Analysis of daily river flow fluctuations using Empirical Mode Decomposition and arbitrary order Hilbert spectral analysis

Yongxiang Huang, François G Schmitt, Zhiming Lu, Yulu Liu

Analysis of daily river flow fluctuations using Empirical Mode Decomposition and arbitrary order Hilbert spectral analysis

Yongxiang Huang[b,a,*], François G. Schmitt[b], Zhiming. Lu[a], Yulu Liu[a]

[a] *Shanghai Institute of Applied Mathematics and Mechanics, Shanghai University, 200072 Shanghai, China*

[b] *Université des Sciences et Technologies de Lille - Lille 1, CNRS, Laboratory of Oceanology and Geosciences, UMR 8187 LOG, 62930 Wimereux, France*

Abstract

In this paper we presented the analysis of two long time series of daily river flow data, 32 years recorded in the Seine river (France), and 25 years recorded in the Wimereux river (Wimereux, France). We applied a scale based decomposition method, namely Empirical Mode Decomposition (EMD), on these time series. The data were decomposed into several Intrinsic Mode Functions (IMF). The mean frequency of each IMF mode indicated that the EMD method acts as a filter bank. Furthermore, the cross-correlation between these IMF modes from Seine river and Wimereux river demonstrated correlation among the large scale IMF modes, which indicates that both rivers are likely to be influenced by the same maritime climate event of Northern France. As a confirmation we found that the large scale parts have the same evolution trend. We finally applied arbitrary order Hilbert spectral analysis, a new technique coming from turbulence studies and time series analysis, on the flow discharge of Seine river. This new method provides an amplitude-frequency representation of the original time series, giving a joint

*Corresponding author. Tel.: +33 321 992935; Fax: +33 321 992901.

Email addresses: `yongxianghuang@gmail.com` (Yongxiang Huang), `francois.schmitt@univ-lille1.fr` (François G. Schmitt), `zmlu@shu.edu.cn` (Zhiming. Lu), `ylliu@staff.shu.edu.cn` (Yulu Liu)

pdf $p(\omega, \mathcal{A})$. When marginal moments of the amplitude are computed, one obtains an intermittency study in the frequency space. Applied to river flow discharge data from the Seine river, this shows the scaling range and characterizes the intermittent fluctuations over the range of scales from 4.5 to 60 days, between synoptic and intraseasonal scales.

Key words: River flow, Empirical Mode Decomposition, Hilbert Spectral Analysis, scaling, intermittency

---

1. Introduction

A better understanding of river flow fluctuations is of sharp practical importance, e.g. for ecosystem studies (transport properties), and for flood understanding and forecasting. River flows fluctuate on many different scales: at small scales, river turbulence induces stochastic fluctuations and at larger scales (from days to years) the river flow fluctuations are the result of complex nonlinear interactions between rainfall processes, topography and geography (Schumm, 2005). They are also impacted by solar forcing and other large scale variations of the climate system (Mauas et al., 2008). Daily river flow time series thus show fluctuations possessing stochastic properties, as well as deterministic forcing resulting from seasonal or annual meteorological and climatic cycles.

Since Hurst (Hurst, 1951) revealed the long-range dependent property in river flow, associated to a scaling property, researchers have tried different methods to characterize the (multi)scaling properties in river flows (Hurst et al., 1965; Tessier et al., 1996; Pandey et al., 1998; Jánosi and Gallas, 1999; Kantelhardt et al., 2003, 2006; Livina et al., 2003a,b; Koscielny-Bunde et al., 2006; Mauas et al., 2008). Below we quickly review the approaches undertaken in these studies.

Tessier et al. (1996) analysized the relation between rainfall and river flow of 30 rivers and basins in France. They used the double trace moment technique to characterize the multifractal properties. They found that a scaling break occurs at a scale about 16 days. They argued that the rain field itself is the source of the river flow, therefore typical scales in the rain field will also be present in the river flow.

Dahlstedt and Jensen (2005) investigated the Danube and the Mississippi river flows and levels by using finite-size-scaling hypothesis (Aji and Goldenfeld, 2001). They considered the river flow basin size L from different locations. They characterized the multiscaling properties of river flow and level records by considering the relative and general relative scaling (or Extended-Self-Similarity and Generalized Extended-Self-Similarity in the turbulent community). They found that the Fourier spectrum may be different from location to location due to the size effect of the basin area.

More recently, several authors applied the so-called detrended fluctuation analysis (DFA) and its multifractal version to describe the scaling and multiscaling properties of river flows (Kantelhardt et al., 2003; Livina et al., 2003a,a; Kantelhardt et al., 2006; Koscielny-Bunde et al., 2006; Livina et al., 2007; Zhang et al., 2008a,b). Livina et al. (2003a,b) argued that the climate is strongly forced by the periodic variations of the Earth with respect to the state of the solar system. The seasonal variations in the solar radiation cause periodic changes in temperature and precipitations, which eventually lead to a seasonal periodicity of river flows. The Fourier and structure function analyses are impacted by this strong periodicity (Livina et al., 2003a,b; Kantelhardt et al., 2003; Koscielny-Bunde et al., 2006). According to these authors, the DFA approach is an efficient method to eliminate the trend effects.

Koscielny-Bunde et al. (2006) found that the Hurst number H varies from river to river between $0.55 \sim 0.95$ in a non-universal manner independent of the size of the basin. They found that at large time scales, $F_q(s)$ scales as $s^{h(q)}$, and they further proposed a simple function form with two parameters a and b, $h(q) = 1/q - [\ln a^q + b^q]/[q \ln(2)]$ to describe the scaling exponent $h(q)$ of all moments (Kantelhardt et al., 2003). Kantelhardt et al. (2006) also found that the Hurst number H estimated from 99 precipitation and 42 river runoff records data are not consistent with the hypothesis that the scaling is universal with an exponent close to 0.75 (Hurst et al., 1965; Peters et al., 2002).

We consider here a method devoted to deal with any nonlinear time series, which has never been applied to river flow data. In this paper, we apply the empirical mode decomposition (EMD) and the arbitrary order Hilbert spectral analysis (HSA) (Huang et al., 2008), which is an extended version of Hilbert-Huang transform, on river flow discharge fluctuations data. The arbitrary order HSA is a new methodology, which provides the joint pdf of the instantaneous frequency ω and the amplitude $\mathcal{A}$, to characterize the scale invariant properties directly in amplitude-frequency space (Huang et al., 2008). We first introduce the EMD and arbitrary order HSA methodology in section 2. We then present two long records of river flow discharge data from the Seine river and the Wimereux river in section 3. The analysis results are presented in section 4. Section 5 is a discussion and section 6 summarizes the main results of this paper.

