

# MULTISCALE TIME SERIES ANALYSIS: SCALING AND EXTREMES, STRUCTURE FUNCTIONS AND EMPIRICAL MODE DECOMPOSITION

Time series conference, 19-21  
September 2012



F. G. Schmitt<sup>1)</sup>, Y.X. Huang<sup>12</sup>

<sup>1)</sup> DR CNRS, director of Lab. Oceanology and Geosciences, Wimereux, France

<sup>2)</sup> Shanghai Institute of Applied Mathematics and Mechanics, Shanghai University, China

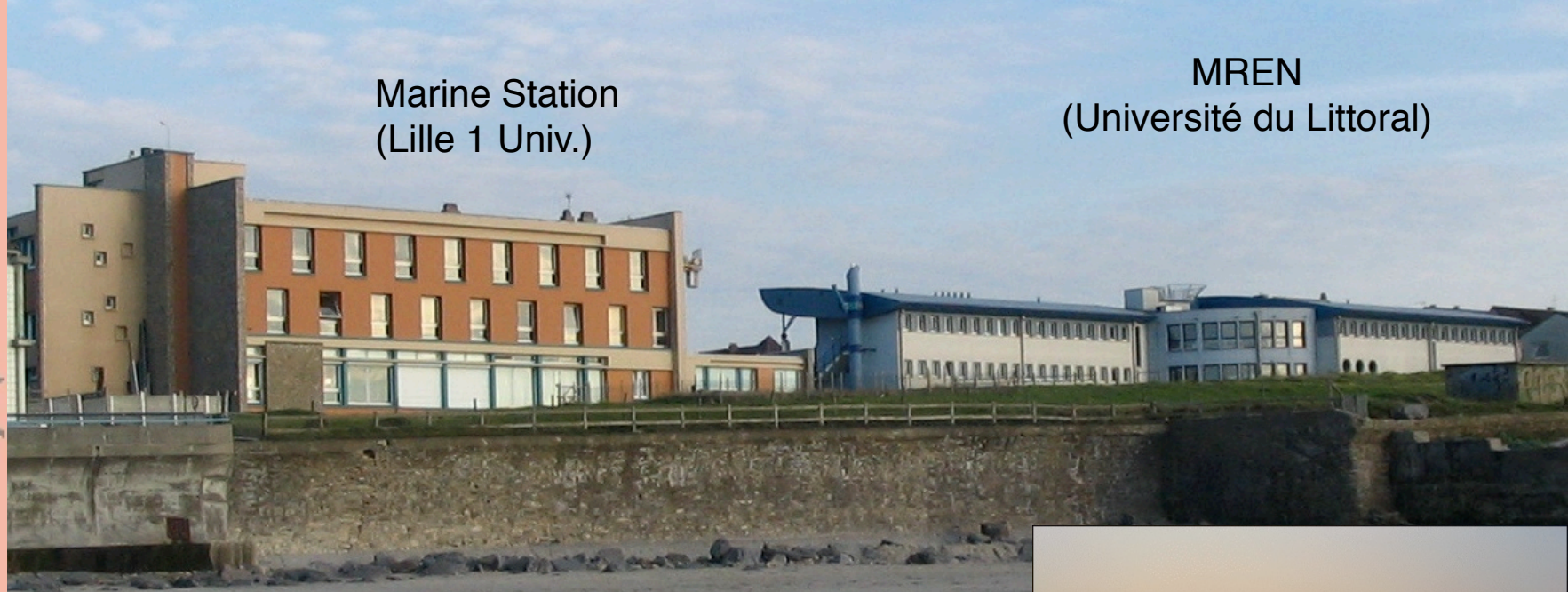


Shanghai University



Université  
Lille1  
Sciences et Technologies





# Laboratory of Oceanology and Geosciences - LOG Wimereux



**Old station (1874) but new CNRS laboratory created in 2008:**

- 35 researchers
- 30 PhD students
- 100 members on site





## Interdisciplinary oceanology studies: physics, biology, geomorphology

### **Activities as a marine station:**

- *Research*
- *Teaching*
- *Hosting (students, researchers)*
- *Observation*

### **Interdisciplinary research:**

- *observation/experimentation/modelization*
- *specialized in coastal research*
- *from bacteria to satellites*
- *from microscales to climate*



# Outline

- \* Motivations
- \* Structure functions as a classical tool in turbulence for scaling processes
- \* Arbitrary-Order Hilbert-Spectral-Analysis
- \* Examples of application

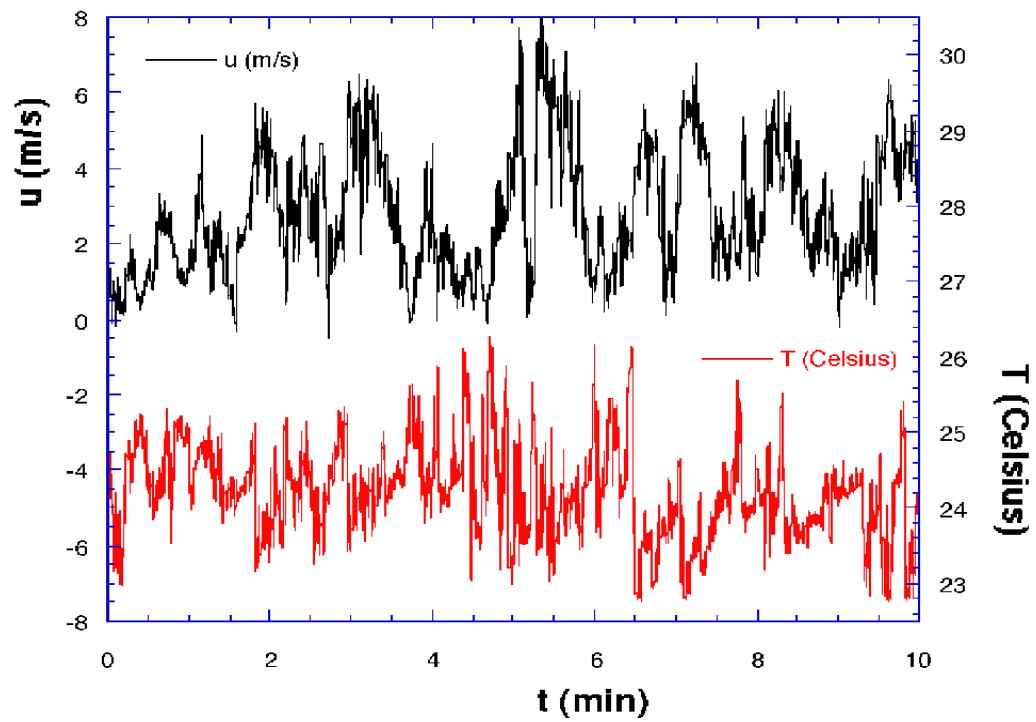


# Motivations (1)

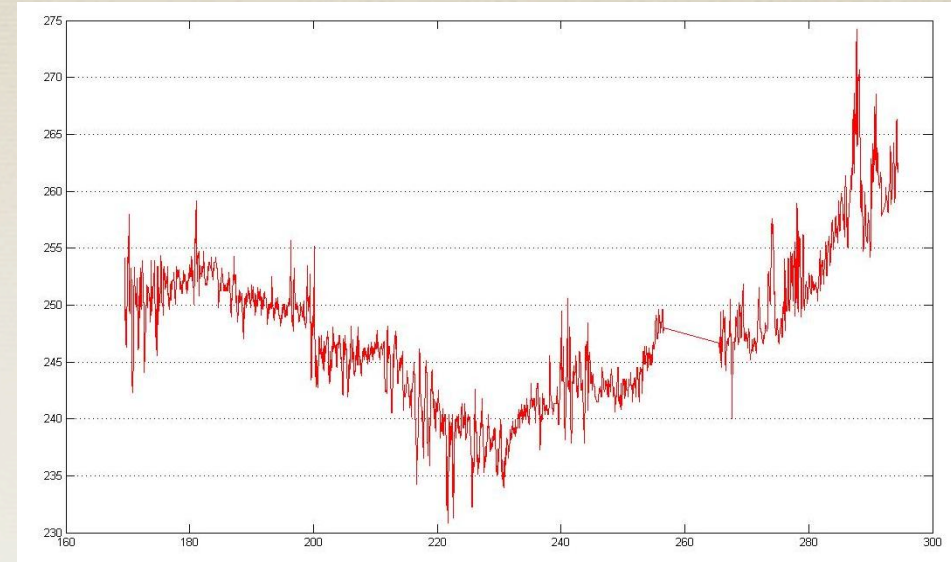
- In oceanography many studies are descriptive
  - Other studies rely on models (too often considered as virtual reality)
  - Other sciences have more rapid progress: comes from universality in their results
- > we should also search for universal relations in oceanography (and geosciences in general)



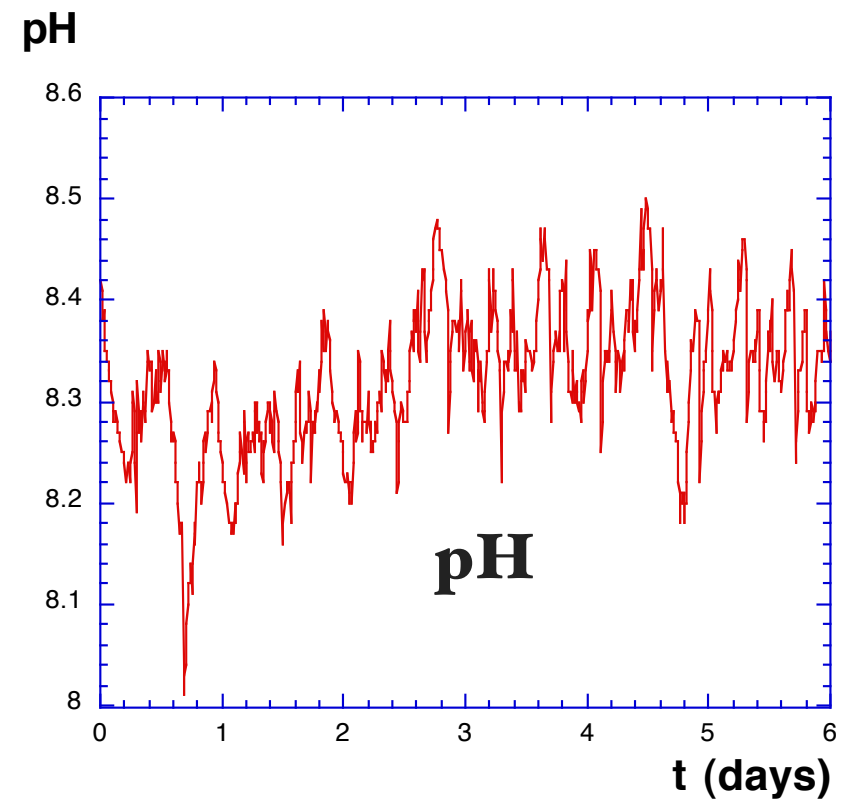
## Velocity



## Temperature



## Dissolved oxygen



Need a stochastic framework; turbulence influences



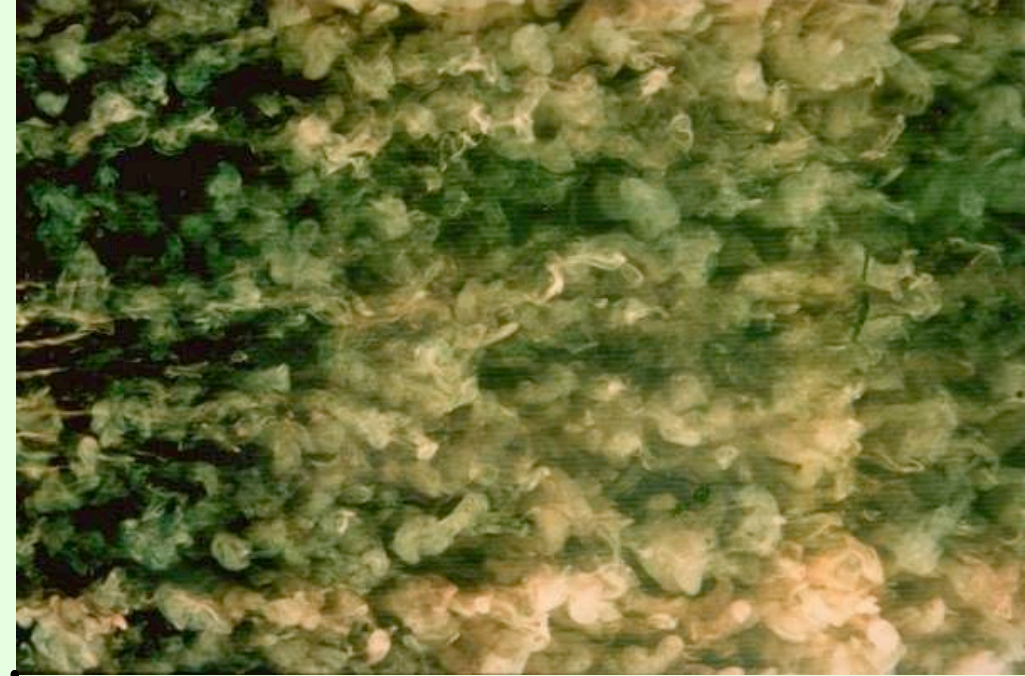
# Motivations (2)

- Search for stochastic universality
- Inspiration by turbulence theory
- Example: Kolmogorov 1941  $-5/3$  power law spectrum; other slopes found for other hypotheses and cases

—> provides a possible generic framework to studying variability at all scales in oceanography (and geosciences), search for stochastic universality



# Fully developed turbulence



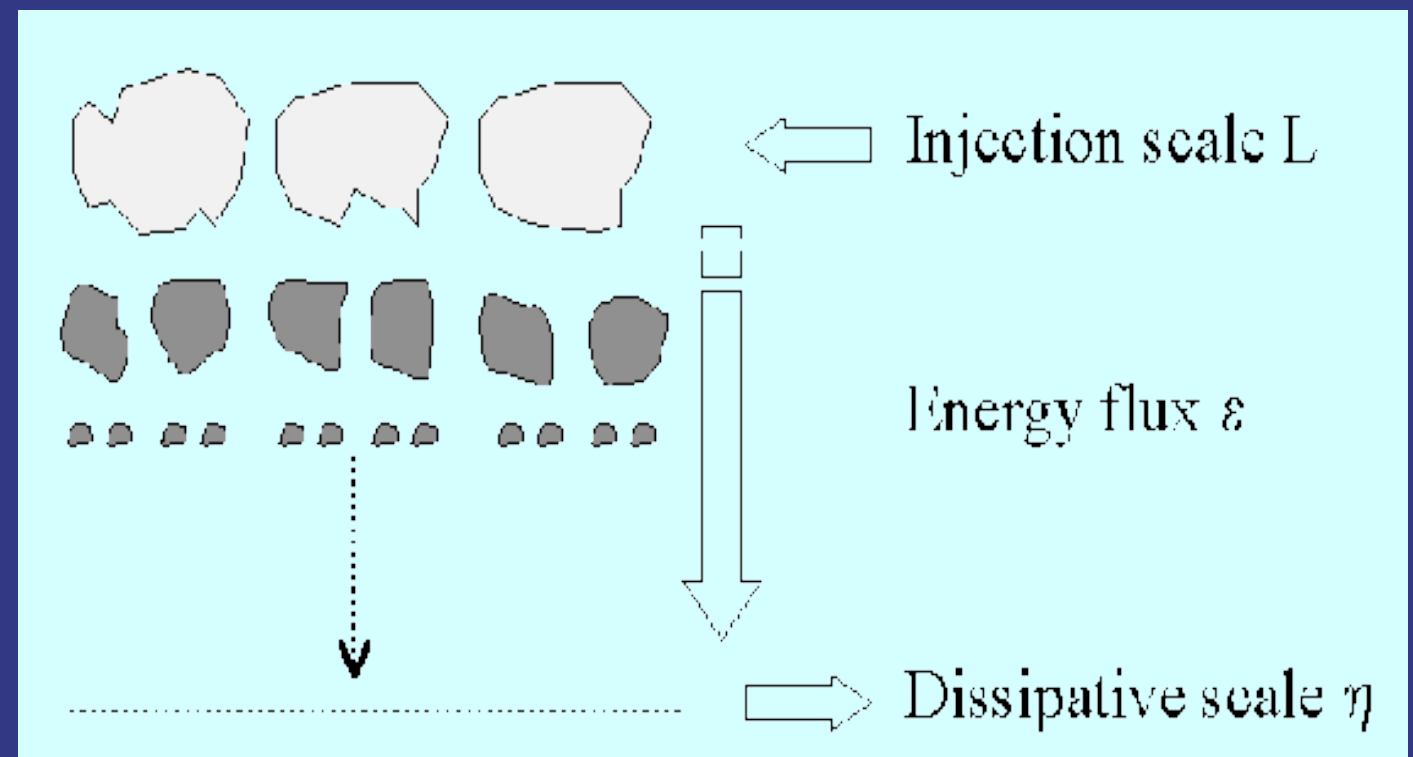
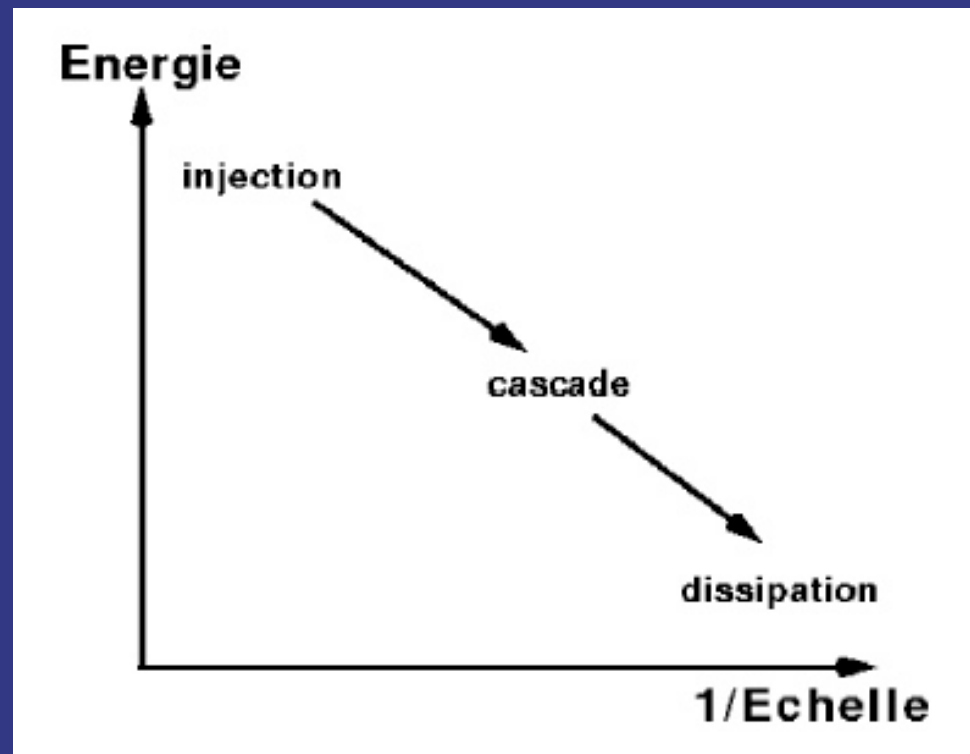
- **Characterized by:**
  - Large Reynolds number  $Re = UL / \nu$
  - High variability on a large range of scales
  - Dissipative system
  - Chaos, loss of predictability
- **One of the last domains of classical physics which is not solved**
- **Large number of applications**
- **Deterministic Navier Stokes equations but stochastic methods to address the problem**



# Experimental origin of multiplicative cascades

## Richardson and Kolmogorov: energy cascade

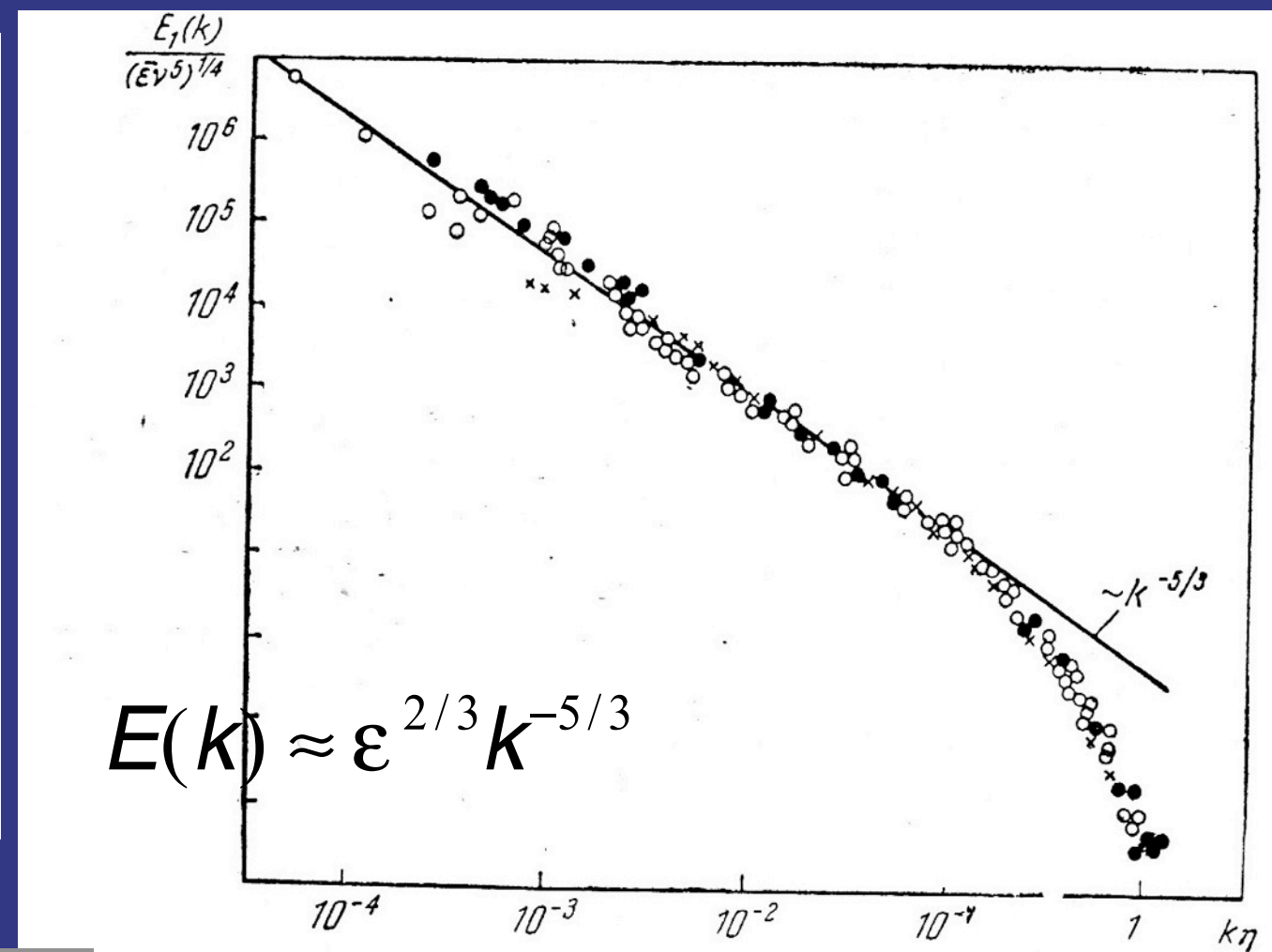
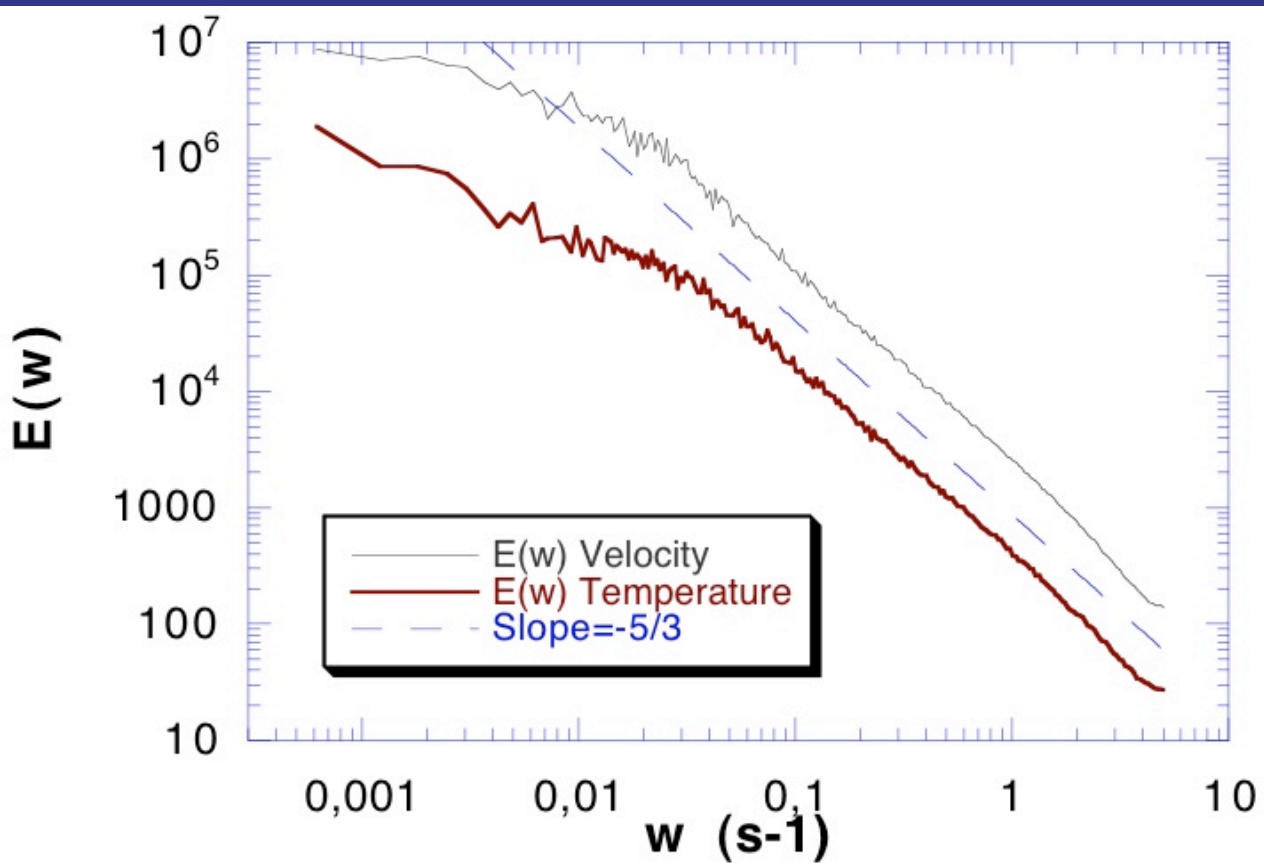
- Richardson (1922): energy cascade from large to small scales
- Kolmogorov (1941): dimensional analysis, leading to a scaling power spectra for velocity fluctuations in  $k^{-5/3}$





