Modelling internet round-trip time data

Keith Briggs

Keith.Briggs@bt.com

http://research.btexact.com/teralab/keithbriggs.html

University of York 2003 July 18

typeset 2003 July 15 13:55 in LaTeX2e on a linux system

Outline

- motivation

Outline

- motivation
- data

Outline

- motivation
- data
- theory

Outline

- motivation
- data
- theory
- model fitting

Motivation

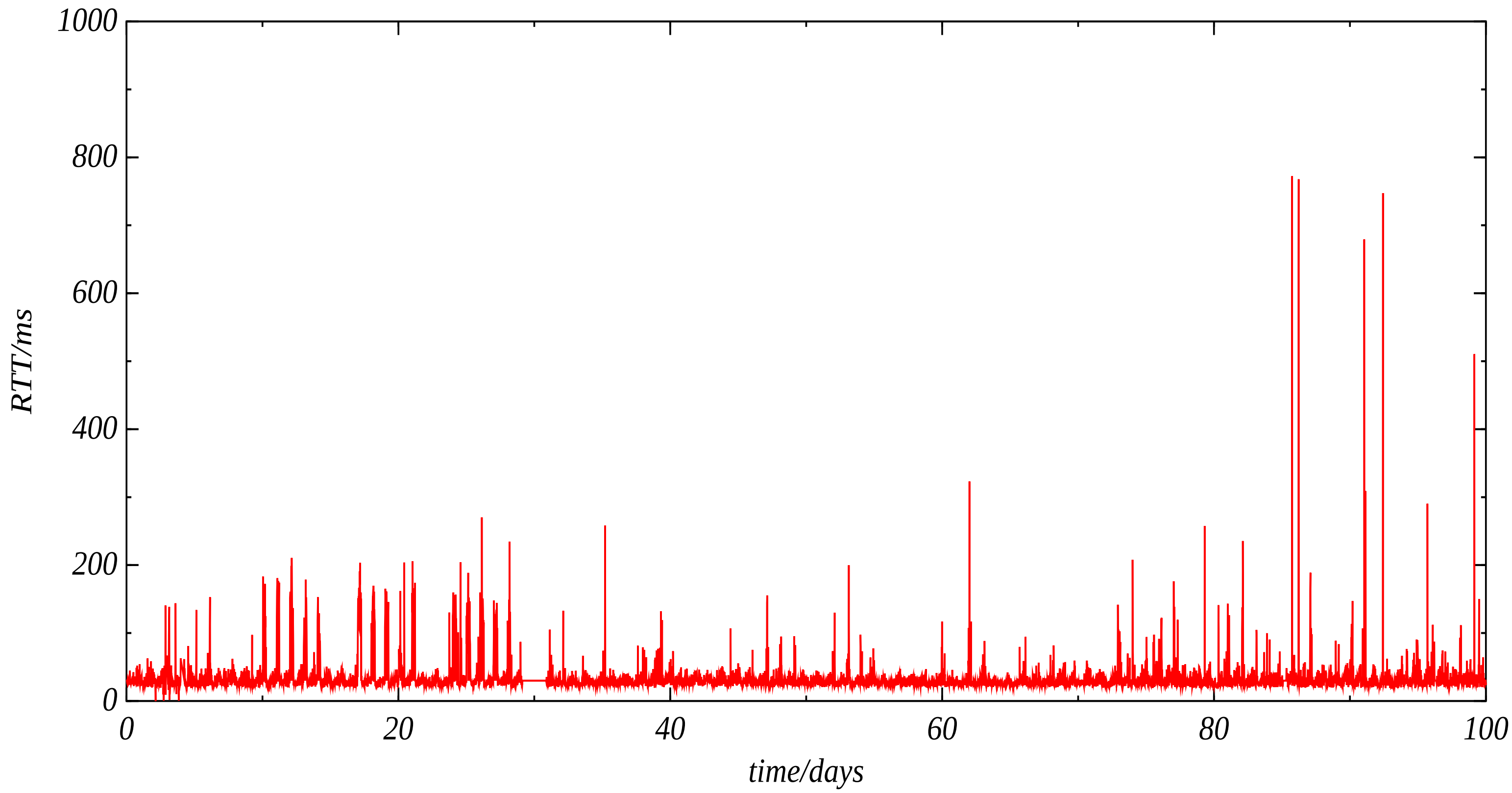
- internet as a complex system
- round-trip time (RTT) data forms an intriguing time series
- successful models would allow:
  - ▷ *forecasting*
  - ▷ *simulation*
  - ▷ *understanding*

Motivation

- internet as a complex system
- round-trip time (RTT) data forms an intriguing time series
- successful models would allow:
  - ▷ *forecasting*
  - ▷ *simulation*
  - ▷ *understanding*

