Navigating Scaling: Modelling and Analysing

P. Abry\(^{(1)}\),

P. Gonçalvès\(^{(2)}\)

\(^{(1)}\) SISYPH, CNRS, Ecole Normale Supérieure Lyon, France

\(^{(2)}\) INRIA Rhône-Alpes, On Leave at IST-ISR, Lisbon

In Collaborations with:


Wavelet And Multifractal Analysis, Cargèse, France, July 2004.
**Navigating Scaling: Modelling And Analysing**

- **Report Date**: 07 JAN 2005
- **Report Type**: N/A
- **Dates Covered**: -
- **Title and Subtitle**: Navigating Scaling: Modelling And Analysing
- **Performing Organization**: SISYPH, CNRS, Ecole Normale Supérieure Lyon, France
- **Abstract**: See also ADM001750, Wavelets and Multifractal Analysis (WAMA) Workshop held on 19-31 July 2004. The original document contains color images.

**Distribution/Availability Statement**
Approved for public release, distribution unlimited

**Supplementary Notes**
See also ADM001750, Wavelets and Multifractal Analysis (WAMA) Workshop held on 19-31 July 2004. The original document contains color images.

**Security Classification of**: unclassified

**Limitation of Abstract**: UU

**Number of Pages**: 64

**Name of Responsible Person**: unclassified
SCALING PHENOMENA?

- **Detection:** Scaling? What does it mean? Non-stationarity?
- **Identification:** Relevant Stochastic Models?
- **Estimation:** Relevant Parameter Estimation?
- **Side Issues:**
OUTLINE

I. INTUITIONS, MODELS, TOOLS
   I.1 INTUITIONS, DEFINITION, APPLICATIONS
   I.2 STOCHASTIC MODELS: SELF-SIMILARITY VS MULTIFRACTAL
   I.3 MULTIRESOLUTION TOOLS, AGGREGATION, INCREMENTS
   I.4 WAVELETS, CONTINUOUS, DISCRETE

II. SECOND ORDER ANALYSIS, SELF SIMILARITY AND LONG MEMORY
   II.1 RANDOM WALKS, SELF SIMILARITY, LONG MEMORY,
   II.2 2ND ORDER WAVELET STATISTICAL ANALYSIS,
   II.3 ESTIMATION, ESTIMATION PERFORMANCE,
   II.4 ROBUSTNESS AGAINST NON STATIONARITIES,

III. HIGHER ORDER ANALYSIS, MULTIFRACTAL PROCESSES
   III.1 MULTIPLICATIVE CASCADES, MULTIFRACTAL PROCESSES,
   III.2 HIGHER ORDER WAVELET STATISTICAL ANALYSIS,
   III.3 FINITENESS OF MOMENTS,
   III.4 ESTIMATION, ESTIMATION PERFORMANCE,
   III.5 NEGATIVE ORDERS,
   III.6 BEYOND POWER LAWS.
IRREGULARITIES, VARIABILITIES

SCALING OR NON STATIONARITIES?
SCALING?
SCALING ?

Trafic (LAN) Ethernet --- Densite Spectrale de Puissance

$\log_{10} (Frequence (Hz))$ vs $\log_{10} (Nombre Octets)$
SCALING!

- **Definition:**
  Non Property: No characteristic scale.
  Non Gaussian, Non Stationary, Non Linear

- **Evidence:**
  The whole resembles to its part, the part resembles to the whole.

- **Analysis:**
  Rather than for a characteristic scale, look for a relation, a mechanism, a cascade between scales.
Scaling: Operational Definitions

- **Multiresolution Quantity:**
  \[ T_X(a, t) \quad \text{(e.g., Wavelet Coef.)}. \]

- **Power Laws:**
  \[ \mathbb{E}|T_X(a, t)|^q = c_q |a|^\zeta(q), \]
  \[ \frac{1}{n} \sum_{k=1}^{n} |T_X(a, t_k)|^q = c_q |a|^\zeta(q), \]
  - for a range of scales \( a \),
  - for a range of orders \( q \),
  - scaling exponents \( \zeta(q) \).

- **Beyond Power Laws: Warped Inf. Div. Cascades**
  \[ \mathbb{E}|T_X(a, t)|^q = C_q |a|^\zeta(q) = C_q \exp(\zeta(q) \ln a) \]
  \[ \mathbb{E}|T_X(a, t)|^q = C_q \exp(\zeta(q)n(a)) \]
  → visit Pierre Chainais’s Poster
UBIQUTY!

- Hydrodynamic Turbulence,
- Physiology, Biological Rythms (Heart beat, walk),
- Geophysics (Faults Repartition, Earthquakes),
- Hydrology (Water Levels),
- Statistical Physics (Long Range Interactions),
- Thermal Noises (semi-conductors),
- Information Flux on Networks, Computer Network Traffic,
- Population Repartition (local: cities, global: continent),
- Financial Markets (Daily returns, Volatily, Currencies Exchange Rates),
- ...
ANALYSING TOOL 1 : AGGREGATION

COMPARE DATA AGAINST A BOX, THEN VARY $a$

$$T_X(a, t) = \frac{1}{aT_0} \int_t^{t+aT_0} X(u)du$$

AVERAGE

WORKS ONLY FOR POSITIVE TIME SERIES, DENSITY
ANALYSING TOOL 2: INCREMENTS

COMPARE DATA AGAINST A DIFFERENCE OF DELTA FUNCTIONS, THEN VARY $a$

$$T_X(a, t) = X(t + a\tau_0) - X(t)$$

DIFFERENCE

INCREMENTS OF HIGHER ORDERS OR GENERALISED $N$-VARIATIONS

- Order 2: $T_X(a, t) = -X(t + 2a\tau_0) + 2X(t + a\tau_0) - X(t)$,
- Order $N$: $T_X(a, t) = \sum_{p=0}^{N} (-1)^p a_p X(t + p a\tau_0)$,
where $\sum_{p=0}^{N} (-1)^p a_p p^k \equiv 0$, $k = 0, \ldots, N - 1$. 