# Experimental origin of multiplicative cascades

Experimental validation of the -5/3 law of Kolmogorov

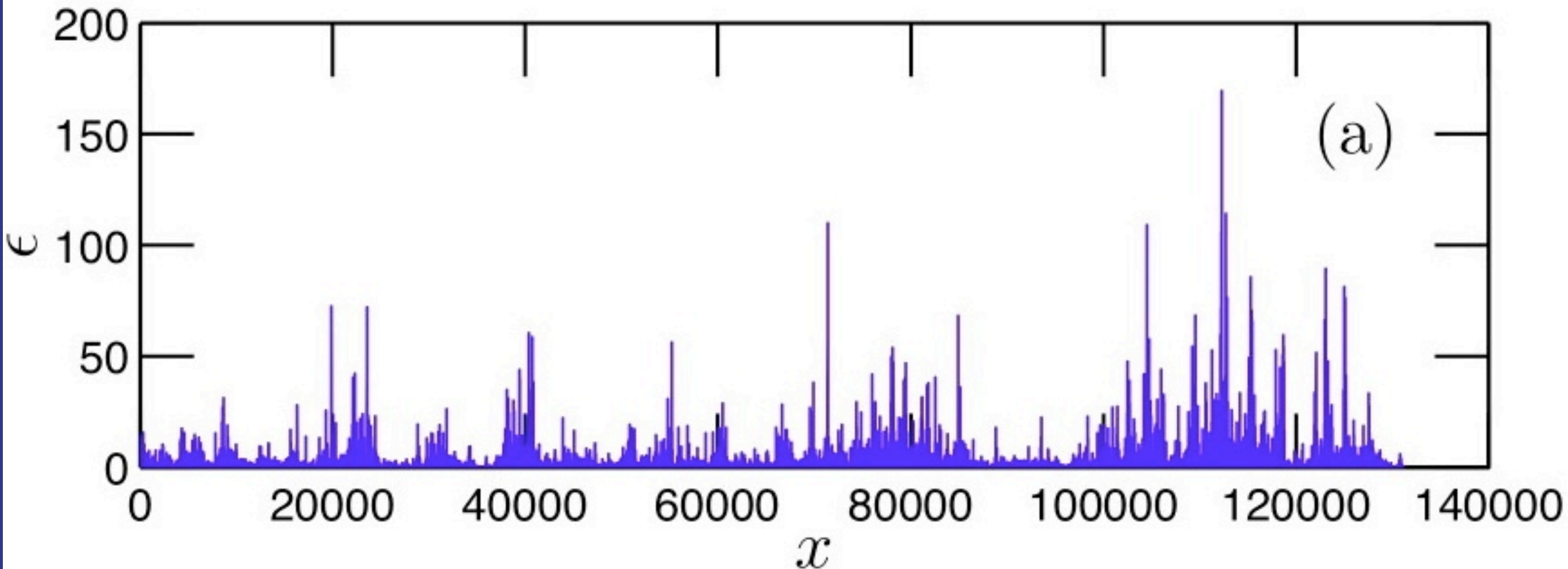


Checked in many situations since the 1960s

Fig. 76 Normalized longitudinal velocity spectrum. x—Sandborn and Marshall; ●—Grant, Stewart, and Moilliet; ○—Pond, Stewart, and Burling.



# Intermittency



1949-1962: discovery of intermittency in turbulence: fluctuations of the velocity difference are **bursty**

**Many pikes ; large intensities ; long-range correlated**



# Framework for intermittency studies

Scaling nonstationary process with stationary increments:  
Statistical moments of structure functions

$$\Delta X_T = X(t+T) - X(t) \quad \langle (\Delta X_T)^q \rangle \approx T^{\zeta(q)}$$

$\zeta(q)$  : moment function

(also, in probability theory, a second  
characteristic function)

$$\zeta(q) = q/2 \quad \text{Brownian motion}$$

$$\zeta(q) = qH; \quad 0 < H < 1 \quad \text{fractional Brownian motion}$$

Statistical scale invariance of velocity fluctuations

Link with power spectrum:  $\beta = 1 + \zeta(2) \quad E_x(f) \approx f^{-\beta}$



# Framework for intermittency studies

## Recipe for multifractal analysis using structure functions

1. Estimate  $\Delta X_T = X(t+T) - X(t)$  versus  $T$  for many different values of  $T$
2. Log-log plot and estimation of the value of  $\zeta(q)$  as the slope of the straight line obtained
3. Perform this for many different values of  $q$  and obtain a function  $\zeta(q)$

If  $\zeta(q)$  is linear: monofractal;

If  $\zeta(q)$  is nonlinear: multifractal. its parameters can be estimated

First Parameter:  $H = \zeta(1)$

Other parameters depend on the model chosen:  
Lognormal, log-Lévy, log-Poisson (She-Lévêque)...

$$\zeta(q) = Aq - Bq^2$$

$$\zeta(q) = Aq - Bq^\alpha$$

$0 \leq \alpha \leq 2$

$$\zeta(q) = Aq - B\beta^q$$

$0 \leq \beta \leq 1$



# A general feature of high frequency coastal Environmental Data

- \* Stochastic variability on a large range of scales
- \* Turbulent-like small-scale stochastic fluctuations
- \* Large-scale deterministic period (tidal and daily cycles...)
- \* Traditional methodologies **such as structure functions** fail to detect the correct scaling property because of the strong forcing



# Arbitrary-order Hilbert Spectral Analysis

A powerful multiscale analysis  
method



# Arbitrary-Order HSA

Empirical Mode Decomposition – EMD

+

Hilbert Spectral Analysis – HSA

↓

Hilbert-Huang Transform – HHT

⇓

Arbitrary Order Hilbert Spectral Analysis



# Empirical Mode Decomposition (EMD)

- A new analysis technique proposed by Huang et al. (1998, 1999)
- Decomposes a signal into a sum of modes, without leaving the time domain
- The mode functions form a basis, nearly orthogonal
- Each mode is localized in frequency space: act as a filter bank
- Can be used for detrending or denoising time series
- Can be applied to nonlinear and non-stationary data, even with relatively small number of datapoints
- Complementary to Fourier or Wavelet analysis



This approach was first proposed by Norden Huang (NASA) in 1998 and 1999 in oceanography to analyze water waves

PROCEEDINGS OF THE ROYAL SOCIETY OF ENGINEERING SCIENCES

Home | Current Issue | Past Issues | Submit | Subscribe | Alerts

Access licensed by Universite des Sciences et Technologies de Lille

## The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis

Norden E. Huang<sup>1</sup>, Zheng Shen<sup>2</sup>, Steven R. Long<sup>3</sup>, Manli C. Wu<sup>4</sup>, Hsing H. Shih<sup>5</sup>, Quanan Zheng<sup>6</sup>, Nai-Chyuan Yen<sup>7</sup>, Chi Chao Tung<sup>8</sup> and Henry H. Liu<sup>9</sup>

+ Author Affiliations

### Abstract

A new method for analysing nonlinear and non-stationary data has been developed. The key part of the method is the 'empirical mode decomposition' method with which any complicated data set can be decomposed into a finite and often small number of 'intrinsic mode functions' that admit well-behaved Hilbert transforms. This decomposition method is adaptive, and, therefore, highly efficient. Since the decomposition is based on the local characteristic time scale of the data, it is applicable to nonlinear and non-stationary processes. With the Hilbert transform, the 'intrinsic mode functions' yield instantaneous frequencies as functions of time that give sharp identifications of imbedded structures. The final presentation of the results is an energy-frequency-time distribution, designated as the Hilbert spectrum. In this method, the main conceptual innovations are the introduction of 'intrinsic mode functions' based on local properties of the signal, which make the instantaneous frequency meaningful; and the introduction of the instantaneous frequencies for complicated data sets, which eliminate the need for spurious harmonics to represent nonlinear and non-stationary signals. Examples from the numerical results of the classical nonlinear equation systems and data representing natural phenomena are given to demonstrate the power of this new method. Classical nonlinear system data are especially interesting, for they serve to illustrate the roles played by the nonlinear and non-stationary effects in the energy-frequency-time distribution.

« Previous | Next Article »  
Table of Contents

### This Article

doi:  
10.1098/rspa.1998.0193  
Proc. R. Soc. Lond. A 8  
March 1998 vol. 454 no.  
1971 903-995

- » Abstract Free
- Full Text (PDF) Free to you

### Services

- Email this article to a friend
- Alert me when this article is cited
- Alert me if a correction is posted
- Article Usage Statistics
- Similar articles in this journal
- Similar articles in Web of Science
- Download to citation manager
- Permissions

- + Citing Articles
- + Google Scholar
- + PubMed
- + Social Bookmarking

- 2500 citations for this 1998 paper
- Hundreds of papers applying the new method to various fields
- ocean, atmosphere, signal processing, mechanical engineering, climate studies, earthquakes, biomedical studies...
- but still no exact mathematical results



# EMD Algorithm

- (1) Identify all local maximum (resp. minimum) extrema of  $x(t)$
- (2) Interpolate maximum (resp. minimum) by cubic spline to form upper (reps. lower) envelop  
 $e_{\max}$        $e_{\min}$
- (3) Compute the average  $m(t) = (e_{\max} + e_{\min}) / 2$
- (4) Extract the detail  $d(t) = x(t) - m(t)$
- (5) Iterate on the residual  $m(t)$

$$x(t) = \sum_{i=1}^N C_i(t) + r_n$$

## Advantages of EMD

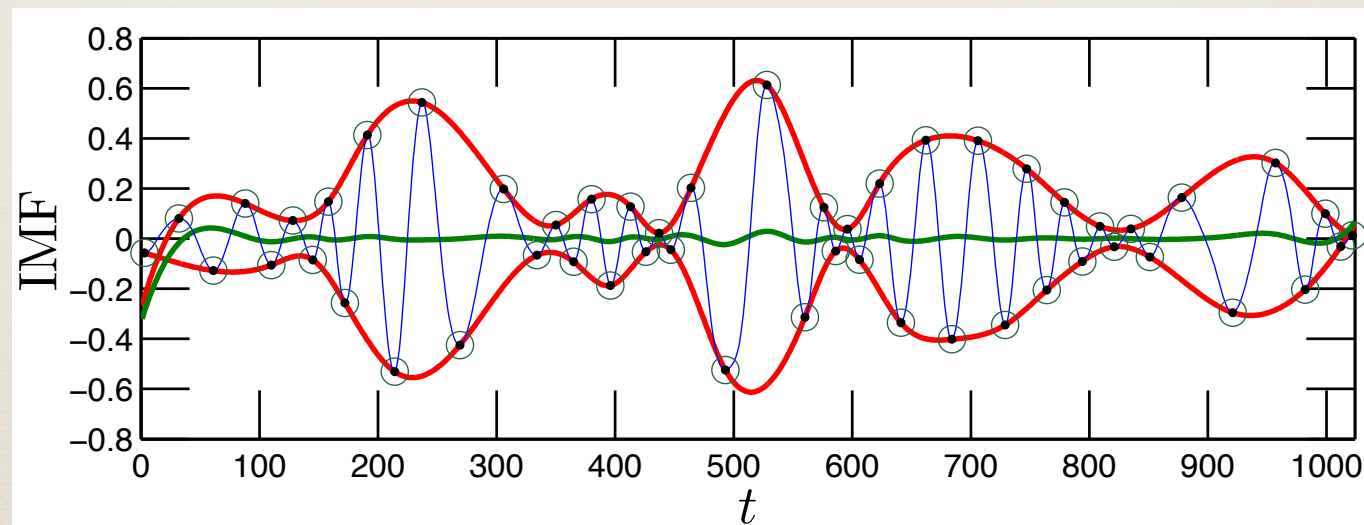
- Locality: characteristic scale is defined as the distance between successive local maxima (or minima)
- Complete self-adpativeness: no basis (function) assumption a priori



# Empirical Mode Decomposition

Intrinsic Mode Function (IMF)  $\rightarrow$  mono-component signal

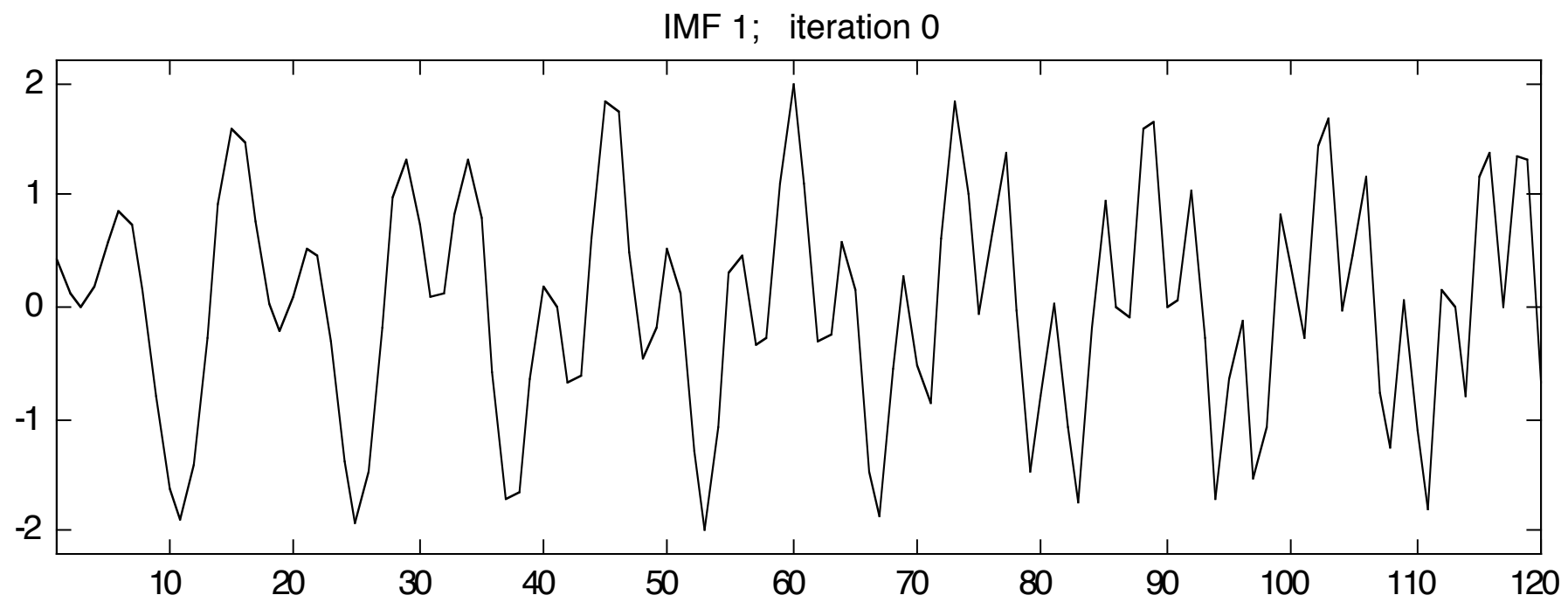
- the difference between the number of local extrema and the number of zero-crossings must be zero or one;
- the running mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.



