

# Time Series Chaos Detection and Assessment via Scale Dependent Lyapunov Exponent

Livio Fenga<sup>1,2</sup>

<sup>1</sup> ISTAT (National Institute of Statistics)

<sup>2</sup> UCSD (University of California San Diego)

Correspondence: Livio Fenga, ISTAT (National Institute of Statistics) Via Cesare Balbo, 16 - 00184, Rome, Italy. E-mail: fenga@istat.it

Received: August 2, 2016 Accepted: August 22, 2016 Online Published: October 17, 2016

doi:10.5539/ijsp.v5n6p1 URL: <http://dx.doi.org/10.5539/ijsp.v5n6p1>

## Abstract

Many dynamical systems in a wide range of disciplines – such as engineering, economy and biology – exhibit complex behaviors generated by nonlinear components which might result in deterministic chaos. While in lab-controlled setups its detection and level estimation is in general a doable task, usually the same does not hold for many practical applications. This is because experimental conditions imply facts like low signal-to-noise ratios, small sample sizes and not-repeatability of the experiment, so that the performances of the tools commonly employed for chaos detection can be seriously affected. To tackle this problem, a combined approach based on wavelet and chaos theory is proposed. This is a procedure designed to provide the analyst with qualitative and quantitative information, hopefully conducive to a better understanding of the dynamical system the time series under investigation is generated from. The chaos detector considered is the well known Lyapunov Exponent. A real life application, using the Italian Electric Market price index, is employed to corroborate the validity of the proposed approach.

**Keywords:** deterministic chaos, economic time series, Lyapunov Exponent, multiresolution analysis

## 1. Introduction

Nature is per se a non-linear entity, whose measurable expressions are non-linear as well. Real-world Data Generating Processes (DGPs) – except for trivial cases or lab-controlled experiments – hardly ever do possess features compatible with a linear framework. In particular, for time series produced by non linear dynamical systems, non linearity can unfold in many different ways, giving rise to various, mainly complex phenomena as a result, such as deterministic chaos, long memory, non stationarity, fractal and multi-fractal behavior. Deterministic chaos is among the most interesting features that can arise in such a framework. In particular, two topics have become crucial in several research fields – such as engineering, meteorology, cosmology and medicine – i.e. its discrimination from random noise and the estimation of its level. In fact, dealing with complicated structures – as in the case of weak RADAR signals (see for example Harman et al., 2006; Liu et al., 2007), noisy satellite images (Olsen et al., 2009) or complex biological data Freeman, 1992 and reference therein) – requires the analyst to have a picture as clear as possible on whether chaotic components are present in the dynamical system at hand as well as on their extent. However, it should be emphasized that, in its short life (it dates back to the early 70s), chaos theory has impacted a wide range of scientific activities, including the non physical ones, such as economics (see e.g. Peters, 1994; Brock et al., 1989; Puu, 2013), psychology (Guastello, 2013), sociology (Eve et al., 1997) and archaeology (Chadwick, 1998).

Deterministic chaos is characterized by sensitivity to initial conditions, which can be detected by measuring the rate of divergence between two trajectories starting from nearby states. Lyapunov Exponent (LE) – is one popular choice routinely employed to this end. This mathematical tool has been conceived and designed to perform in the field of physical science, where the full knowledge of the underlying stochastic process – in terms of both signal-to-noise ratio (SNR) and determining equations – guarantees the control of the system investigated. However, many times evidences of chaos in a given experimental system has to be evaluated on empirical basis, namely by direct observation of the empirical data. Regardless of the field of application, two typical features of such a framework are the lack of knowledge of the set of rules governing the system as well as of the quantity of noise present (and its exact nature). While the former makes system control procedures in general difficult to apply, the latter might have even more dangerous consequences. As an ubiquitous and unavoidable entity inherent with the experimental nature of the data, noise can have a catastrophic impact on virtually all the stages of chaos detection and assessment. Either in the form of environmental fluctuations or limited experimental resolution, noise can introduce significant amount of uncertainty in the system and chaos can go undetected as a result. Another serious problem is related to the fact that the estimation of the Lyapunov exponent, being performed

on trajectories of finite length, can be severely biased due to their insufficient duration.