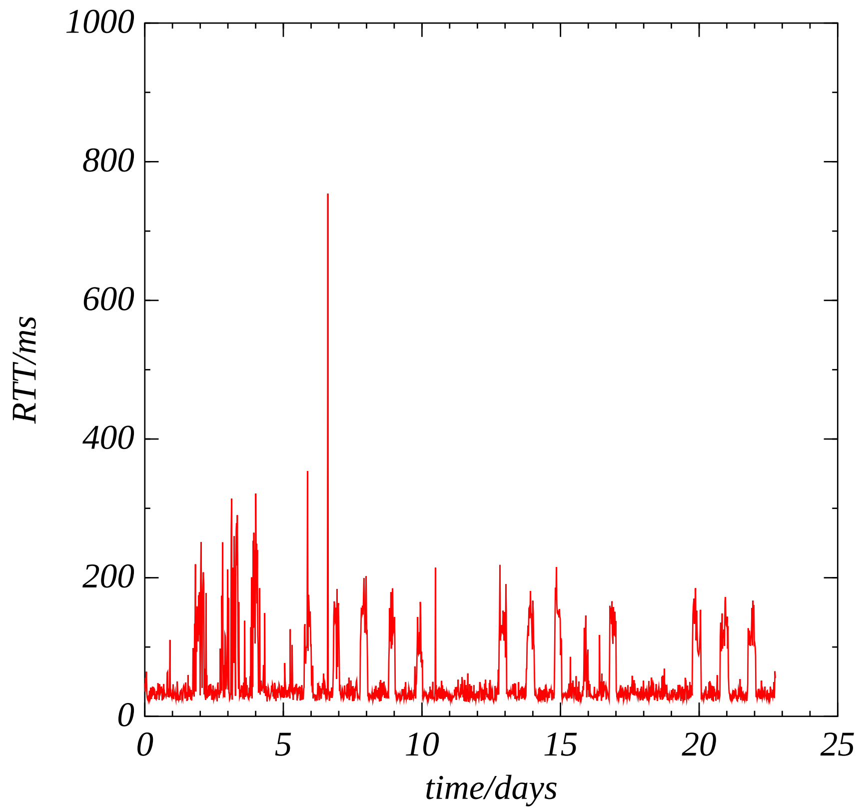
- any model used should incorporate features believed to exist in the data in a *natural* way

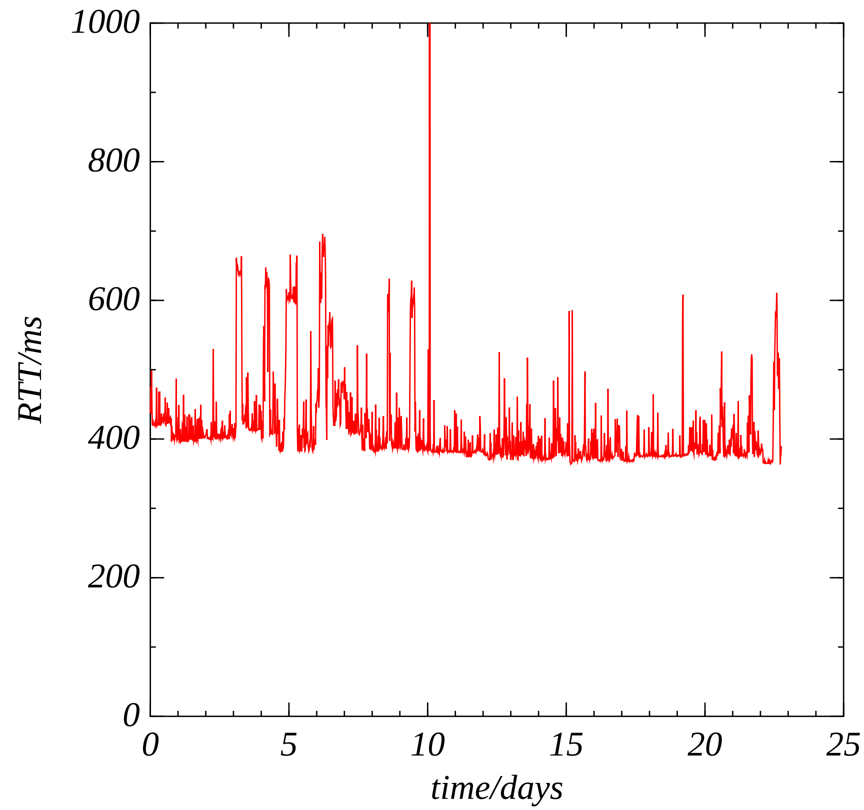
Raw data 1



100 days of typical raw data, from `www.edinburgh.ac.uk`

Raw data 2





`www.edinburgh.ac.uk` and `www.chem.uwa.edu.au`

Long-range dependence?

- definition:

$$\lim_{k \to \infty} \rho(k)\, k^{2(1-H)} = \text{constant}$$

- $0 < \alpha = 2(1-H) < 1$

Long-range dependence?

- definition:

$$\lim_{k \to \infty} \rho(k)\, k^{2(1-H)} = \text{constant}$$

- $0 < \alpha = 2(1-H) < 1$
- useless!
  - ▷ *for stationary process only*
  - ▷ *large-k limit*
  - ▷ *H cannot be estimated in practice even when it exists and is known*

Long-range dependence?