[12]
ANALYSING TOOL: MULTI Resolution Analysis

- **MULTIRESOLUTION QUANTITIES:**

\[ X(t) \rightarrow T_X(a, t) = \langle f_{a,t} | X \rangle, \quad f_{a,t}(u) = \frac{1}{a} f_0(\frac{u-t}{a}) \]

**AGGREGATION**

\[ f_0(u) = (\beta_0) \]

\[ = \frac{1}{aT_0} \int_t^{t+aT_0} X(u)du \]

**BOX, AVERAGE**

**INCREMENTS**

\[ f_0(u) = (I_0) \]

\[ = X(t + a\tau_0) - X(t) \]

**DIFFERENCE**
**Analyzing Tool: MultiResolution Analysis**

- **MultiResolution Quantities:**
  \[ X(t) \rightarrow T_X(a,t) = \langle f_{a,t} | X \rangle, \quad f_{a,t}(u) = \frac{1}{a} f_0(\frac{u-t}{a}) \]

- **Choices for Mother Functions:** \( f_0, \)

  - **Aggregation**
    \[ f_0(u) = (\beta_0)^N \]
    \[ = \frac{1}{aT_0} \int_t^{t+aT_0} X(u) du \]
  
  - **Box, Average**

  - **Increments**
    \[ f_0(u) = (I_0)^N \]
    \[ = X(t + a\tau_0) - X(t) \]
  
  - **Difference**

  - **Wavelets**
    \[ f_0(u) = \psi_{0,N} \]
    \[ = \int X(u) \frac{1}{a} \psi_0(\frac{u-t}{a}) \]
  
  - **Average, Difference**
**WAVELETS AND SCALING: KEY INGREDIENTS**

- **DILATION OPERATOR,** \( \frac{1}{|a|} \psi_0(\frac{t}{|a|}) \)

- **NUMBER OF VANISHING MOMENTS,**
  \[
  N \geq 1, \int t^k \psi_0(t) dt \equiv 0, \quad k = 0, 1, \ldots, N - 1.
  \]
**Wavelet Transforms**

- **Mother-Wavelet and "Basis"**: \( \int \psi_0(u) du = 0, \quad \psi_{a,t}(u) = \frac{1}{|a|} \psi_0\left(\frac{u-t}{a}\right) \)

- **Wavelet Coefficients**: Continuous WT \( T_X(a, t) = \langle X, \psi_{a,t} \rangle \)

**Modulus Maxima WT**

**Skeleton**: Maxima Lines

**And Discrete WT**

\( d_X(j, k) = T_X(a = 2^j, t = 2^j k) \)
OUTLINE

I. INTUITIONS, MODELS, TOOLS
   I.1 INTUITIONS, DEFINITION, APPLICATIONS
   I.2 STOCHASTIC MODELS: SELF-SIMILARITY VS MULTIFRACTAL
   I.3 MULTiresolution TOOLS, AGGREGATION, INCREMENTS
   I.4 WAVELETS, CONTINUOUS, DISCRETE

II. SECOND ORDER ANALYSIS, SELF SIMILARITY AND LONG MEMORY
   II.1 RANDOM WALKS, SELF SIMILARITY, LONG MEMORY,
   II.2 2ND ORDER WAVELET STATISTICAL ANALYSIS,
   II.3 ESTIMATION, ESTIMATION PERFORMANCE,
   II.4 ROBUSTNESS AGAINST NON STATIONARITIES,

III. HIGHER ORDER ANALYSIS, MULTIFRACTAL PROCESSES
   III.1 MULTIPLICATIVE CASCADES, MULTIFRACTAL PROCESSES,
   III.2 HIGHER ORDER WAVELET STATISTICAL ANALYSIS,
   III.3 FINITENESS OF MOMENTS,
   III.4 ESTIMATION, ESTIMATION PERFORMANCE,
   III.5 NEGATIVE ORDERS,
   III.6 BEYOND POWER LAWS.
MOD. TOOL 1: RAND. WALKS AND SELF SIMILARITY

RANDOM WALK: \( X(t + \tau) = X(t) + \delta_{\tau}X(t) \)

Steps or Increments

STATISTICAL PROPERTIES OF THE STEPS:
- **A1**: Stationary,
- **A2**: Independent,
- **A3**: Gaussian,
  \[ \Rightarrow \text{Ordinary Random Walk, Ordinary Brownian Motion,} \]
  \[ \Rightarrow \mathbb{E}X(t)^2 = 2D|t|, \text{ Einstein relation,} \]
  \[ \Rightarrow \mathbb{E}|X(t)|^q = 2D|t|^{q/2}, \quad q > -1. \]

ANOMALIES:
  \[ \Rightarrow \mathbb{E}X(t)^2 = 2D|t|^\gamma, \]
  \[ \Rightarrow \mathbb{E}X(t)^2 = \infty. \]

SELF SIMILAR RANDOM WALKS:
- **B1**: Stationary,
- **B2**: Self Similarity
**Modelling Tool 1: Self-Similarity**

- **Definition:**
  \[ \delta_\tau X(t) \overset{fdd}{=} c^H \delta_{\tau/c} X(t/c), \ \forall c > 0, \ \text{Dilation Factor}, \ 1 > H > 0 : \text{Self-Similarity Exponent} \]

- **Interpretations:**
  - Covariance under Dilation (Change of Scale);
  - The Whole and the SubPart (Statistically) Undistinguishable;
  - No Characteristic Scale of Time.