In the present paper, it is shown how chaos detection and assessment can be pursued in empirical setups, i.e. characterized by sensitivity to external noise and poor data resolution, through the estimation of the Lyapunov Exponent within the framework of multi-resolution approximation (MRA). The rest of the paper is organized as follows: in Section 1.1 the goals pursued are more precisely articulated and in Section 1.2 wavelet theory is introduced. Its particularization within Chaos theory is discussed in Section 2, where also the analytic tools employed for chaos detection are introduced. The employed variance decomposition technique is illustrated in Section 2.2. Finally, the empirical experiment is illustrated in Sections 3.

### 1.1 The Goal

As already pointed out, the objective of the present paper is to undertake the analysis of a given time series to check for the presence of deterministic chaos and, if so, to provide an estimate – either quantitative and qualitative – of its degree. In more details, the proposed approach combines chaos and wavelet theories with the three-fold purposes of a) detecting deterministic chaos b) discriminating it from random noise and c) quantifying chaos at different time scales. By having a clearer picture of the embedded structures of the time series at hand, one is more likely to assess the predictive capabilities of a given forecasting tool or to conclude that a portion of it (either in time or frequency), or even the whole time series is unpredictable. In particular, wavelet decomposition can help the analyst choose the more predictable components and to design ad hoc forecasting procedures. Finally, a variance decomposition algorithm is employed as a proxy of the amount of chaos present at a given decomposition level. In more details, point a) is dealt with by repeating the LE estimation at each scale: by probing small length resolutions, chaos (and possibly other non-linear phenomena) detection as well as better understanding of its nature can be gained. Regarding the issue sub b), MRA low-pass filtering capabilities can be useful to discriminate pure random noise and chaotic components whereas point c) is pretty straightforward, as it consists in applying variance decomposition (W-ANOVA, i.e. Wavelet Analysis of Variance) procedures at each scale.

### 1.2 Signal Decomposing Procedure

Dynamical decomposition of the observed time series is performed through a mathematical transform of the type MRA, which is induced by well localized functions: the wavelets, formalized in (1) and (2). MRA consists of a hierarchical sequence of nested subspaces  $\{V_j\}_{j \in \mathbb{Z}}$  progressively approximating the Hilbert space  $L^2(\mathbb{R})$  of all the squared integrable functions satisfying the following properties:

- a)  $\dots V_{-1} \subset V_0 \subset V_1 \subset L^2(\mathbb{R})$  for all  $j \in \mathbb{Z}$ ;
- b)  $\cup_{j=-\infty}^{+\infty} V_j$  is dense in  $L^2(\mathbb{R})$  and  $\cap_{j=-\infty}^{+\infty} V_j = \{0\}$ ;
- c)  $f(x) \in V_j \iff f(2x) \in V_{j+1}$ , for all  $j \in \mathbb{Z}$ ;
- d)  $f(x) \in V_j \rightarrow f(x - k) \in V_0 \quad \forall k \in \mathbb{Z}$
- e)  $\exists \phi(x) \in V_0$ , s.t. the set  $\{\phi(x - k) | k \in \mathbb{Z}\}$  is a Riesz basis for  $V_0$ .

The scaling function  $\phi$  (the father wavelet), constitutes the scale function for a given MRA. In other words, scaling by  $2^j$  generates basis functions for the space  $V_j$  so that, given that  $(\dots V_{j-1} \subset V_j \subset V_{j+1} \dots)$ , the scaling equation can be expressed as (1), under the condition of a proper choice of the coefficients  $\{c_k; k \in \mathbb{Z}\}$ ; by linearly combining the scaled father wavelet with a suitably chosen set of the coefficients  $\{b_k; k \in \mathbb{Z}\}$ , i.e.

$$\langle \phi(x - k), \psi(x - l) \rangle = 0; \quad l, k \in \mathbb{Z},$$

the mother wavelet  $\psi$  is obtained (2).

$$\phi(x) = \sum_{k \in \mathbb{Z}} c_k \phi(2x - k), \tag{1} \quad \psi(x) = \sum_{k \in \mathbb{Z}} b_k \phi(2x - k). \tag{2}$$