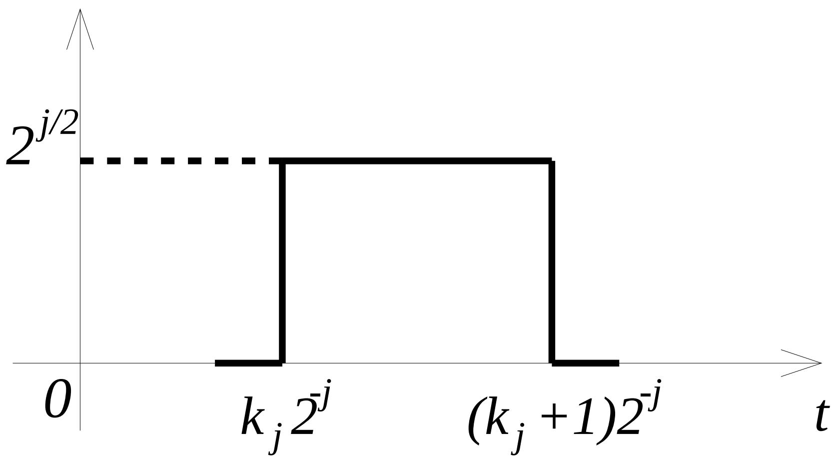
- definition:

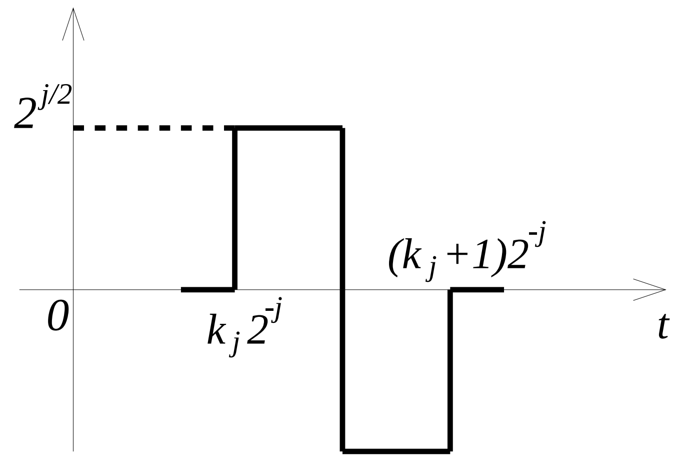
$$\lim_{k\to\infty} \rho(k)\, k^{2(1-H)} = \text{constant}$$

- $0 < \alpha = 2(1-H) < 1$
- useless!
  - *for stationary process only*
  - *large-k limit*
  - *H cannot be estimated in practice even when it exists and is known*
  - *tries to reduce complex phenomena to a single number*

Wavelet transform 1

- I use the Haar basis - left: scaling function ϕ; right: wavelet function ψ





-

$$\psi_{j,k}(t) = 2^{j/2}\,\psi(2^j t - k)$$
$$\phi_{j,k}(t) = 2^{j/2}\,\phi(2^j t - k)$$

Wavelet transform 2

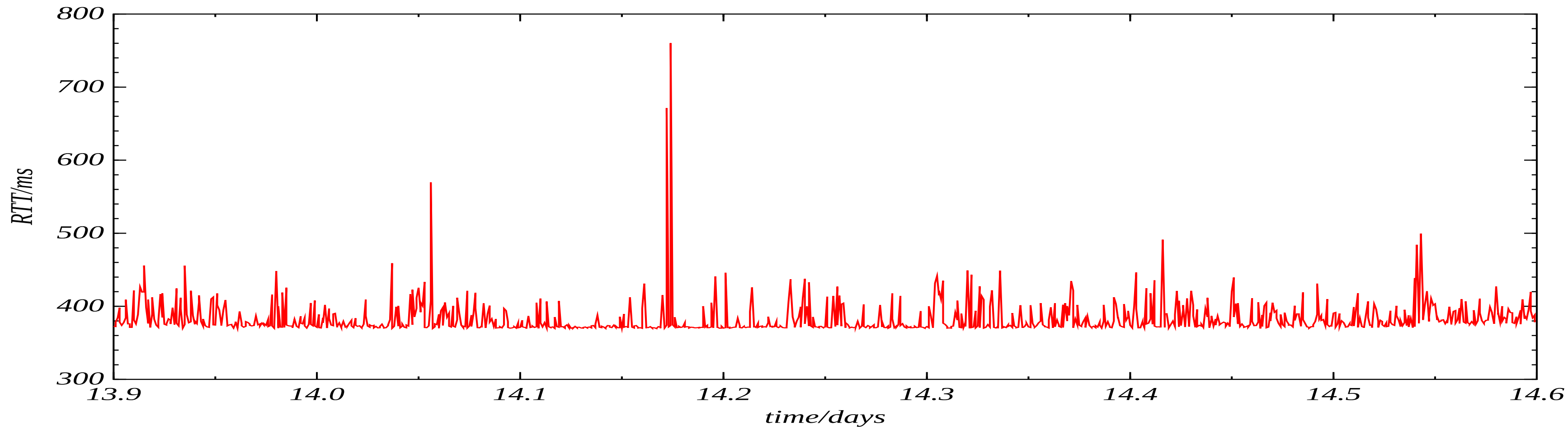
$$f_t = \sum_k U_{0,k}\ \phi_{0,k}(t) + \sum_{j=0}^{J} \sum_k W_{j,k}\ \psi_{j,k}(t)$$