2. Methodology

2.1. Empirical Mode Decomposition

The starting point of the EMD is that most of the signals are multi-components, which means that there exist different scales simultaneously (Cohen, 1995; Huang et al., 1998, 1999). The signal can be considered as a superposition of fast oscillations to slower ones at a very local level (Rilling et al., 2003; Flandrin and Gonçalvès, 2004). Time series analysis methods generally consider a characteristic scale explicitly or implicitly. For example, the Fourier analysis characterizes the scale by the length of one period of sine (or cosine) wave. Then an integration operator is applied to extract the components information. Fourier analysis is thus an energy based method: only when the component contains enough energy, it can be detected by such method (Huang et al., 1998; Huang, 2005). The characteristic scale for the present EMD approach is defined as the distance between two successive maxima (or minima) points. This scale based definition gives the EMD a very local ability (Huang et al., 1998, 1999). According to the above definition of a characteristic scale, the so-called Intrinsic Mode Function (IMF) is then proposed to approximate the mono-component signal, which satisfies the following two conditions: (i) the difference between the number of local extrema and the number of zero-crossings must be zero or one; (ii) the running mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

The Empirical Mode Decomposition algorithm is proposed to extract the IMF modes from a given time series (Huang et al., 1998, 1999; Flandrin et al., 2004). The first step of the EMD algorithm is to identify all the local maxima (resp. minima) points for a given time series $x(t)$. Once all the local maxima points are identified, the upper envelope $e_{\max}(t)$ (resp. lower envelope $e_{\min}(t)$) is constructed by a cubic spline interpolation. The mean between these two envelopes is defined as $m_1(t) = (e_{\max}(t) + e_{\min}(t))/2$. The first component is estimated by $h_1(t) = x(t) - m_1(t)$. Ideally, $h_1(t)$ should be an IMF as expected. In reality, however, $h_1(t)$ may not satisfy the condition to be an IMF. We take $h_1(t)$ as a new time series and repeat the shifting process j times, until $h_{1j}(t)$ is an IMF. We thus have the first IMF component $C_1(t) = h_{1j}(t)$ and the residual $r_1(t) = x(t) - C_1(t)$ from the data $x(t)$. The shifting procedure is then repeated on residuals until $r_n(t)$ becomes a monotonic function or at most has one local extreme point. This means that no more IMF can be extracted from $r_n(t)$. Thus, with this algorithm we finally have $n-1$ IMF modes with one residual $r_n(t)$. The original data $x(t)$ is then rewritten as

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

A stopping criterion has to be introduced in the EMD algorithm to stop the shifting process (Huang et al., 1998, 1999; Rilling et al., 2003; Huang et al., 2003; Huang, 2005). However, this is beyond our topic here: for more details about the EMD method, we refer to Huang et al. (1998, 1999); Rilling et al. (2003); Huang et al. (2003); Flandrin et al. (2004); Flandrin and Gonçalvès (2004); Huang (2005); Rilling and Flandrin (2008).

2.2. Hilbert Spectra Analysis

Having obtained the IMF modes from the EMD algorithm, one can apply the associated Hilbert Spectral Analysis (HSA) (Cohen, 1995; Long et al., 1995; Huang et al., 1998, 1999) to each IMF component C_i to extract the energy-time-frequency information from the data. The Hilbert transform is written as

$$\tilde{C}_i(t) = \frac{1}{\pi} P \int_0^{\infty} \frac{C_i(t')}{t - t'} \,\mathrm{d}t' \tag{2}$$

where P means the Cauchy principle value (Cohen, 1995; Long et al., 1995). From this, we can construct the analytic signal, $\mathbb{C}_i(t)$, defined as

$$\mathbb{C}_i(t) = C_i(t) + j\tilde{C}_i(t) = \mathcal{A}_i(t)\mathrm{e}^{j\theta_i(t)} \tag{3}$$

in which $\mathcal{A}_i(t) = |\mathbb{C}_i(t)| = [C_i(t)^2 + \tilde{C}_i^2(t)]^{1/2}$ and $\theta_i(t) = \arctan(\tilde{C}_i(t)/C_i(t))$. Hence the corresponding instantaneous frequency can be defined as

$$\omega_i = \frac{\mathrm{d}\theta_i(t)}{\mathrm{d}t} \tag{4}$$

The original signal is then finally represented as (excluding the residual $r_n(t)$)

$$x(t) = \mathrm{RP} \sum_{i=1}^{N} \mathcal{A}_i(t) e^{j\theta_i(t)} = \mathrm{RP} \sum_{i=1}^{N} \mathcal{A}_i(t) e^{j\int \omega_i(t)\,\mathrm{d}t} \tag{5}$$

where RP means real part. This approach allows Frequency-Modulation and Amplitude-Modulation simultaneously (Huang et al., 1998, 1999). Then the Hilbert spectrum $H(\omega, t) = \mathcal{A}^2(\omega, t)$ is introduced, representing the energy in a time frequency space, and we define the Hilbert marginal spectrum as

$$h(\omega) = \int_0^{\infty} H(\omega, t) \,\mathrm{d}t \tag{6}$$

This is similar to the Fourier spectrum, since its corresponds to the energy associated to the frequency (Huang et al., 1998, 1999).

The combination of EMD and HSA is also sometimes called Hilbert-Huang transform. It provides an alternative powerful tool to analysize non-stationary and nonlinear time series. The main advantage of EMD over traditional approaches is its complete self-adaptiveness and its very local ability both in physical space and frequency space. Therefore it is especially suitable for nonlinear and nonstationary time series analysis (Huang et al., 1998, 1999). Since its introduction, this method has attracted a large interest (Huang, 2005). It has been shown to be an efficient method to separate a signal into a trend and small scale fluctuations on a dyadic bank (Wu and Huang, 2004; Flandrin et al., 2004; Flandrin and Gonçalvès, 2004); it has also been applied to many fields including physiology (Su et al., 2008), geophysics (Jánosi and Müller, 2005), climate studies (Solé et al., 2007), mechanical engineering (Chen et al., 2004), acoustics (Loutridis, 2005), aquatic environment (Schmitt et al., 2007, 2008) and turbulence (Huang et al., 2008), to quote a few.

2.3. Arbitrary Order Hilbert Spectral Analysis

HSA represents the energy-time-frequency information $H(\omega, t)$ at a very local level. We can then define the joint pdf $p(\omega, \mathcal{A})$ of the instantaneous frequency ω and the amplitude $\mathcal{A}$ for all the IMF modes (Long et al., 1995; Huang et al., 2008). Thus the corresponding Hilbert marginal spectrum is rewritten as

$$h(\omega) = \int_0^\infty p(\omega, \mathcal{A}) \mathcal{A}^2 \, \mathrm{d}\mathcal{A} \tag{7}$$

The above equation is no more than the second order statistical moment. In a recent paper we have generalized the above definition into arbitrary order moment, which is written as (Huang et al., 2008)

$$\mathcal{L}_q(\omega) = \int_0^\infty p(\omega, \mathcal{A})\mathcal{A}^q \,\mathrm{d}\mathcal{A} \tag{8}$$

where $q \geq 0$. In case of scale invariance we can write

$$\mathcal{L}_q(\omega) \sim \omega^{-\xi(q)} \tag{9}$$

where $\xi(q)$ is the corresponding scaling exponent. Due to the integration, $\xi(q) - 1$ corresponds to $\zeta(q)$ the scaling exponent of structure functions, which is classically written as

$$\langle \Delta x_\tau^q \rangle \sim \tau^{\zeta(q)} \tag{10}$$

where $\Delta x_\tau = |x(t+\tau) - x(t)|$ is the amplitude of the increments at scale τ.

We provide here some comments on the arbitrary order HSA methodology. If one represents the structure function analysis in Fourier space, one may find that it measures the scale invariance by an indirect way, which is influenced by the trend or strong large scales. The increment operates the data on very local level in the physical domain, but nevertheless, it is still a global operator in the frequency domain. On the contrary, the present methodology has completely self-adaptiveness and very local ability both in the physical and frequency domains (Huang et al., 2008).