A typical IMF mode from EMD decomposition

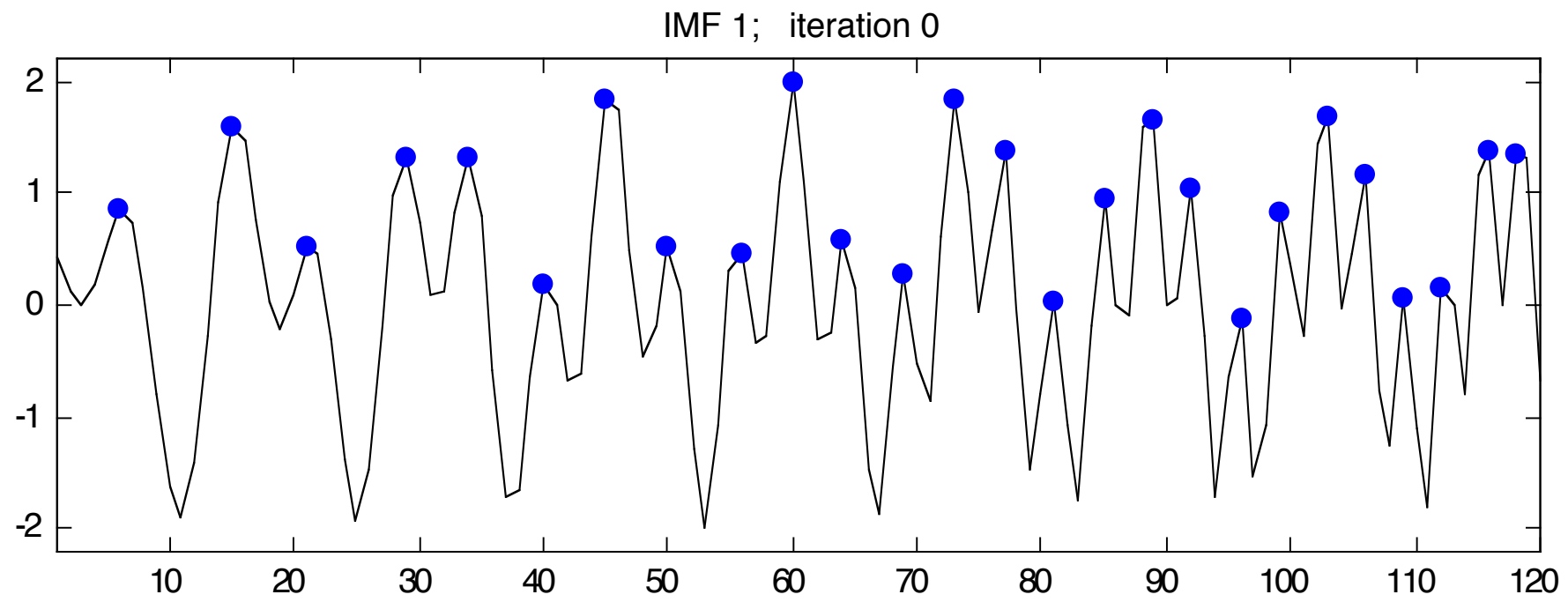


**Example** (taken from Flandrin, <http://perso.ens-lyon.fr/patrick.flandrin/>)



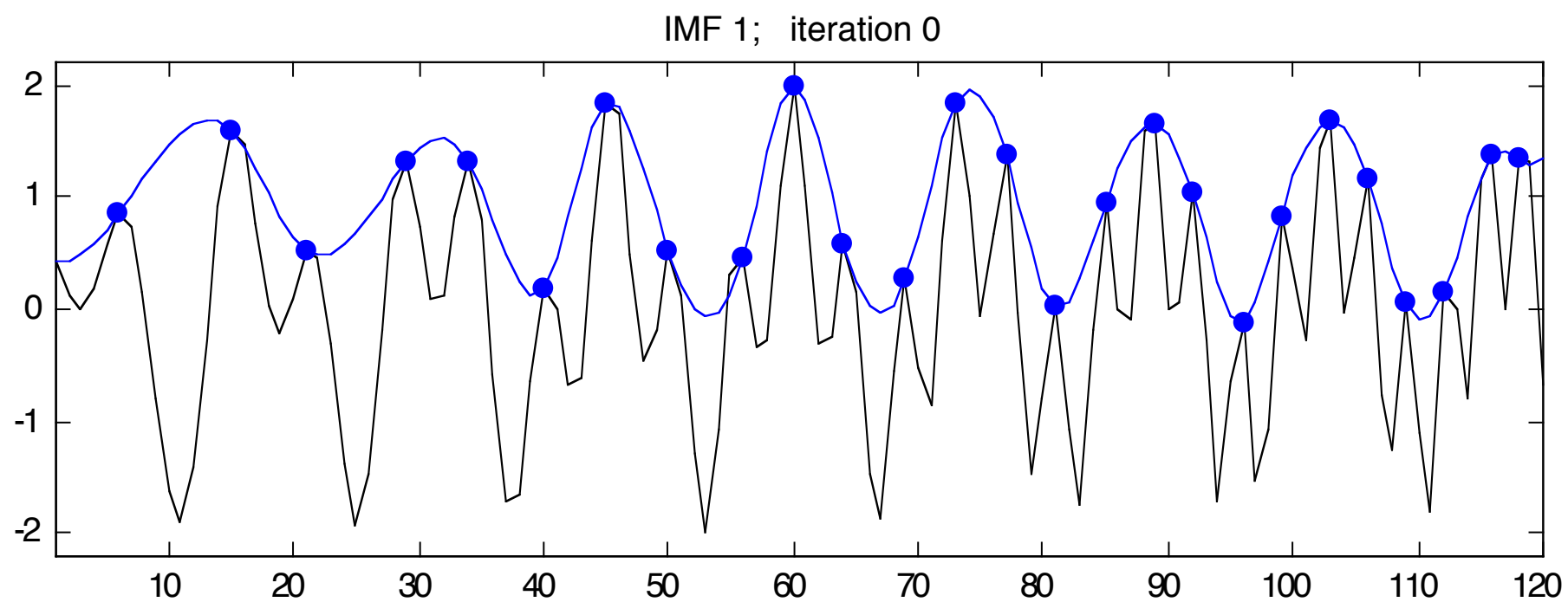


**Example** (taken from Flandrin, <http://perso.ens-lyon.fr/patrick.flandrin/>)

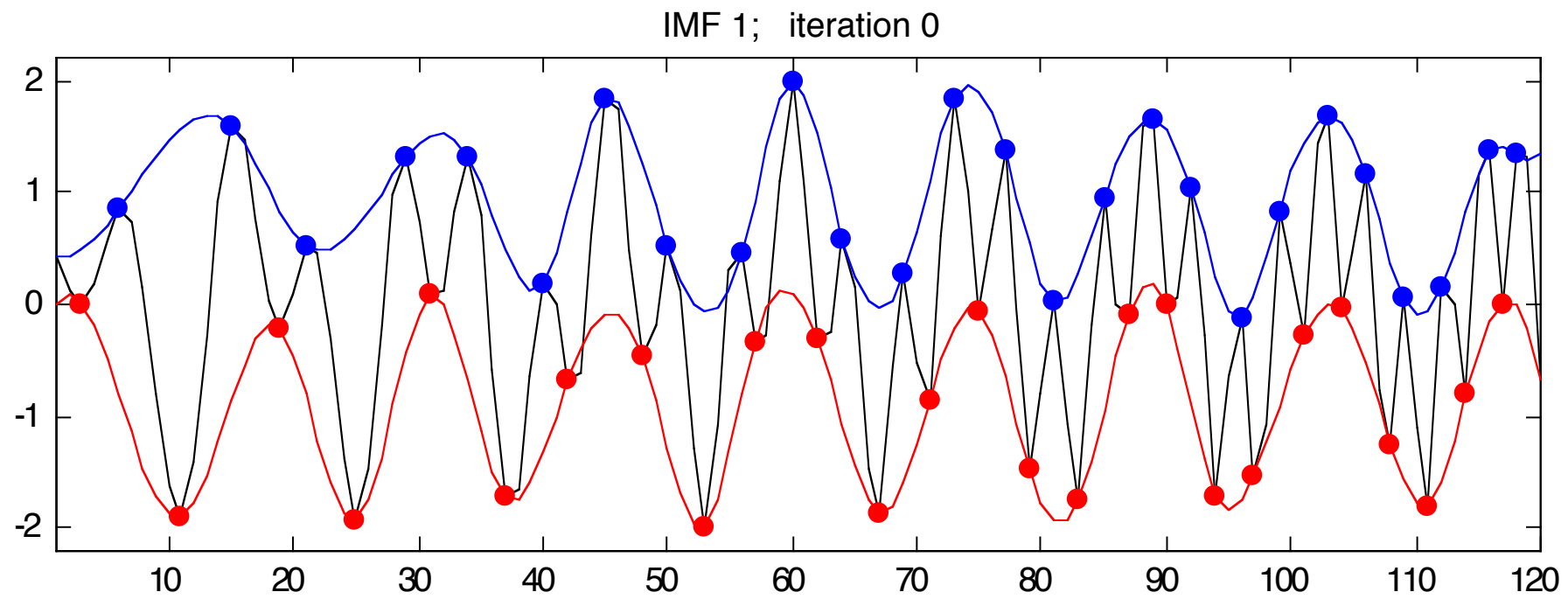




**Example** (taken from Flandrin, <http://perso.ens-lyon.fr/patrick.flandrin/>)

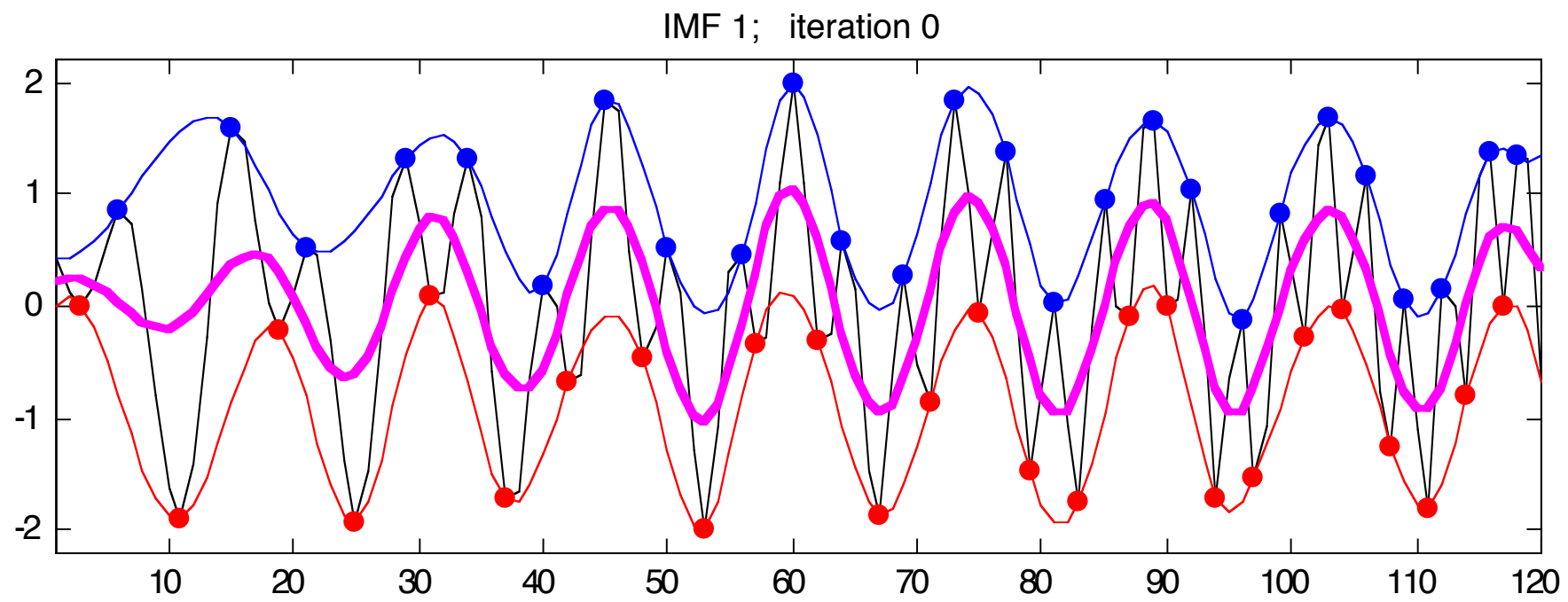


**Example** (taken from Flandrin, <http://perso.ens-lyon.fr/patrick.flandrin/>)

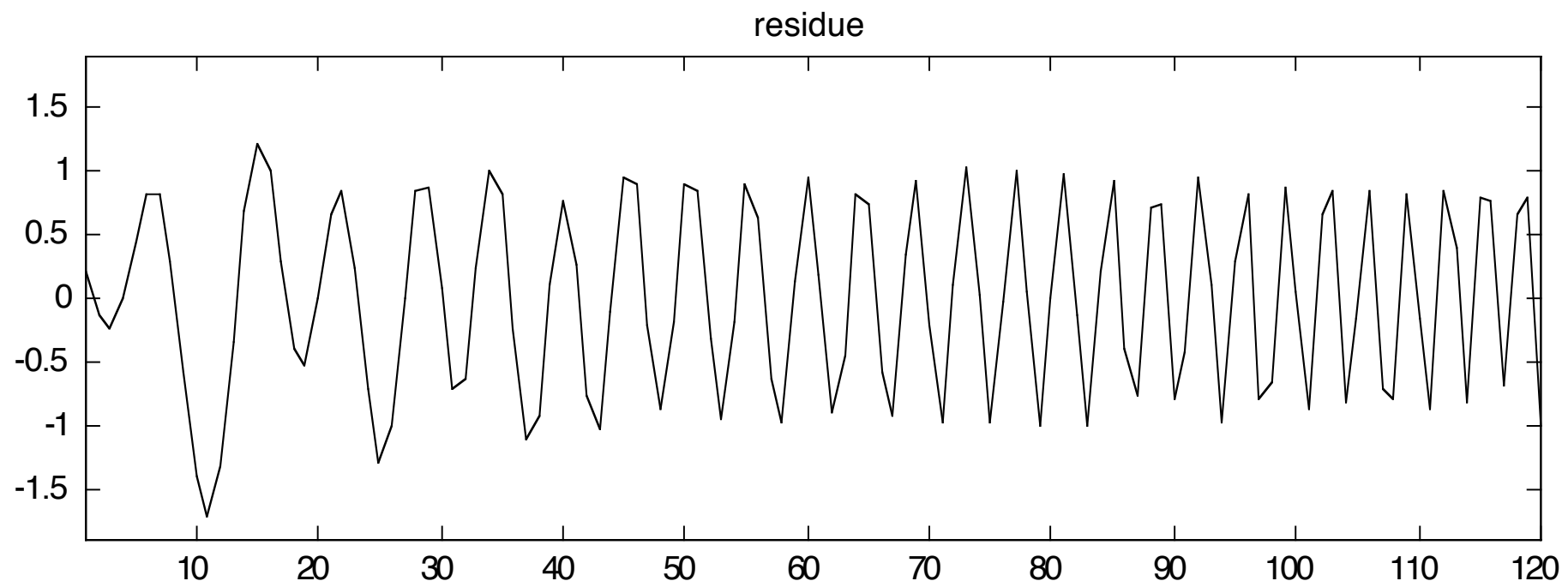
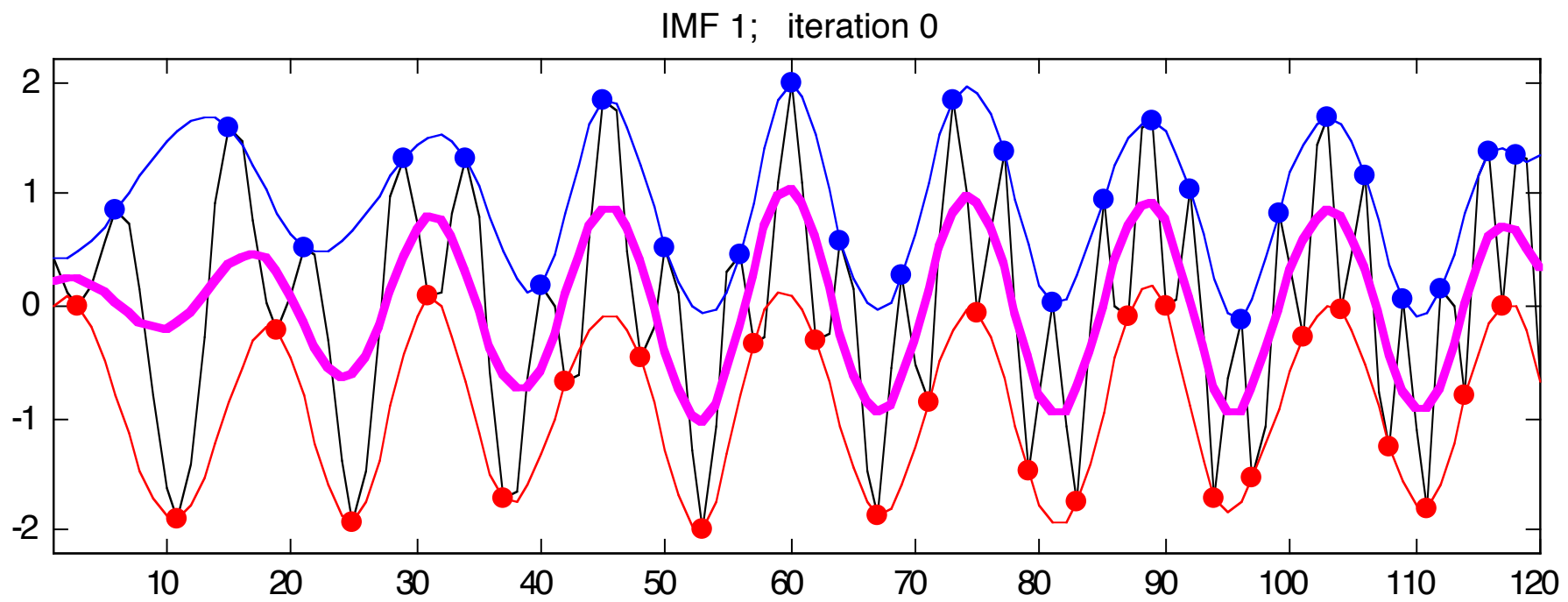




**Example** (taken from Flandrin, <http://perso.ens-lyon.fr/patrick.flandrin/>)

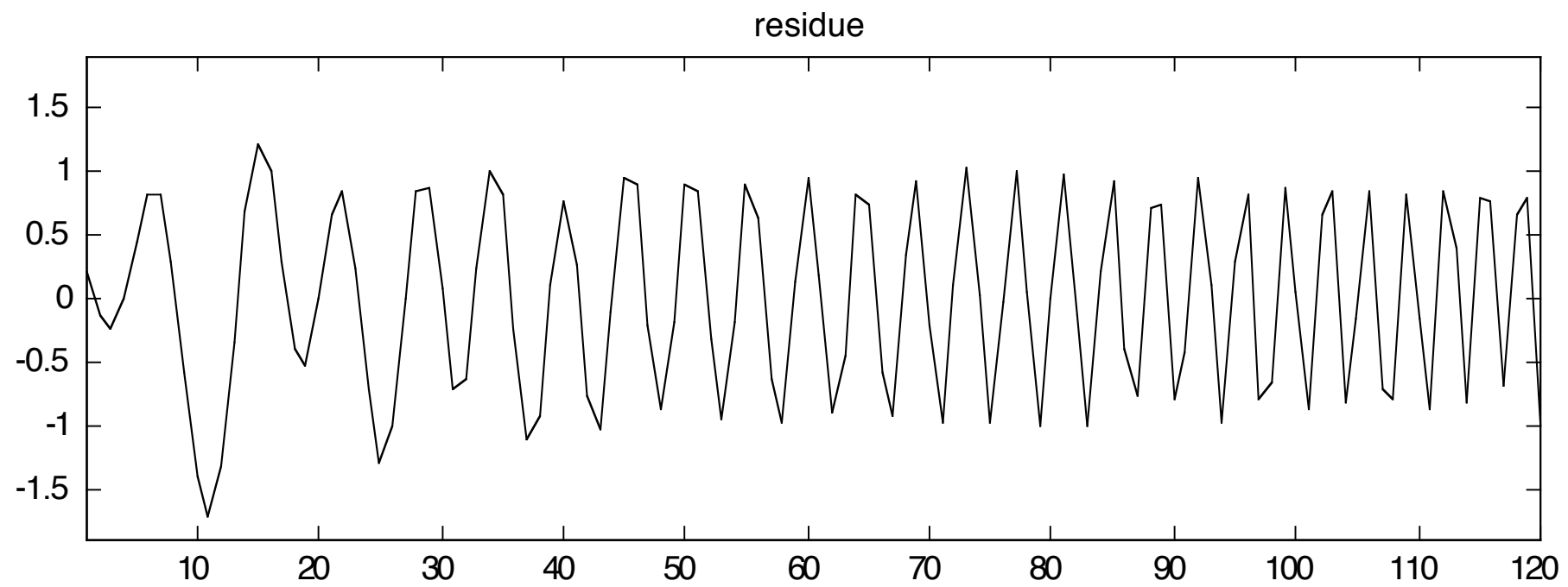
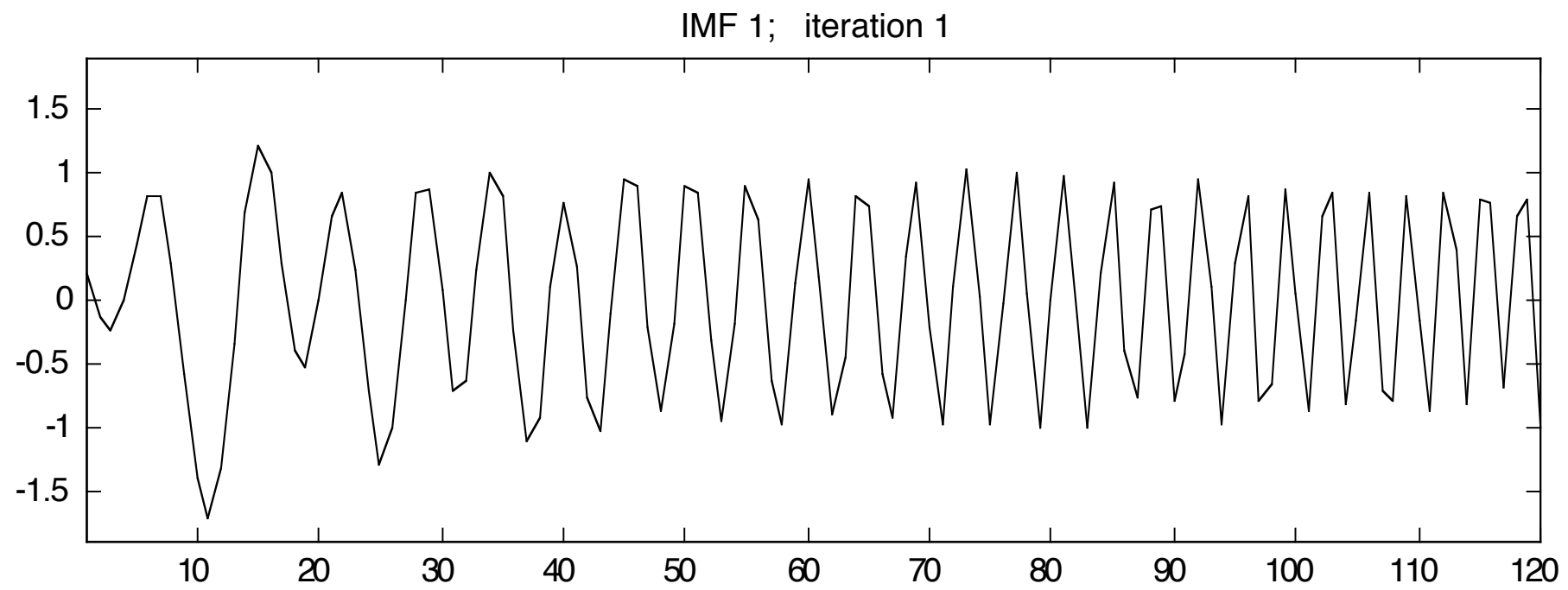


**Example** (taken from Flandrin, <http://perso.ens-lyon.fr/patrick.flandrin/>)

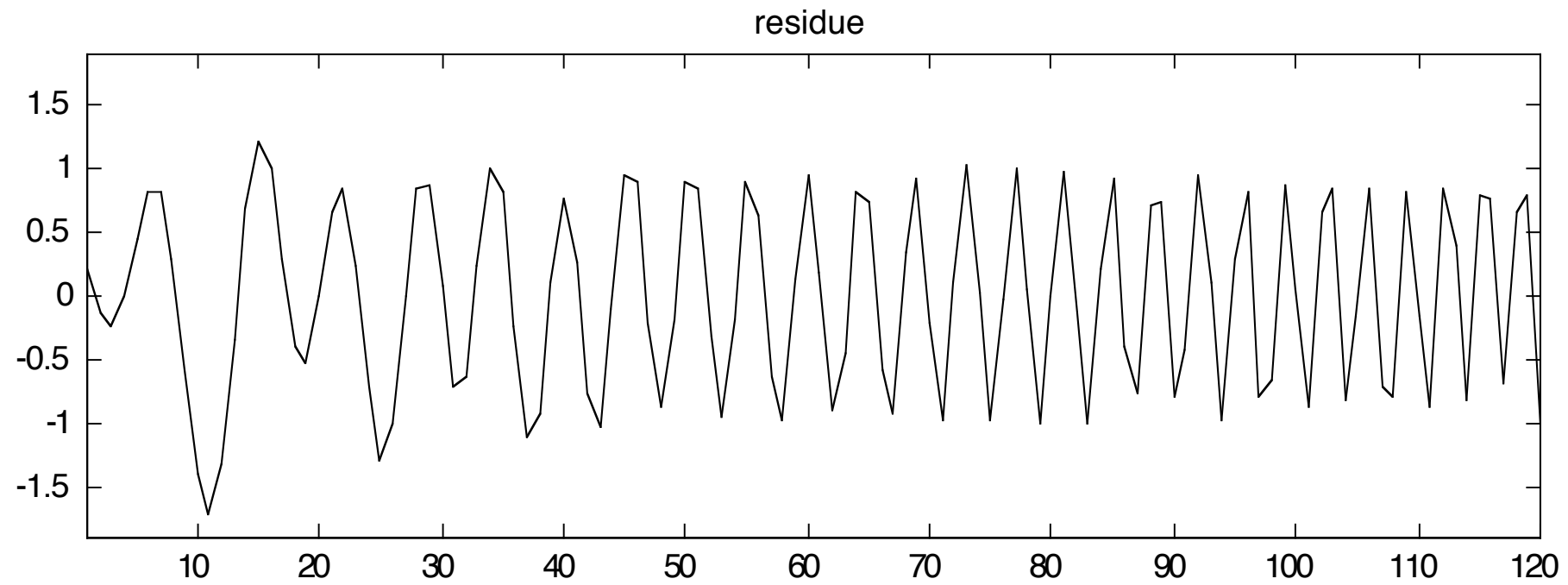
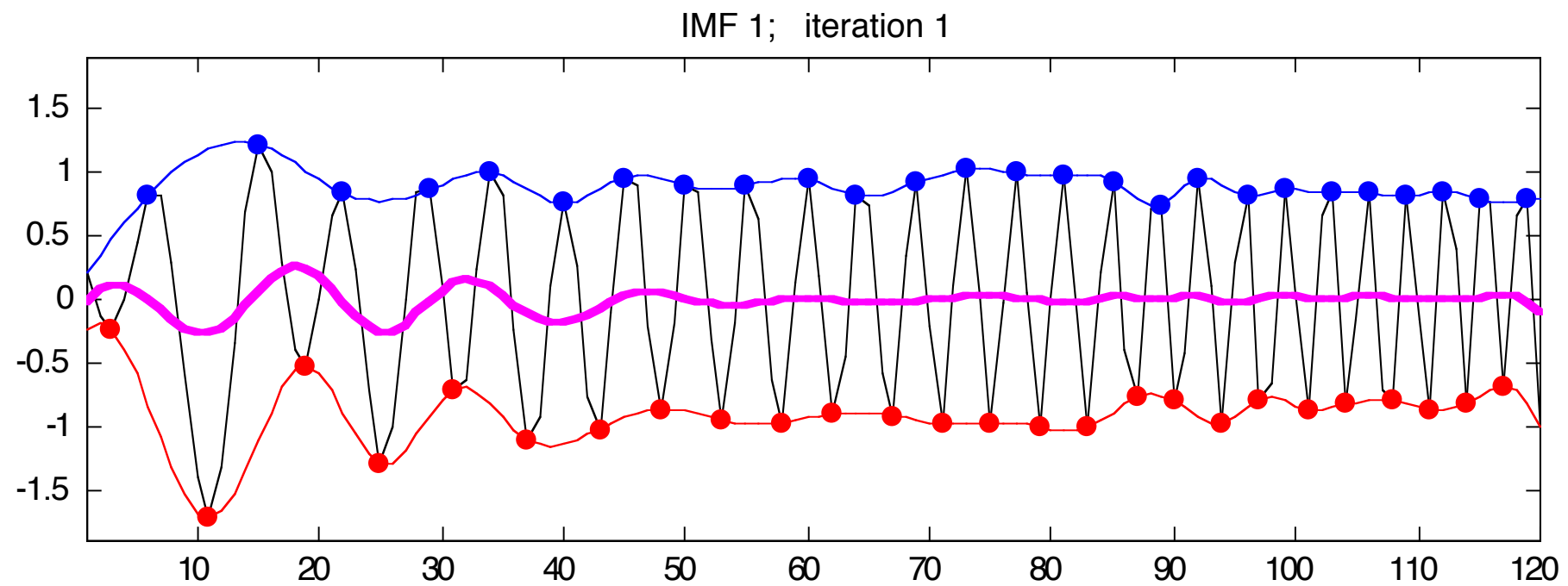




**Example** (taken from Flandrin, <http://perso.ens-lyon.fr/patrick.flandrin/>)

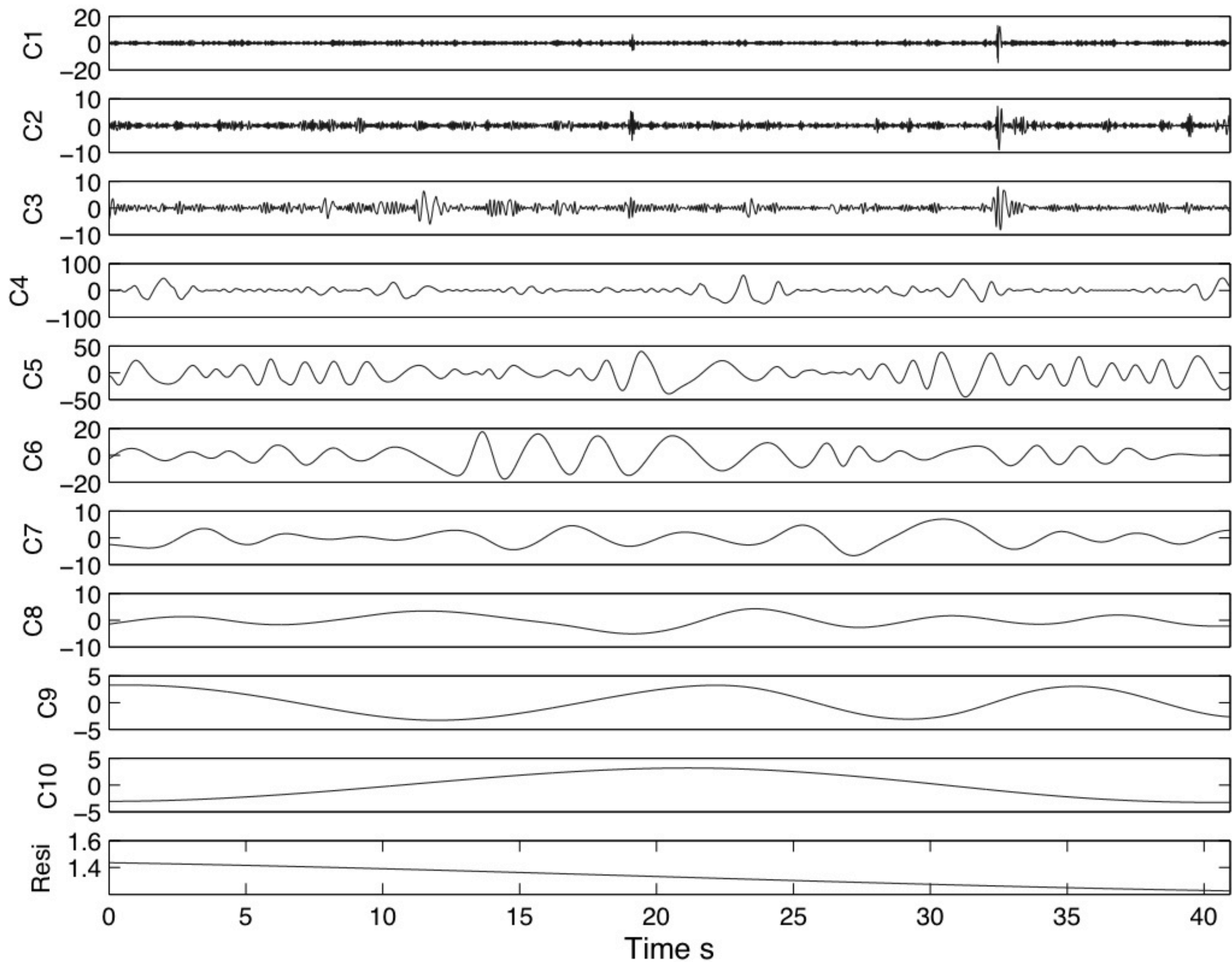


**Example** (taken from Flandrin, <http://perso.ens-lyon.fr/patrick.flandrin/>)





# Decomposition



# Hilbert Huang Algorithm

## Hilbert Spectral Analysis

After decomposition, the original signal is rewritten as

$$x(t) = \sum_{i=1}^n C_i + r_n$$

Hilbert transform :

$$C_i^H(t) = \frac{1}{\pi} P \int_{-\infty}^{+\infty} \frac{C_i(t')}{t - t'} dt'$$

where  $P$  means the principle value.



# Hilbert Huang Algorithm

## Analytical Signal

Then the so-called *Analytical Signal* is then written as

$$C_i^A(t) = C_i(t) + jC_i^H(t) = \mathcal{A}_i(t)e^{j\theta_i(t)}$$

where

$$\begin{cases} \mathcal{A}_i(t) = \sqrt{C_i^2 + C_i^{H2}} \\ \theta_i(t) = \tan^{-1}\left(\frac{C_i^H}{C_i}\right) \end{cases}$$

Therefore we can define the *Instantaneous Frequency (IF)* as

$$\omega_i = \frac{d\theta_i(t)}{dt}$$

# Hilbert Huang Algorithm

## Hilbert Spectral Analysis

The original signal (exclude the residual)

$$x(t) = RP \sum_{i=1}^N \mathcal{A}_i(t) e^{j\theta_i(t)} = RP \sum_{i=1}^N \mathcal{A}_i(t) e^{j \int \omega_i(t)}$$

where *RP* means real part.

## Compared with Fourier transform

Amplitude Modulation ( $\mathcal{A}(t)$ ) + Frequency Modulation ( $\omega(t)$ )



# Hilbert Huang Algorithm

## Hilbert-Huang Transform

We can extract the joint probability density function  $\mathcal{P}(\omega, \mathcal{A})$ , where  $\omega$  is the instantaneous frequency and  $\mathcal{A}$  modulation. Then the Hilbert marginal spectrum is redefined as

$$h(\omega) = \int_0^{+\infty} \mathcal{P}(\omega, \mathcal{A}) \mathcal{A}^2 d\mathcal{A}$$

Above procedure is the classical Hilbert Spectral Analysis (HSA).

