT: FARIMA$[3, d, 0]$ and ARMA$[7, 6]$
- non-nested model selection: ARMA$[7, 6]$
- but: 5 parameters are compatible with 0!
- consider also FARIMA$[3, d, 0]$

Große Vils

Vilsbiburg/Große Vils

Map

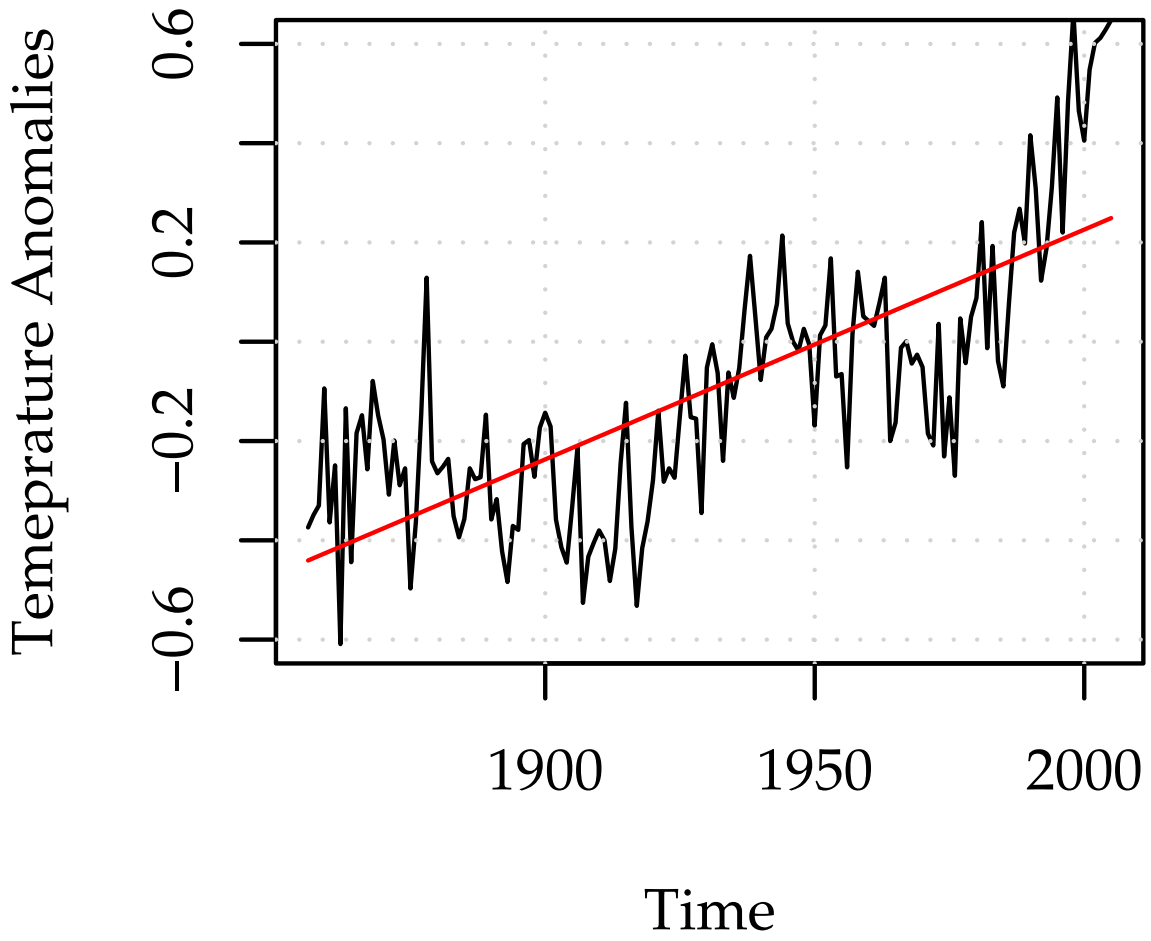
- approx. 62 years daily discharge
- FARIMA$[p, d, q]$/ARMA$[p, q]$
- LRT: FARIMA$[3, d, 0]$ and ARMA$[7, 6]$
- non-nested model selection: ARMA$[7, 6]$
- but: 5 parameters are compatible with 0!
- consider also FARIMA$[3, d, 0]$