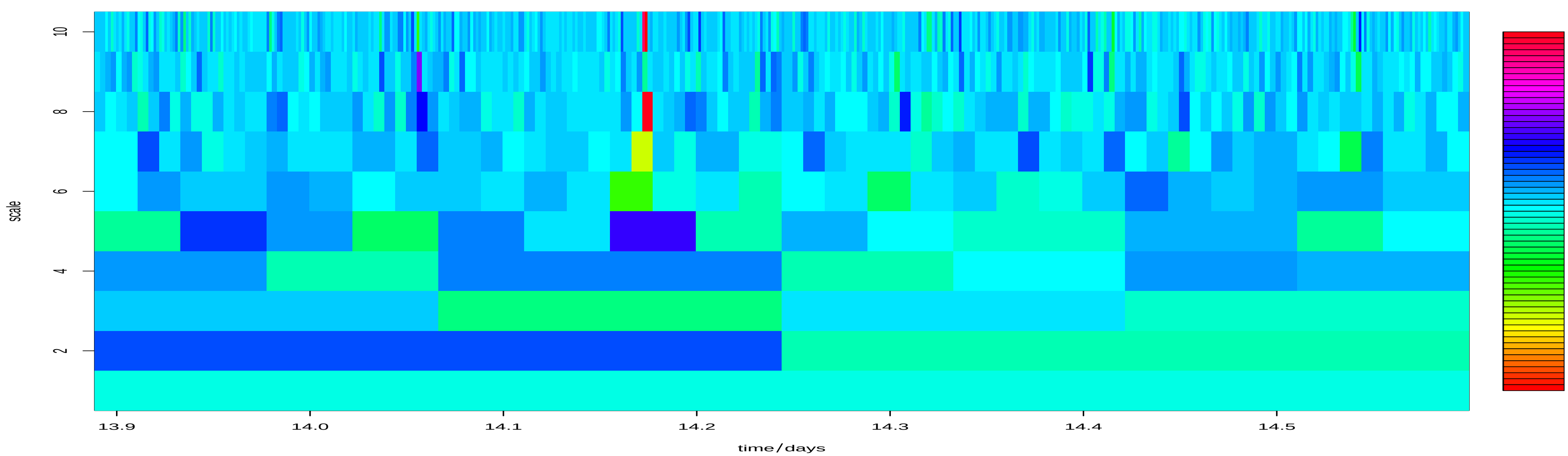
wavelet coefficients, $W_{j,k}$, and *scaling coefficients*, $U_{j,k}$, are defined by

$$W_{j,k} = \sum_{t=1}^{T} f_t\ \psi_{j,k}(t)$$

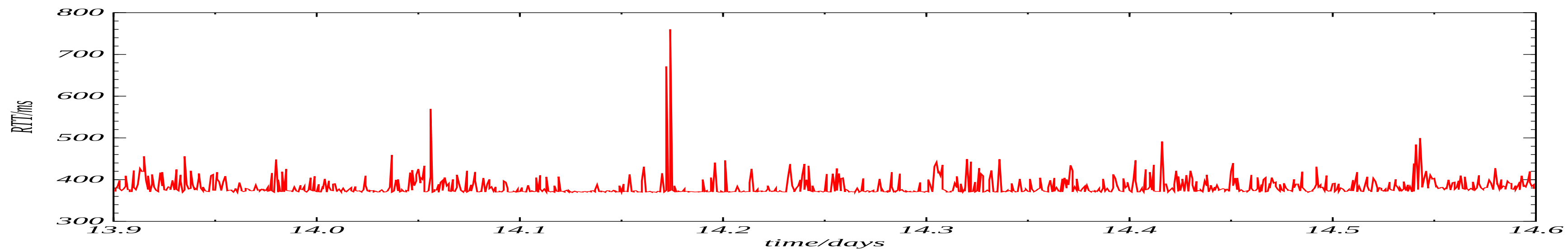
$$U_{j,k} = \sum_{t=1}^{T} f_t\ \phi_{j,k}(t)$$