EMD + HSA  $\rightarrow$  Hilbert-Huang Transform (HHT)

Our contribution:

**Arbitray Order Hilbert Spectral Analysis**



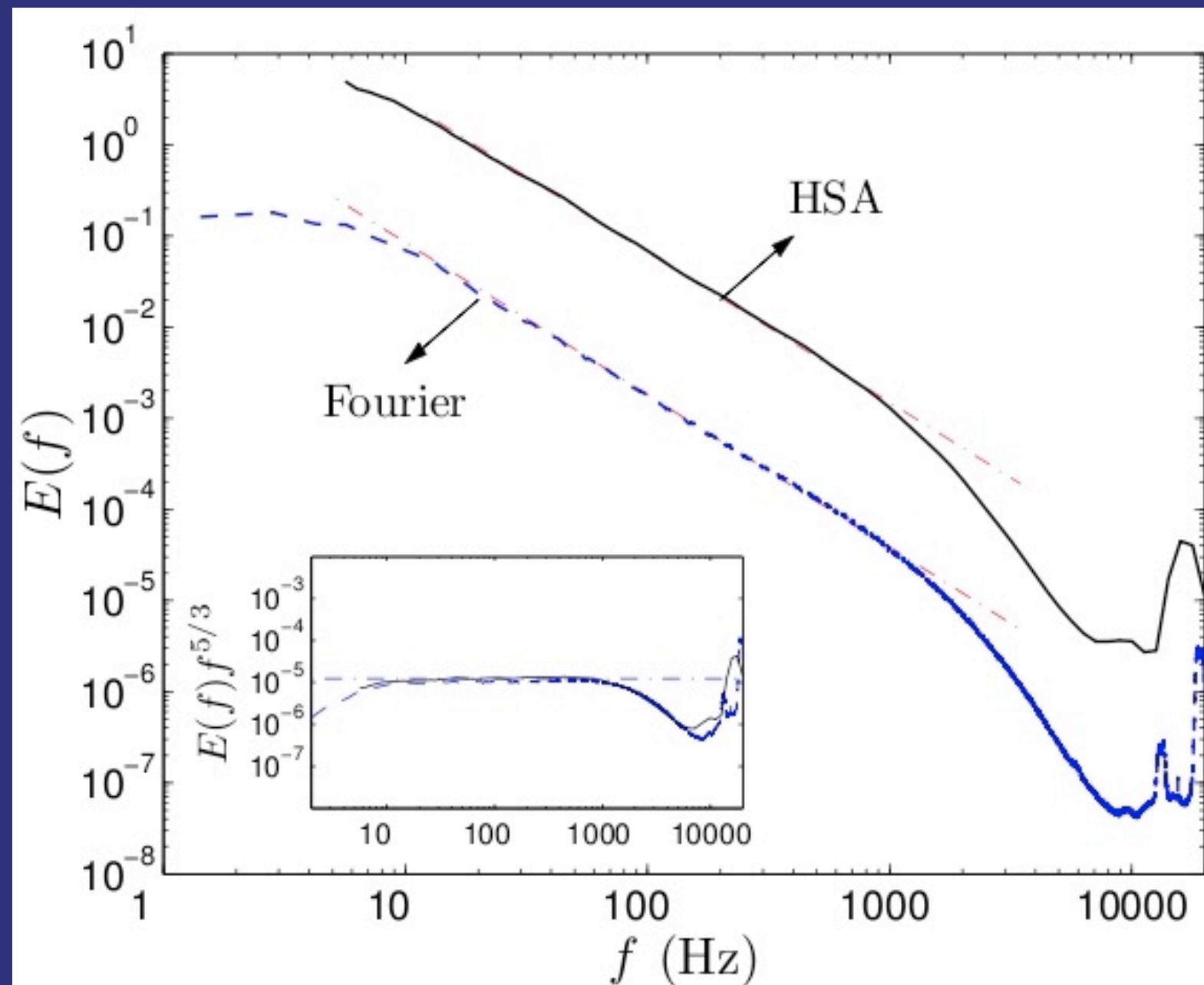
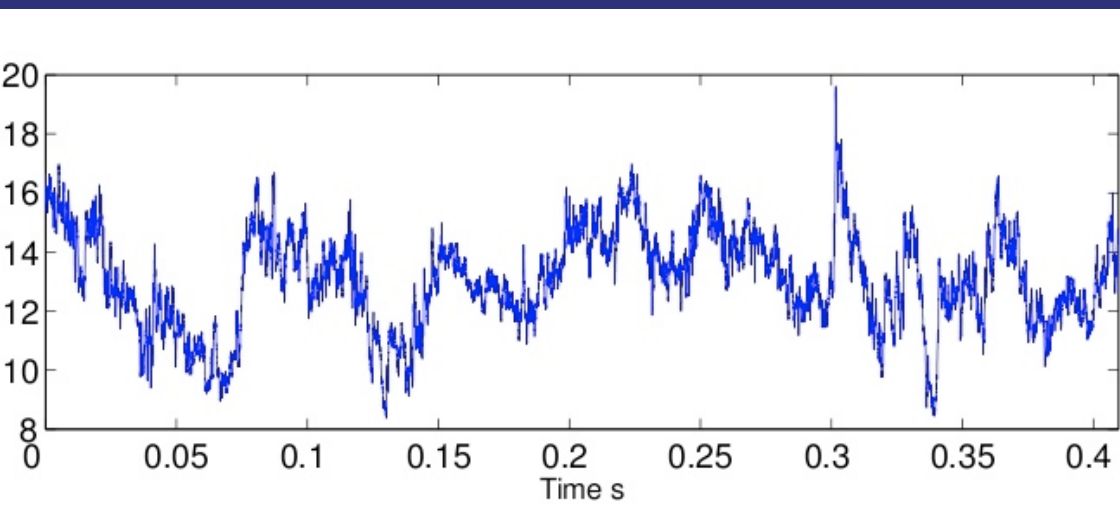
# Marginal energy spectrum in Hilbert frame

Local extraction of amplitude and frequency information

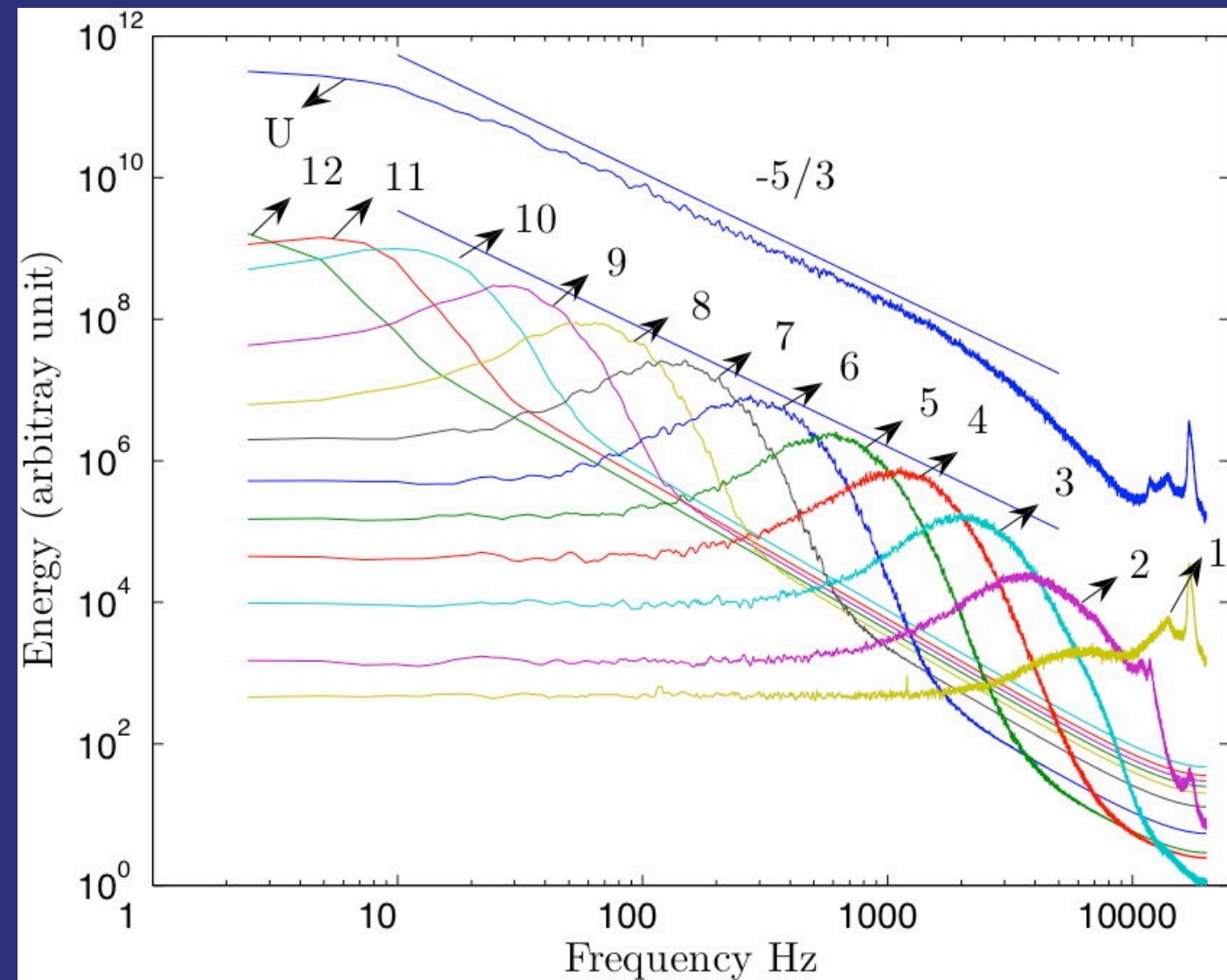
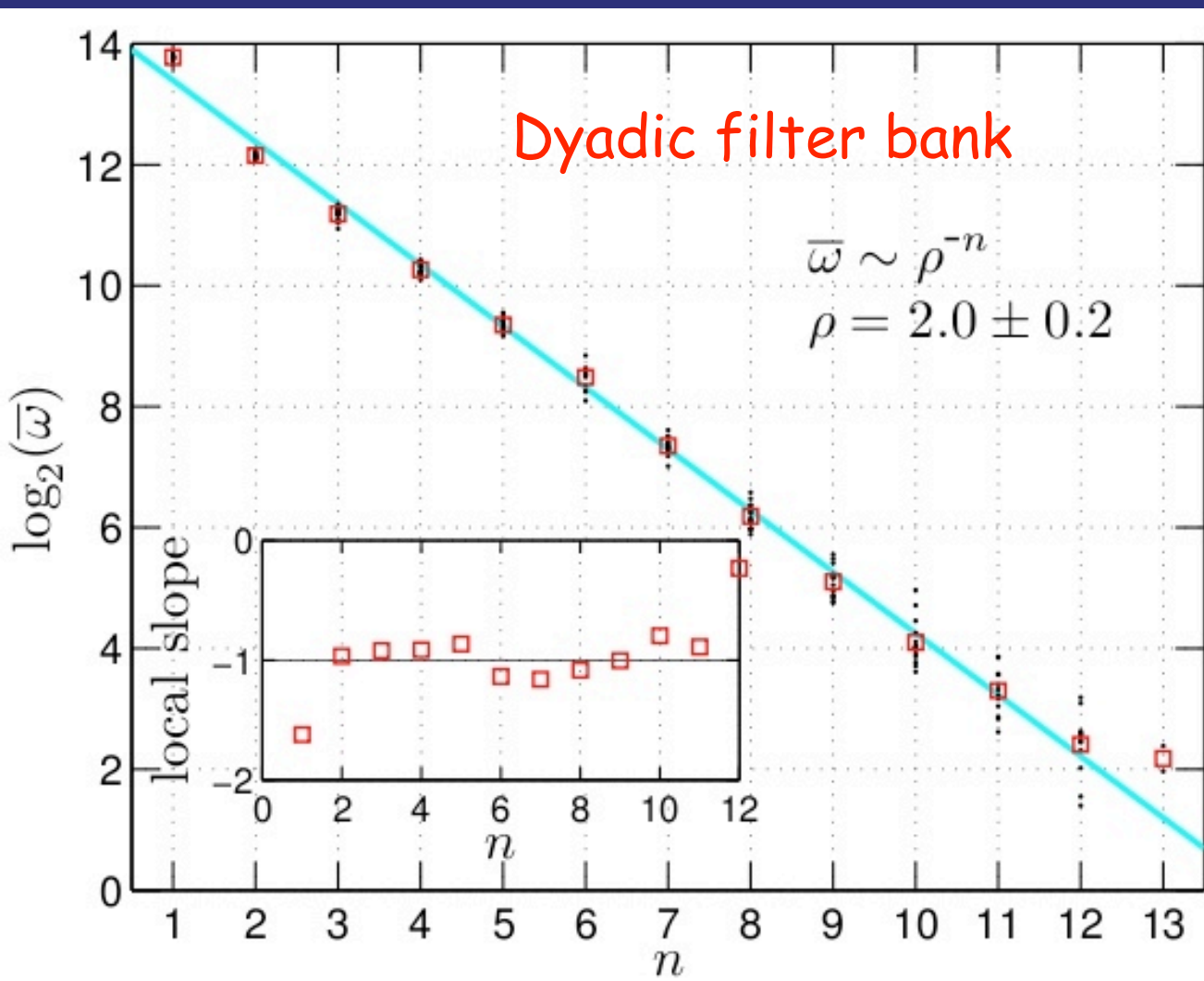
$$\{A(t), \omega(t)\}$$

$$h(\omega) = \int_0^{\infty} H(\omega, t) dt$$

Turbulence 5/3 spectrum:



# EMD as a filter bank for turbulence data



Spectra of successive modes: narrow banded



# Generalisation: arbitrary order Hilbert spectral analysis

Local extraction of amplitude and frequency information

$$\{A(t), \omega(t)\}$$

$$h(\omega) = \int_0^{\infty} H(\omega, t) dt = \int_0^{\infty} p(\omega, A) A^2 dA$$

Generalized to moment of order  $q > 0$ :

$$L_q(\omega) = \int_0^{\infty} p(\omega, A) A^q dA$$

For a scaling process one expect using dimensional scaling arguments:

$$L_q(\omega) \approx \omega^{-\xi(q)}$$

$$\xi(q) = 1 + \zeta(q)$$

# Expressions for scaling processes

For a scaling process one expect using dimensional scaling arguments:

$$L_q(\omega) \approx \omega^{-\xi(q)}$$

$$\xi(q) = 1 + \zeta(q)$$

For a monofractal process

$$\xi(q) = 1 + qH$$

For a scaling lognormal process

$$\xi(q) = 1 + \frac{q}{2} - \frac{\mu}{2} \left( \frac{q^2}{4} - \frac{q}{2} \right)$$

For general multifractal processes, nonlinear function  $\xi(q)$



# Verification for fractional Brownian motion

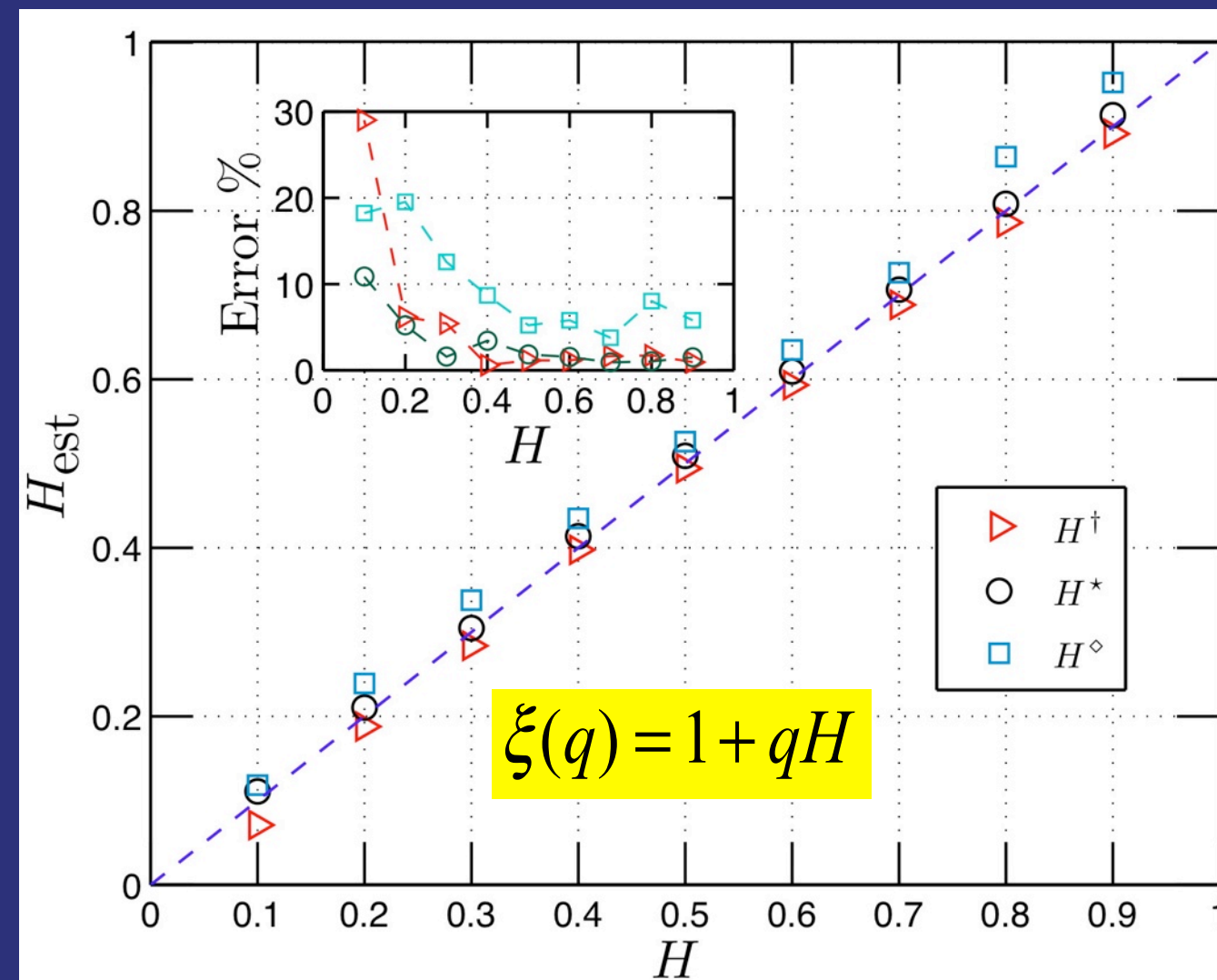
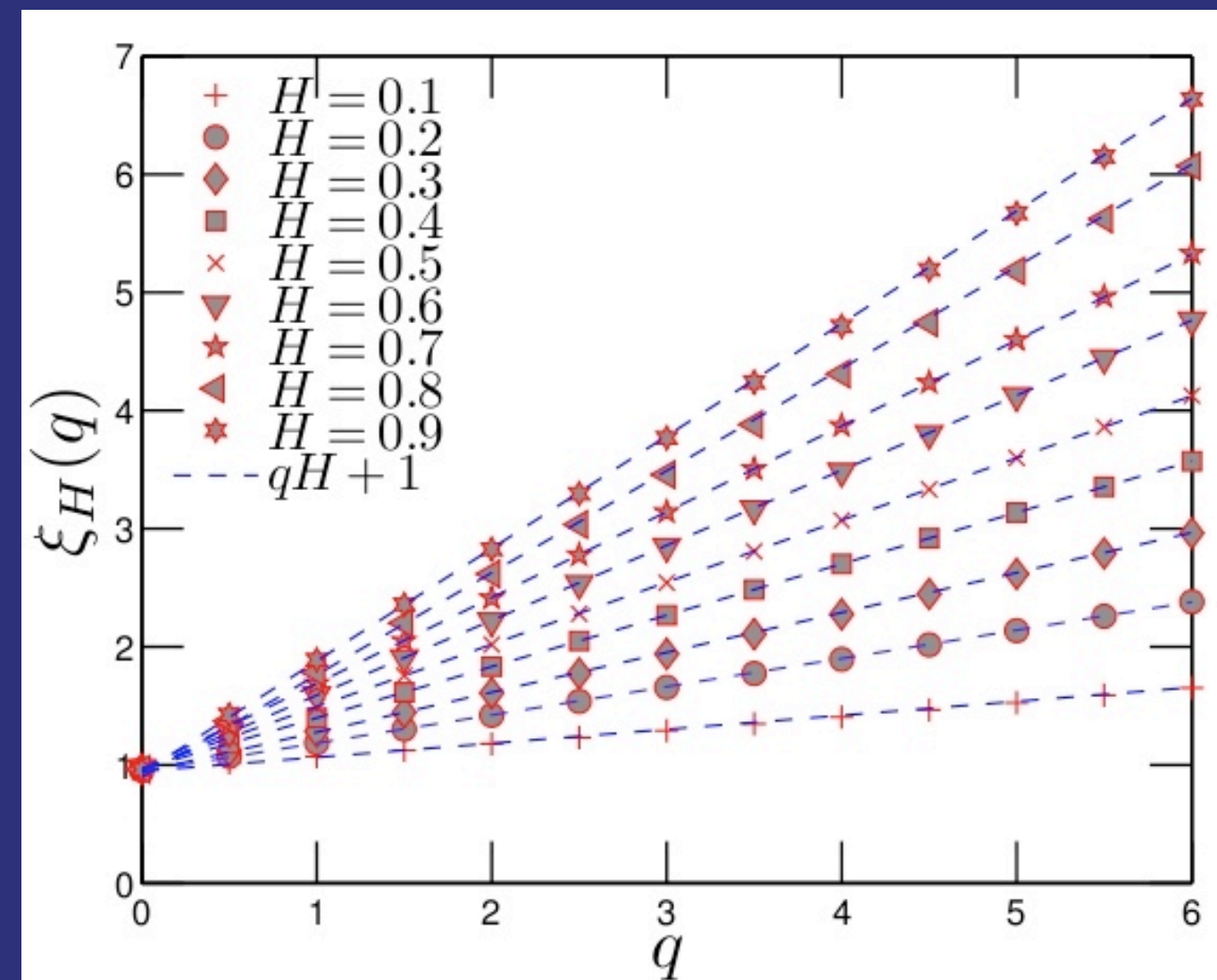
$$\text{Cov}(X(t), X(t')) = \frac{1}{2} \left( |t|^{2H} + |t'|^{2H} - |t - t'|^{2H} \right) \text{Var}(X(1))$$

Scaling with

$$\beta = 1 + 2H$$

$$\zeta(q) = qH$$

$$\zeta(q) = qH$$



# Verification for synthetic multifractal data

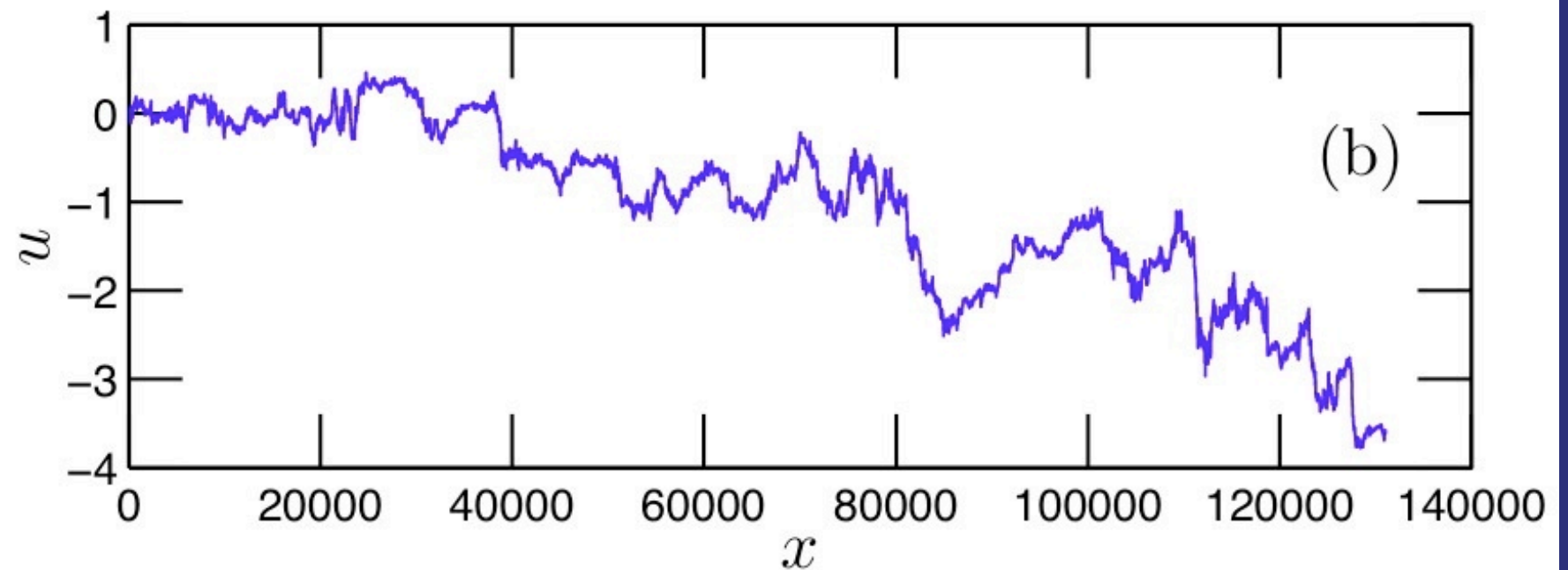
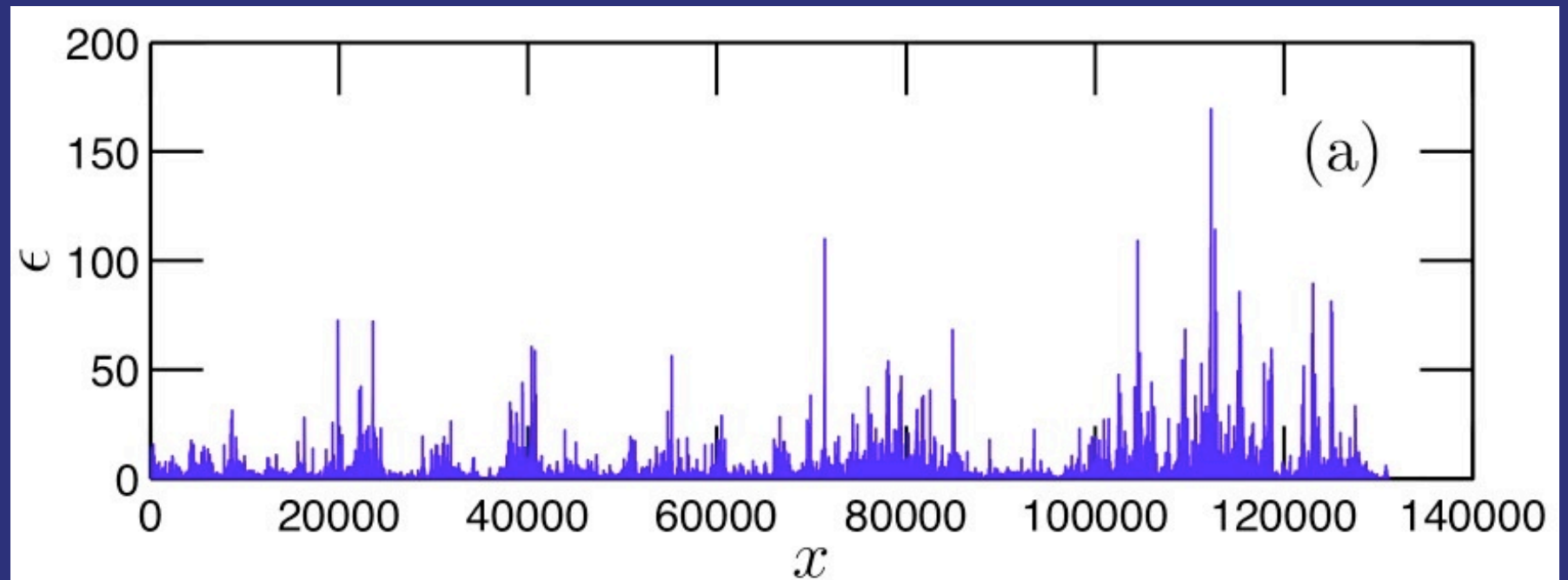
Synthetic multifractal series with lognormal statistics