- **Implications:**
  - Non Stationarity process with stationary increments
  - \[ \mathbb{E} \left| X(t + a\tau_0) - X(t) \right|^q = C_q |a|^{qH}, \]
  - \[ \forall a > 0, \ \forall c > 0, \ \forall q / \mathbb{E} |X(t)|^q < \infty, \]
  - A Single Scaling Exponent \( H \).
  - Additive Structure,
MOD. TOOL 1 (bis): LONG RANGE DEPENDENCE

● DEFINITIONS:

- Let \( X \) be a 2nd stationary process with,
- Covariance: \( c_X(\tau) = \mathbb{E}X(t)X(t + \tau) \)
- Spectrum: \( \Gamma_X(\nu) \)

\[
c_X(\tau) = c_\tau |\tau|^{-\beta}, \quad 0 < \beta < 1, \quad |\tau| \to +\infty
\]
\[
\Gamma_X(\nu) = c_f |\nu|^{-\alpha}, \quad 0 < \alpha < 1, \quad |\nu| \to 0
\]

With \( \alpha = 1 - \beta \) and \( c_f = 2(2\pi) \sin((1 - \gamma)\pi/2)c_\tau \).

● CONSEQUENCES:

- \( \sum_A^{+\infty} c_X(\tau) d\tau = +\infty, \; A > 0, \)
- No Characteristic Scale,
- Aggregation: \( T_X(a, t) = \frac{1}{aT_0} \int_t^{t+aT_0} X(u) du, \)
  \( \Rightarrow \text{VAR} T_X(a, t) \sim C a^{\alpha-1}, \; a \to +\infty, \)
- Increments of Self.-Sim. Proc. (with \( H > 1/2 \))
  are Long Range Dep. (with \( \alpha = 2H - 1 \)).
Wavelets and Self-Similar Processes with Stationary Increments - Summary

(Flandrin et al., Tewfik and Kim)

- **P1**: \( \{d_X(j, k), k \in \mathbb{Z}\} \) Stationary Sequences for each Scale \( 2^j \).
  \( N \geq 1 \)

- **P2**: **Self-Similarity** : Dilation
  \( \{X(t)\} \overset{d}{=} \{c^H X(t/c)\} \Rightarrow \{d_X(0, k)\} \overset{d}{=} \{2^{-jH} d_X(j, k)\} \)

- **P3**: **Marginal Dist.** \( P_j(d) = \frac{1}{\beta_0} P_{j'}(\frac{d}{\beta_0}), \quad \beta_0 = \left(\frac{2^{j'}}{2^j}\right)^H \)

- **P4**: \( \{d_X(j, k)\} \) **Short Range Dependent if** \( N > H + 1/2 \).
  \( |2^j k - 2^{j'} k'| \to +\infty, \quad |\text{Cov } d_X(j, k)d_X(j', l)| \leq D|2^j k - 2^{j'} k'|^{2(H-N)}, \quad N \geq 1 \) and Dilation
Wavelets and Long Range Dependence (Flandrin)

\[ H = 0.15 \]

\[ H = 0.5 \]

\[ H = 0.95 \]
WAVELETS AND SELF-SIMILAR PROCESSES WITH STATIONARY INCREMENTS - SUMMARY

- **P1**: \(\{d_X(j, k), k \in \mathbb{Z}\}\) Stationary Sequences for each Scale \(2^j\). \(N \geq 1\)

- **P2**: Self-Similarity: Dilation

  \(\{X(t)\} \overset{d}{=} \{c^H X(t/c)\} \Rightarrow \{d_X(0, k)\} \overset{d}{=} \{2^{-jH}d_X(j, k)\}\)

- **P3**: Marginal Dist.

  \(P_j(d) = \frac{1}{\beta_0} P_j'\left(\frac{d}{\beta_0}\right), \quad \beta_0 = \left(\frac{2^j}{2^j}\right)^H\)

- **P4**: \(\{d_X(j, k)\}\) Short Range Dependent if \(N > H + 1/2\).

  \(|2^j k - 2^j' k'| \to +\infty, \quad |\text{Cov} \, d_X(j, k) d_X(j', k')| \leq D |2^j k - 2^j' k'|^{2(H-N)}, \quad N \geq 1\) and Dilation

\(\Rightarrow\) Idealisation: \(d_X(j, k)\) Independent Variables.

\(\Rightarrow\) Interpretations:

\(X(t) = \sum_k a_X(J, k) \varphi_{J,k}(t) + \sum_{j=1, \ldots, J,k} d_X(j, k) \psi_{j,k}(t)\).