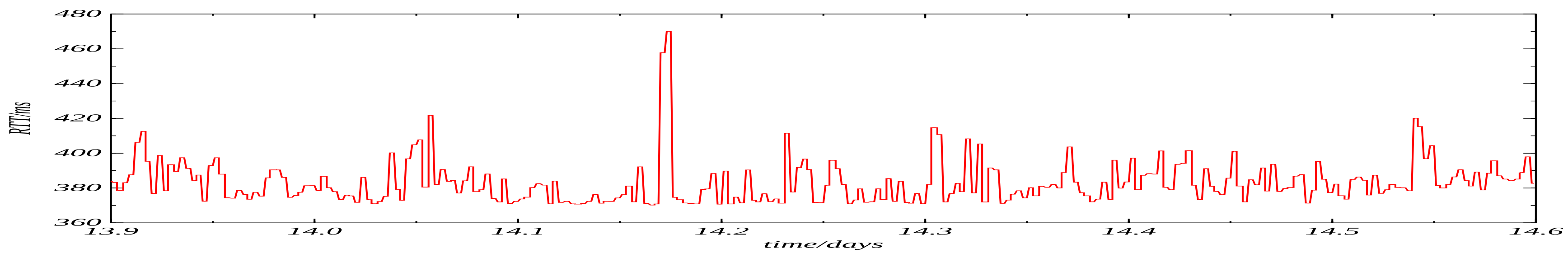
Example wavelet transform 1

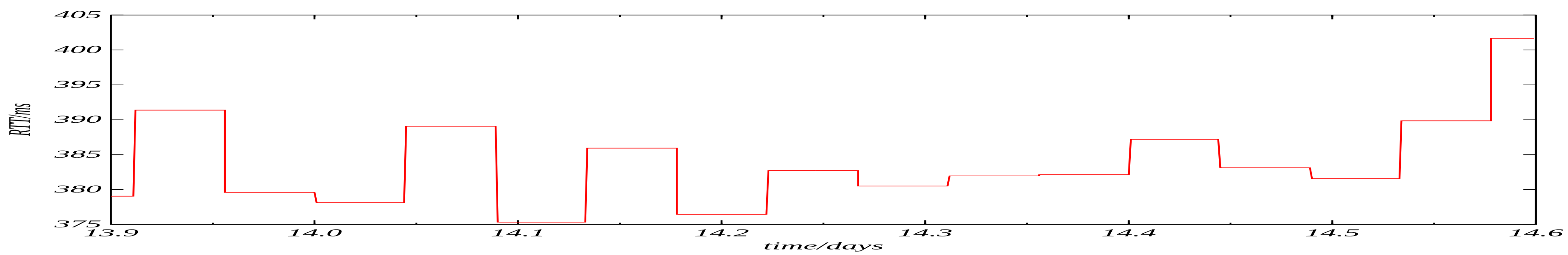




Example wavelet transform 2







Information at scales J, $J-1$ and $J-4$

Multifractal spectrum

- g is Lipschitz α at x_0 if α is the supremum of those a such that in a neighbourhood of x_0

$$\left| \, g(x) - p_{\lfloor \beta \rfloor}(x) \, \right| = \mathcal{O}(\left| \, x - x_0 \, \right|^a)$$

where $p_{\lfloor \alpha \rfloor}$ is a polynomial of degree $\lfloor \alpha \rfloor$

Multifractal spectrum

- g is Lipschitz α at x_0 if α is the supremum of those a such that in a neighbourhood of x_0

$$\left| \, g(x) - p_{\lfloor \beta \rfloor}(x) \, \right| = \mathcal{O}(\mid x - x_0 \mid^{a})$$

  where $p_{\lfloor \alpha \rfloor}$ is a polynomial of degree $\lfloor \alpha \rfloor$

- Let $f(\alpha)$ be the Hausdorff dimension of the set of points where the Lipschitz exponent is α. This is the *multifractal spectrum* of g

Multifractal spectrum

- g is Lipschitz α at x_0 if α is the supremum of those a such that in a neighbourhood of x_0

$$\mid g(x) - p_{\lfloor\beta\rfloor}(x) \mid = \mathcal{O}(\mid x - x_0 \mid^a)$$

  where $p_{\lfloor\alpha\rfloor}$ is a polynomial of degree $\lfloor\alpha\rfloor$

- Let $f(\alpha)$ be the Hausdorff dimension of the set of points where the Lipschitz exponent is α. This is the *multifractal spectrum* of g

- example:
  - ▷ *fractal Brownian motion has zero mean and Gaussian increments s. t. the mean square increment at lag Δ is proportional to $|\Delta|^{2H}$*
  - ▷ *this is a monofractal - $f(\alpha) = \delta(H)$*
  - ▷ *however, estimates of f from a finite sample will not show this delta function behaviour*

The multifractal formalism 1

- to estimate $f(\alpha)$ from discrete data, use partition function $\tau(q)$

The multifractal formalism 1

- to estimate $f(\alpha)$ from discrete data, use partition function $\tau(q)$
- For a process $X(t)$, define the *structure function*, $S_j(q)$, by

$$S_j(q) = \sum_{k_j} (2^{-j/2} U_{j,k_j})^q$$

where U_{j,k_j} are Haar scaling coefficients for $X(t)$

The multifractal formalism 1

- to estimate $f(\alpha)$ from discrete data, use partition function $\tau(q)$
- For a process $X(t)$, define the *structure function*, $S_j(q)$, by

$$S_j(q) = \sum_{k_j} (2^{-j/2} U_{j,k_j})^q$$

where U_{j,k_j} are Haar scaling coefficients for $X(t)$

- The partition function is then defined as

$$\tau(q) = -\lim_{j\to\infty} \frac{1}{j} \log S_j(q)$$

The multifractal formalism 1

- to estimate $f(\alpha)$ from discrete data, use partition function $\tau(q)$
- For a process $X(t)$, define the *structure function*, $S_j(q)$, by

$$S_j(q) = \sum_{k_j} (2^{-j/2} U_{j,k_j})^q$$

where U_{j,k_j} are Haar scaling coefficients for $X(t)$

- The partition function is then defined as

$$\tau(q) = -\lim_{j\to\infty} \frac{1}{j} \log S_j(q)$$

- Next define

$$f_L(\alpha) = \inf_{q\in\mathbb{R}} (q\alpha - \tau(q))$$

The multifractal formalism 2

- The multifractal formalism shows that

$$f(\alpha) \leqslant f_L(\alpha)$$

  The Legendre transform of the partition function is the concave hull of $f(\alpha)$. $f_L(\alpha)$ is known as the *Legendre spectrum*