The Physics behind

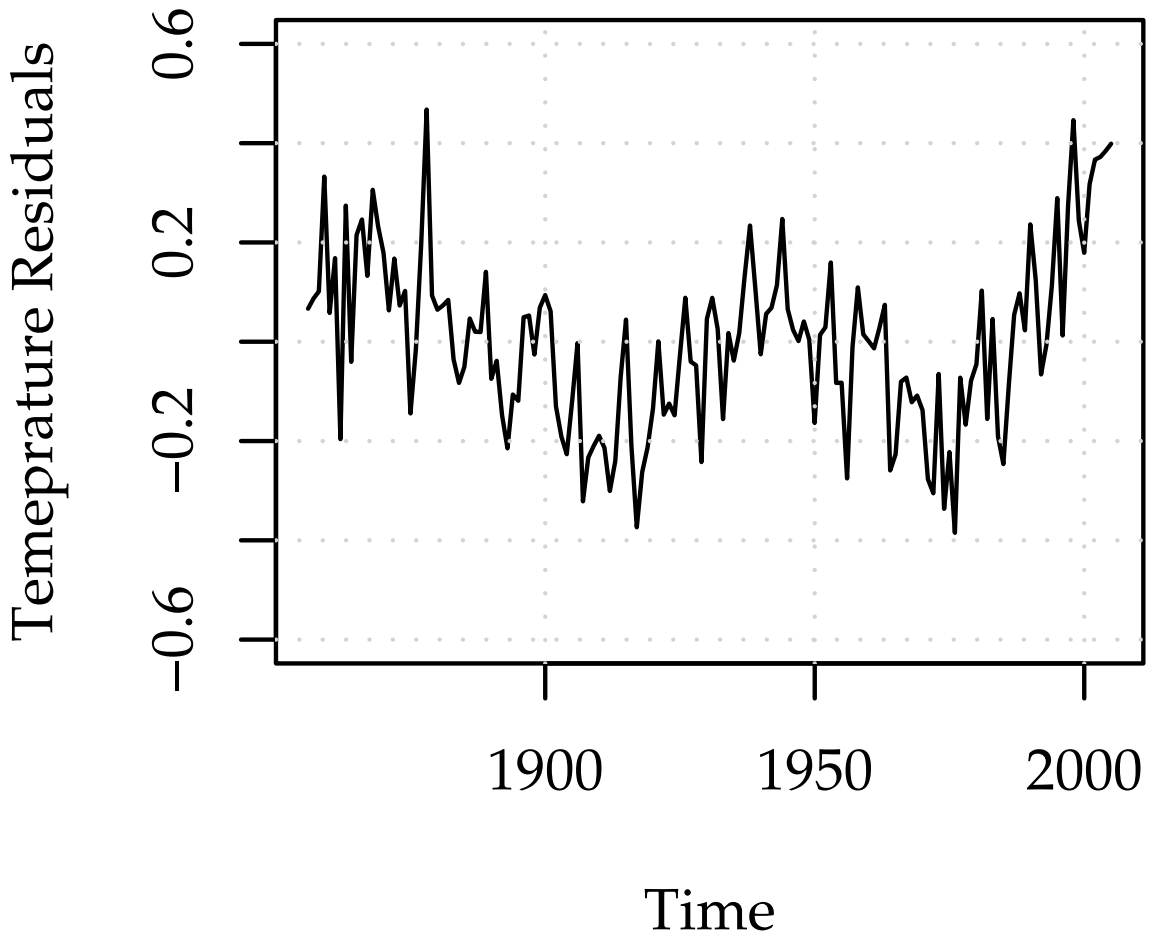
LRD in river discharge

- superposition of processes with different time scales
  - various contributaries
  - supported by complex ARMA models
- changes on large time scales lead to large scale variations, which appear similar to LRD
  - climate change
  - river bed modification
  - land use chage
  - ...