\(\Rightarrow\) Implications:

\(\mathbb{E}|d_X(j, k)|^q = \mathbb{E}|d_X(0, k)|^q 2^{jqH} \quad \forall q/\mathbb{E}|d_X(0, k)|^q < \infty.\)
**Wavelets and Long Range Dependence**

- **Spectral Analysis:**
  Let $X$ be a 2nd Order stationary process,
  Let $\Psi$ be the FT of $\psi$ with central frequency $\nu_0$ and bandwith $\Delta\nu_0$.

  \[ E|d_X(j, k)|^2 = \int \Gamma_X(\nu)|\Psi(2^j \nu)|d\nu \]
  \[ \simeq 2^{-j}\Gamma_X(2^{-j}\nu_0) \text{ within bandwith } 2^{-j}\Delta\nu_0. \]

- **Let $X$ be Long Range Dependent:**
  - **Power Law:** $\Gamma_X(\nu) = cf|\nu|^{-\alpha}, 0 < \alpha < 1, |\nu| \to 0$
  - **Power Law:** $E|d_X(j, k)|^2 \sim C2^{j(\alpha-1)}, j \to +\infty,$

- $\{d_X(j, k)\}$ **Short Range Dependent if** $N > \alpha - 1$.

  $|2^j k - 2^{j'} k'| \to +\infty,$  $|Cov d_X(j, k)d_X(j', k')| \leq D|2^j k - 2^{j'} k'|^{\alpha-1-2N},$

  $N \geq 1$ and Dilation
**2nd Order Wavelet Statistical Analysis**

Abry, Gonçalvès, Flandrin

**Principles:**
- **Ideas:**
  
  \[ P1 \Rightarrow E|d_X(j, k)|^2 = C_2 2^{2H} \]
  \[ \Rightarrow \log_2 E|d_X(j, k)|^2 = j 2H + \beta_q, \]

- **Problems:** Estimate \( E|d_X(j, k)|^2 \) from a single finite length observation?

- **Solution:**
  
  \[ P2 \text{ et } P3 \Rightarrow \text{Statistical Averages } \Rightarrow \text{Time Averages}, \]
  \[ S_2(j) = (1/n_j) \sum_{k=1}^{n_j} |d_X(j, k)|^2 \]

**Log-Scale Diagrams:** \( \log_2 S_2(j) \) vs \( \log_2 2^j = j \)
2ND ORDER WAVELET-BASED STATISTICAL ANALYSIS FOR SELF-SIMILARITY

\[ \alpha = 2.57 \quad (1 \leq j \leq 10) \]
2nd Order Wavelet-based Statistical Analysis for Long Range Dependence

\[ \alpha = 0.55 \]
\[ c_f = 4.7 \]
\[ 4 \leq j \leq 10 \]
**Wavelets and 2nd-order Scaling: Estimation**

- **Dyadic Grid (Discrete Wavelet Transform):**
  \[ a_j = 2^j, \quad t_{j,k} = k2^j, \]

- **Structure Function (Time Average):**
  \[ Y_j = \left( \frac{1}{2} \log_2 S_2(2^j) \right) = \frac{1}{2} \log_2(1/n_j) \sum_{k=1}^{n_j} |dX(j,k)|^2 \]

- **Definition:**
  \[ Y_j \text{ versus } \log_2 2^j = j, \]
  \[ \hat{H} = \sum_{j=j_1}^{j_2} w_j Y_j. \]

  \[ \text{WHERE } \sum_j jw_j \equiv 1, \quad \sum_j w_j \equiv 0, \text{ WITH } w_j \equiv \frac{B_0 j - B_1}{B_0 B_2 - B_1^2}, \]
  \[ \text{AND } \quad p = 0, 1, 2, \quad B_p = \sum_j j^p / a_j, \quad a_j \text{ ARBITRARY NUMBERS.} \]

- **What Are the Performance of Such an Estimator?**
  When applied to a Self-Similar. or LRD Process
WAVELETS AND 2ND-ORDER SCALING: ESTIMATION

Abry, Gonçalvès, Flandrin, Abry, Veitch

- **ASSUME**:  
  - i) $X$ GAUSSIAN,  
  - ii) IDEALISATION: EXACT INDEPENDENCE.

- **BIAS**:  
  \[
  \mathbb{E} \log_2 S_2(j) = \log_2 \mathbb{E} S_2(j) + \Gamma'(n_j/2) - \log_2(n_j/2). 
  \]

  \[ 
  \Rightarrow \mathbb{E} \hat{H} = H + \frac{1}{2} \sum_j w_j g_j, 
  \]

- **VARIANCE**:  
  - \[ \text{Var} \hat{H} = \frac{1}{4} \sum_j w_j^2 \sigma_j^2, \]
  - \[ \min \text{Var} \hat{H} \implies a_j \propto \text{Var} \log_2 S_2(j) \]
  - \[ \text{Var} \log_2 S_2(j) \simeq C/n_j \simeq 2^j C/n, \]

  \[ 
  \Rightarrow \text{VAR} \hat{H} \simeq \left( (\log_2 e)^2 \left( \sum_j w_j^2 2^j \right) \right) / n, 
  \]

  \[ 
  \Rightarrow \text{ANALYTICAL (APPROXIMATE) CONFIDENCE INTERVAL (DOES NOT DEPEND ON UNKNOWN H)}. 
  \]

- **ACTUAL PERFORMANCES**: NEGLIGIBLE BIAIS, EXTREMELY CLOSE TO MLE.