By dilations and translations of  $\psi(x)$ , the space of functions  $\Psi(x)$  is generated, i.e.  $\Psi = \{\psi_{c,b}(x) = |c|^{-1/2} \psi(\frac{x-b}{c}), c, b \in \mathbb{R}, c \neq 0\}$ , with  $\psi$  obeying to the following conditions: a)  $\int_{-\infty}^{\infty} \psi(t) dt = 0$ ; b)  $\int_{-\infty}^{\infty} |\psi(t)| dt < \infty$ ; c)  $\int_{-\infty}^{\infty} \frac{|\hat{\psi}(\xi)|}{|\xi|} dt < \infty$ ,  $\hat{\psi}(\xi)$  being the Fourier transform of  $\psi(t)$ ; d)  $\int_{-\infty}^{\infty} t^j \psi(t) dt = 0, \quad j = 0, 1, \dots, r-1$ , under the conditions that it must exist at least a  $r \geq 1$  and  $\int_{-\infty}^{\infty} t^r \psi(t) dt < \infty$ . By projecting the data onto shifted and translated transformations of the mother and father wavelets, the sets of wavelet and scaling coefficients are respectively obtained, i.e.

$$d_{j,k} = \int_{\mathbb{R}} \phi_{j,k}x(t)dt \tag{3} \qquad s_{J,k} = \int_{\mathbb{R}} \psi_{J,k}x(t)dt. \tag{4}$$

Here, each set of coefficients, usually referred to as crystal, is linked to a spacial scale  $j$ , whereas every single coefficient, called atom, accounts for a particular location. In practice, wavelet coefficients  $\{d_{j,k}; j = 1, 2, \dots, J\}$  in (3) account for and represent progressively finer and finer details whereas smooth dynamics at the coarsest scale are captured by the crystal  $\{s_{J,k}\}$  (4). MODWT is the filtering approach achieved used to perform MRA. Here the wavelet and scale coefficients are given by:

$$d_{j,t} = \frac{1}{2^{j/2}} \sum_{l=0}^{L_j-1} \tilde{h}_{j,l}, X_{t-l} \text{ mod } N, \qquad S_{J,t} = \frac{1}{2^{J/2}} \sum_{l=0}^{L_J-1} \tilde{g}_{J,l}, X_{t-l} \text{ mod } N, \tag{5}$$

where  $\{\tilde{h}_{j,l}\}$  and  $\{\tilde{g}_{j,l}\}$  are the length  $L$ , level  $j$ , wavelet and scaling filters, obtained by rescaling their DWT counterparts, i.e.  $\{h_{j,l}\}$  and  $\{g_{j,l}\}$ , as follows:  $\tilde{h}_{j,l} = \frac{h_{j,l}}{2^{j/2}}$  and  $\tilde{g}_{j,l} = \frac{g_{j,l}}{2^{j/2}}$ . Here, the sequences of coefficients  $\{h_{j,l}\}$  and  $\{g_{j,l}\}$  are approximate filters: the former, of the type band-pass, with nominal pass-band  $f \in [\frac{1}{4}\mu_j, \frac{1}{2}\mu_j]$ , and the latter of the type low-pass, with a nominal pass-band  $f \in [0, \frac{1}{4}\mu_j]$ , with  $\mu_j$  denoting the scale. Considering all the  $J = J^{max}$  sustainable scales, MRA wavelet representation of the given time series  $x_t$ , in the  $L^2(\mathbb{R})$  space, can be expressed as follows:

$$x(t) = \sum_k s_{J,k}\phi_{J,k}(t) + \sum_k d_{J,k}\psi_{J,k}(t) + \sum_k d_{J-1,k} + \psi_{J-1,k}(t) + \dots + \sum_k d_{j,k}\psi_{j,k}(t) \dots + \sum_k d_{1,k}\psi_{1,k}(t), \tag{6}$$

with  $k$  taking integer values from 1 to the length of the vector of wavelet coefficients related to the component  $j$ .

### 1.2.1 Practicalities

In an empirical set up, boundary conditions can be selected on the basis of the characteristics of the time series and/or as a part of a preliminary ad hoc investigation on a guess and check basis. On the contrary, the choice of the wavelet function and its length  $L$  is in general critical and strictly system specific. In what follows, the 4<sup>th</sup> order Daubechies least asymmetric wavelet filter (known also as symmlets) of length  $L = 8$ , usually denoted by LA(8), has been made use of. This type of filter, commonly adopted in several applications, has been chosen for its salient properties, that is: near symmetry in the midpoint and (approximately) linear phase. MRA has been performed using Maximum Overlapping Discrete Wavelet Transform (MODWT) algorithm. Technically, it is a filtering approach aimed at modifying the observed series  $x_t$ , by artificially introducing an extension of it, so that the unobserved samples  $\{x_{t \in \mathbb{Z}^-}\}$  are assigned the observed values  $X_{T-1}, X_{T-2}, \dots, X_0$ . This method, considers the series as it were periodic<sup>1</sup>, and is known as using circular boundary conditions. For more details on MODWT, the reader is referred to Percival and Walden (2006) and Percival (2002).