The multifractal formalism 2

- The multifractal formalism shows that

$$f(\alpha) \leqslant f_L(\alpha)$$

The Legendre transform of the partition function is the concave hull of $f(\alpha)$. $f_L(\alpha)$ is known as the *Legendre spectrum*

- Let us assume that our RTT data is a sample of an underlying continuous process. Assume further that the observable scaling behaviour of $S_j(q)$ is continued beyond the finest measured scale to the limit $j \to \infty$. That is,

$$S_j(q) \approx 2^{-j\tau(q)}$$

over $j = j_1, ..., j_2$, where $j_1, j_2 \in [0, J]$

The multifractal formalism 2

- The multifractal formalism shows that

$$f(\alpha) \leqslant f_L(\alpha)$$

  The Legendre transform of the partition function is the concave hull of $f(\alpha)$. $f_L(\alpha)$ is known as the *Legendre spectrum*

- Let us assume that our RTT data is a sample of an underlying continuous process. Assume further that the observable scaling behaviour of $S_j(q)$ is continued beyond the finest measured scale to the limit $j \to \infty$. That is,

$$S_j(q) \approx 2^{-j\tau(q)}$$

  over $j = j_1, ..., j_2$, where $j_1, j_2 \in [0, J]$

- $\tau(q)$ can be estimated from the gradient of a plot of $\log_2 S_j(q)$ against j over a finite range of scales

Riedi's multifractal wavelet model

- Riedi and others have proposed a stochastic variant of this type of multifractal model

Riedi's multifractal wavelet model

- Riedi and others have proposed a stochastic variant of this type of multifractal model
- We let the wavelet coefficients from one scale to the next be iid random variables
  - ▷ *for example, from the Beta distribution*

Riedi's multifractal wavelet model

- Riedi and others have proposed a stochastic variant of this type of multifractal model
- We let the wavelet coefficients from one scale to the next be iid random variables
  - ▷ *for example, from the Beta distribution*
- Fitting to data then involves estimating the parameters in the distribution