- **CONCEPTUAL AND PRACTICAL SIMPLICITY**: MATLAB CODE AVAILABLE.
WAV. AND 2ND-ORDER SCALING: ROBUSTNESS

Superimposed Trends

\[ Y(t) = X(t) + T(t) \Rightarrow d_Y(j, k) = d_X(j, k) + d_T(j, k) \]

- If \( T(t) \) Polynomial of degree \( P \), then \( d_T \equiv 0 \) when \( N > P \),
- If \( T(t) \) smooth trend, then the \( d_T \) decrease as \( N \) increases.

Vary \( N \)!
WAV. AND 2ND-ORDER SCALING: ROBUSTNESS

Superimposed Trends - Ethernet Data (Veitch, Abry)

Logscalet Diagram, N=2

Full trace: $\alpha = 0.60$

Part I: $\alpha = 0.62$

Part II: $\alpha = 0.58$
WAV. AND 2ND-ORDER SCALING: ROBUSTNESS

Constancy along time of Scaling laws (Veitch, Abry)
SELF-SIMILARITY

- **SELF-SIMILARITY:**
  \[
  \mathbb{E}|d_X(j, k)|^q = C_q(2^j)^{qH}
  \]
  - Power Laws,
  - \(\forall 2^j\) (for all scales),
  - \(\forall q \mathbb{E}|d_X(j, k)|^q < \infty\),
  - A single parameter \(H\)
  - Additive Structure.

- ?

- ?

- ?
BEYOND SELF-SIMILARITY

- **Self-Similarity:**
  \[ E|d_X(j, k)|^q = C_q(2^j)^qH \]
  - Power Laws,
  - \( \forall 2^j \) (for all scales),
  - \( \forall q \) \( \forall E|d_X(j, k)|^q < \infty \),
  - A single parameter \( H \)
  - Additive Structure.

- **MultiFractal**
  \[ E|d_X(j, k)|^q = C_q(2^j)\zeta(q) \]
  - Power Laws,
  - \( \forall 2^j < L \), (for fine scales only, in the limit \( 2^j \rightarrow 0 \),)
  - \( \forall q \)
  - A whole collection of scaling parameter \( \zeta(q) \)
  - Multiplicative Structure.

- ?
I. INTUITIONS, MODELS, TOOLS
   I.1 INTUITIONS, DEFINITION, APPLICATIONS
   I.2 STOCHASTIC MODELS: SELF-SIMILARITY VS MULTIFRACTAL
   I.3 MULTiresolution tools, Aggregation, Increments
   I.4 Wavelets, Continuous, Discrete

II. SECOND ORDER ANALYSIS, SELF SIMILARITY AND LONG MEMORY
   II.1 Random Walks, Self Similarity, Long Memory,
   II.2 2nd Order Wavelet Statistical Analysis,
   II.3 Estimation, Estimation Performance,
   II.4 Robustness against Non Stationarities,

III. HIGHER ORDER ANALYSIS, MULTIFRACTAL PROCESSES
   III.1 Multiplicative Cascades, Multifractal Processes,
   III.2 Higher Order Wavelet Statistical Analysis,
   III.3 Finiteness of Moments,
   III.4 Estimation, Estimation Performance,
   III.5 Negative Orders,
   III.6 Beyond Power Laws.
**Definition:**
- **Split Dyadic Intervals** $I_{j,k}$ into two,
- **I.I.D. Multipliers** $W_{j,k}$
- $Q_J(t) = \prod\{ (j,k): 1 \leq j \leq J, t \in I_{j,k} \} W_{j,k}$

**Implications:**
- **Local Holder Exponent**, 
- **MultiFractal Sample Paths**, **MultiFractal Spectrum** $D(h)$
- **Cascades, Multiplicative Structure**, 
- $\sum_k \left( \frac{1}{a} \int_{t_k}^{t_{k+1}} X(u)du \right)^q = C_q |a|^{q} \zeta_q$, **Fine Scales** $a \to 0$,
- **Multiple Exponents** $\zeta_q$,
- **No Characteristic Scale**, 
- $\zeta_q = - \log_2 |E W|^q$, **Non Linear in** $q$. 

[36]
**Modelling Tool 2: Multiplicative Cascades**

<table>
<thead>
<tr>
<th>Yaglom, Mandelbrot</th>
<th>Barral, Mandelbrot</th>
<th>Schmmitt et al., Bacry et al., Chainais et al.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mandelbrot's Cascade (CMC)</td>
<td>Compound Poisson Cascade (CPC)</td>
<td>Infinitely Divisible Cascade (IDC)</td>
</tr>
<tr>
<td>- IID $W$,</td>
<td>- IID $W$,</td>
<td>- Continuous Infinitely</td>
</tr>
<tr>
<td>- Dyadic Grid,</td>
<td>- Point Process,</td>
<td>Divisible Measure $M$,</td>
</tr>
</tbody>
</table>

\[
Q_r(t) = \prod W_{j,k},
\]

\[
\varphi(q) = -\log_2 \mathbb{E} W^q,
\]

\[
\mathbb{E} \left| A(t + a\tau_0) - A(t) \right|^q = c_q \left| a \right|^{q+\varphi(q)},
\]

**For a range of $q$s,** $\mathbb{E} \left| A(t + a\tau_0) - A(t) \right|^q = c_q \left| a \right|^{q+\varphi(q)},$

**Resolution Depth < Scale < Integral Scale,** $a_m = r < a < a_M = L$. 

\[\text{[37]}\]
**Density:** \[ Q_r(t) = \prod W_{j,k} \]
\[
\mathbb{E} \left( \frac{1}{a} \int_t^{t+a\tau_0} Q_r(u) du \right)^q = c_q a^{\varphi(q)},
\]