## 2. The Wavelet Phase-space Reconstruction and the Scale Dependent Chaos Detector

The Lyapunov-based investigation is defined by considering an ensemble of trajectories, which requires a suitable reconstruction of the phase space of the underlying dynamical system, say  $\mathbf{x}_{n+1} = \mathcal{F}(\mathbf{x}_n)$ . This is observed through a function  $y = \mathcal{H}(\mathbf{x})$ , whose elements are the scalar measurements  $S(n)$  at a different time, i.e.

$$\mathbf{y}_n = [S(n), S(n + \tau), S(n + 2\tau), \dots, S(n + (m - 1)\tau)], \tag{7}$$

with  $m$  being the embedding dimension and  $\tau$  the time delay. Takens embedding theorem allows us to use the delay coordinates (7), in virtue of the two dynamical systems arising from this setup, i.e.  $\mathbf{y}_{n+k} = \mathcal{H}(\mathbf{x}_{n+k}) = \mathcal{H}(\mathcal{F}^k(\mathbf{x}_n))$  and  $\mathbf{y}_{n+1} = \mathcal{G}(\mathbf{y}_n)$ . Considering the latter,  $\mathbf{x} \leftrightarrow \mathbf{y}$ , one-to-one type relationship is guaranteed only for a "sufficiently" large  $m$ . Takens theorem particularizes to time dependent processes, the results obtained by Whitney in the case of dimension  $D$  smooth manifold, say  $\mathbf{M}$ , whose dynamic can be embedded in  $\mathbb{R}^{2D+1}$  and immersed in  $\mathbb{R}^{2D}$ . In essence, Takens

<sup>1</sup>However, it should be emphasized that, even though such an approach suffices in many cases, it shows weakness when non-periodic signals are affected by discontinuities, as in the case, for example, of certain deseasonalized economic indicators. To alleviate the problem, a common strategy is to introduce an artificial extension of the time series, by doubling up its original sample size through boundary conditions of the type reflection, i.e. the unobserved values  $x_{-1}, x_{-2}, \dots, x_{-T}$  are assigned the values observed at  $x_0, x_1, \dots, x_{T-1}$ . Let  $\{x_t^c\}$  the artificially extended time series generated according to

$$\begin{cases} \{X_t^c\} = \{X_t\}; & t = 0, \dots, T - 1 \\ \{X_t^c\} = \{X_{2T-1-t}\}; & t = T, \dots, 2T - 1, \end{cases}$$

the circular boundary conditions, are then re-expressed on  $\{X_t^c\}$  to obtain new wavelet and scaling filters, respectively expressed by:

$$d_{j,t} = \frac{1}{2^{j/2}} \sum_{l=0}^{L_j-1} \tilde{h}_{j,l}, X_{t-l}^c \text{ mod } 2N, \qquad S_{j,t} = \frac{1}{2^{j/2}} \sum_{l=0}^{L_j-1} \tilde{g}_{j,l}, X_{t-l}^c \text{ mod } 2N.$$

proved that – given a diffeomorphism  $\mathfrak{M}$  defining the set of trajectories on  $M$ , a smooth observation function  $y : M \rightarrow \mathbb{R}$  generates an embedding of  $M$  in  $2m + 1$  dimensions under the transformation  $\tilde{\mathfrak{M}}_{\mathfrak{M},y} : M \rightarrow \mathbb{R}^{2m+1}$ , where  $\tilde{\mathfrak{M}}_{\mathfrak{M},y}(\mathbf{x}) = \langle y(\mathbf{x}), y(\mathfrak{M}(\mathbf{x})), y(\mathfrak{M}^2(\mathbf{x})), \dots, y(\mathfrak{M}^{2m}(\mathbf{x})) \rangle$ . Here, each of the elements  $\langle y(\mathbf{x}), y(\mathfrak{M}(\mathbf{x})), y(\mathfrak{M}^2(\mathbf{x})), \dots, y(\mathfrak{M}^{2m}(\mathbf{x})) \rangle$  represents the time-shifted observations of the dynamics induced by  $\mathfrak{M}$  on  $M$ . An in-depth and detailed discussion of this subjected can be found in Deyle and Sugihara (2011) and Noakes (1991).