**Multifractal random walk**

$$u(x) = \int_0^x \varepsilon(x')^{1/2} dB(x')$$

$$\langle |\Delta u_\tau|^q \rangle \approx \tau^{\zeta(q)}$$

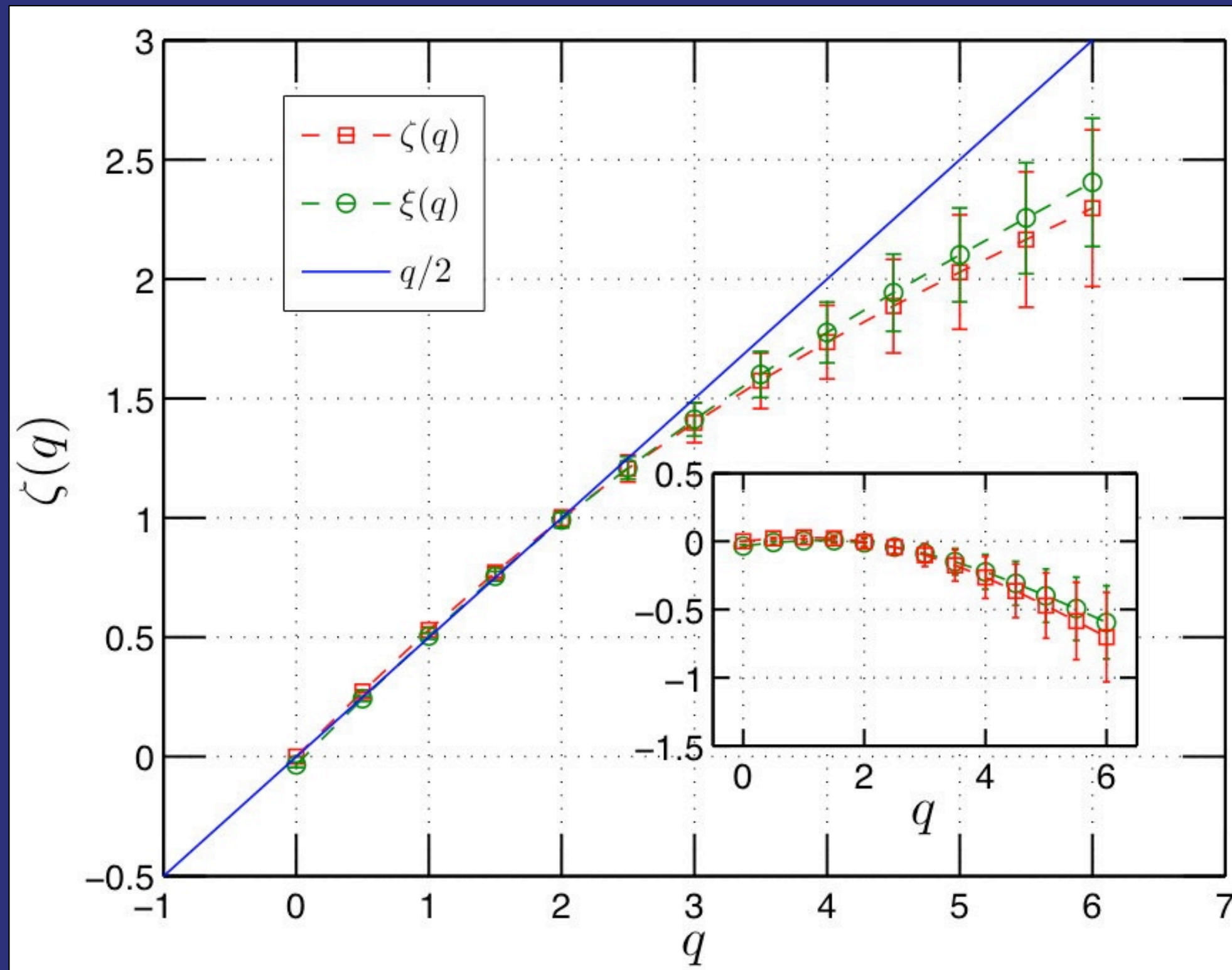
$$\zeta(q) = \frac{q}{2} - \frac{\mu}{2} \left( \frac{q^2}{4} - \frac{q}{2} \right)$$



# Verification for synthetic multifractal data

Scaling exponents for 70,000 realizations, with  $\mu=0.25$ .

The scaling exponents provided by the two methods are in good agreement with each other and with the theory



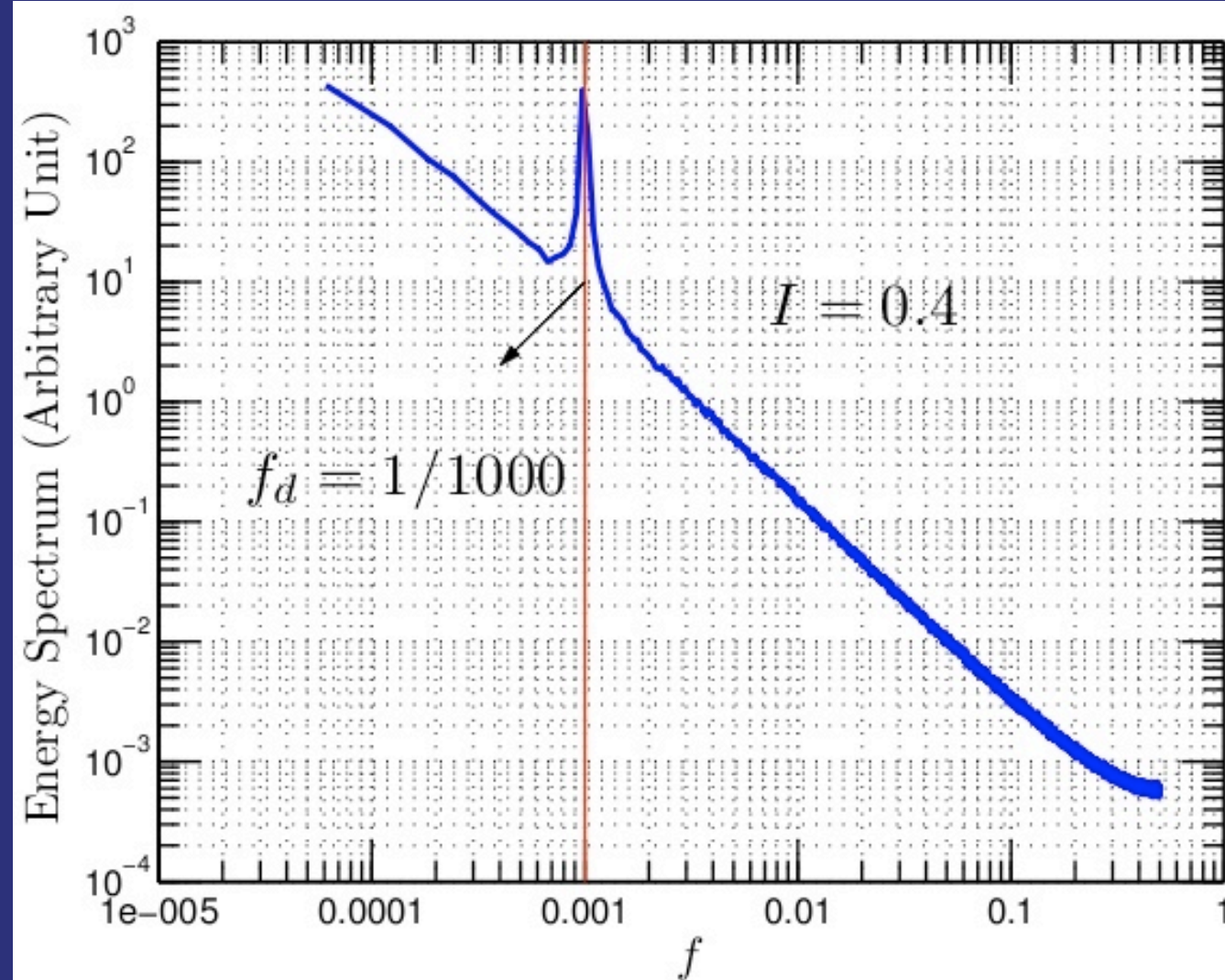
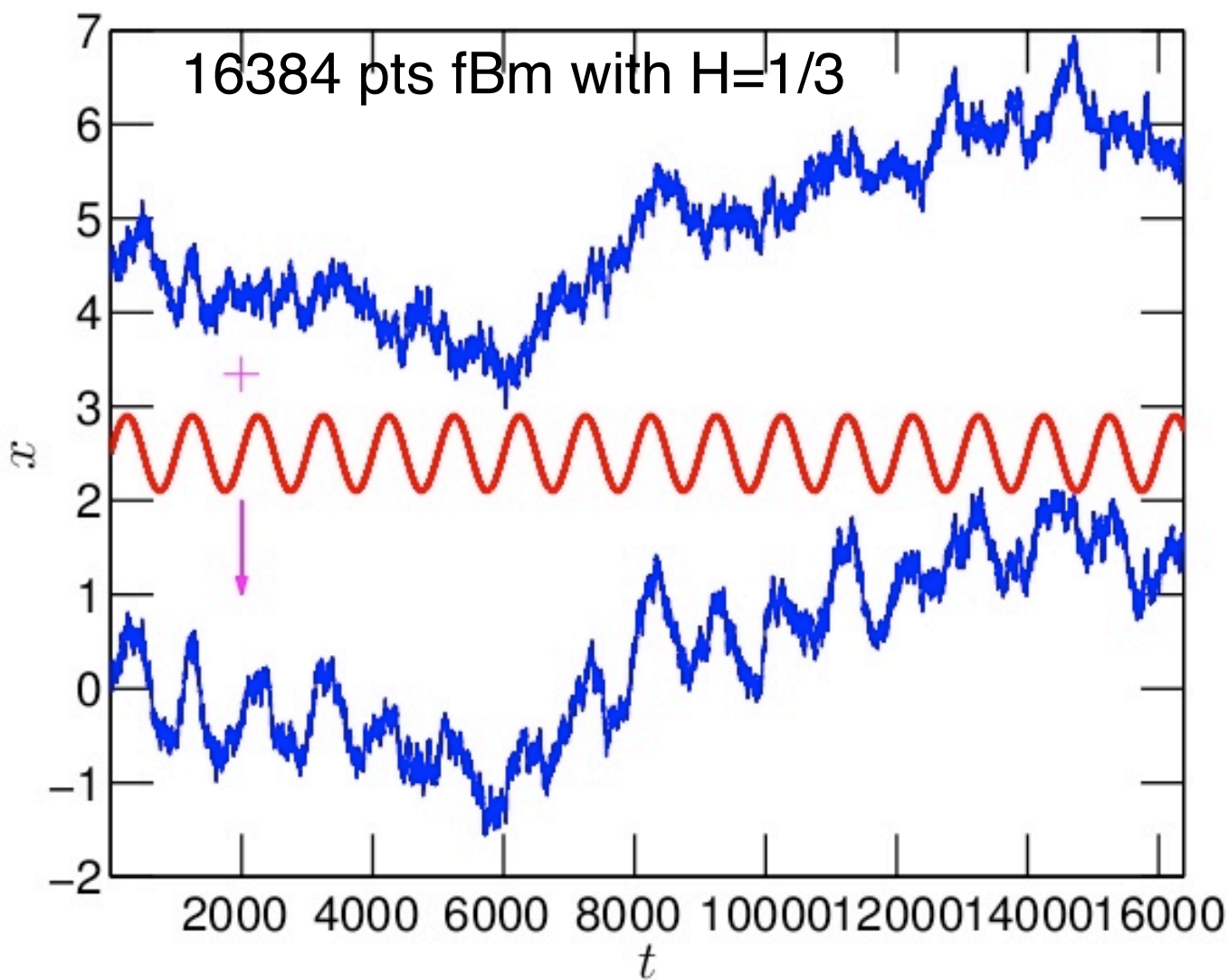
Validation of the method for multifractal signals



- Generalized EMD-HSA method can be used to estimate monofractals and multifractal exponents.
- Seems to provide parameters with a better precision than the classical structure functions.
- Estimation of multifractal exponents in frequency space

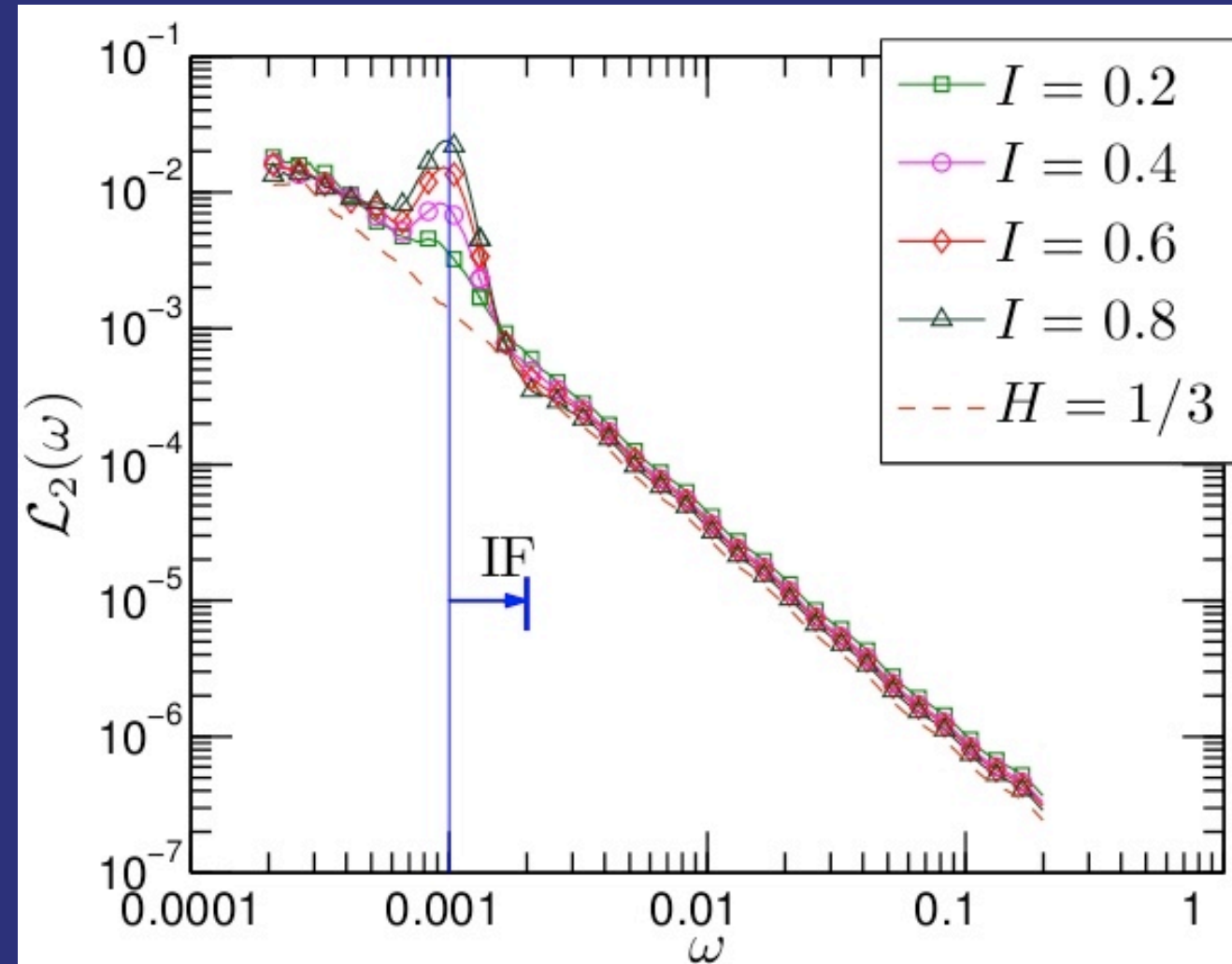
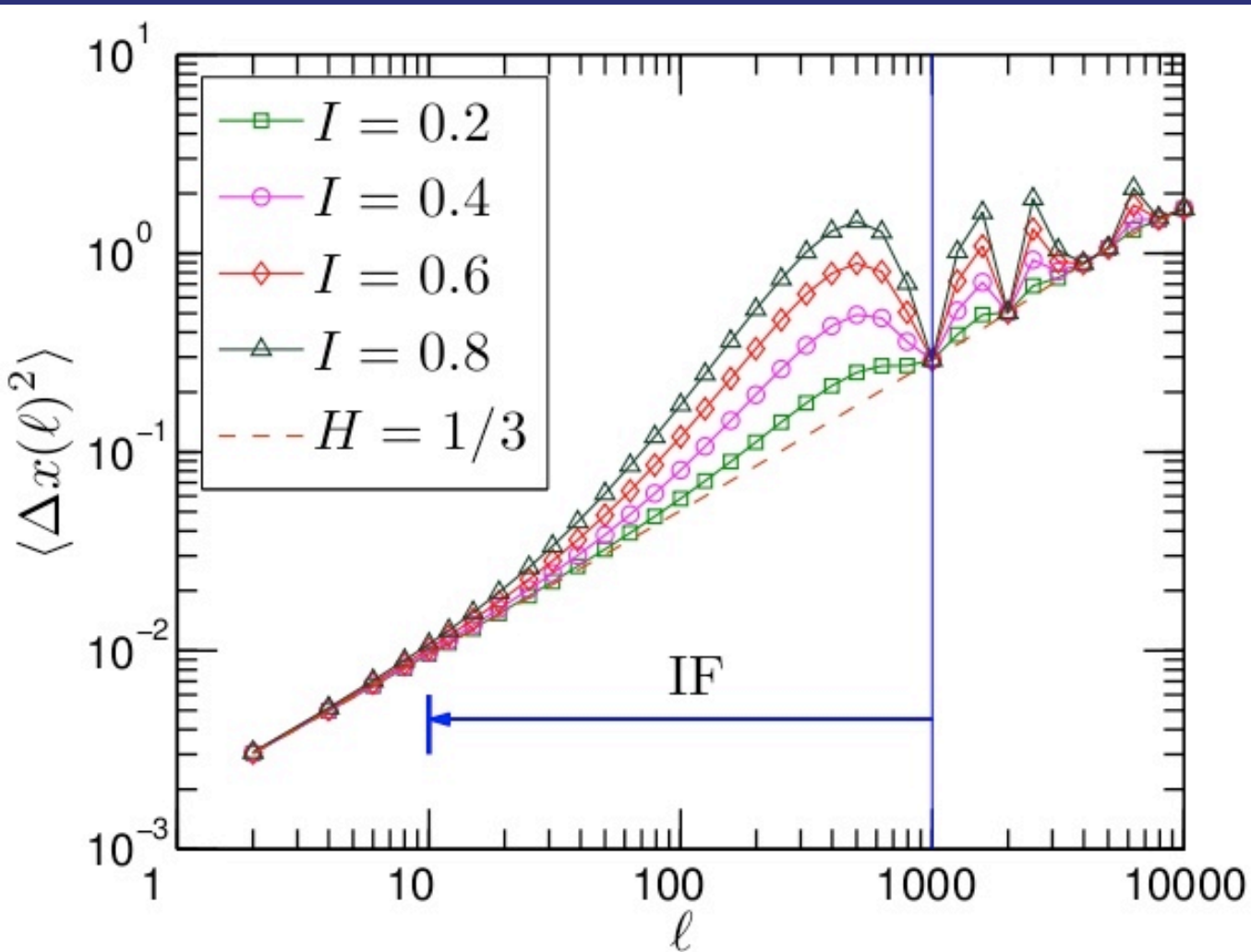
# Influence of large scale forcing

Many data series from the real world, e.g. geophysics, present intermittency mixed with **large-scale deterministic forcing** (often from astronomic origin)



$$X'(t) = \frac{X(t)}{Var(X)} + I \sin(2\pi f_d t)$$

# Influence of large scale forcing



Much smaller influence range of large scales for arbitrary order HSA than for structure functions



- Structure function analysis (or the variant MFDA) are strongly influenced by large scales
- Not adequate to study scaling properties when there is large scale forcing
- EMD-arbitrary order HSA is a good alternative in these cases, to retrieve scaling and intermittent properties

## **Example of analysis:**

- MAREL high frequency data
- bivalve microclosing behavior
- irradiance high frequency time series
- phytoplankton abundance time series



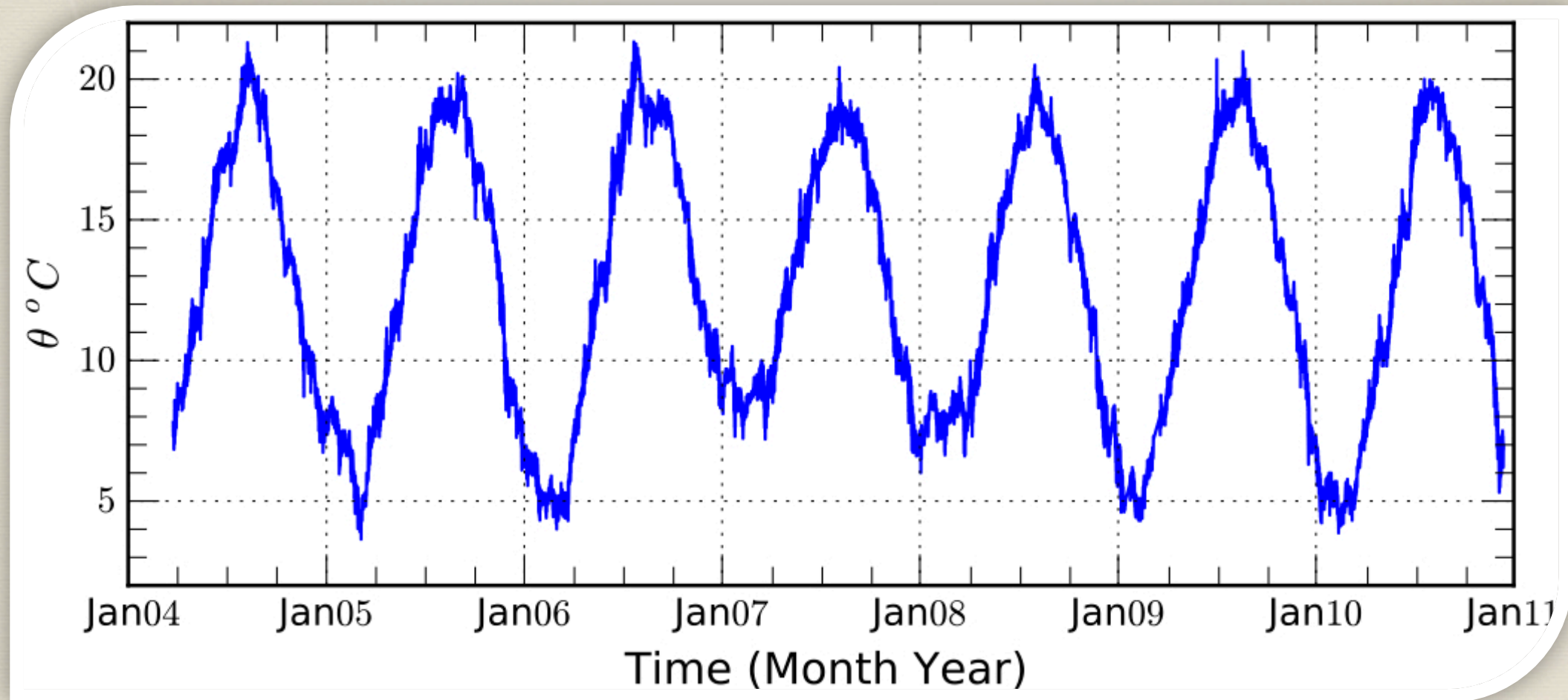
# MAREL Data



A photo showing the MAREL system in Boulogne-sur-mer. Operated by Ifremer. Automatic measurements of many parameters (T, S, pH, DO, etc.) every 20 minutes