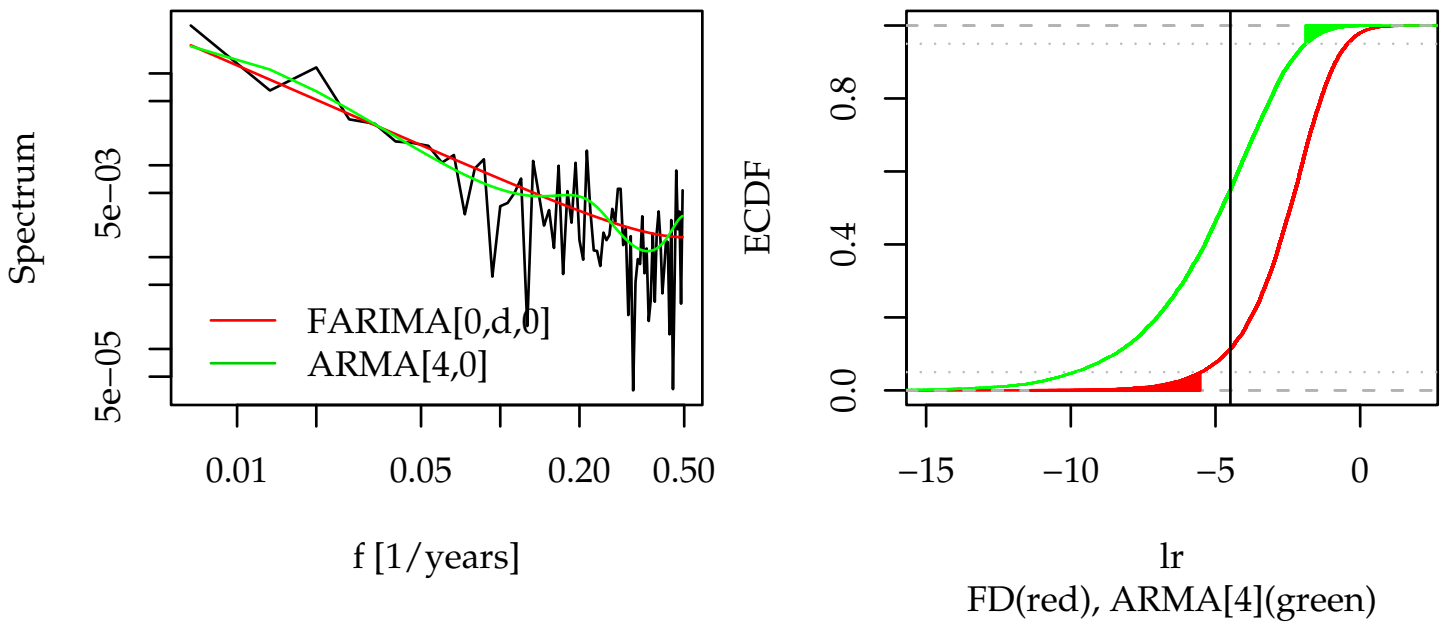
Northern Hemisphere Mean Temperature



Northern Hemisphere Mean Temperature



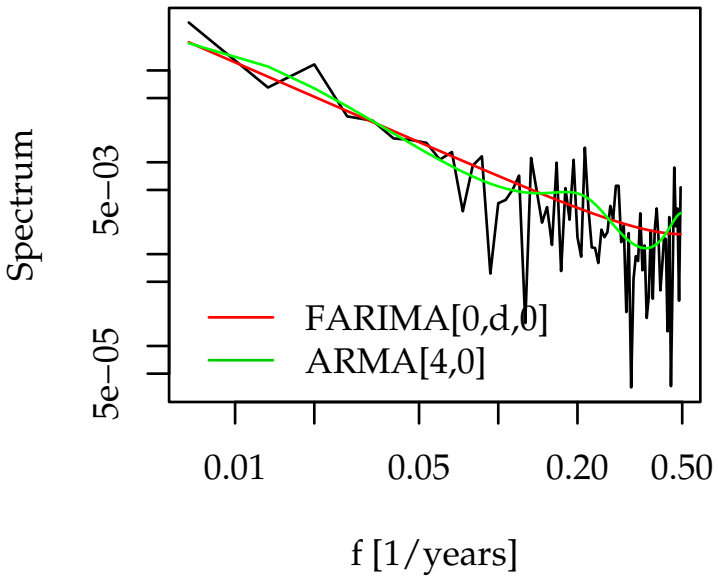
Northern Hemisphere Mean Temperature

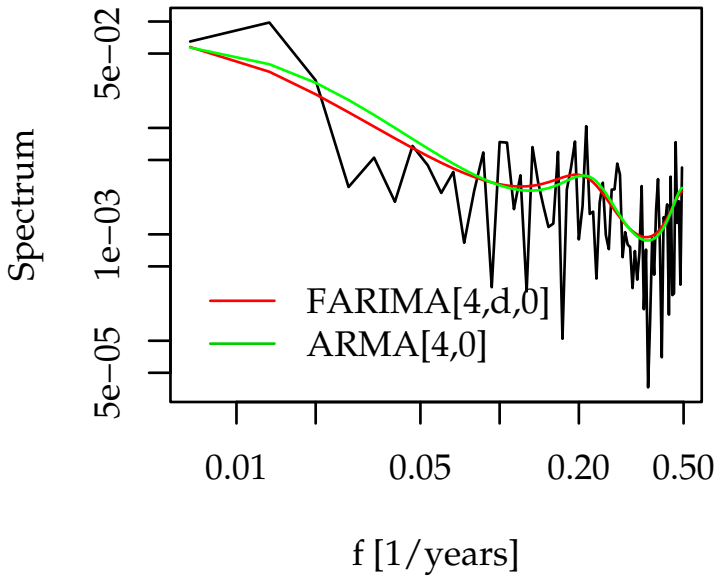


Result

- anomalies(left): FARIMA$[0, d, 0]$, $\hat{d} = 0.62(0.06)$
- residuals(right): ARMA$[4, 0]$ best model
- problem: linear trend assumption

Northern Hemisphere Mean Temperature





Result

- anomalies(left): FARIMA$[0, d, 0]$, $\hat{d} = 0.62(0.06)$
- residuals(right): ARMA$[4, 0]$ best model
- problem: linear trend assumption

Northern Hemisphere Mean Temperature

Many Radiative Forcing Factors

- greenhouse gases
- aerosols (volcanos, anthropogenic)
- variable solar forcing
- etc

Include Physics in the Model

- detection and attribution with finger printing (Santer, 1994; Hasselmann, 1997)

Summary

- LRD alters statistical inference (mean or quantile estimators, test for trends and change points, etc)
- various detection strategies
  - parametric (FARIMA, Whittle)
  - semi-parametric (log-periodogram)
  - heuristic (DFA)
- comparative parametric more robust against false detection of LRD
- think of the physics!

Links

for lecture notes, talks and software

E2C2 Web Page

- `http://www.e2c2.ipsl.fr`

My Web Page

- `http://www.agnld.uni-potsdam.de:~hrust`

R Web Page

- `http://www.r-project.org`

You never walk alone

- R - packages:
  - fracdiff
    - MLE estimation of FARIMA
  - longmemo
    - Whittle estimation of FARIMA and fGn
  - farima
    - Whittle estimation of FARIMA
    - LRT, AIC
    - simulation based model selection
    - shared folders: `farima_1.0-8.tar.gz` (Linux/Unix/Mac?), `farima_1.0-8.tar.gz` (Windows)