3. Seine River and Wimereux River

The Seine river is the third largest river in France. Its length is 776 km, and its basin is 78650 km^2. It is economically important for France, with 25%

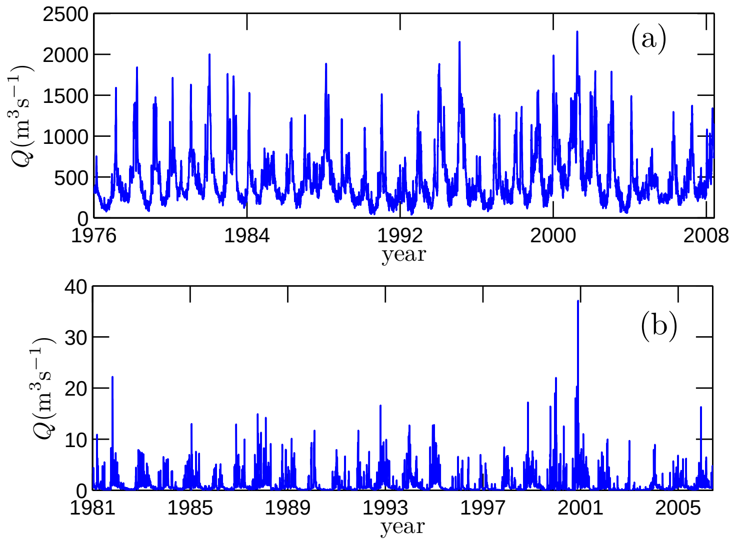

Figure 1: The river flow discharge time series of (a) Seine River, recorded from 1 January 1976 to 28 April 2008, (b) Wimereux river, recorded from 1 January 1981 to 27 May 2006. The data illustrate clear strong annual cycles with huge fluctuations. The total lengths are 11828 and 9278 data points for the Seine river and the Wimereux river, respectively.

of its population as well as 40% of its industry and agriculture concentrated in and around it (Dauvin, 2007). The flow data is provided by the Service de Navigation de la Seine (SNS). This corresponds to daily flow data Q ($\mathrm{m^3 s^{-1}}$), recorded from 1 January 1976 to 28 April 2008. There are 11828 data values, with some missing values due to interruptions for maintenance or because of the failure of measuring devices. Due to the local ability of HSA approach, which is performed through spline interpolation, the missing values in the time series do not change the results, since the method can be applied even for irregular sampling. The data are shown in Fig. 1 (a), demonstrating some large fluctuations at all scales. The mean and standard deviation of the discharge are $488\,\mathrm{m^3 s^{-1}}$ and $349\,\mathrm{m^3 s^{-1}}$, respectively. This figure shows a

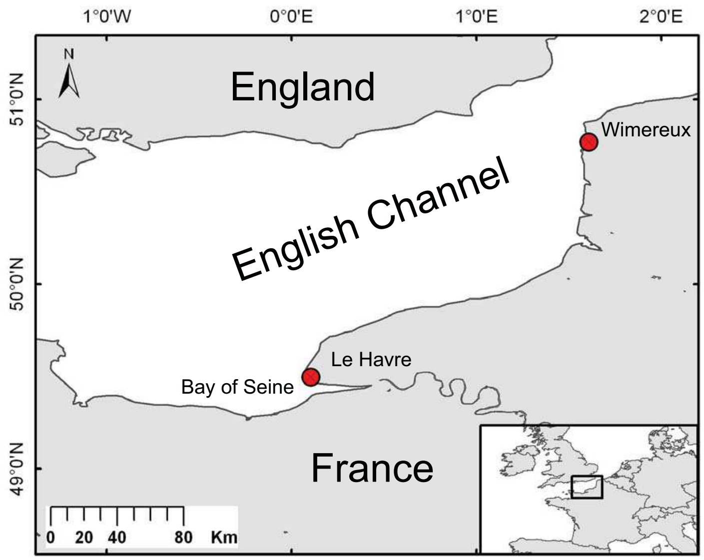

Figure 2: A map showing the location of the Seine river and the Wimereux river, in the eastern English Channel. The distance between them about 300 km.

complex and stochastic behavior, with a visible strong annual cycle.

The Wimereux river is a small river in the North of France. Its length is 22 km, and its basin is 78 km^2. It can have strong fluctuations due to fast increase of the flow in case of heavy rain. The daily flow discharge is recorded from 1 January 1981 to 27 May 2006, with a total length of 9278 points values with some missing, see Fig. 1 (b). The mean and standard deviation of the discharge data are $1.02\,\mathrm{m^3s^{-1}}$ and $1.73\,\mathrm{m^3s^{-1}}$.

Figure 2 shows the location of these two rivers, where the Seine river is represented as a solid line. The Wimereux river is too small to be displayed in the same figure. The difference between these two rivers is clear: the Seine river is a real big one, and the Wimereux river is much smaller and strongly influenced by the local rainfall conditions. The distance between them is about 300 km, see Fig. 2. Both of them are affected by the same large scale climatic factors and belong to the marine west coast climate of Northern France. This climate is found on the west coast of middle latitude regions and can be quite humid. Indeed it is subject to western wind bringing important variability intermittent clouds, important precipitation and temperate temperatures. The direct estimation of the cross correlation between these two recorded data is about 0.256, a value that may be contaminated by the small scale uncorrelated fluctuations. We will apply to these two data sets by the EMD method in the following section.

4. Results

4.1. EMD results

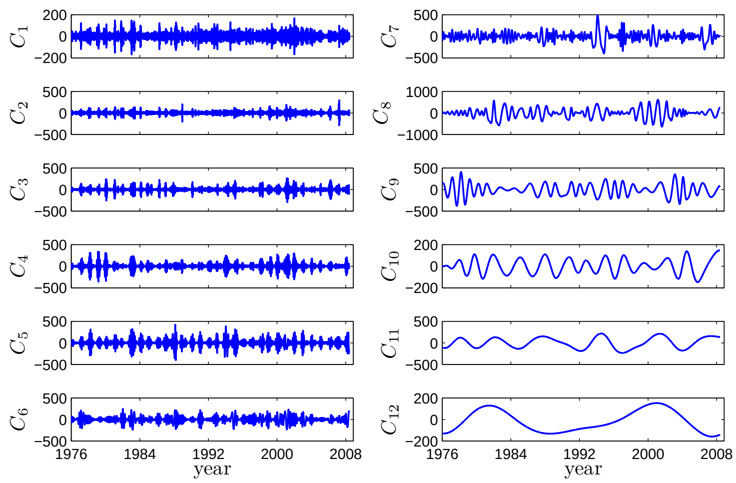

Figure 3: IMF modes (excluding the residual) from EMD for the Seine river. Here the data are taken from 1 January 1976 to 28 April 2008. The characteristic scale is increasing with the mode index number n.

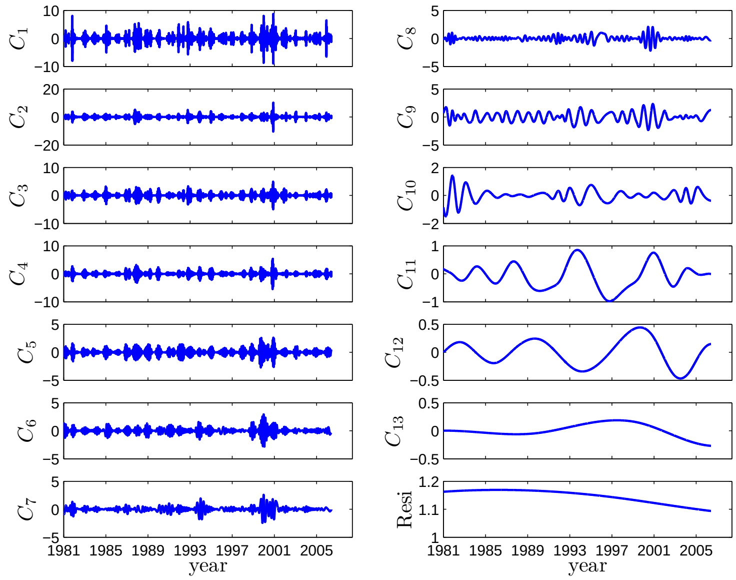

Figure 4: IMF modes from EMD for the Wimereux river. Here the data are taken from 1 January 1981 to 27 May 2006.