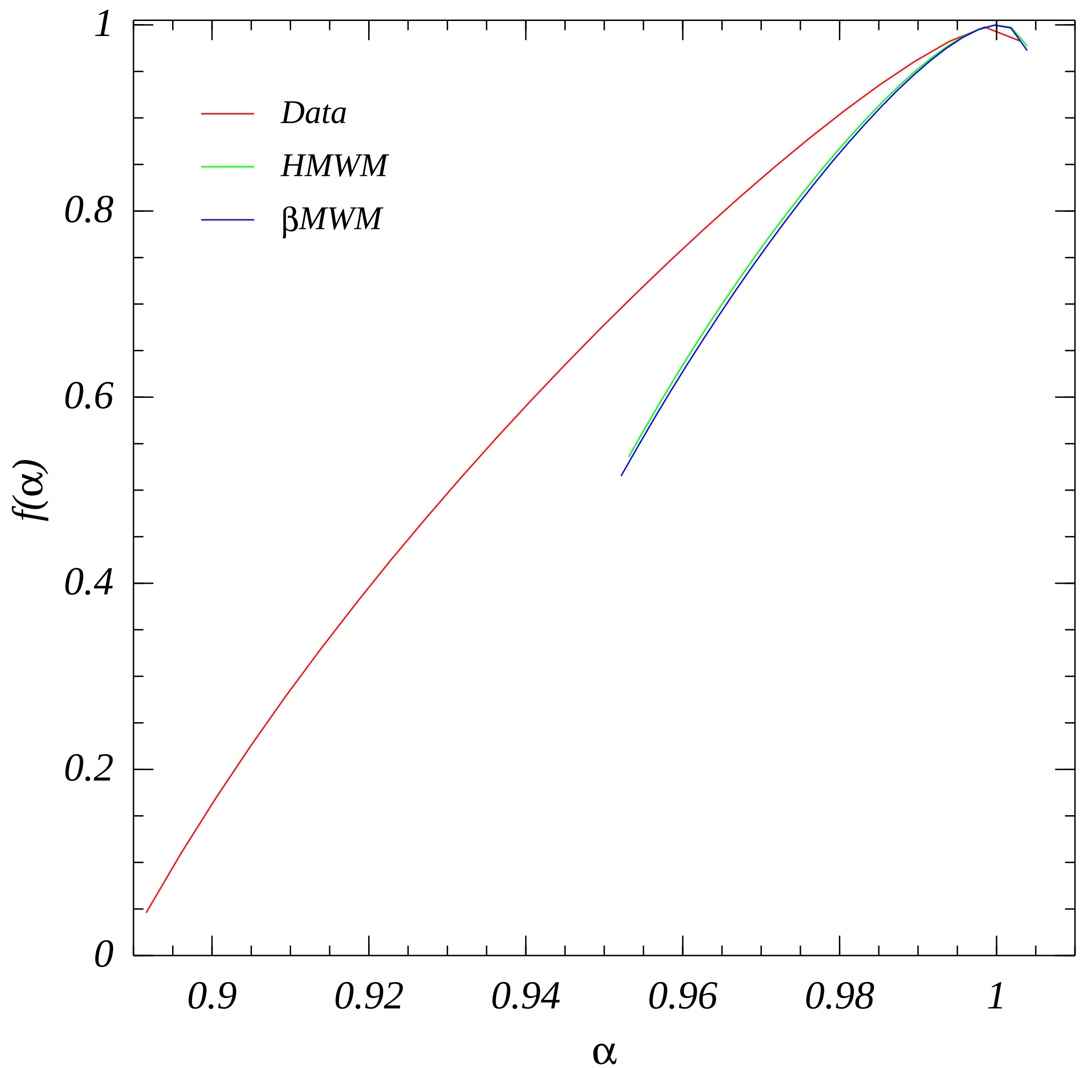
Riedi's multifractal wavelet model

- Riedi and others have proposed a stochastic variant of this type of multifractal model
- We let the wavelet coefficients from one scale to the next be iid random variables
  - ▷ *for example, from the Beta distribution*
- Fitting to data then involves estimating the parameters in the distribution
- simulating involves drawing random variates from the fitted distribution

Riedi's multifractal wavelet model

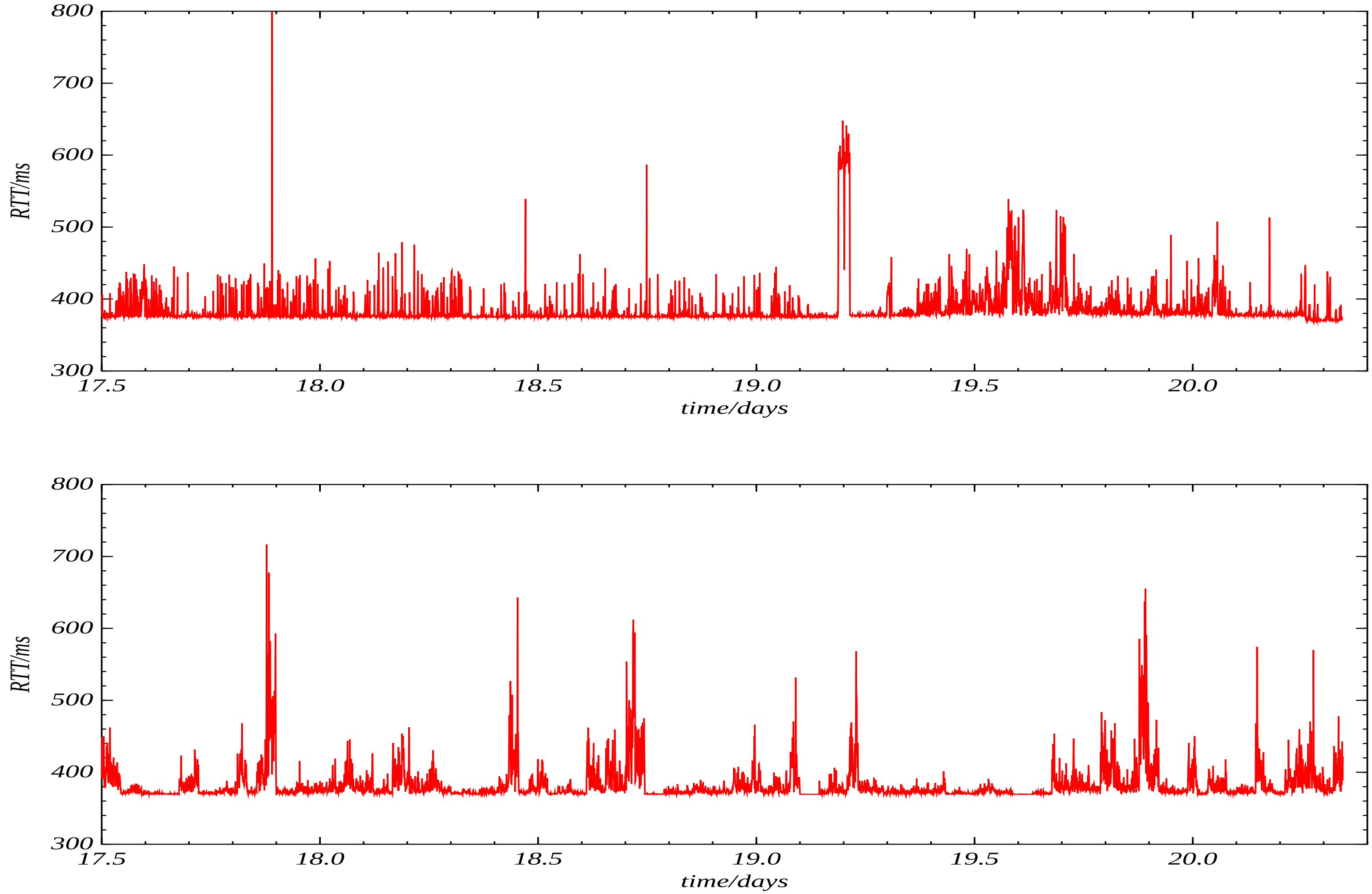
- Riedi and others have proposed a stochastic variant of this type of multifractal model
- We let the wavelet coefficients from one scale to the next be iid random variables
  - ▷ *for example, from the Beta distribution*
- Fitting to data then involves estimating the parameters in the distribution
- simulating involves drawing random variates from the fitted distribution
- determinism and preprocessing
  - Always remove any clear deterministic features from the data first:
    - ▷ *trends*
    - ▷ *periodic components*
    - ▷ *baseline shifts*