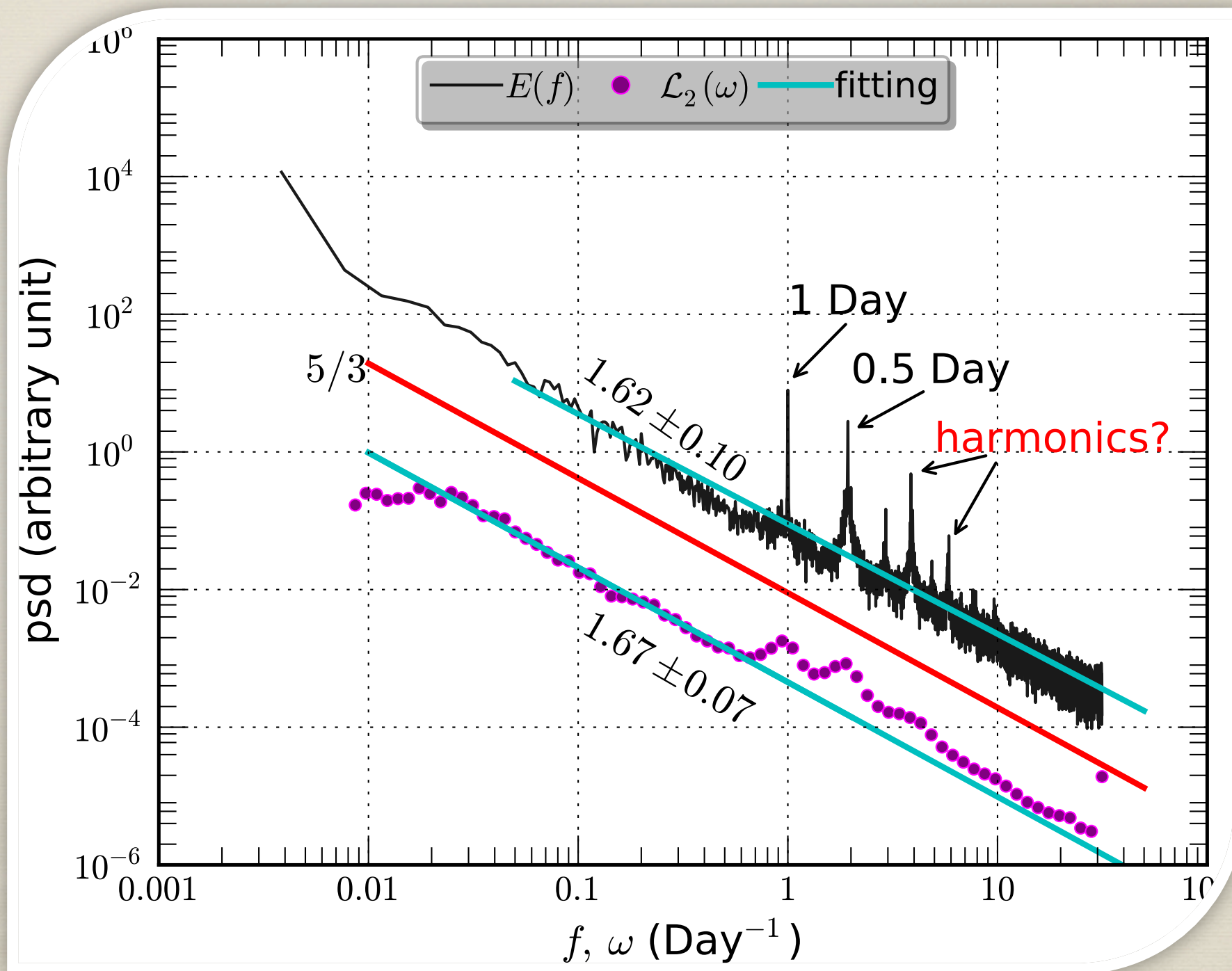
# MAREL Data



The recorded MAREL water temperature data from Mar. 2004 to Dec. 2010.

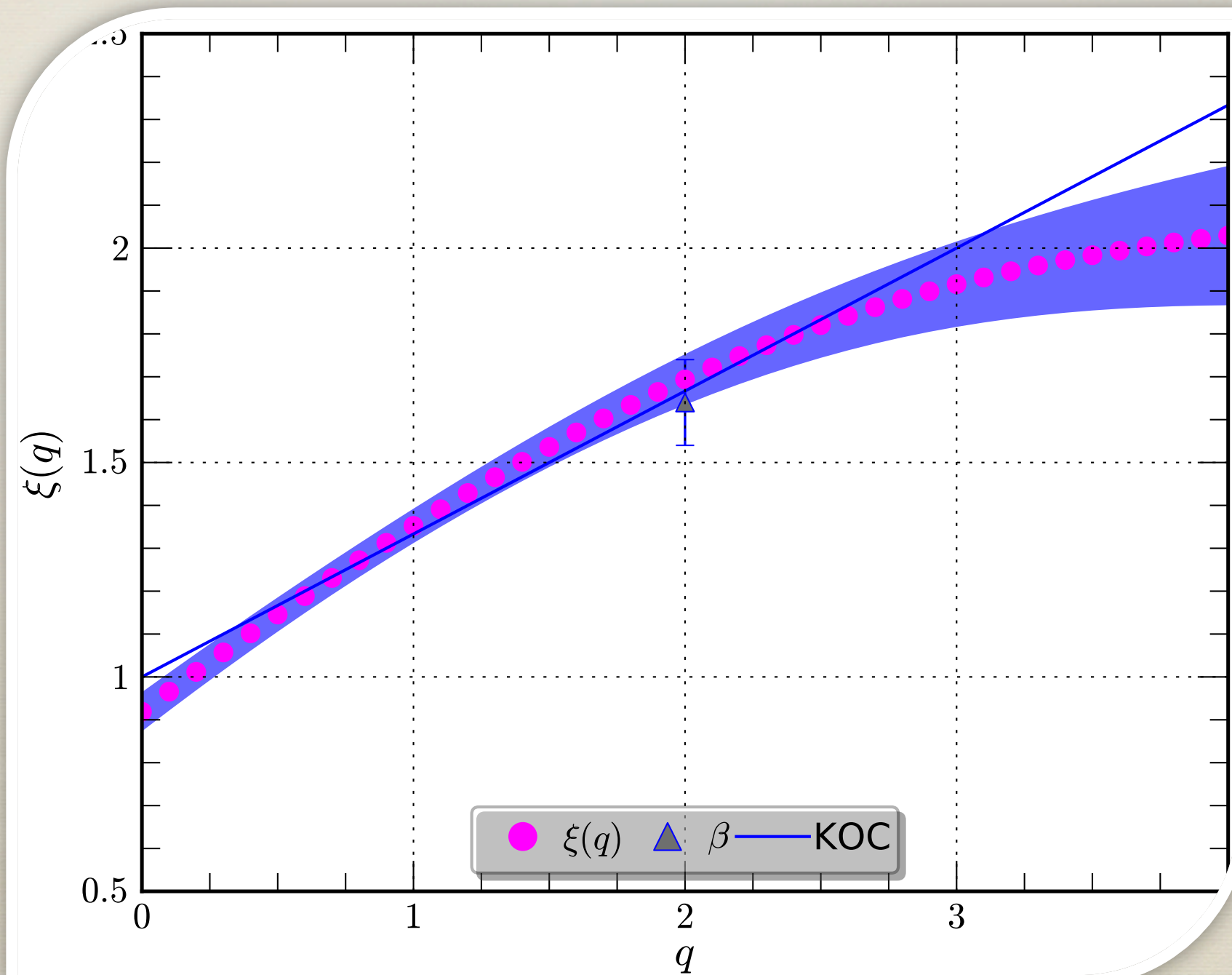


# Power Spectral Density





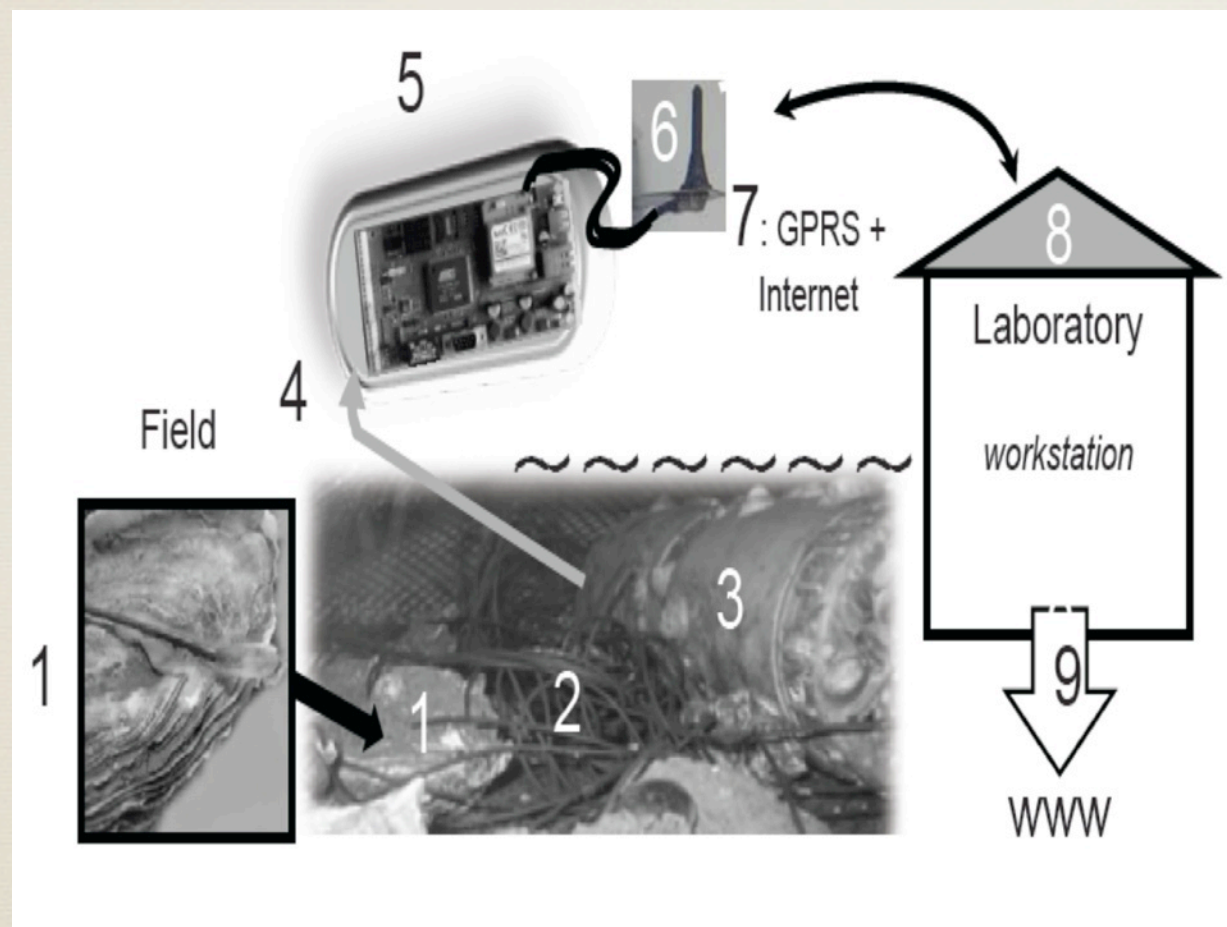
# Scaling exponents extraction using EMD and Hilbert spectral analysis





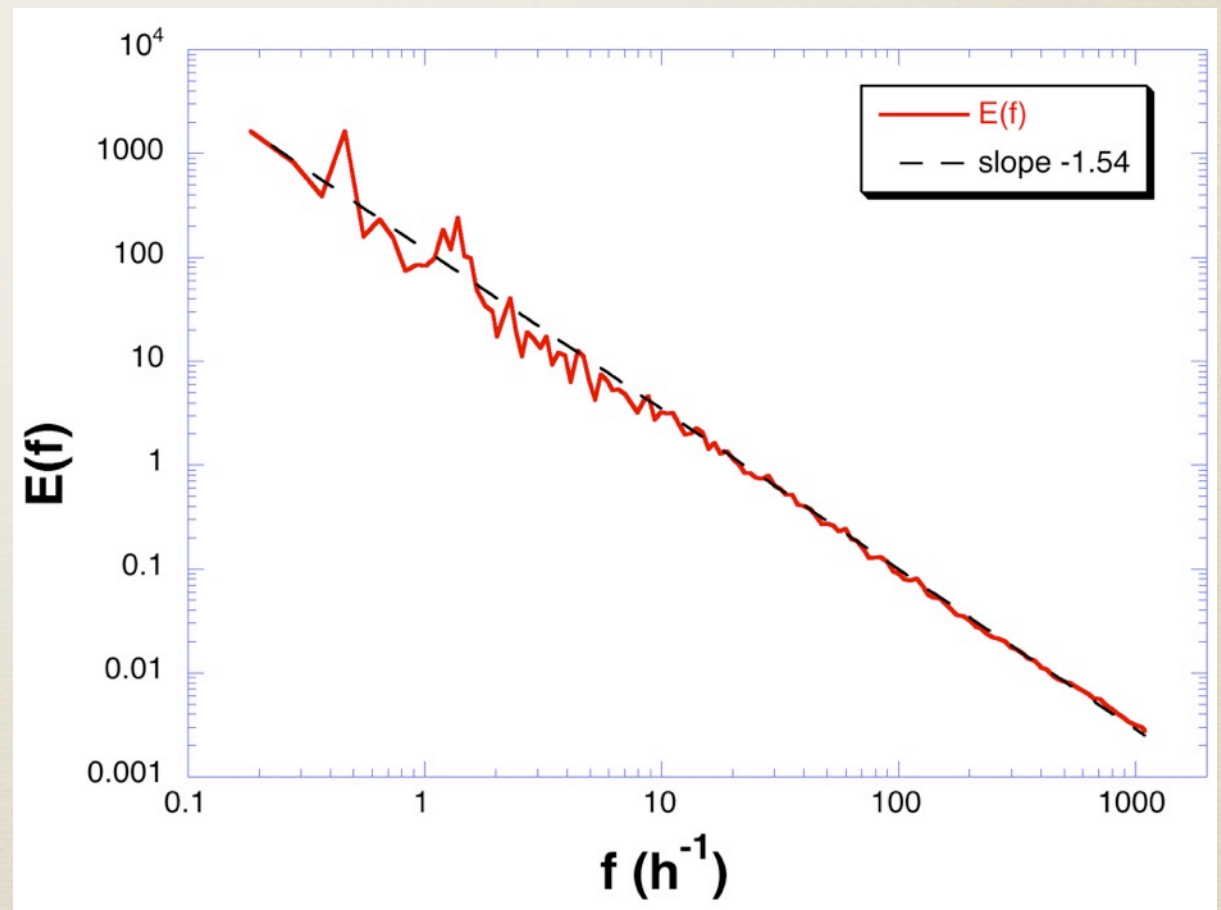
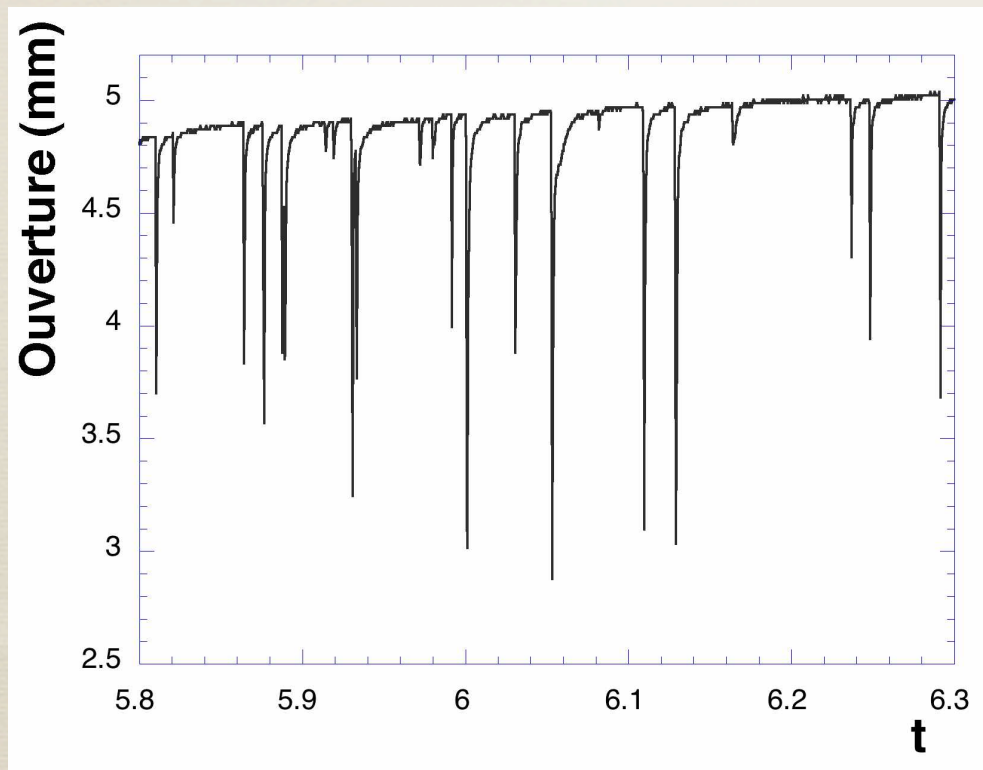
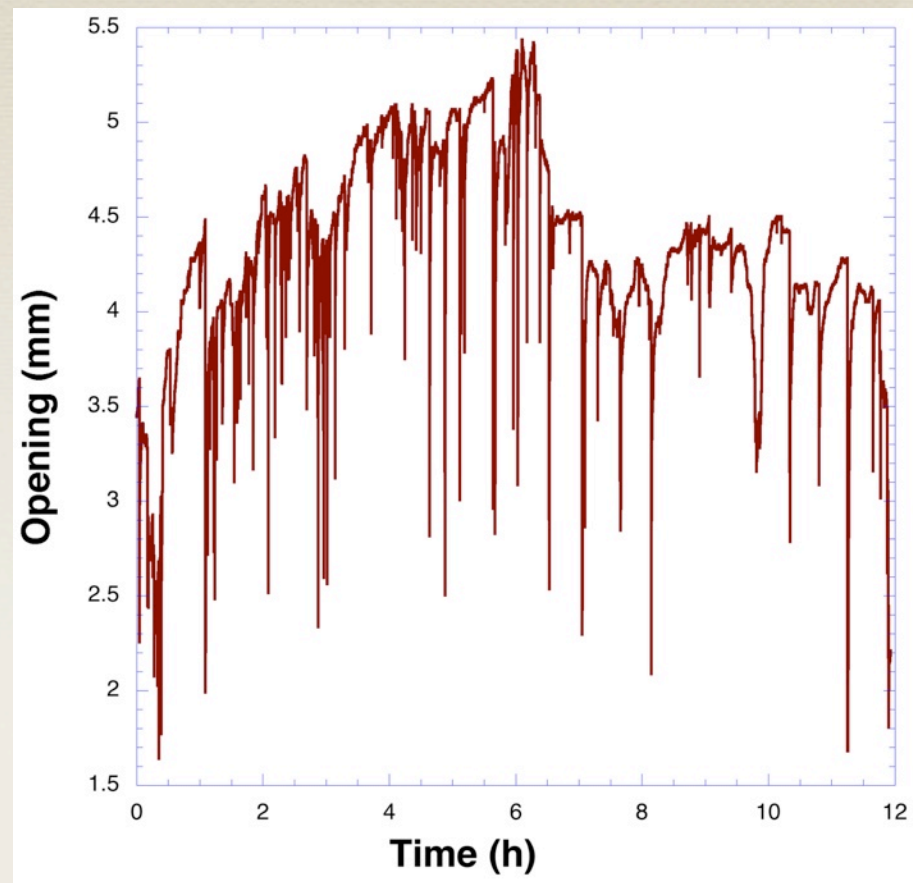
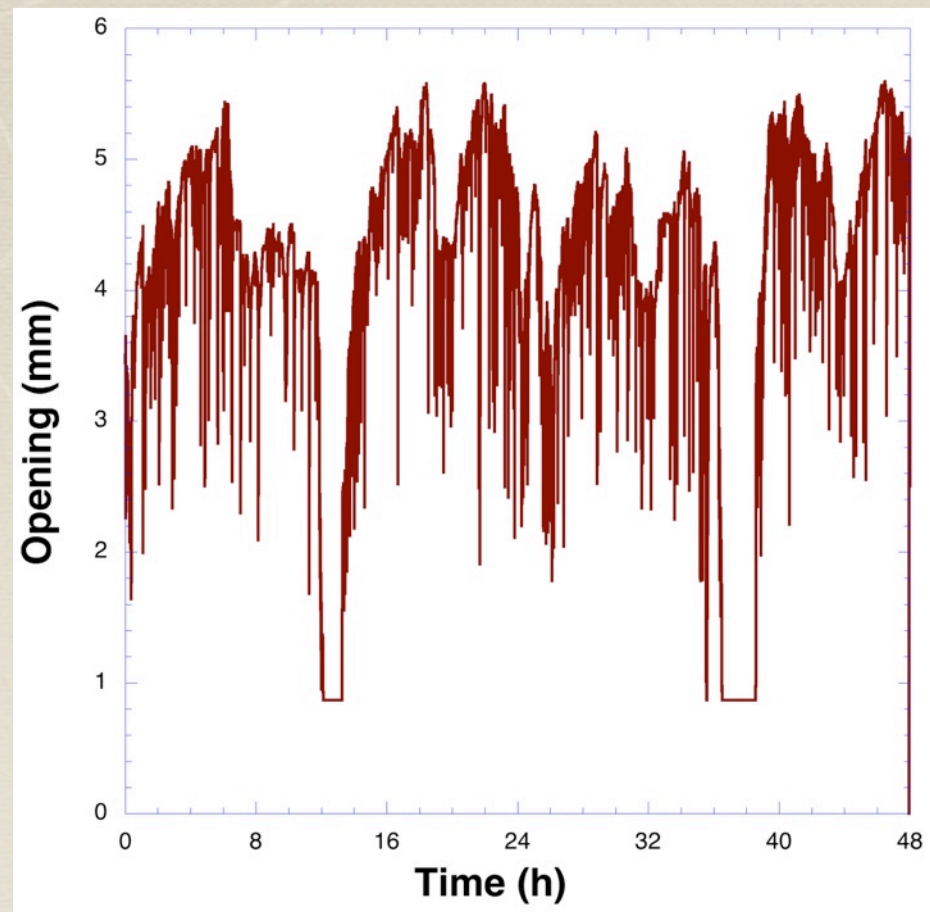
# Bivalve opening time series

collaboration with Arcachon marine laboratory (JC Massabuau, G Durrieu)