After the application of the EMD method, the original data is separated into several IMF modes. We then represent the IMF modes in Fig. 3 and Fig. 4 for the Seine river and the Wimereux river, respectively. For display convenience, we exclude the residual for the Seine river. Graphically, one can see that the characteristic scale is increasing with the mode index n. Let us note that the number of IMF modes is produced by the algorithm and depends on the length and the complexity of the data. In practice, based on the dyadic filter bank property of the EMD method, this number is usually less than $\log_2(N)$, where N is the length of the data (Flandrin and Gonçalvès, 2004; Flandrin et al., 2004; Wu and Huang, 2004; Huang et al., 2008). First, we estimate the mean frequency $\overline{\omega}$ of each IMF mode. We use the following three definitions of mean frequency $\overline{\omega}$. The first one is proposed

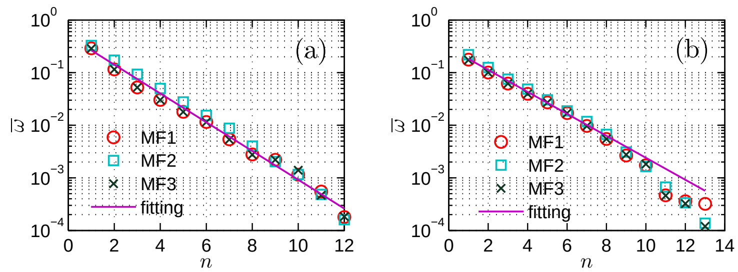

Figure 5: Representation of the mean frequency $\overline{\omega}$ vs the mode index n in log-linear view: (a) Seine river, (b) Wimereux river, where the mean frequency $\overline{\omega}$ are estimated by using Eqs. (11) (○), (12) (□) and (13) (×), respectively. An exponential law is observed for each representation. The straight line is the least square fit of the data.

Table 1: The mean period (in days) of each IMF mode (excluding the residual) of the Seine river and the Wimereux river, respectively. Here the mean period is estimated as $\overline{T} = 1/\overline{\omega}$, where $\overline{\omega}$ is calculated by Eq. (11). The 8^{th} and 9^{th} IMF modes of the Seine river and Wimereux river, respectively, are close to the annual cycle.

IMF	1	2	3	4	5	6	7	8	9	10	11	12	13
Seine	3	8	19	33	55	86	185	358	452	869	1823	5551	
Wimereux	5	9	16	25	36	58	103	182	376	574	2149	2785	3125

by Huang (Huang et al., 1998), which is written as

$$\overline{\omega}_i = \frac{\int_0^\infty f S_i(f)\,\mathrm{d}f}{\int_0^\infty S_i(f)\,\mathrm{d}f} \tag{11}$$

where $S_i(f)$ is Fourier spectrum of $\mathbb{C}_i$. It is an energy weighted average in Fourier space. The second one is given by Flandrin (Flandrin et al., 2004; Flandrin and Gonçalvès, 2004), which is written as

$$\overline{\omega}_i = \frac{\mathbb{N}^0 - 1}{\mathbb{L}^0} \tag{12}$$

where $\mathbb{N}^0$ is the zero-crossing number, and $\mathbb{L}^0$ is the distance between the first and last zero-crossing. The third one is introduced here for the first

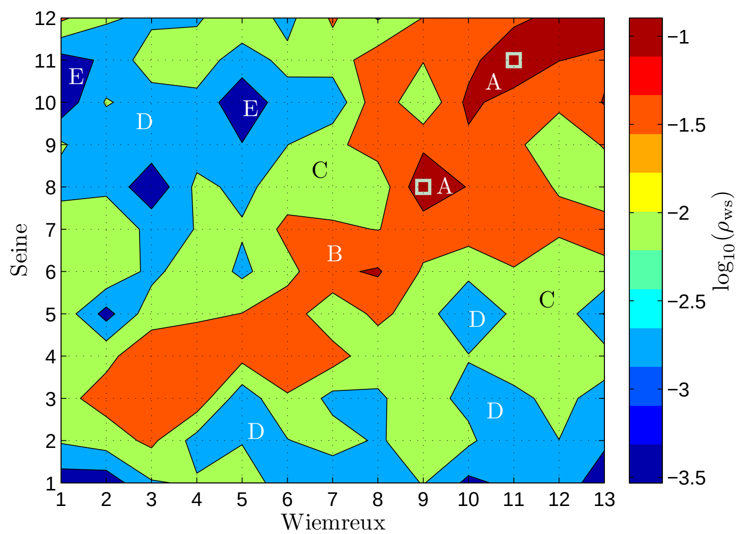

Figure 6: Representation of the cross-correlation $\rho_{\rm ws}$ between IMF modes from the Seine and Wimereux rivers. The data span is taken from 1 January 1981 to 27 May 2006 for both series. For convenience, we consider the coefficient value $\log_{10}(\rho_{\rm ws}(i,j))$. As expected, the annual cycle shows a strong correlation with a coefficient $\rho_{\rm ws}(9,8) = 0.426$. The coefficient of the most correlated modes is $\rho_{\rm ws}(11,11) = 0.579$. These two strong correlations are then marked by $\square$.

time, and is defined as

$$\overline{\omega}_i = \frac{\int_0^\infty \omega h_i(\omega)\,\mathrm{d}\omega}{\int_0^\infty h_i(\omega)\,\mathrm{d}\omega} \tag{13}$$

where $h_i(\omega)$ is the Hilbert marginal spectrum for the $i^{\rm th}$ mode. This definition is similar to the first one: it is an energy weighted measurement of the mean frequency in Hilbert space. We then represent the mean frequency $\overline{\omega}$ estimated by these three definitions (11) ($\bigcirc$), (12) ($\square$) and (13) ($\times$) for each mode in Fig. 5 for (a) the Seine river, and (b) the Wimereux river. One can see that the two energy weighted estimators give almost the same mean frequency. However, they are slightly smaller than the zero-crossing