**Measure:** \[ A(t) = \lim_{r \to 0} \int_0^t Q_r(u) du, \]
\[
\mathbb{E} |A(t + a\tau_0) - A(t)|^q = c_q |a|^{q+\varphi(q)},
\]

**Fractional Brownian Motion in Multifractal Time:**
\[ V_H(t) = B_H(A(t)), \]
\[
\mathbb{E} |V_H(t + a\tau_0) - V_H(t)|^q = c_q |a|^{qH+\varphi(qH)},
\]

**Multifractal Random Walk:**
\[ Y_H(t) = \int_t^t Q_r(s) dB_H(s), \]
\[
\mathbb{E} |Y_H(t + a\tau_0) - Y_H(t)|^q = c_q |a|^{qH+\varphi(q)}.
\]
**Higher-Order Wavelet Statistical Analysis**

**Principles:**

- **Ideas:**
  \[ P1 \Rightarrow \mathbb{E}|d_X(j, k)|^q = \mathbb{E}|d_X(0, k)|^q 2^{jq} \]
  \[ \Rightarrow \log_2 \mathbb{E}|d_X(j, k)|^q = j \zeta q + \beta_q, \]

- **Problems:** Estimate \( \mathbb{E}|d_X(j, k)|^q \) from a single finite length observation?

- **Solution:**
  \[ P2 \text{ et } P3 \Rightarrow \text{Statistical Averages } \Rightarrow \text{Time Averages,} \]
  \[ S_q(j) = (1/n_j) \sum_{k=1}^{n_j} |d_X(j, k)|^q \]

**Log-Scale Diagrams:**
\[ \log_2 S_q(j) \text{ vs } \log_2 2^j = j \]
LOGSCALE DIAGRAMS - MULTI FRACTAL PROC.

Graphs showing the relationship between Octave and various values of q, with different scales for the y-axis.
WAV. AND HIGHER-ORDER SCALING: ESTIMATION

- **Dyadic Grid (Discrete Wavelet Transform):**
  \[ a_j = 2^j, \quad t_{j,k} = k2^j, \]

- **Structure Functions (Time Average):**
  \[ S_q(j) = (1/n_j) \sum_{k=1}^{n_j} |d_X(j,k)|^q \]

- **Definition:**
  \[ Y_{j,q,n} = \log_2 S_n(2^j, q; f_0) \text{ versus } \log_2 2^j = j, \]
  \[ \hat{\zeta}(q, n) = \sum_{j=j_1}^{j_2} w_{j,q} Y_{j,q,n}. \]
  Non Weighted: \( a_j = \text{cste} \)

- **What Are the Performance of Such Estimators?**
  When applied to MultiFractal Processes.
TEST FOR THE FINITENESS OF MOMENTS

GONÇALVÈS, RIEDI

THEOREM:

Let $X$ be a RV with characteristic function $\chi(s) := \mathbb{E}\exp\{isX\}$. If $\mathcal{H}_{\Re\chi} := \sup\{\alpha > 0 : |\Re\chi(s) - P_\alpha(s)| \leq C|s|^\alpha\}$, is the local Hölder regularity of $\Re\chi$ at the origin, then

$$\mathbb{E}|X|^q < +\infty \forall q \leq q_c^+$$

and $\mathcal{H}_{\Re\chi} \leq q_c^+ \leq \lceil \mathcal{H}_{\Re\chi} \rceil + 1$.

ESTIMATOR:

$$\{X_k\}_{k=1,\ldots,n}, \text{ n i.i.d RVs, set}$$

$$W(a) := n^{-1} \sum_{k=1}^{n} \Psi(a.X_k)$$

with $\Psi$ a real and semi-definite Fourier transform of a sufficiently regular wavelet $\psi$. Then

$$\mathcal{H}_{\Re\chi} = \limsup_{a \to 0^+} \frac{\log |W(a)|}{\log a}.$$
ESTIMATING THE PARTITION FUNCTION SUPPORT
METHODOLOGY

- **Numerical Synthesis of Processes:**
  - Accumulate $nbreal$ numerical replications with length $n$ samples.

- **Apply Scaling Exponents Estimators:**
  - Compute $\hat{\zeta}(q, n(l))$ for each replication,
  - Average over repl. to obtain the statistical performance of $\hat{\zeta}(q, n)$

- **Asymptotic Behaviours:**
  - The cascade depth increases for a given number of Integral Scales.
  - …,

[Diagram showing a scatter plot with points $(t_i, r_i)$ enclosed in a shaded triangle.]
**Methodology**

- **Numerical Synthesis of Processes:**
  - Accumulate \(nbreal\) numerical replications with length \(n\) samples.