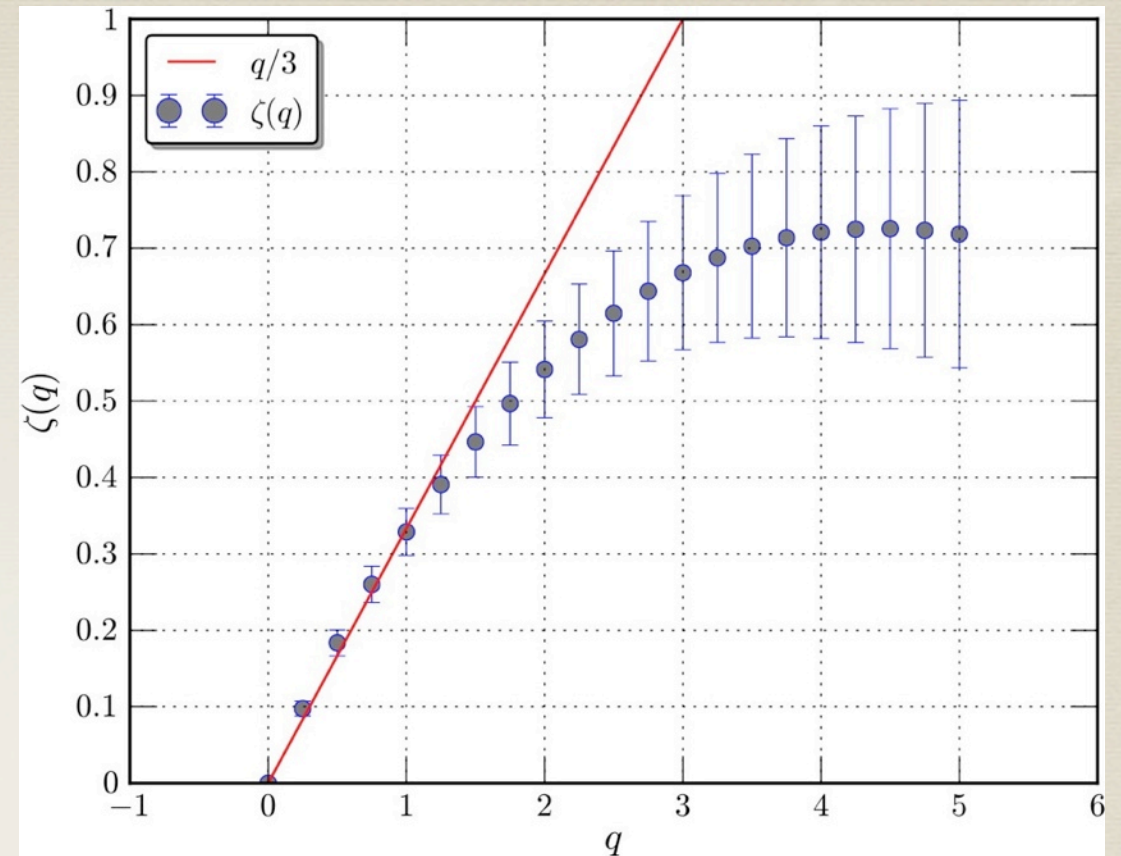
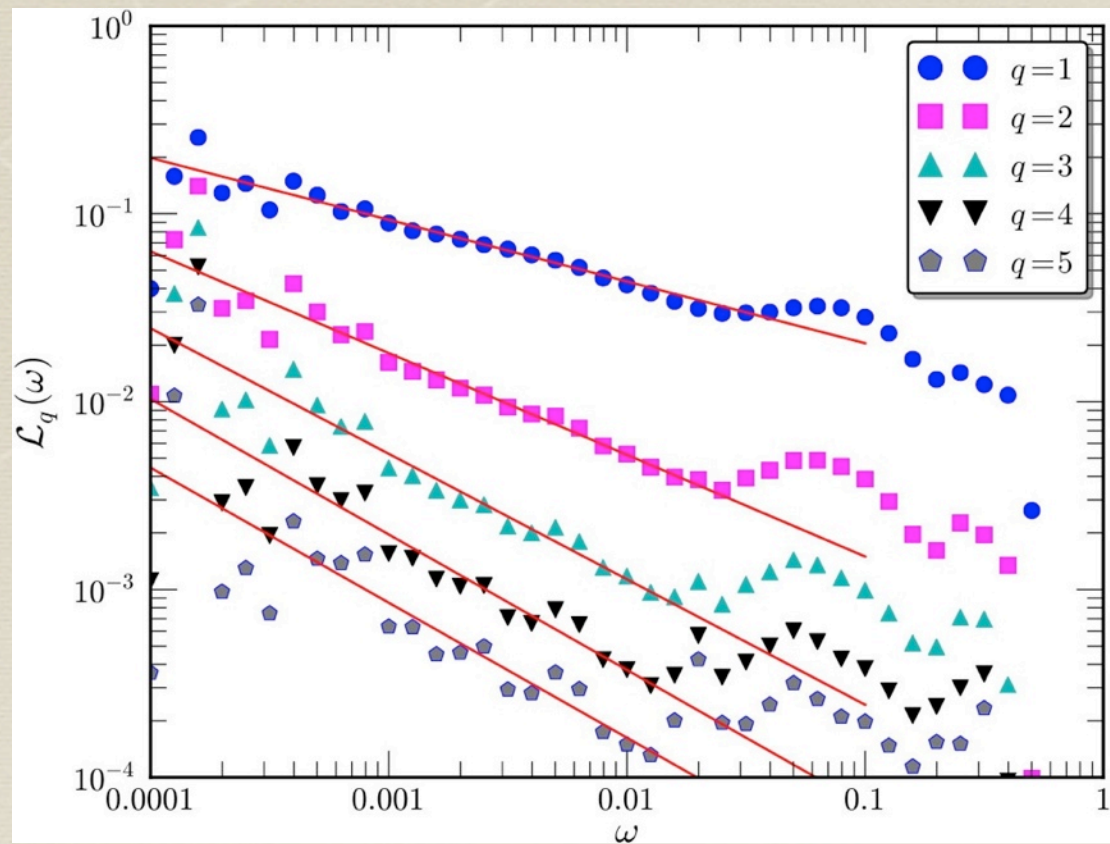


High frequency bivalve opening time series (every 1.6s)

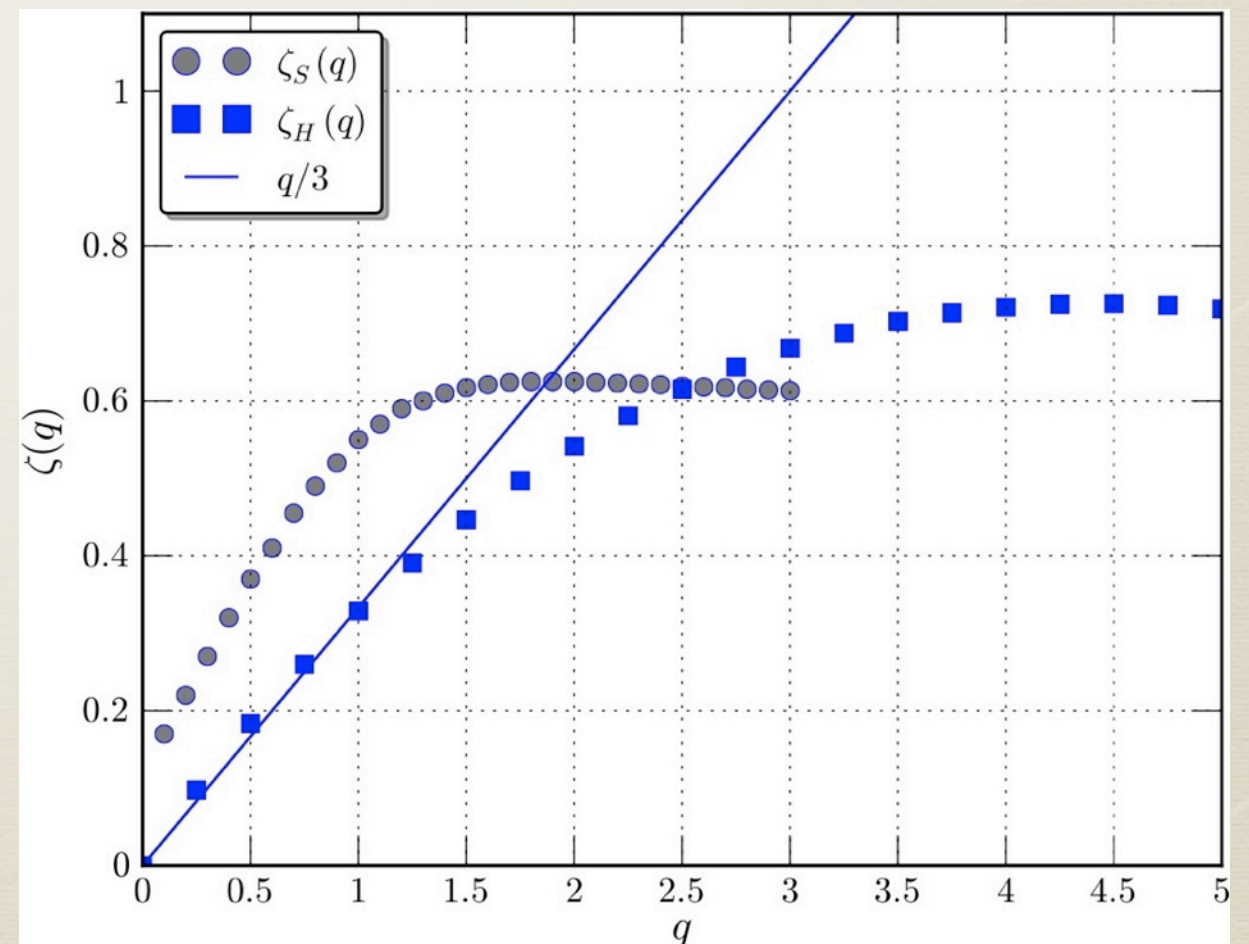








Extraction of  
 intermittency  
 parameters  
 HSA works better than  
 structure functions





# High frequency irradiance time series

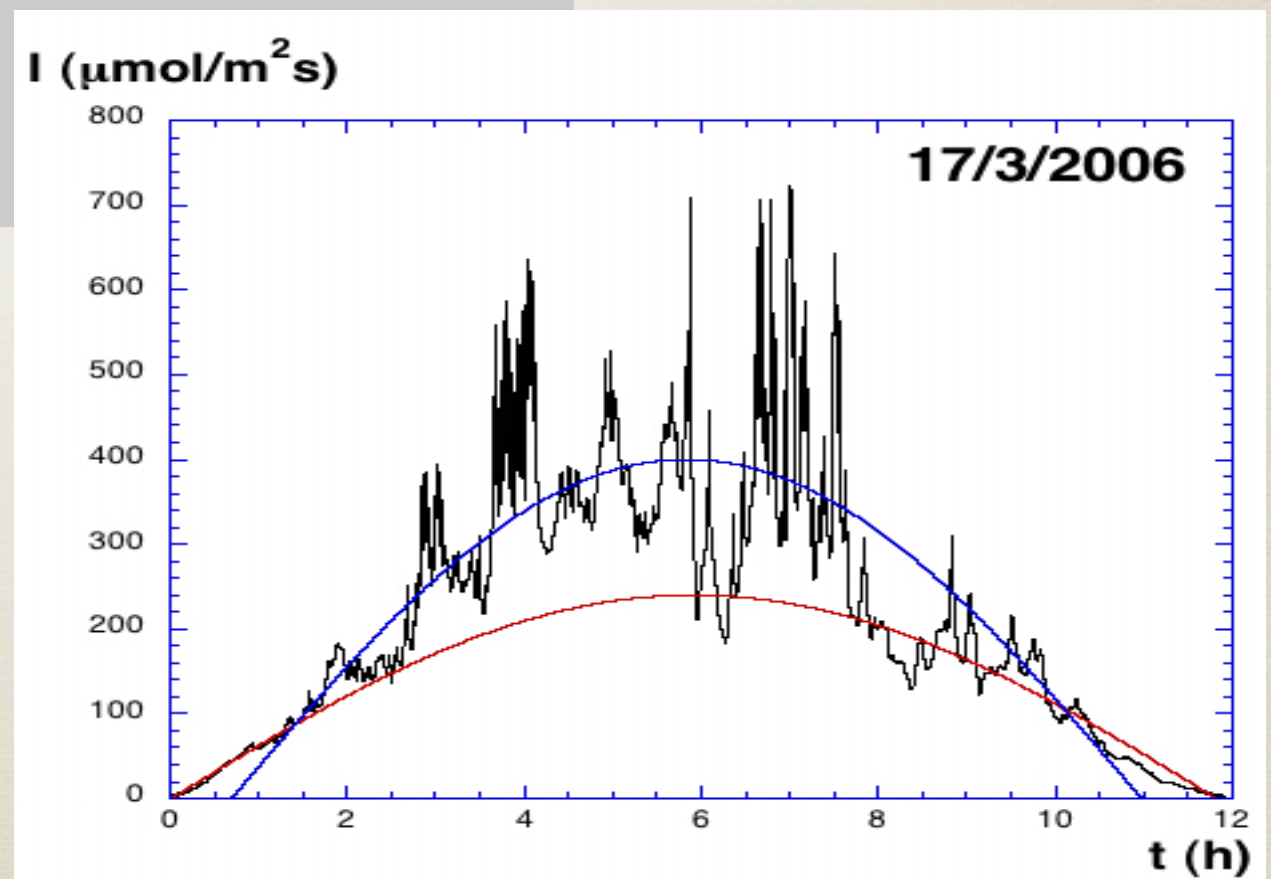


Huge intermittent fluctuations

For the clear sky fit,  $I_0$  is unknown: cannot be determined from the data

$$I = \begin{cases} I_0 \left( c_1 - c_2 \cos \frac{\pi t}{12} \right) & \text{if } c_1 > c_2 \cos \frac{\pi t}{12} \\ 0 & \text{if } c_1 \leq c_2 \cos \frac{\pi t}{12} \end{cases}$$

Strong high frequency fluctuations of irradiance; not smooth as sine-model predicts; there are consequences for phtoplankton production models





# High frequency irradiance time series

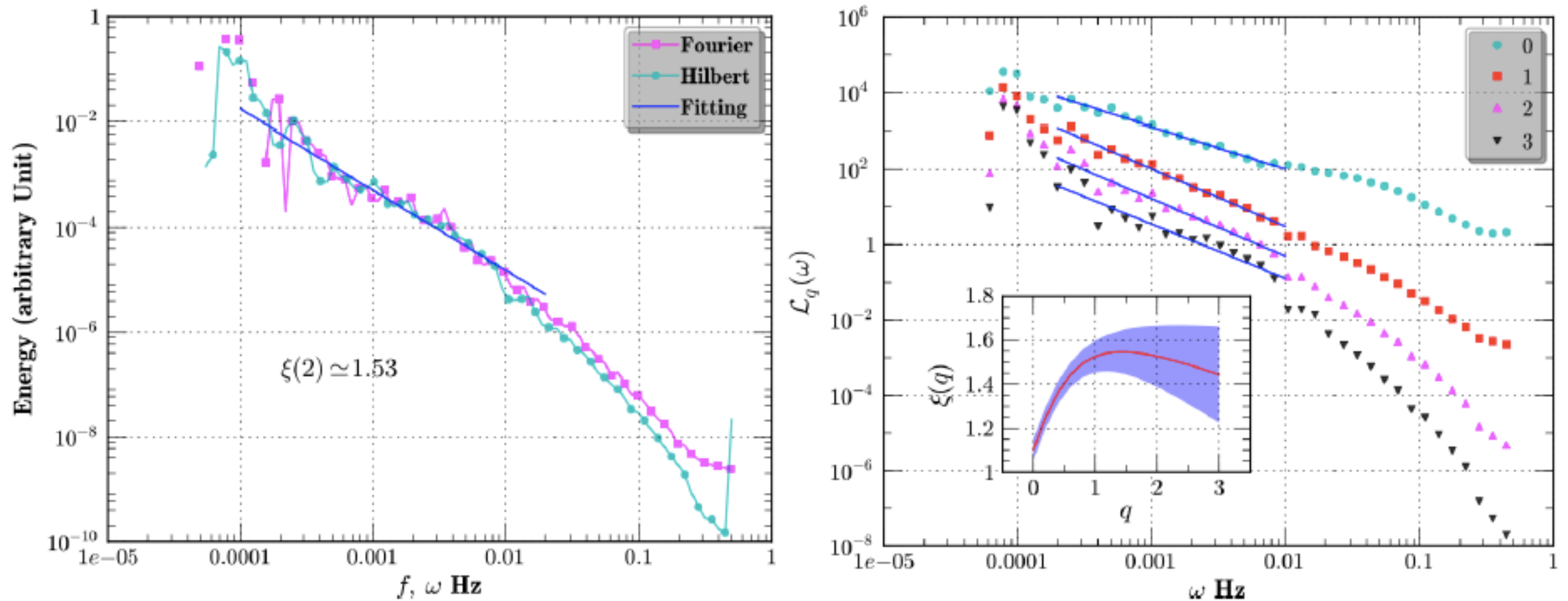


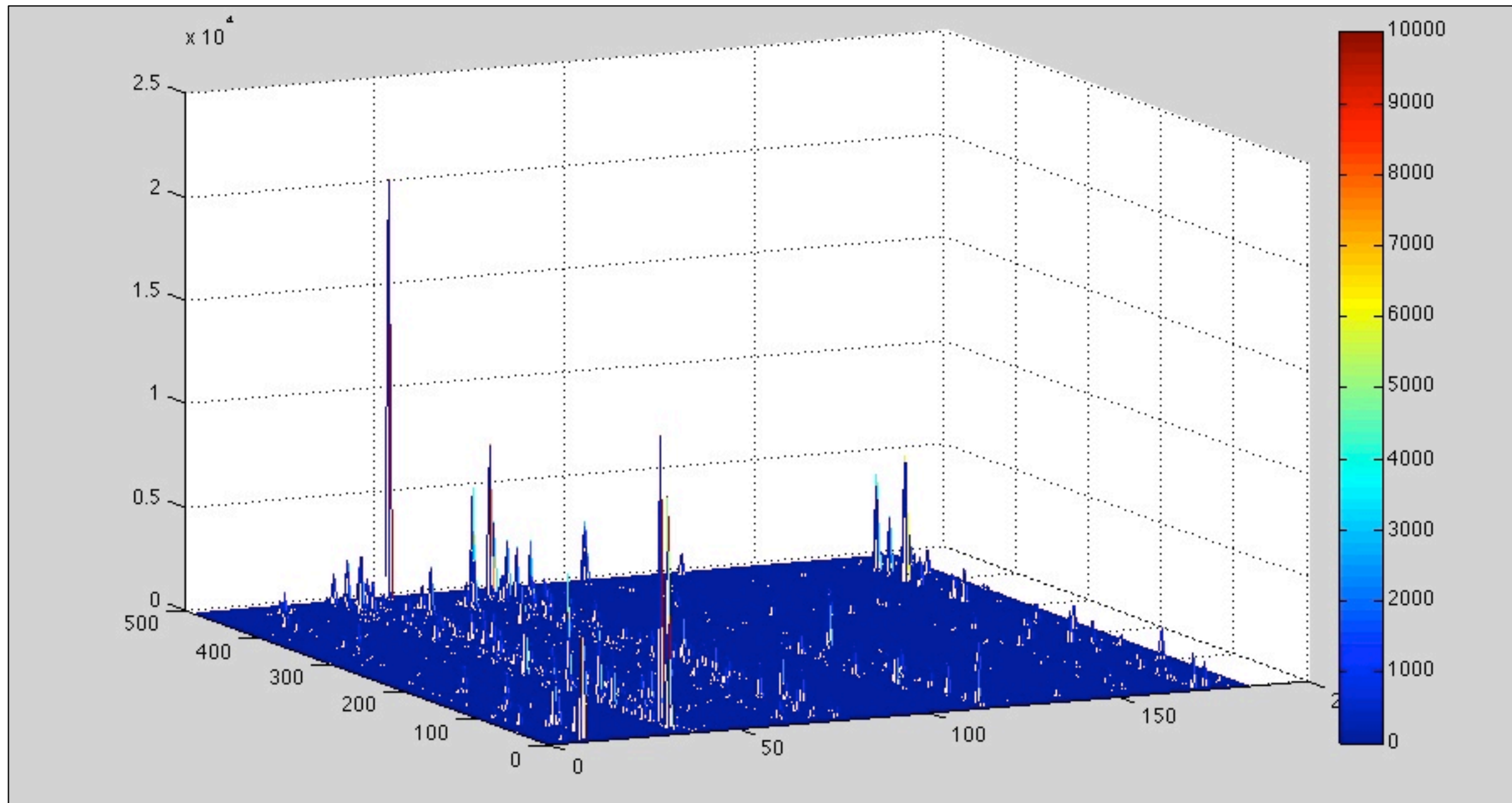
Fig.4. Gauche : spectre d'énergie en Fourier et en espace de Hilbert des données d'irradiance ; droite : invariance d'échelle des différents moments estimés en espace de Hilbert et en inset, la fonction d'invariance d'échelle  $\zeta(q)$  non linéaire





# Phytoplankton abundance data, lake Geneva

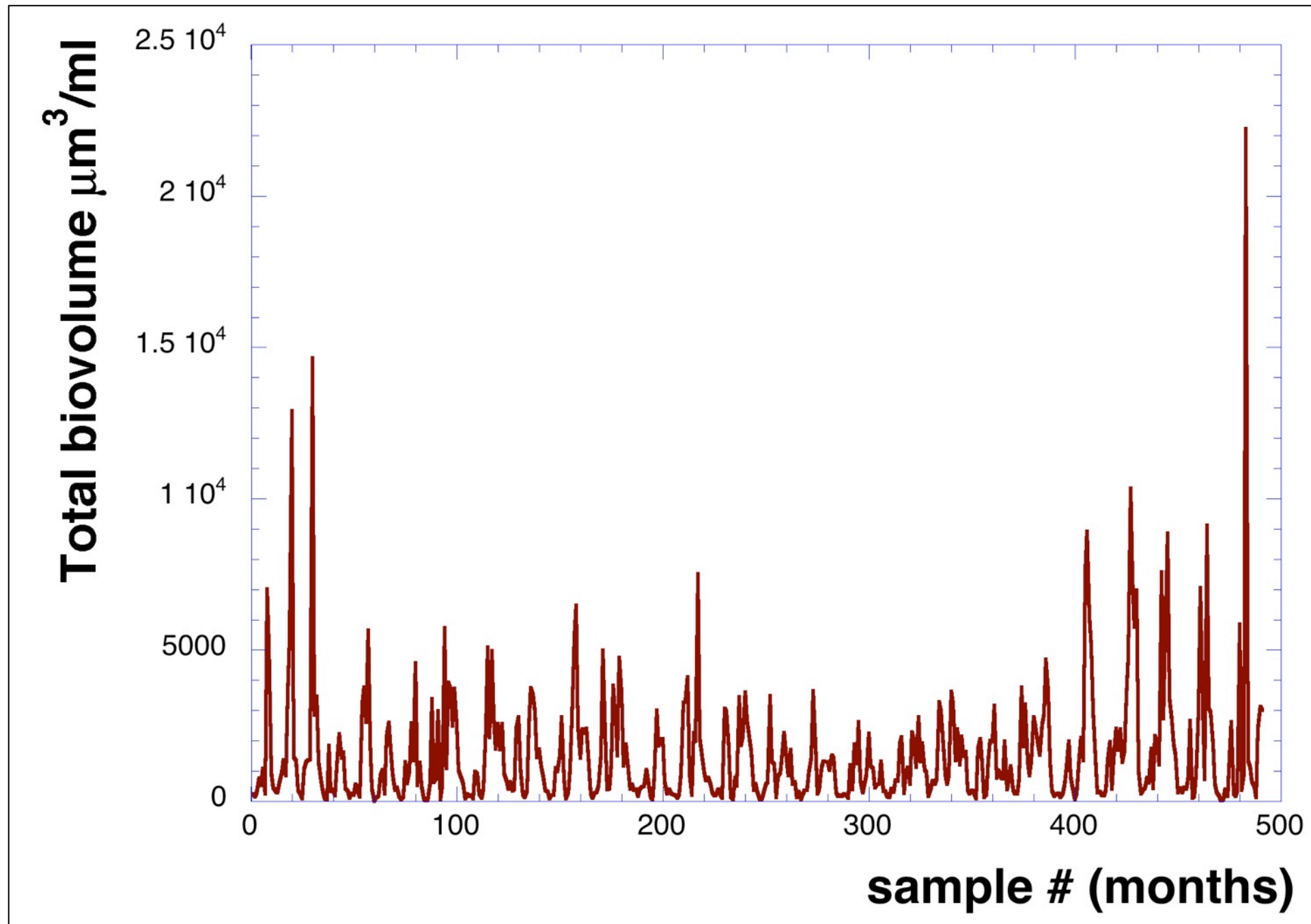
- Phytoplankton sampling from 1974 to 2000 at lake Geneva, monthly sampling
- 491 different sampling times, 184 species identified



Bursty dynamics; many zeroes (90%)

# Biovolume dynamics and statistics

Bursty, intermittent dynamics of the total biovolume

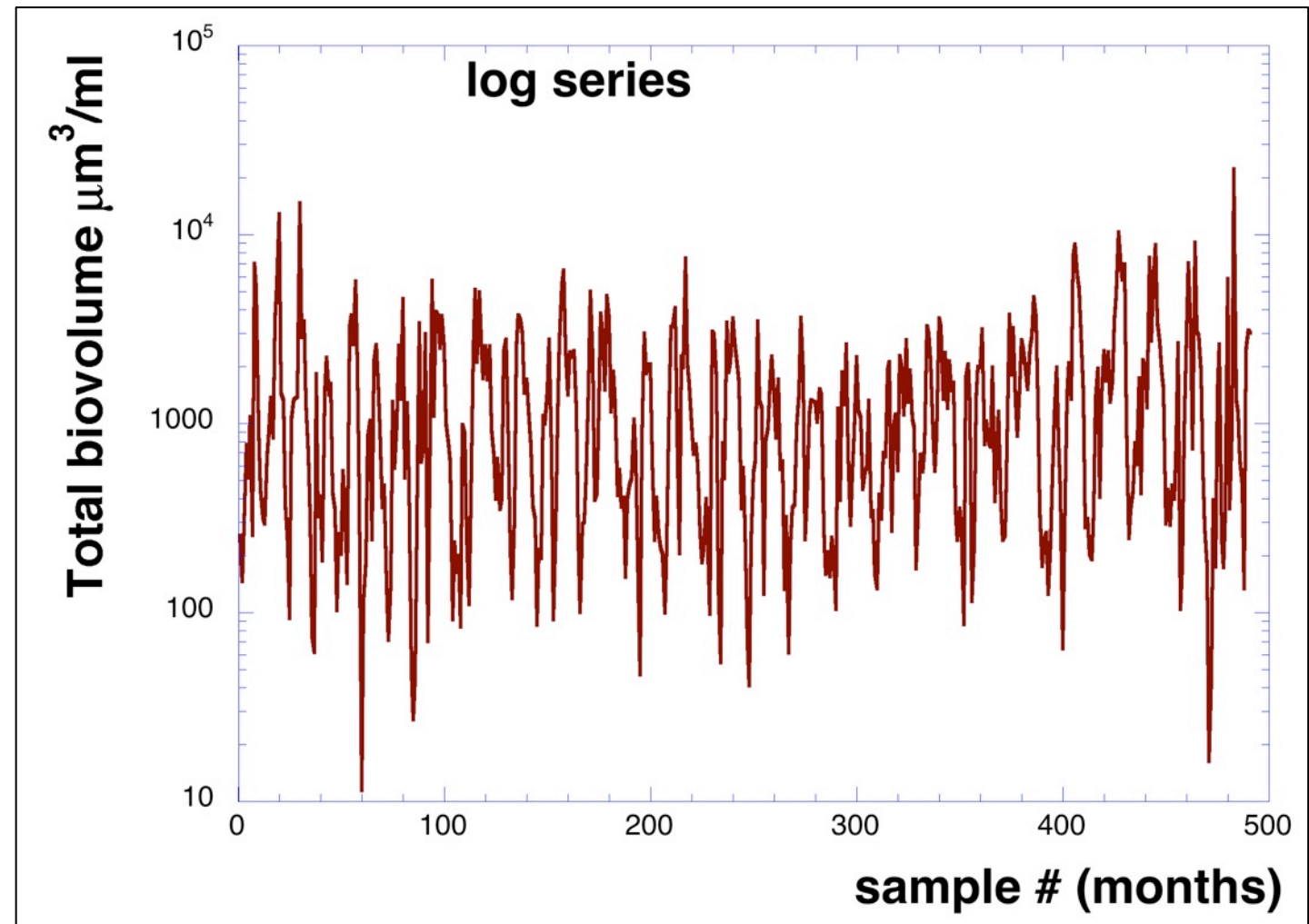
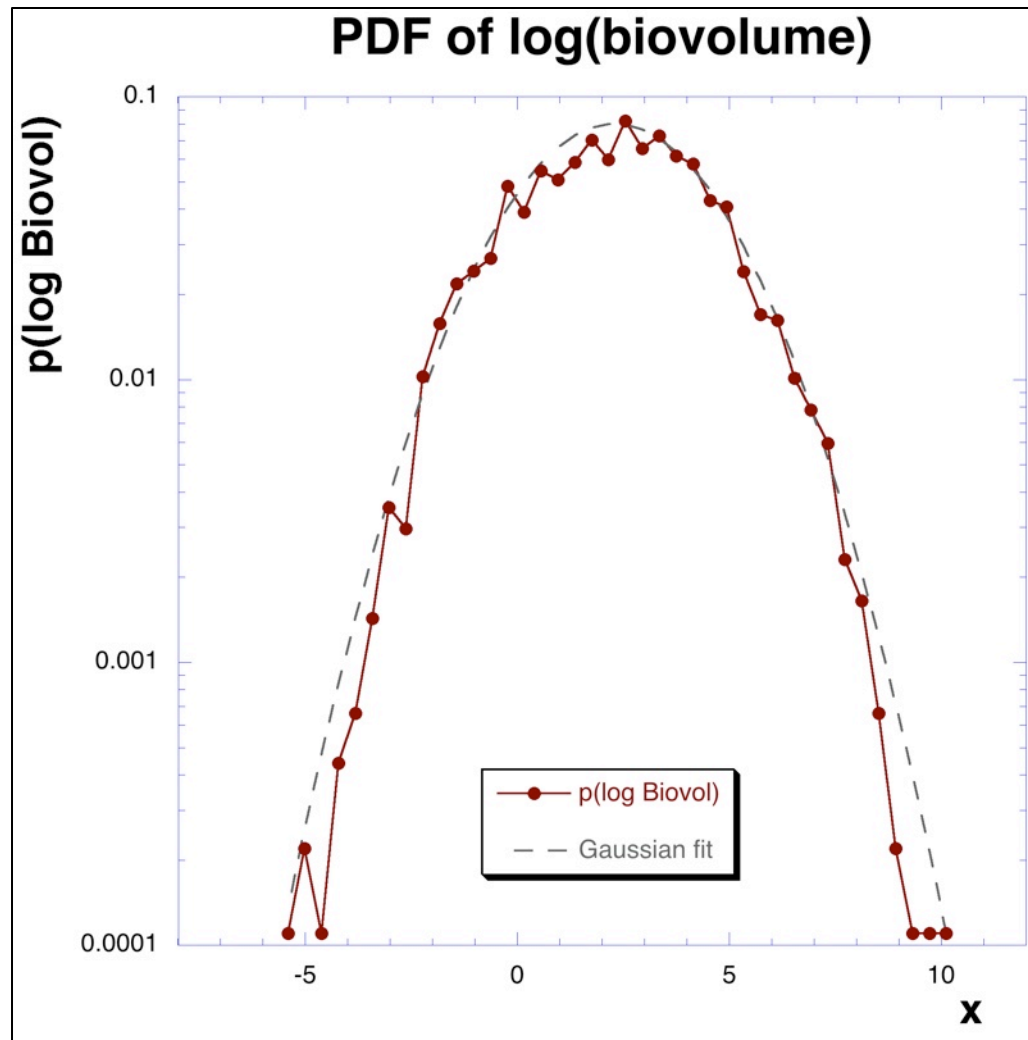