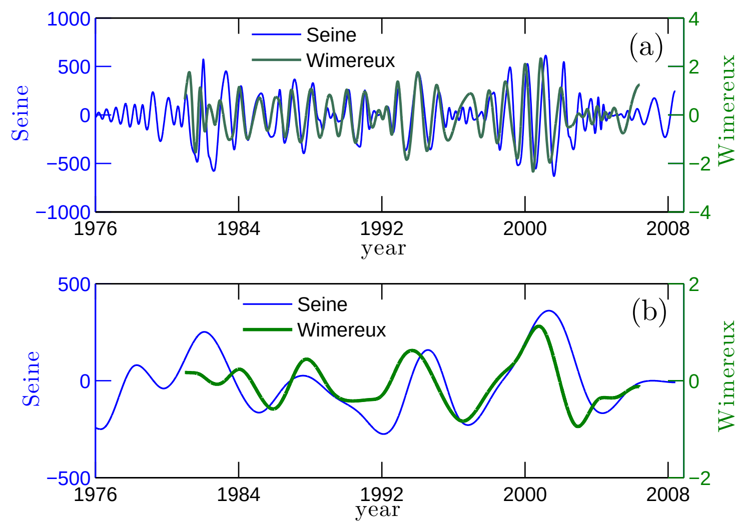

Figure 7: Most correlated IMF modes: (a) the annual cycle mode for the Seine river (thin solid line) and the Wimereux river (thick solid line), (b) the reconstruction of the large scale part for the Seine river (thin solid line) and the Wimereux river (thick solid line). We took the IMF modes $11 \sim 12$ from the Seine river and $11 \sim 13$ from the Wimereux river, which means periods larger than 3 years, to reconstruct the large scale part. Graphically, they have the same evolution trend on range 1 January 1981 to 28 May 2006.

based estimator. Graphically, all these three estimators suggest the following exponential law

$$\overline{\omega}(n) \sim \gamma^{-n} \tag{14}$$

where $\gamma_{\mathrm{s}} \simeq 1.88$, $\gamma_{\mathrm{w}} \simeq 1.62$ are estimated by using the least square fitting for the Seine river and the Wimereux river, respectively. This result implies that the mean frequency of a given mode is γ times larger than the mean frequency of next one. We notice that these values are significantly different from 2, which would correspond to a dyadic filter bank, which are reported for white noise (Wu and Huang, 2004), fractional Gaussian noise (Flandrin et al., 2004; Flandrin and Gonçalvès, 2004) and turbulence time series (Huang et al., 2008). However, it still indicates that the EMD algorithm acts a filter bank here.

We list the mean period $\overline{T}$ (in days) in Table 1, where $\overline{T} = 1/\overline{\omega}$. Since the three above mentioned mean frequency estimators give almost the same value, we thus only present the value estimated by Eq. (11). One can find that the EMD approach captures the annual cycle, which is the 8^{th} and 9^{th} mode for the Seine river and Wimereux river, respectively. Both rivers belong to the same climate and it is expected that large scale modes are correlated. However, the data at daily scale are not (the cross-correlation at this scale is 0.256); this is due to the influence of small scales. The cross-correlation between two IMF modes is defined as

$$\rho_{\text{ws}}(i,j) = \frac{\langle C_{\text{w},i} C_{\text{s},j} \rangle}{\langle C_{\text{w},i}^2 \rangle^{1/2} \langle C_{\text{w},i}^2 \rangle^{1/2}} \tag{15}$$

where $\langle \cdot \rangle$ means ensemble average. The corresponding cross-correlation $\rho_{\text{ws}}(i,j)$ is then plotted in Fig. 6, where the most correlated modes are marked by $\Box$. The large scale modes are correlated as expected. More precisely, we observe a larger cross-correlation between the annual cycle modes, $\rho_{\text{ws}}(9,8) = 0.426$, and the most correlation coefficient is $\rho_{\text{ws}}(11,11) = 0.579$, with mean periods of about 6 and 8 years for the Seine river and the Wimereux river, respectively.

We then replot the annual cycle for the Seine river (thin solid line) and Wimereux river (thick solid line) in Fig. 7 (a). One can find that their shapes are almost the same on the range from 1 January 1981 to 28 May 2006. We also reconstruct the large scale signal from those modes, with mean period larger than 3 years, 11^{th} and 12^{th} from the Seine river (thin solid line), and 11^{th} to 13^{th} from the Wimereux river (thick solid line). The result is shown in Fig. 7 (b). Graphically, they have almost the same shape and evolution trend.

4.2. Arbitrary order HSA results

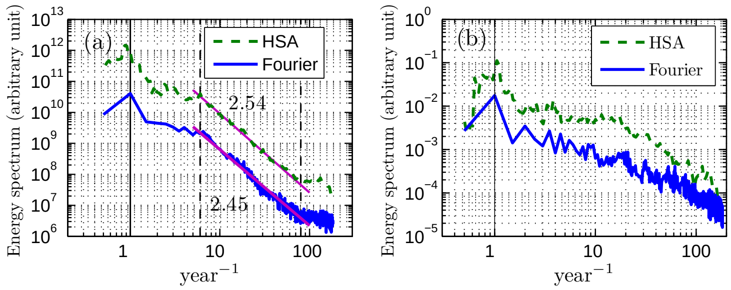

Figure 8: Comparison of the Hilbert marginal spectrum (dashed line) and Fourier spectrum (solid line) for (a) the Seine river, (b) the Wimereux river. For the Seine river, a power law behaviour is observed on the range $6 < \omega < 80\,\text{year}^{-1}$, or 4.5 $\sim$ 60 days: this range is marked by the vertical dashed lines. The scaling values are 2.54 and 2.45 for Hilbert spectrum and Fourier spectrum, respectively. The vertical solid line indicates the annual cycle.

In order to characterize the intermittent properties of river flow fluctuations, we consider here HSA and arbitrary order HSA analysis. We firstly compare the Hilbert marginal spectrum (dashed line) and Fourier spectrum (solid line) in Fig. 8 for (a) the Seine river, and (b) the Wimereux river to identify the power law range, where the scale invariance holds. For the Seine river, both methods capture the annual cycle (vertical solid line) and show power law behaviour on the range $6 < \omega < 80$ year^{-1} or from 4.5 to 60 days, with scaling exponent 2.54 and 2.45, respectively. The power law range is between synoptic and intraseasonal scales (Zhang, 2005). The latter may be linked to the Madden-Julian Oscillation (MJO), since some connection be-

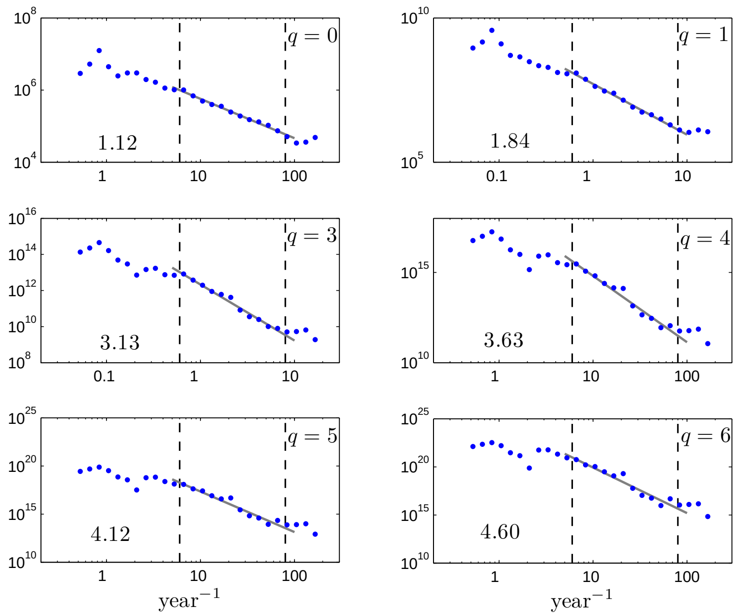

Figure 9: Representation of arbitrary order Hilbert marginal amplitude spectra $\mathcal{L}_q(\omega)$ for the Seine river, where $q = 0$, 1, 3, 4, 5 and 6. A power law behaviour is observed in all cases on the range $6 < \omega < 80\,\mathrm{year}^{-1}$. The vertical dashed lines indicate the power law range. The corresponding scaling values are shown in each figure.

tween and the North Atlantic Oscillation (NAO) and MJO have been found (Cassou, 2008). For the Wimereux river, the power law range is less clear. We therefore only apply below the arbitrary order HSA analysis on the Seine river.