- **Apply Scaling Exponents Estimators:**
  - Compute \(\hat{\zeta}(q, n)(l)\) for each replication,
  - Average over repl. to obtain the statistical performance of \(\hat{\zeta}(q, n)\)

- **Asymptotic Behaviours:**
  - The cascade depth increases for a given number of Integral Scales.
  - The number of Integral Scales increases for a given cascade depth,
LINEARISATION EFFECT: $\hat{\zeta}(q)$

LASHERMES, ABRY, CHAINAIS

CPC $Q_r$ $EI(1)$

CPC $V_H$ $EIII(3)$

$q > q_o$, $\hat{\zeta}(q, n) = \alpha_o + \beta_o q$, $q_o, \alpha_o, \beta_o$ ARE RV.
LINEARISATION EFFECT: LEGENDRE TRANSFORM

\[ D(h) = d + \min_q(qh - \zeta(q)), \] (d EUCLIDIEN DIMENSION OF SPACE).

\[ \text{CPC } Q_r \ E1(1) \]

\[ \text{CPC } V_H \ EIII(3) \]

ACCUMULATION POINTS: \( D_o(h_o) \), WITH \( D_o = d - \alpha_o, \ h_o = \beta_o, \)
\( D_o, h_o \) ARE RV.
**LIN. EFFECT: ASYMPTOTIC BEHAVIOURS**

- **Given resolution, increasing number of integral scales,**

- **Given number of integral scales, increasing resolution,**
**LINEARISATION EFFECT: CONJECTURE**

- **Critical Points:**
  \[
  \begin{align*}
  D_*^{\pm} &= 0, \\
  D(h_*^\pm) &= 0, \\
  h_*^\pm &= (d\zeta(q)/dq)_{q=q_*^{\pm}}.
  \end{align*}
  \]

- **Results:**

  **EI:**
  \[
  \begin{align*}
  \hat{\zeta}(q, n) &= d - D_o^- + h_o^- q 
  \quad \rightarrow \quad d - D_*^- + h_*^- q, \quad q \leq q_*^-, \\
  \zeta(q, n) &= d - D_o^+ + h_o^+ q 
  \quad \rightarrow \quad \zeta(q), \quad q_*^- \leq q \leq q_*^+, \\
  \hat{\zeta}(q, n) &= d - D_o^- + h_o^- q 
  \quad \rightarrow \quad d - D_*^+ + h_*^+ q, \quad q_*^+ \leq q.
  \end{align*}
  \]

  **EII & III:**
  \[
  \begin{align*}
  \hat{\zeta}(q, n) &= d - D_o^- + h_o^- q 
  \quad \rightarrow \quad \zeta(q), \quad 0 < q \leq q_*^+, \\
  \hat{\zeta}(q, n) &= d - D_o^+ + h_o^+ q 
  \quad \rightarrow \quad d - D_*^+ + h_*^+ q, \quad q_*^+ \leq q.
  \end{align*}
  \]

- **Illustration:**

---

[49]
LINEARISATION EFFECT: COMMENTS

WHEN DOES THE LINEARISATION EFFECT EXIST?

- for all types of cascades: CMC, CPC, IDC,
- for all types of processes: $Q_r, A, V_H, Y_H$,
- for all numbers of vanishing moments: $N \geq 1$,
- for all MRA-based estimators: Wavelets, Increments, Aggregation,
- can be worked out for $q < 0$,
- extends to dimension higher than $d > 1$. 
EXTENSION: STANDARD WT VERSUS WTMM (1/3).
EXTENSION: 2D MULTIPlicative CASCADE (2/3).
EXTENSION: 3D MULTIPLICATIVE CASCADE (3/3).

3D CMC (LOG NORMAL), EI(1) COMPARED TO A 2D SLICE.
LINEARISATION EFFECT: COMMENTS

WHEN DOES THE LINEARISATION EFFECT EXIST?
− for all types of cascades: CMC, CPC, IDC,
− for all types of processes: $Q_r$, $A$, $V_H$, $Y_H$,
− for all numbers of Vanishing Moments: $N \geq 1$,
− for all MRA-based estimators: Wavelets, Increments, Aggregation,
− can be worked out for $q < 0$,
− extends to dimension higher than $d > 1$.

WHAT THE LINEARISATION EFFECT IS NOT:
− a low performance estimation effect.
− a finite size effect: The critical parameters do not depend on $n$,
  be it the number of integral scales,
  or the depth (or resolution) of the cascades.
− a finiteness of moments effect,
  \[ q^-_c < 0 < 1 < q^+_c, \quad q - 1 + \varphi(q) = 0, \]
  \[ q^-_c < q^-_* < 0 < 1 < q^*_+ < q^+_c, \]

WHAT THE LINEARISATION EFFECT MIGHT BE:
− Multiplicative Martingales?
− Ossiander, Waymire 00, Kahane, Peyrière 75, Barral, Mandelbrot 02.
LINEARISATION EFFECT: PICTURE

- **TWO POWER-LAWS, TWO FUNCTIONS OF** $q$:
  
  - **BARE CASCADE:**
    $$\mathbb{E}Q_r(t)^q = r^{\varphi(q)}, \quad q \in \mathcal{R}.$$  
  
  - **DRESSED CASCADE:**
    $$\begin{align*}
    \mathbb{E}T^q_{Q_0}(t, a; \beta_0) &= c_q |a|^{\zeta(q)}, & q &\in [q_c^-, q_c^+], \\
    \mathbb{E}T^q_{Q_0}(t, a; \beta_0) &= \infty, & \text{ELSE}, &
    \end{align*}$$
    WITH:
    $$\begin{align*}
    \zeta(q) &= 1 + q h_-^*, & q &\in [q_c^-, q_*^-], \\
    \zeta(q) &= \varphi(q), & q &\in [q_*^-, q_*^+], \\
    \zeta(q) &= 1 + q h_+^*, & q &\in [q_*^+, q_c^+].
    \end{align*}$$

- **CONFUSION BETWEEN** $\varphi(q)$ **AND** $\zeta(q)$:
  