Total biovolume (all species)



# Biovolume dynamics and statistics

## Logarithm of total biovolume



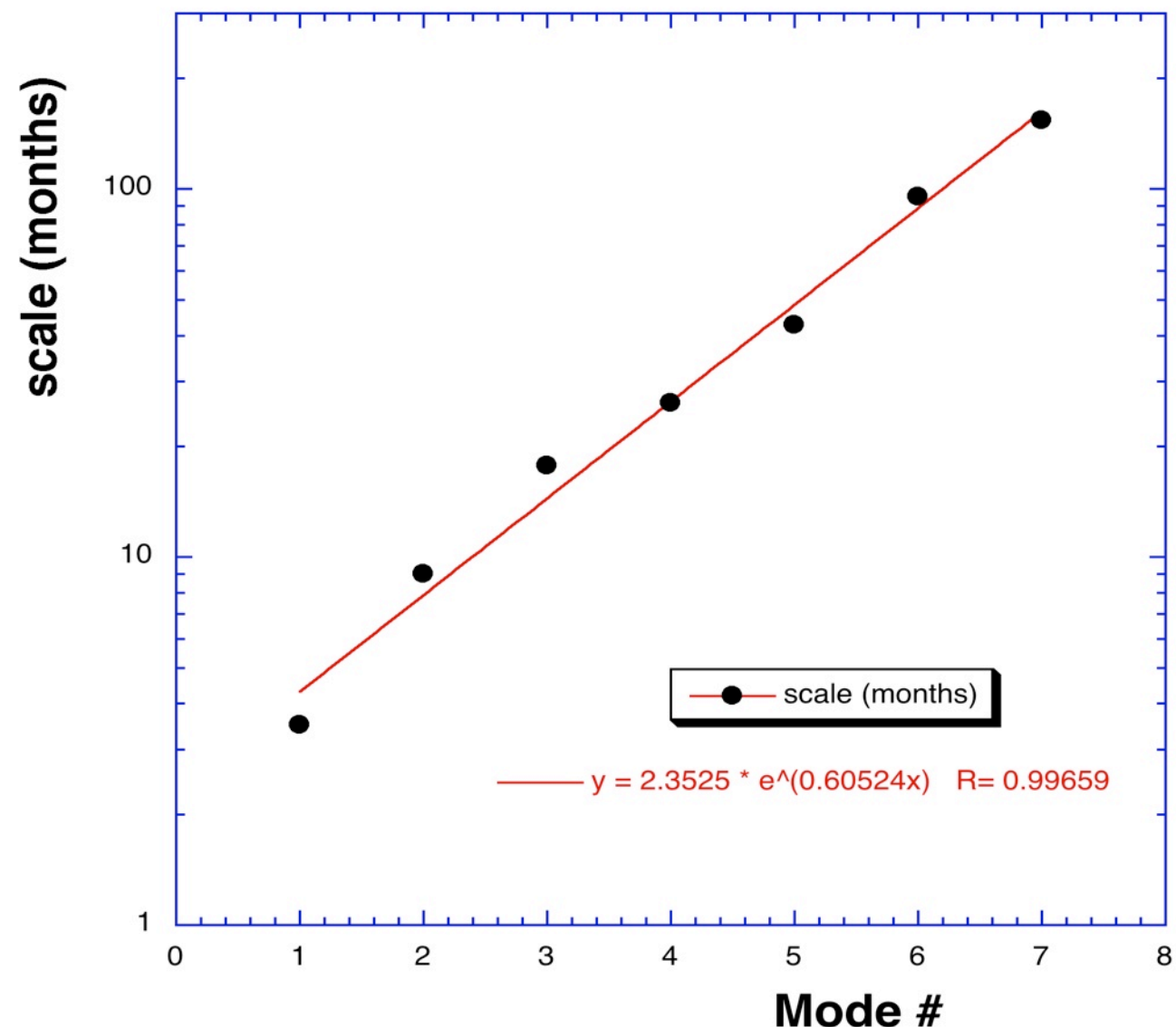
Stationary and intermittent series

Probability density function close to Gaussian: lognormal process

# Mean mode frequency: filter bank

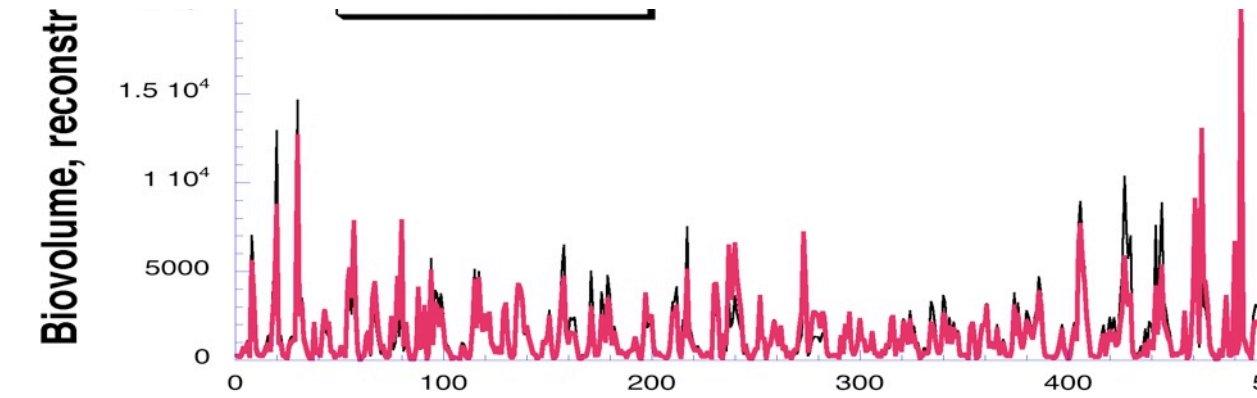
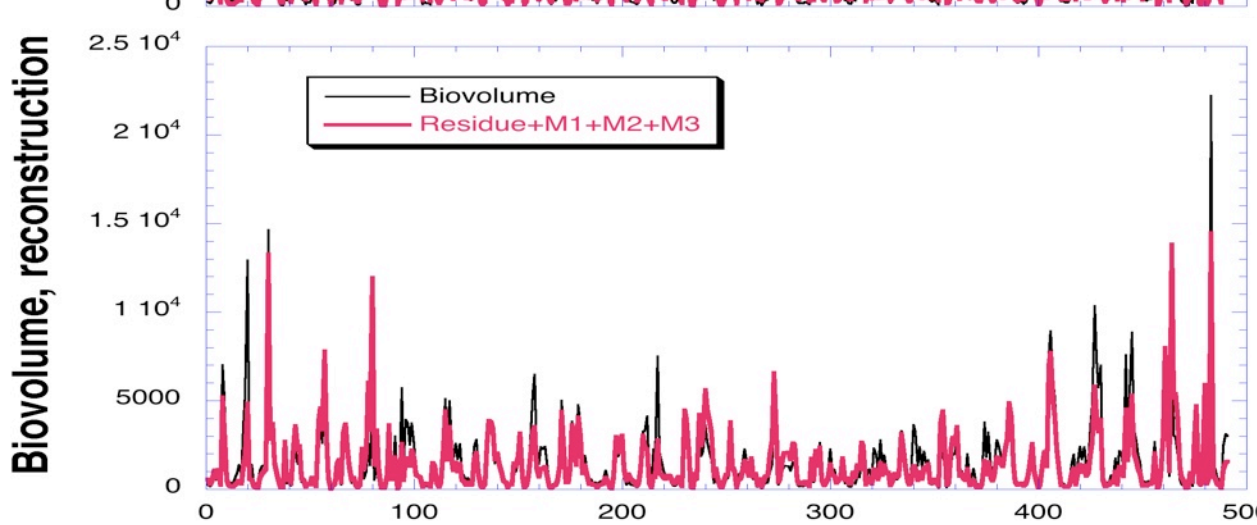
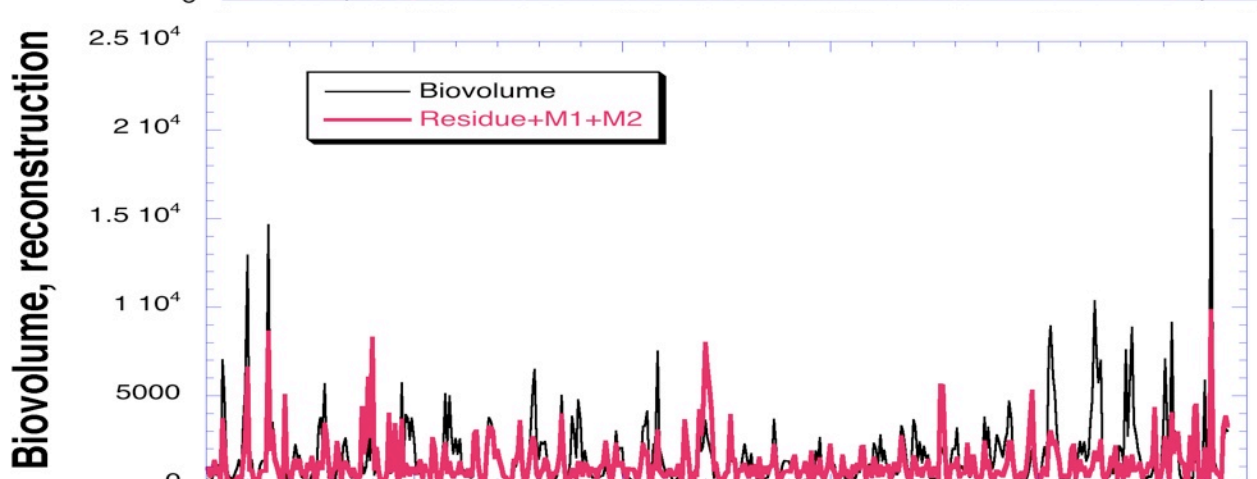
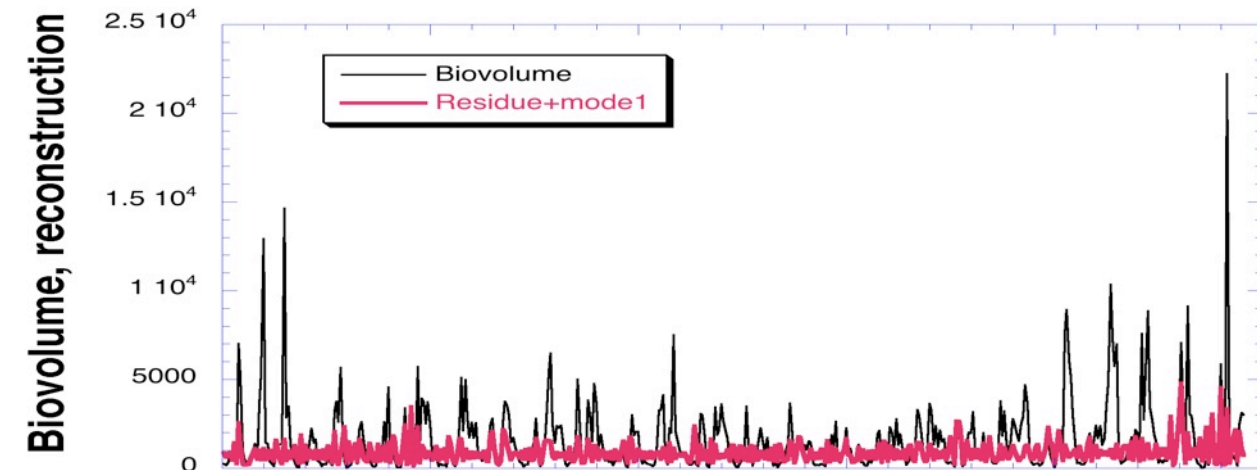
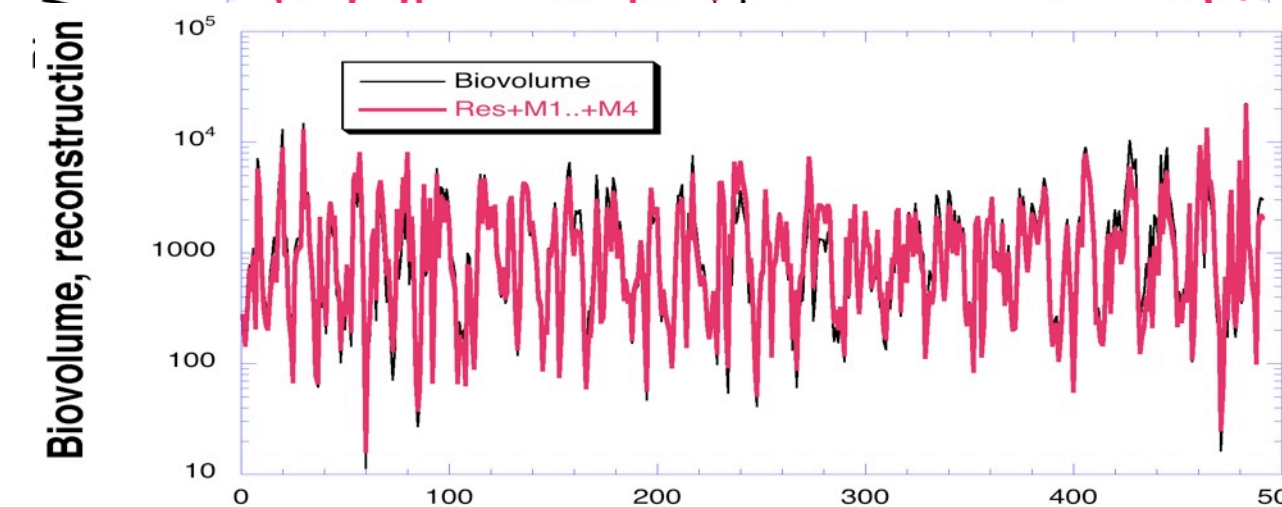
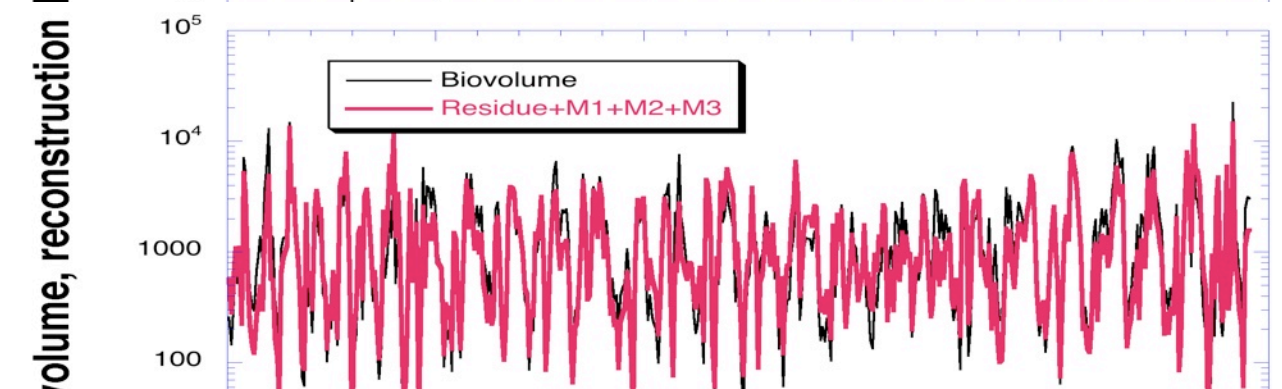
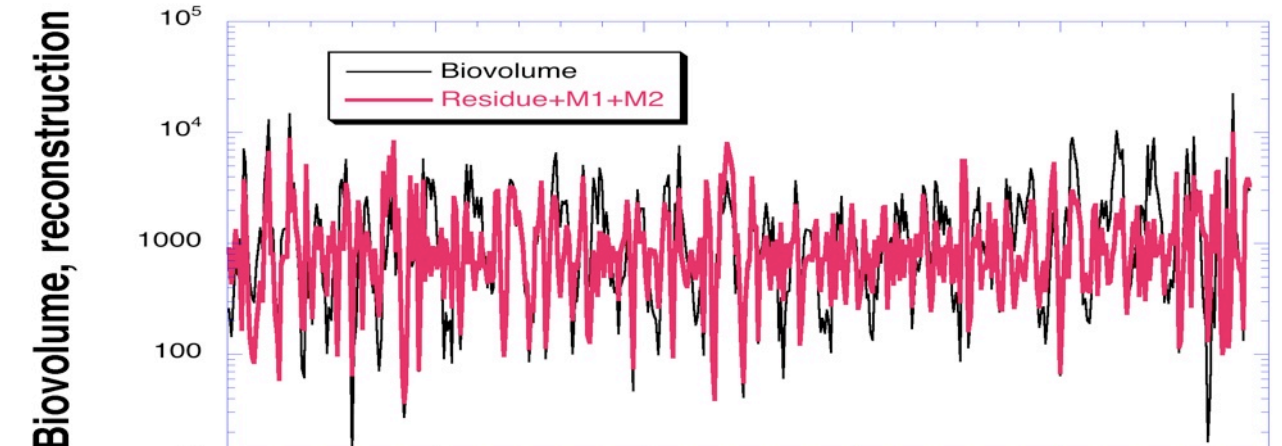
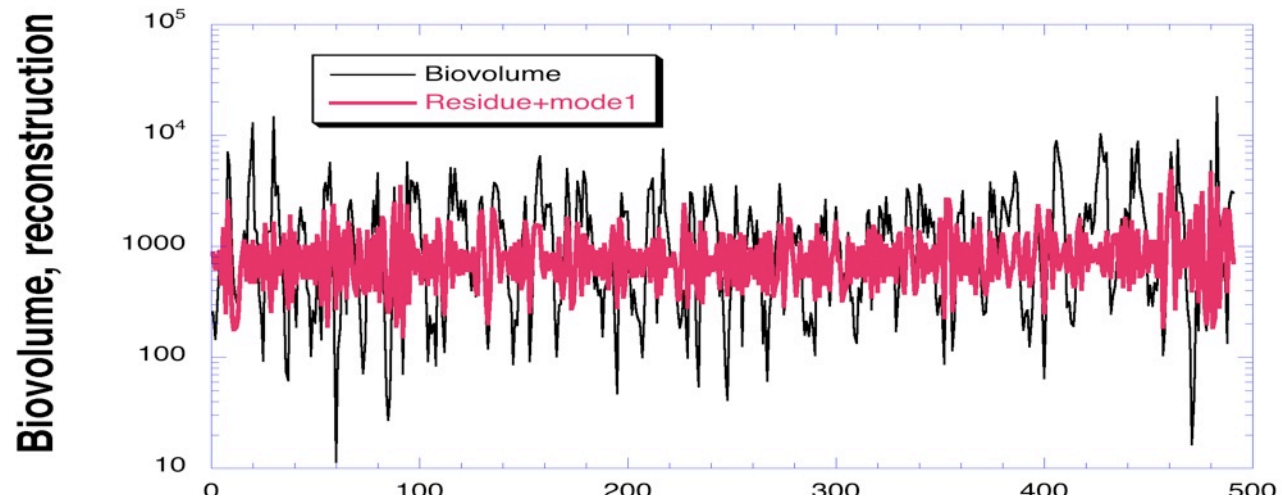
- 7 modes for the decomposition
- Each mode has a smaller characteristic frequency, or **larger characteristic time scale**
- exponential decay of the mean frequency for each mode: **filter bank with scale factor 1.83**

Mode #	Mean scale
1	3.5 months
2	9 months
3	18 months
4	26 months
5	3.5 yr
6	7.9 yr
7	12.8 yr





# Multiplicative mode reconstruction





# Conclusion

- \* Search for universality: **exact and generalizable results** to characterize stochastic fluctuations in geosciences
- \* Scaling regimes/deterministic forcing
- \* Need of **new methods** to study these fields and their **universality**
- \* A Hilbert-based methodology is proposed here
- \* Several high frequency time series shown here as examples

**More info: [www.fg-schmitt.fr](http://www.fg-schmitt.fr)**



# Publications

- Dur, G., F. G. Schmitt, S. Souissi: Analysis of high frequency temperature time series in the Seine estuary from the Marel autonomous monitoring buoy, *Hydrobiologia*, 588, 59-68, 2007.
- Schmitt, F. G., J. C. Molinero, S. Zongo Brizard, Nonlinear dynamics and intermittency in a long term copepod time series, *Comm. Nonlin. Sc. Num. Sim.* 13, 407-415, 2008.
- Schmitt FG, G. Dur, S. Souissi, S.B. Zongo, Statistical properties of turbidity, oxygen and pH fluctuations in the Seine river estuary (France), *Physica A*, 387, 6613-6623, 2008.
- Huang Y., F. G. Schmitt, Z. Lu, Y. Liu, An amplitude-frequency study of turbulent scaling intermittency using Hilbert spectral analysis, *EPL* 84, 40010, 2008.
- Huang, Y. X., Schmitt, F. G., Lu, Z. M., Liu, Y. L.: Analyse de l'invariance d'échelle de séries temporelles par la décomposition modale empirique et l'analyse spectrale de Hilbert *Traitement du Signal* 25, 481-492, 2008.
- Schmitt FG, Y Huang, Z. Lu, Y. Liu, N. Fernandez, Analysis of turbulent fluctuations and their intermittency properties in the surf zone using empirical mode decomposition, *Journal of Marine Systems* 77, 473-481, 2009.
- Huang Y.X., Schmitt FG, Lu, ZM, Liu YL, « Analysis of daily river flow fluctuations using empirical mode decomposition and arbitrary order Hilbert spectral analysis », *Journal of Hydrology* 373, 103-111, 2009.
- Huang YX, Schmitt FG, Lu, ZM, Fougairolles, P., Gagne, Y., Liu YL: Second order structure functions in fully developed turbulence, *Physical Review E* 82, 026319, 2010.
- Huang, Y., F.G. Schmitt, J.-P. Hermand, Y. Gagne, Z. M. Lu, Y.L. Liu, Arbitrary order Hilbert spectral analysis for time series possessing scaling statistics: a comparison study with detrended fluctuation analysis and wavelet leaders, *Physical Review E* 84, 016208, 2011.
- Zongo, S. B, Schmitt, F. G: Scaling analysis of pH fluctuations in coastal waters of the English Channel, *Nonlinear Processes in Geophysics* 18, 829-839, 2011.
- Schmitt, F. G., M. De Rosa, G. Durrieu, M. Sow, P. Ciret, D. Tran, J.-C. Massabuau: Statistical analysis of bivalve high frequency microclosing behavior: scaling properties and shot noise modeling, *International Journal of Bifurcation and Chaos* 21, 3565-3576, 2011.

**More info: [www.fg-schmitt.fr](http://www.fg-schmitt.fr)**