Relevant for the present analysis is the result from Sauer and Yorke (1993) on whether filtering procedure could affect proper embedding. Their central result is that by applying filters of the type Finite Impulse Response (FIR) embedding procedures, in general, would not be compromised, under the assumption that a sufficient number of independent observables is available. This setup is consistent with the adopted MRA approach – other than the assumption, previously stated, of time series of "sufficient" length. Regarding the first condition, a basis of the type Riesz can be transformed into an orthogonal basis through the transfer function  $\{g(\cdot)\}$  (5), which entirely determines the scaling function  $S_{J,t}$  (5). The  $g$ -induced scaling function  $S_t$  is compactly supported if and only if  $\{g(\cdot)\}$  has a finite number of non zero coefficients, that is  $g(\cdot)$  is a FIR filter.

However, in order for the function  $g(\cdot)$  to generate multiresolution approximations, some conditions need to be satisfied (see, for example Merry & Steinbuch, 2005). They are:

$\phi \in L^2(\mathbb{R})$  is an integrable scaling function iff the Fourier series

$$g[n] = \langle \frac{1}{\sqrt{2}} \phi(t/2), \phi(t - n) \rangle$$

satisfies

- (i)  $\forall \omega \in \mathbb{R}, |\hat{g}(\omega)|^2 + |\hat{g}(\omega + \pi)|^2 = 2;$
- (ii)  $\hat{g}(0) = \sqrt{2}.$

On the other hand, if

- (a)  $\hat{g}(\omega)$  is periodic (period =  $2\pi$ );
- (b)  $\hat{g}(\omega)$  of the type  $C^1$  in a neighborhood of  $\omega = 0$ ;
- (c) conditions (i, ii) above hold;
- (d)  $\inf_{\omega \in [-\frac{\pi}{2}, \frac{\pi}{2}]} |\hat{g}(\omega)| > 0,$

then  $\hat{S}(\omega) = \prod_{p=1}^{+\infty} \frac{\hat{g}(2^{-p}\omega)}{\sqrt{2}}$  is the Fourier transform of a scaling function  $S \in L^2(\mathbb{R})$ .

Finally, the application of Lyapunov exponent at different resolution levels finds its theoretical justification in the fact that (see e.g. Lamarque & Malasoma, 1996 ) given a scalar function  $f : \mathbb{R} \rightarrow \mathbb{R}$ , and a wavelet function  $\phi(x)$  (1), the  $f$ -wavelet transformation  $Q$  is given by

$$Q_{\phi,a,b}(f) = \frac{1}{a} \int_{-\infty}^{+\infty} f(t) \phi(\frac{x-b}{a}) dx, \tag{8}$$

with  $b$  being the focal point of the mathematical microscope  $\phi(\cdot)$  and  $\frac{1}{a}$  the magnifying factor. The quantification of the separation ( $\delta$ ) over time ( $\Delta_t$ ) of couple of neighbor trajectories starting in  $\delta_0 = \|\mathbf{x}_{t_1} - \mathbf{x}_{t_2}\|$ , i.e.  $\delta_{\Delta_t} \approx \|\mathbf{x}_{t_1 + \Delta_t} - \mathbf{x}_{t_2 + \Delta_t}\|$ , giving rise to the Lyapunov exponent  $\lambda$ , i.e.

$$\delta_{\Delta_t} \approx \delta_0 e^{\lambda \Delta_t}, \tag{9}$$

is done with reference to the function  $Q(\cdot)$ .

Therefore, the Lyapunov exponent is now expressed as

$$Q(\cdot) = S(\Delta_t) = \frac{1}{N} \sum_{t_0=t_1}^{t_N} \ln \left( \frac{1}{|U(\mathbf{x}_{t_0})|} \sum_{\mathbf{x}_t \in U(\mathbf{x}_{t_0})} |\mathbf{x}_{t_0 + \Delta_t} - \mathbf{x}_t + \Delta_t| \right), \tag{10}$$

where  $U(\mathbf{x}_{t_0})$  defines the neighborhood centered on  $\mathbf{x}_{t_0}$ .

In the empirical analysis presented in Section 3,  $x_0$  counts the data point falling within a radius  $\epsilon = \frac{1}{10} \sqrt{\sigma^2(Q(\cdot))}$ .

### 2.1 Embedding Dimension and Time Delay Estimation

In what follows, the analytic tools employed for the optimal LE estimation are presented. They are designed to capture patterns in the probabilistic structure in the linear (autocorrelation function) and non linear (mutual information) case, whereas the estimation of the time delay is provided by the false nearest neighbor method.