Example data analysis



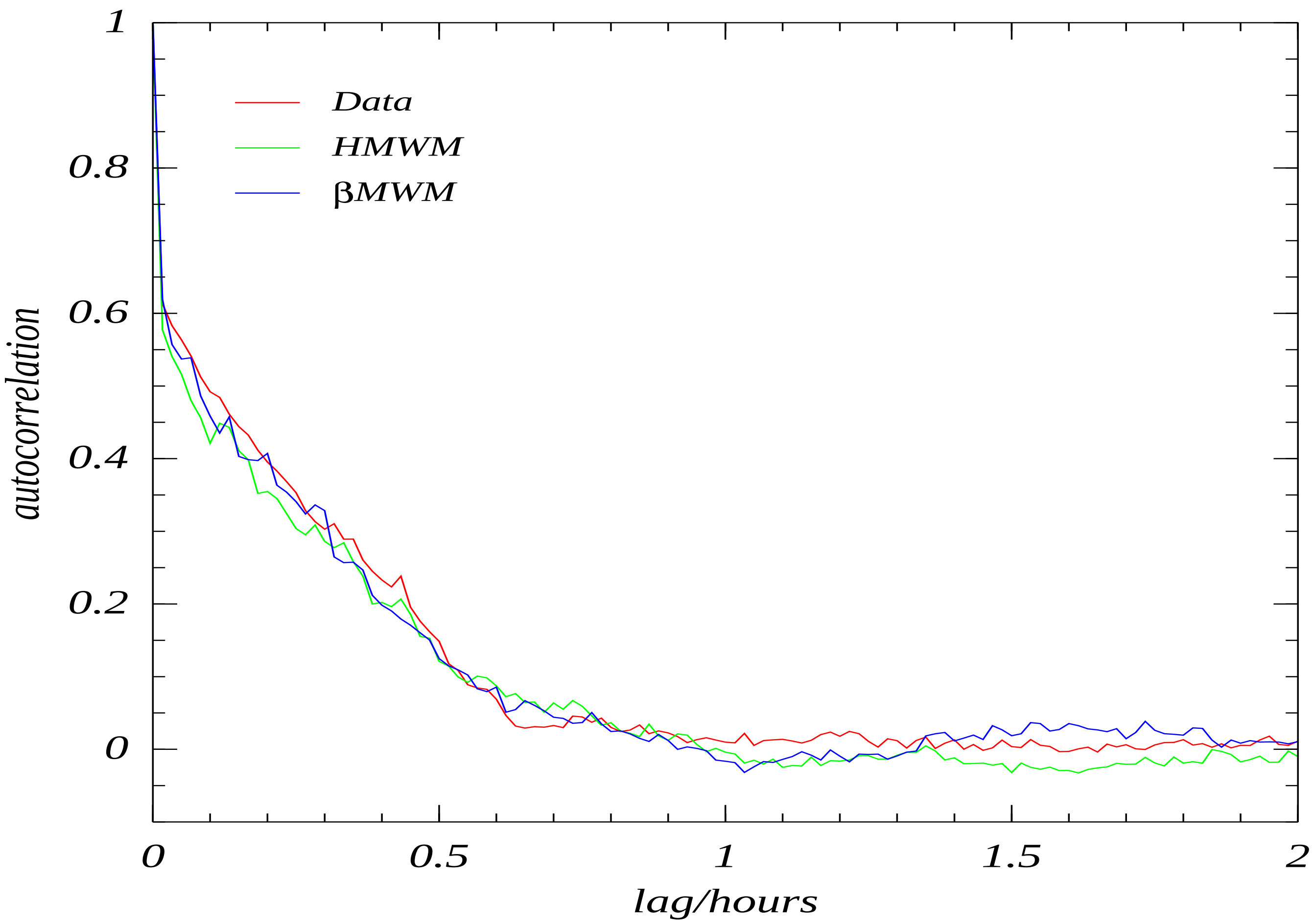
HMWM, βMWM and `www.chem.uwa.edu.au` multifractal spectra

Realizations



2^{12} points from `www.chem.uwa.edu.au` (top) compared with a βMWM realisation (bottom)

Autocorrelation



`www.chem.uwa.edu.au`, HMWM, and βMWM autocorrelation

Conclusion and references

- There are many open questions in this type of work and few concrete models

Conclusion and references

- There are many open questions in this type of work and few concrete models
- Fitting and parameter estimation is a major problem

Conclusion and references

- There are many open questions in this type of work and few concrete models
- Fitting and parameter estimation is a major problem
- Full report and bibliography:
    - ▷ *Analysis and simulation of internet round-trip times*