Since we are concerned with the scaling property in the above range, we thus divide the entire time series into 16 segments, each one has 2×365 points, 2 years each. The arbitrary order Hilbert marginal spectra are shown in Fig. 9, where $q = 0$, 1, 3, 4, 5 and 6. Power law behaviour is then observed in all cases on range $6 < \omega < 80\,\mathrm{year}^{-1}$, graphically. The corresponding scaling exponents $\xi(q)$ are estimated on this range by using least square

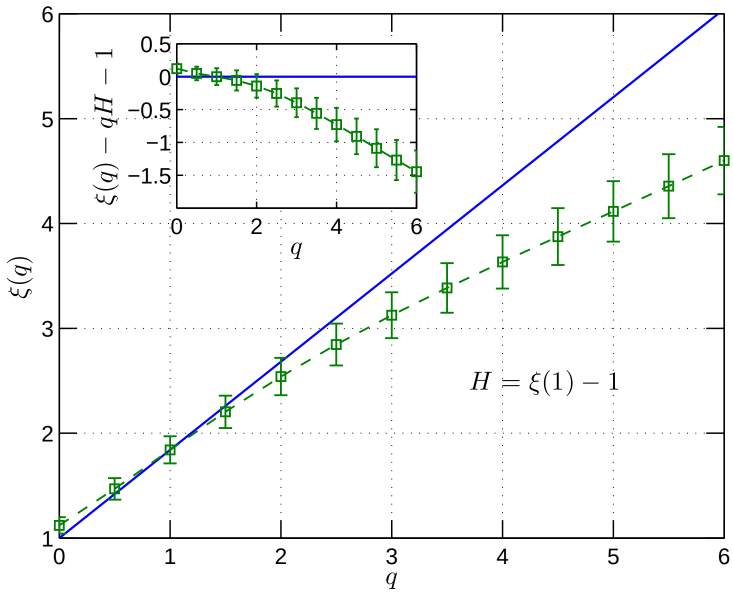

Figure 10: Scaling exponents $\xi(q)$ ($\bigcirc$) for the Seine river. The inset shows the departure from the reference line $qH + 1$, where $H = \xi(1) - 1$. The shape of these scaling exponents is concave, which indicates the small scale intermittency nature of river flow.

fitting with 95% confidence limit, Fig. 10 shows the scaling exponents $\xi(q)$ ($\bigcirc$). Graphically, this curve is concave, which indicates the multifractal properties of the river flow discharge (Pandey et al., 1998; Kantelhardt et al., 2003, 2006). For comparison, we also show a reference line $qH+1$ (solid line), where $H = \xi(1) - 1 = 0.84 \pm 0.08$, which corresponds to the mono-scaling case. The departure from this reference mono-scaling line is then shown in inset.

5. Discussions

We compare the above observation with the conventional structure function analysis, the traditional way to extract the scaling exponents. We plot

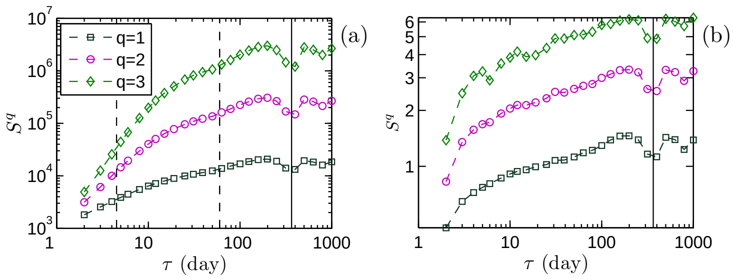

Figure 11: Structure function for (a) the Seine river, and (b) the Wimereux river, where $q = 1$ (□), 2 (○) and 3 (◇). The vertical dashed lines indicate the range $4.5 \sim 60$ days. The annual cycle influence is also indicated by the solid line.

the result in Fig. 11, where $q = 1$ (□), 2 (○) and 3 (◇), respectively. Some scaling portion are visible on these figures, of a relatively limited amplitude. To reveal the scale invariance more clearly, we consider the Extended Self-Similarity (ESS) properties, a relative scaling expressed as

$$\langle \Delta x^q \rangle \sim \langle \Delta x \rangle^{\psi(q)} \tag{16}$$

where in case of scaling (Eq. 10), we have $\zeta(q) = H\psi(q)$. Eq (16) can be used to estimate more accurately the exponents $\psi(q)$. The ESS is verified for the Seine river on range $2 < \tau < 60$ days, see Fig 12. Figure 13 shows the ESS result for the Wimereux river. Graphically, it is scaling and is rather scattered. We then show the relative scaling exponents $\psi(q)$ and the normalized scaling exponents $(\xi(q) - 1)/(\xi(1) - 1)$ in Fig 14. In the mono-scaling case and when there is no large scale forcing, they should collapse on a solid line $\psi(q) = q$. The same approach is applied to the Wimereux river. In this case the HSA approach is not displaying any clear scaling range. We thus use the ESS approach and compare the resulting curve $\psi(q)$ to the one obtained from the Seine river. The Wimereux river scaling

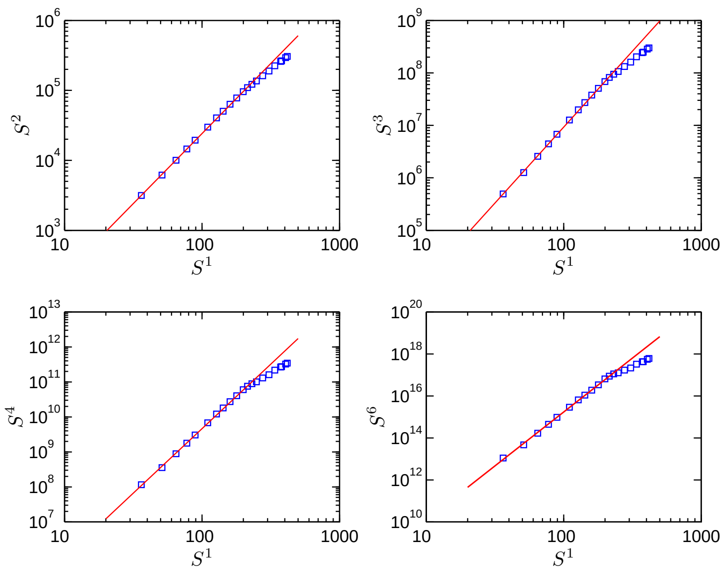

Figure 12: Extended self-similarity test of the Seine river on range $2 < \tau < 300$ day. The relative scaling is very well captured for all moments.

exponents are saturating at $\psi(q = 1)$, and the curve is quite different from the Seine river. This shows that the Wimereux river is more intermittent than the Seine river: which may come from the fact that its catchment basin is much smaller, hence its discharge variation can be more rapid. This may also be an effect of strong oscillations that reduce the multifractal degree (see Telesca et al. (2004b); Bolzan et al. (2009)). It is also interesting to see in the same graph the difference between the HSA based exponents and structure function's exponents for the Seine river. The discrepancy can be interpreted as coming from the influence of the periodic component in the time series. Indeed we have shown elsewhere (Huang et al., 2009) that the influence of periodic components is stronger on structure function than on HSA exponents, which can be linked to the fact that EMD acts a filter

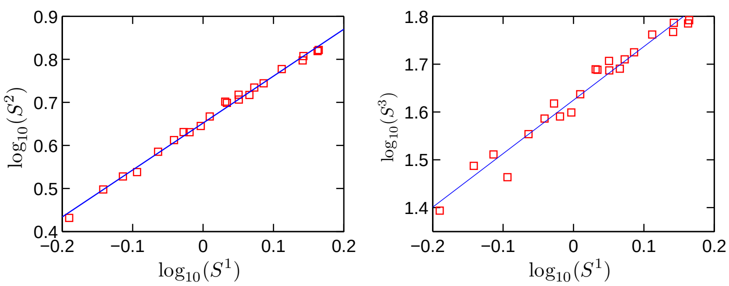

Figure 13: Extended self-similarity test of the Wimereux river on range $2 < \tau < 300$ day.

bank (Flandrin and Gonçalvès, 2004; Flandrin et al., 2004; Huang et al., 2008; Wu and Huang, 2004). Periodic components tend to increase the value of $\zeta(q)$ relative to the real theoretical curve.