#### 2.1.1 Autocorrelation and Mutual Information

The usual autocorrelation function (11–12) and the time delayed mutual information (13–14), as well as visual inspection of delay representations with various lags provide important information about reasonable delay times.

$$\rho(\tau) = \mathbb{E}[(X_t - \mu)(X_{t+\tau} - \mu)], \tag{11}$$

$$\hat{\rho}^j(\tau) = \frac{1}{(T - \tau)\sigma_{X^j}^2} \sum_{t=1}^{T-\tau} (x_t^j - \bar{x}^j)(x_{t+\tau}^j - \bar{x}^j). \tag{12}$$

Here (12) – whose asymptotic confidence intervals are  $\frac{1}{\sqrt{T}} \pm 2 \sqrt{T} \approx \pm 2 \sqrt{T}$  – represents the estimator for (11) at a given resolution level  $j$ . However, more guidance has been gained by means of the mutual information function. This function is able to capture and account for linear and nonlinear correlations; in essence, it expresses the entropy–related concept of measuring the "amount of information" obtained about one random variable, through the other random variable. It is given by:

$$\mathcal{I}^j(\mathcal{X}^j; \mathcal{Y}^j) = \int_{\mathcal{Y}^j} \int_{\mathcal{X}^j} p(\mathbf{x}^j, \mathbf{y}^j) \log_2 \frac{p(\mathbf{x}^j, \mathbf{y}^j)}{p(\mathbf{x}^j)p(\mathbf{y}^j)} d\mathbf{y}^j d\mathbf{x}^j. \tag{13}$$

Its estimator, for the resolution level  $j$ , is given by

$$M^j = - \sum_{i,r} p_r(\tau) \log_2 p_{ir}(\tau) [(p_i \cdot p_r)]^{-1}. \tag{14}$$

#### 2.1.2 Time Delay Estimation

Determination of the proper time delay  $d$  has been done with regard to nearest neighbor of each of the vectors (7) using the  $L_2$  norm. This is a critical and in general a non trivial task: its underestimation is particularly dangerous, as two time delay vectors might show a small distance not as a result of the peculiar system dynamics but due to projection. The employed procedure, conducted at the  $j$ -level, declares a given neighbor of  $\mathbf{y}_{j,k}$ , denoted as  $\mathbf{y}_{j,k}^\circ$ , a false neighbor according to the rate of false nearest neighbors in the reconstructed phase space, i.e.:

$$\frac{\|\mathbf{y}_{j,k} - \mathbf{y}_{j,k}^\circ\|^2 + [\mathbf{x}_{j,k+d} - \mathbf{x}_{j,k+d}^\circ]^2}{\mathbf{B}^2(\mathbf{A})} > \mathbf{A}_{thr}^2.$$

Here,  $\mathbf{A}_{thr} \approx 2$  (see e.g. Aittokallio et al., 1999) is the radius of the attractor whereas

$$\mathbf{B}^2(\mathbf{A}) = \frac{1}{N} \sum_{k=1}^N [\mathbf{x}^j - \mu(\mathbf{x}^j)]^2,$$

with  $\mu(\cdot)$  denoting the mean value computed on all the points.

However, this procedure – designed to test progressively higher dimensionalities until a sufficiently small number is obtained (possibly 0) – is able to provide reliable outcomes in case of noise–free data. On the contrary, when noisy components are embedded in the time series, the likelihood of finding false nearest neighbors increase with the sample size. The use of wavelet theory, in the form of MRA algorithm, is a coping strategy for such a circumstance, as it is able to provide information at different levels which can be less affected by noise.

Spurious temporal correlation, which can seriously bias the estimation of the system embedding dimension, are dealt with by ruling out the set of points closer than some threshold time, i.e. the Theiler window. This quantity has been estimated using the method introduced by Provenzale et al. (1992), which is basically based on a sequence of space-time separation plots employed for the detection of temporal structures in the data. MODWT's translation-invariant property allows the effective alignment of the different events at different resolution levels so that the integrity of the dynamics induced by transient events is preserved.

Finally, being Lyapunov exponent invariant to the embedding dimension  $m$ , under  $m > d$ ,  $S(\Delta_r)$  has been calculated for  $m = 2 \cdot (d), \dots, m + 9$ .