Appendix

Mechanisms / "Explanations" for LRD

- self-similarity
- infinite superposition of processes with various time scales (e.g. AR[1])
- non-linear stochastic DE

"Explanation" for the Hurst-Phenomenon

- pre-asymptotic SRD
- finite superposition of AR[1]
- specific trend shapes
- large periodicities
- heteroskedasticity

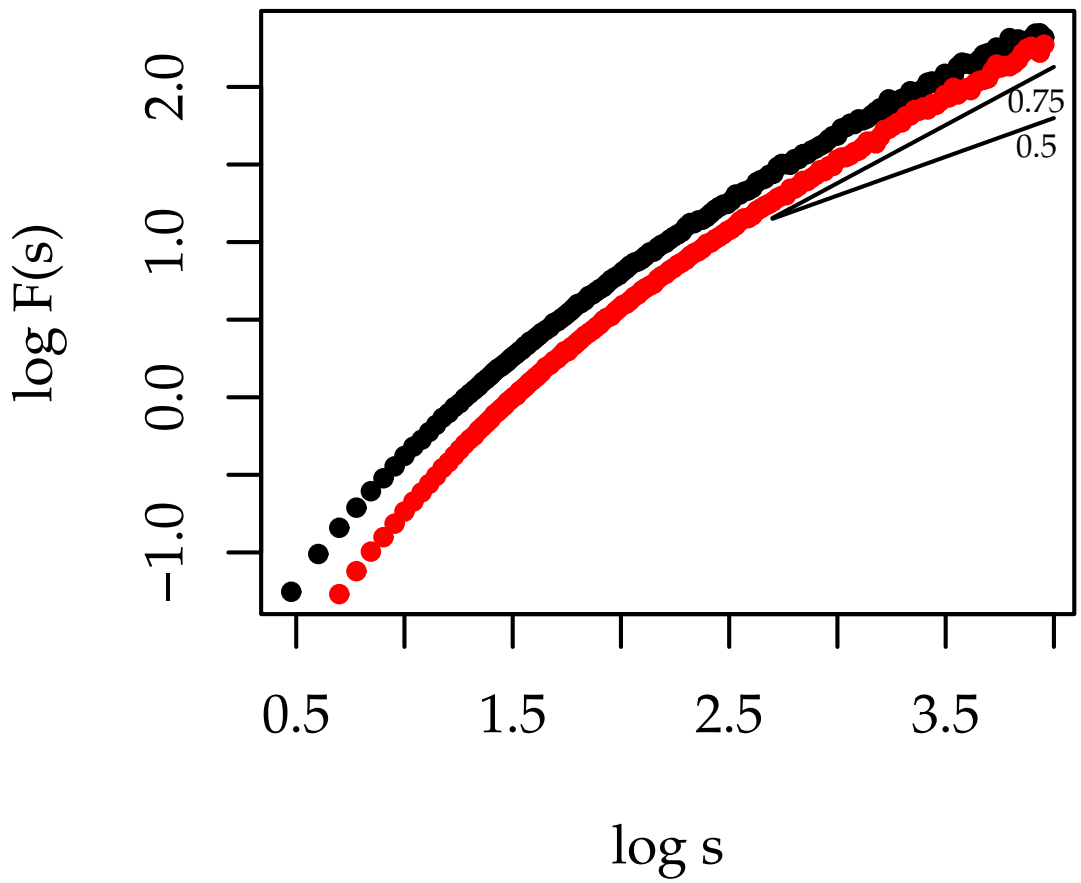
Mechanisms / "Explanations" for LRD

- self-similarity
- infinite superposition of processes with various time scales (e.g. AR[1])
- non-linear stochastic DE

"Explanation" for the Hurst-Phenomenon

- pre-asymptotic SRD
- finite superposition of AR[1]
- specific trend shapes
- large periodicities
- heteroskedasticity

Gauge Achleiten, DFA



- $F(s)$ DFA fluctuation function, s scale
- slope shows $H = d + 0.5$ an

Gauges

Vilsbiburg Achleiten