6. Conclusions

In this paper we applied for the first time the EMD methodology to river flow time series. Using daily river flow discharge data, 32 years recorded in the Seine river (France), and 25 years recorded in the Wimereux river (France), we have shown that the time series can be successfully separated into several IMF modes. Exponential laws for the mean frequency of each mode have been found, with exponents $\gamma_{\rm s} = 1.88$ and $\gamma_{\rm w} = 1.62$ for the Seine river and the Wimereux river, respectively. These values are smaller than 2, the value for dyadic filter bank. Even though, it still confirmed that the EMD algorithm acts as a filter bank for river flow data. Furthermore, strong cross-correlation have been observed between annual cycles and the large scale modes having a mean period larger than 3 years. Based on the correlation analysis results, we have found that the annual cycle mode and the reconstructed large scale part have almost the same evolution trend.

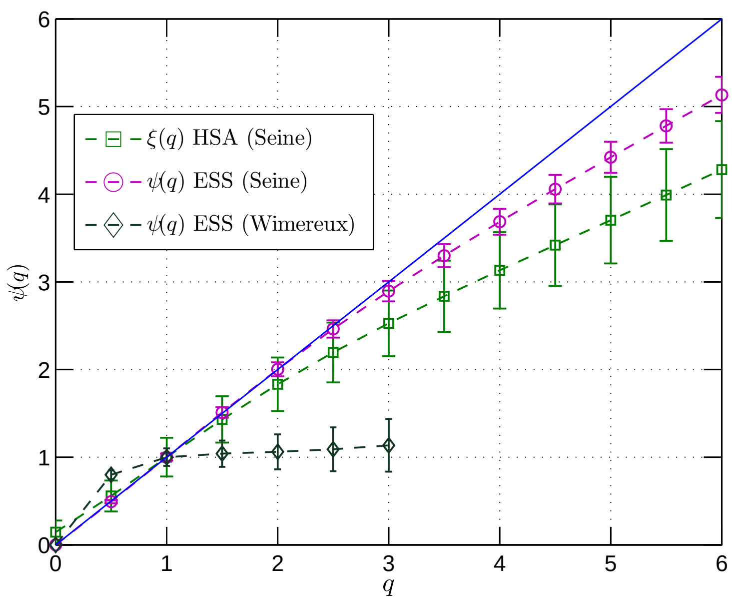

Figure 14: Comparison of the relative scaling exponent $\psi(q)$ ($\square$) and $(\xi(q)-1)/(\xi(1)-1)$ ($\bigcirc$).

We have also characterized the intermittency of the time series over the ranges showing scaling properties. For the Seine river, we observed power laws for the first six order Hilbert marginal spectra on the range $6 < \omega < 80\,\mathrm{year}^{-1}$ or 4.5$\sim$60 days, between synoptic and intraseasonal scales. The corresponding scaling exponents $\xi(q)$ indicate the small scale multifractal nature of the river flow data analyzed here. The differences obtained using the structure functions approach and the frequency based HSA approach have been emphasized, which is especially clear for large order moments associated to the more active fluctuations. We have interpreted this difference as coming from the strong annual cycle which has more influence on structure functions scaling exponents than on the HSA approach. We have also compared the scaling exponents estimated from the ESS method, for the Seine river and

Wimereux river; the much smaller exponents obtained for the Wimereux river express a higher degree of multifractality, which was interpreted as coming from the inertia associated to the large scale basin for the Seine river, whereas small rivers such as the Wimereux river may be more sensitive to local precipitation events.

Several previous studies have considered scaling properties of river flows using other methods such as rescaled range analysis, trace moments, double trace moments, wavelet analysis, multifractal detrended fluctuation analysis (MFDA). We applied here a new method which gives results similar to the classical methods (structure functions, wavelet analysis, MFDA) for fractional Brownian motion or pure multifractal processes Huang et al. (2009). However, we have shown in the same paper that strong deterministic forcing had important influence on classical methods, whereas the HSA approach was much more stable and presented less influence Huang et al. (2009). This method seems hence more appropriate for environmental time series that possess often strong periodic components superposed to scaling regimes. The origin of this stability property is the adaptative and local approach which is at the heart of the HSA method.

We have compared here two rivers of very different size and catchment basin in order to compare their scaling properties. One of the objectives of scaling analyses of river flow time series is indeed to detect some differences among rivers, but also to evaluate some universality, i.e. some general similarity in statistical properties. This was done for normalized pdfs Dahlstedt and Jensen (2005), for river flow volatilities Livina et al. (2003a,b), and for scaling regimes Tessier et al. (1996) or multifractal parameters Pandey et al.

(1998). We hope that the method presented in this paper, which we claim to be well adapted to environmental time series, will help this quest for universal properties of river flow scaling statistics.

Acknowledgments

This work is supported in part by the National Natural Science Foundation of China (No.10672096 and No.10772110) and the Innovation Foundation of Shanghai University. Y. H. is financed in part by a Ph.D. grant from the French Ministry of Foreign Affairs and by part from university of Lille 1. Financial support was provided by the program Seine Aval: this work is a contribution to ZOOSEINE project funded by Seine-Aval IV program. The EMD Matlab codes used in this paper are written by P. Flandrin from laboratoire de Physique, CNRS & ENS Lyon (France): http://perso.ens-lyon.fr/patrick.flandrin/emd.html.

References

Aji, V., Goldenfeld, N., 2001. Fluctuations in Finite Critical and Turbulent Systems. Phys. Rev. Lett. 86 (6), 1007–1010.

Bolzan, M., Rosa, R., Sahai, Y., 2009. Multifractal analysis of low-latitude geomagnetic fluctuations. Ann. Geophys 27, 569–576.

Cassou, C., 2008. Intraseasonal interaction between the Madden–Julian Oscillation and the North Atlantic Oscillation. Nature 455 (7212), 523–527.

Chen, J., Xu, Y. L., Zhang, R. C., 2004. Modal parameter identification of

Tsing Ma suspension bridge under Typhoon Victor: EMD-HT method. J. Wind Eng. Ind. Aerodyn. 92 (10), 805–827.

Cohen, L., 1995. Time-frequency analysis. Prentice Hall PTR Englewood Cliffs, NJ.

Dahlstedt, K., Jensen, H., 2005. Fluctuation spectrum and size scaling of river flow and level. Physica A 348, 596–610.