### 2.2 Wavelet ANOVA

MODWT is a transformation of the type energy conserving, i.e.  $\|X\|^2 = \sum_{j=1}^{J_0} \|\widetilde{W}_j\|^2 + \|\widetilde{V}_{j_0}\|^2$ , therefore a scale dependent analysis of variance can be derived, based on the set of wavelet and scaling coefficients, i.e.:

$$\hat{\sigma}_X^2 = \|X\|^2 - \bar{X}^2 = \frac{1}{N} \sum_{j=1}^{J_0} \|\widetilde{W}_j\|^2 + \frac{1}{N} \|\widetilde{V}_{j_0}\|^2 - \bar{X}^2. \tag{15}$$

Equation (15) enable us to quantify the allocation of the energy roughly attributable to chaotic components across the  $J$  scales

Ruling out the boundary coefficients sets, we have that the wavelet variance

$$\hat{v}_X^2(\tau_j) = \mathbb{E}(\widetilde{W}_{j,t})^2 = \frac{1}{N} \sum_{t=0}^{N-1} \widetilde{W}_{j,t}^2 \tag{16}$$

is time independent under second order stationarity or difference stationary assumptions. The former implies that the extracted sequence (signal)  $\{\tau_{j,t}\}_{t \in \mathbb{Z}^+}$ , that  $\mathbb{E}(X) = \mu$  and  $cov(\tau_{j,t}, \tau_{j,t+k}) = \gamma_{j,k}$ , being  $\mu$  constant and  $\gamma_k$  time independent. Defining the back-shift operator  $L$ , i.e.  $L\tau_t = \tau_{t-1}$  (therefore  $L^n\tau = \tau_{t-n}$ ) and the difference operator (the subscript  $j$  is omitted)  $\nabla^d\tau_t = (1-L)^d\tau_t \quad d = 0, 1, \dots, D$ , the latter implies that the transformed time series  $\nabla^d\tau_t =$  stationary. However, these conditions are usually met at coarser levels, where

$${}^u\hat{v}_X^2(\tau_j) = \mathbb{E}(\widetilde{W}_{j,t})^2 = \frac{1}{M_j} \sum_{t=L_j-1}^{N-1} \widetilde{W}_{j,t}^2, \tag{17}$$

with  $M - j = N - L_j + 1$  being the size of the set of non-boundary coefficients for the  $j$ -level,  $L_j$  the wavefilter length, and  $u$  stands for unbiased. However, by using MODWT, one is forced to use either reflection, or circular boundary conditions. Therefore, the proper estimator is given by:

$${}^{bias}\hat{v}_X^2(\tau_j) = \mathbb{E}(\widetilde{W}_{j,t})^2 = \frac{1}{2N} \sum_{t=0}^{2N-1} \widetilde{W}_{j,t}^2. \tag{18}$$

Confidence intervals for the true wavelet variance can be built for both (17) and (18) on the basis of their asymptotic approximation to a scaled  $\chi^2$  distribution with  $\delta_j$  EDOF (Equivalent Degree Of Freedom), which reflects the correlation structure at different resolution levels  $j$ . Following Percival and Walden (2006), 100%(1 - 2p) confidence intervals are approximated by

$$CI(\hat{v}_X^2(\tau_j)) \approx \left[ \frac{\delta_j \hat{v}_X^2(\tau_j)}{Q_{\delta_j}(1-p)}, \frac{\delta_j \hat{v}_X^2(\tau_j)}{Q_{\delta_j}(p)} \right], \tag{19}$$

with

$$\delta_j = \max \left\{ \frac{M_j}{2^j}, 1 \right\} \quad \text{or} \quad \delta_j = \max \left\{ \frac{N}{2^j}, 1 \right\},$$

according to whether the case (17) or (18) is considered.

### 3. Empirical Experiment

In order to corroborate the validity of the presented approach, an empirical analysis has been conducted on the electricity market. Commonly denominated Italian Power Exchange (IPEX), the Electricity Market, is where producers, consumers and wholesale customers enter into hourly electricity purchase and sale contracts. Here, accepted demand bids are remunerated at the National Single Price (PUN), which therefore represents the purchase price.

The employed data set consists in the time series of PUNs, denominated henceforth  $X_t$ , which is freely and publicly available at the website

<http://www.mercatoelettrico.org/En/download/DatiStorici.aspx>.