  - **MULTIPLICATIVE CASCADE:** $\varphi(q), \quad q \in \mathcal{R}$, 
  
  - **SCALING EXPONENTS:** $\zeta(q), \quad q \in [q_c^-, q_c^+]$. 
LINEARISATION EFFECT: SKETCHED VIEWS

Moments

\[ E A_{\tau}(t)^q = \begin{array}{c} q_c^- \quad q_*^- \quad -1 \quad 0 \quad 1 \quad q_*^+ \quad q_c^+ \\ \end{array} \]

Estimated \( \zeta(q,n) \)

\[ \begin{array}{c} \text{EI} \quad 1+q_h^- \\ \end{array} \]

Estimated \( \zeta(q,n) \)

\[ \begin{array}{c} \text{EI II & E III} \\ \end{array} \]

\[ \begin{array}{c} \zeta(q) \quad 1+q_h^+ \\ \end{array} \]
LINEARISATION EFFECT: IMPACTS AND IMPORTANCE

CONSEQUENCES: RECAST THE USUAL GOALS:

- Estimate the Integral Scale and the Resolution of the Cascade,
  ⇒ i.e., Find a Scaling Range \([a_m, a_M]\]
- Estimate the Critical Parameters \(D^\pm, h^\pm, q^\pm\),
- Estimate the \(\zeta(q)\) for \(q \in [q^-, q^+]\),
  → VISIT B. LASHERMES’S POSTER.

IMPORTANCE OF THE LINEARISATION EFFECT:

- Discrimination of MF Models based on \(\hat{\zeta}(q, n)\),
- Discrimination between monoFractal and MultiFractal,
NEGATIVE VALUES OF $q$S

DIFFICULTIES?

- Finiteness? $S_q(j) = (1/n_j) \sum_{k=1}^{n_j} |d_X(j, k)|^q < \infty$?
- Numerical Instability? $d_X(j, k) \simeq 0 \rightarrow |d_X(j, k)|^q = \infty$?
- Theory? Full Multifractal Spectrum?

SOLUTIONS?
NEGATIVE VALUES OF $qS$ - SOLUTION 1

AGGREGATION: \[ T_X(a, t) = \frac{1}{aT_0} \int_{t}^{t+aT_0} X(u) \, du \]

APPLIES ONLY TO POSITIVE DATA (MEASURE)
NEGATIVE VALUES OF $qs$ - SOLUTION 2

WT MODULUS MAXIMA (ARNEODO ET AL.)

$$L_X(a, t_k) = \text{SUP}_{a' < a} \left| T_X(a', t_k(a')) \right|$$

COMPUTATIONALLY EXPENSIVE

[60]
**NEGATIVE VALUES OF $qS$ - SOLUTION 3**

**Wavelet Leaders:** (Jaffard et al.)

\[ d_X(j, k) \rightarrow L_X(j, k) = \text{SUP}_{j' < j} d_X(j', 2^{-j'}) \]

**Computationally Efficient and Excellent Statistical Performance**
BEYOND POWER LAWS

● SELF-SIMILARITY:
  \[ \mathbb{E}|d_X(j, k)|^q = C_q(2^j)^{qH} = C_q \exp(qH \ln 2^j) \]
- POWER LAWS,
- \( \forall 2^j \) (FOR ALL SCALES),
- \( \forall q/\mathbb{E}|d_X(j, k)|^q < \infty \),
- A SINGLE PARAMETER \( H \)
- ADDITIVE STRUCTURE.

● MULTI-FRACTAL
  \[ \mathbb{E}|d_X(j, k)|^q = C_q(2^j)^{\zeta(q)} = C_q \exp(\zeta(q) \ln 2^j) \]
- POWER LAWS,
- \( \forall 2^j < L \), (FOR FINE SCALES ONLY, IN THE LIMIT \( 2^j \to 0 \),)
- \( \forall q \),
- A WHOLE COLLECTION OF SCALING PARAMETER \( \zeta(q) \)
- MULTIPLICATIVE STRUCTURE.

● BEYOND POWER LAWS : WARPED INF. DIV. CASCADES
  \[ \mathbb{E}|d_X(j, k)|^q = C_q(2^j)^{qH} = C_q \exp(qH \ln 2^j) \]
  \( \mathbb{E}|d_X(j, k)|^q = C_q(2^j)^{\zeta(q)} = C_q \exp(\zeta(q) \ln 2^j) \)
  \( \mathbb{E}|d_X(j, k)|^q = C_q \exp(\zeta(q) n(2^j)) \)

→ VISIT PIERRE CHAINAIS’S POSTER
CONCLUSIONS AND REFERENCES

ANALYSING SCALING IN DATA?
— THINK WAVELET
  — EFFICIENCY,
  — PRACTICAL AND CONCEPTUAL ADEQUATION AND SIMPLICITY,
  — ROBUSTNESS AGAINST NON STATIONARITIES,
  — EASY TO USE, LOW COST, REAL TIME ON LINE.

MODELLING SCALING IN DATA?
— THINK SELF SIMILARITY VERSUS MULTIPLICATIVE CASCADES,
— AND POSSIBLY ADD LONG MEMORY.
— ALSO SCALING MAY NOT BE POWER LAWS

REFERENCES AND RESOURCES, VISIT:
— perso.ens-lyon.fr/patrice.abry
— inrialpes.fr/is2/∼pgoncalv
— www.cubinlab.ee.mu.oz.au/∼darryl
— fraclab
— www.isima.fr/∼chainais