Dauvin, J. C., 2007. Paradox of estuarine quality: Benthic indicators and indices, consensus or debate for the future. Mar. Pollut. Bull. 55 (1-6), 271–281.

Flandrin, P., Gonçalvès, P., 2004. Empirical Mode Decompositions as Data-Driven Wavelet-Like Expansions. Int. J. Wavelets, Multires. Info. Proc. 2 (4), 477–496.

Flandrin, P., Rilling, G., Gonçalvès, P., 2004. Empirical mode decomposition as a filter bank. IEEE Sig. Proc. Lett. 11 (2), 112–114.

Huang, N. E., 2005. Hilbert-Huang Transform and Its Applications. World Scientific, Ch. 1. Introduction to the hilbert huang trasnform and its related mathmetical problems, pp. 1–26.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N., Tung, C. C., Liu, H. H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. R. Soc. London 454 (1971), 903–995.

Huang, N. E., Shen, Z., Long, S. R., et al., 1999. A new view of nonlinear water waves: The Hilbert Spectrum . Annu. Rev. Fluid Mech. 31 (1), 417–457.

Huang, N. E., Wu, M. L., Long, S. R., Shen, S. S. P., Qu, W., Gloersen, P., Fan, K. L., 2003. A confidence limit for the empirical mode decomposition and Hilbert spectral analysis. Proc. R. Soc. London 459 (2037), 2317–2345.

Huang, Y., Schmitt, F. G., Lu, Z., Liu, Y., 2008. An amplitude-frequency study of turbulent scaling intermittency using hilbert spectral analysis. Europhys. Lett. 84, 40010.

Huang, Y., Schmitt, F. G., Lu, Z., Liu, Y., 2009. Analyse de l'invariance d'échelle de séries temporelles par la décomposition modale empirique et l'analyse spectrale de hilbert. Traitement du Signal (in press).

Hurst, H., Black, R., Simaika, Y., 1965. Long-term storage: an experimental study.

Hurst, H. E., 1951. Long-term Storage Capacity of Reservoirs. Trans. Am. Soc. Civ. Eng. 116, 770–808.

Jánosi, I., Gallas, J., 1999. Growth of companies and waterlevel fluctuations of the river Danube. Physica A 271 (3), 448–457.

Jánosi, I., Müller, R., 2005. Empirical mode decomposition and correlation properties of long daily ozone records. Phys. Rev. E 71 (5), 56126.

Kantelhardt, J., Rybski, D., Zschiegner, S., Braun, P., Koscielny-Bunde, E., Livina, V., Havlin, S., Bunde, A., 2003. Multifractality of river runoff and precipitation: comparison of fluctuation analysis and wavelet methods. Physica A 330 (1-2), 240–245.

Kantelhardt, J., Koscielny-Bunde, E., Rybski, D., Braun, P., Bunde, A., Havlin, S., 2006. Long-term persistence and multifractality of precipitation and river runoff records. J. Geophys. Res. 111.

Koscielny-Bunde, E., Kantelhardt, J., Braun, P., Bunde, A., Havlin, S., 2006. Long-term persistence and multifractality of river runoff records: Detrended fluctuation studies. J. Hydrol. 322 (1-4), 120–137.

Livina, V., Ashkenazy, Y., Braun, P., Monetti, R., Bunde, A., Havlin, S., 2003a. Nonlinear volatility of river flux fluctuations. Phys. Rev. E 67 (4), 42101.

Livina, V., Ashkenazy, Y., Kizner, Z., Strygin, V., Bunde, A., Havlin, S., 2003b. A stochastic model of river discharge fluctuations. Physica A 330 (1-2), 283–290.

Livina, V., Kizner, Z., Braun, P., Molnar, T., Bunde, A., Havlin, S., 2007. Temporal scaling comparison of real hydrological data and model runoff records. J. Hydrol. 336 (1-2), 186–198.

Long, S. R., Huang, N. E., Tung, C. C., Wu, M. L., Lin, R. Q., Mollo-Christensen, E., Yuan, Y., 1995. The Hilbert techniques: an alternate approach for non-steady time series analysis. IEEE Geoscience and Remote Sensing Soc. Lett. 3, 6–11.

Loutridis, S. J., 2005. Resonance identification in loudspeaker driver units: A comparison of techniques. Appl. Acoust. 66 (12), 1399–1426.

Mauas, P. J. D., Flamenco, E., Buccino, A. P., 2008. solar forcing of the stream flow of a continental scale south american river. Phys. Rev. Lett. 101, 168501.

Pandey, G., Lovejoy, S., Schertzer, D., 1998. Multifractal analysis of daily river flows including extremes for basins of five to two million square kilometres, one day to 75 years. J. Hydrol. 208 (1-2), 62–81.

Peters, O., Hertlein, C., Christensen, K., 2002. A complexity view of rainfall. Phys. Rev. Lett. 88 (1), 018701.

Rilling, G., Flandrin, P., Gonçalvès, P., 2003. On empirical mode decomposition and its algorithms. IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing.

Rilling, G., Flandrin, P., 2008. One or two frequencies? The empirical mode decomposition answers. IEEE Trans. Signal Process.

Schmitt, F. G., Huang, Y. X., Lu, Z. M., Zongo, S. B., Molinero, J. C., Liu, Y. L., 2007. Analysis of nonliner biophysical time series in aquatic environments: scaling properties and empirical mode decomposition. In: Nonlinear Dynamics in Geosciences. edited by A. Tsonis and J. Elsner. Springer, pp. 261–280.

Schmitt, F. G., Huang, Y. X., Lu, Z. M., Liu, Y. L., Fernandez, N., 2008. Analysis of velocity fluctuations and their intermittency properties in the surf zone using empirical mode decomposition. J. Marine Syst. DOI:10.1016/j.jmarsys.2008.11.012.

Schumm, S., 2005. River Variability and Complexity. Cambridge University Press.

Solé, J., Turiel, A., Llebot, J., 2007. Using empirical mode decomposition to correlate paleoclimatic time-series. Nat. Hazard Earth Sys. Sci. 7, 299–307.

Su, Z.-Y., Wang, C.-C., Wu, T., Wang, Y.-T., Feng-Cheng, T., 2008. Instantaneous frequency-time analysis of physiology signals: the application of pregnant women's radial artery pulse signals. Physica A 387, 485–494.

Telesca, L., Lapenna, V., Macchiato, M., 2004b. Mono-and multi-fractal investigation of scaling properties in temporal patterns of seismic sequences. Chaos, Solitons and Fractals 19 (1), 1–15.

Tessier, Y., Lovejoy, S., Hubert, P., Schertzer, D., Pecknold, S., 1996. Multifractal analysis and modeling of rainfall and river flows and scaling, causal transfer functions. J. Geophys. Res. 101, 26427–26440.

Wu, Z., Huang, N. E., 2004. A study of the characteristics of white noise using the empirical mode decomposition method. Proc. R. Soc. London 460, 1597–1611.

Zhang, C., 2005. Madden-Julian Oscillation. Rev. Geophys. 43, 36.

Zhang, Q., Xu, C., Chen, Y., Yu, Z., 2008a. Multifractal detrended fluctuation analysis of streamflow series of the Yangtze River basin, China. Hydrol. Process., DOI: 10.1002/hyp.7119.

Zhang, Q., Xu, C. Y., Liu, C. L., Chen, Y. D., 2008b. Multifractal analysis of streamflow records of the east river basin (pearl river), china. Physica A, DOI: 10.1016/j.physa.2008.11.025.