The reported values are expressed in Euro currency (Euro per Megawatt), whereas the sampling frequency is one hour. The span of time considered in the present study is January 1<sup>st</sup> 2005 –February 29<sup>th</sup> 2016, for a total of 97.848 data. The time series has been differenced twice, with difference of order 1 respectively at lag 24 and 168, therefore the effective sample size is  $T= 97656$ .

In Italy, the electric market has been fully liberalized on July 1<sup>st</sup> 2007, therefore, in order to gain a better insight on the performances of the proposed approach, the whole analysis has been executed having in mind this reference time. In practice, PUN time series has been split in three non-overlapping sub-periods (denominated  $Z_1, Z_2, Z_3$ ), covering the time before and after the liberalization, plus an arbitrary chosen warming up period located in between, that is:

- $Z_1 \equiv [01/01/2005; 06/30/2007]$ , (21671 data);
- $Z_2 \equiv [07/01/2007; 12/31/2012]$ , (48265 data);
- $Z_3 \equiv [01/01/2013; 12/29/2016]$ , (27720 data).

Finally, with  $Z_{tot}$  the full length window will be denoted, i.e.  $Z_{tot} \equiv [01/01/2005; 12/29/2016]$ .

### 4. Outcomes of the Experiment

In Table 1 the LE, computed for the three sub-series with span  $Z_{1,2,3}$  and  $Z_{tot}$  are reported. The presence of chaos is confirmed, for the windows  $Z_{2,3}$ , by a positive LE. The increasing pattern noticeable in the LE values, probably reflect the system behavior transitioning from a non chaotic to a chaotic status, as a result of the greater and greater degrees of freedom introduced into the electric market over the years. However, by inspecting Tables 2 – 4, additional information can be obtained. These are organized according to the six resolution levels – depicted in Figure 1 – the time series  $X_t$  has been broken into. For each of them, other than the LEs, the following parameters are considered: variance estimation ( $\sigma^2(d_j)$ ) as a value and as a ratio to the total variability ( $\frac{\sigma^2(d_j)}{\sigma_X^2}$ ), equivalent degrees of freedom ( $\delta_j$ ) and confidence intervals ( $CI_{low}, CI_{high}$ ) for  $\sigma^2(d_j)$ . Deterministic chaos seems to be attributable in greater part to coarser resolution levels ( $> 2$  days ), i.e.  $d4 - d6$ . In fact, in the case of  $Z_1$  these components (weighing approx 47% of the total variability) are characterized by a negative LE whereas they show a positive LE for  $Z_3$ , where their relative weight is .70. By restricting our attention only on the resolution levels  $j = 5, 6$ , we are able to pinpoint with more precision the resolution levels at which deterministic chaos is generated. In fact, while  $d5$  and  $d6$  in  $Z_1$  (i.e. 34% of total variability) show in both of the cases a negative LE, in  $Z_2$  (47% of the total variability) they become positive as well as in  $Z_3$ . In the last case, however, they represent 54% of the total variation and, all the more so, LE in both the cases show greater magnitudes. Also, it is worth mentioning that the finest MRA levels, i.e.  $j = 1, 2, 3$ , exhibit values and relative importance consistently decreasing as the sub-series progress in time. In particular, the relative variability is 51% for  $Z_1$  and 39% and .29% for  $Z_2$  and  $Z_3$  respectively. This last evidence gives account of the fact that the negative LEs found at levels 2 and 3 are related to a small quota of the overall variability ( $< \frac{1}{4}\sigma_X^2$ ), and therefore cannot counteract the general chaotic behavior of this segment of  $X_t$ . The same reasoning applies to  $Z_1$ , where the same components – representing the 35% of the total variability – even if showing positive LE values, are embedded in the sub-series  $Z_1$  whose overall dynamic is non chaotic.

Table 1. Series  $X_t$ , Lyapunov Exponent estimates for the time windows  $Z_1, Z_2, Z_3, Z_{tot}$

series	$Z_1$	$Z_2$	$Z_3$	$X_t$
Lyapunov Exponent Estimates	-.0005527541	.0005230929	.00476779	.001668556

Table 2. Series  $X_t$ , MRA outcomes for the time frame  $Z_1$

Crystals \ Parameters	Lyapunov exponent	$\sigma^2$	$\frac{\sigma^2(d_j)}{\sigma_x^2}$	$\delta_j$	$CI_{low}$	$CI_{high}$
d1	-.0003111276	39.43	.17	7770.66	38.56	40.32
d2	.0001053132	43.16	.17	4781.97	41.83	44.56
d3	.0003320897	44.73	.18	2441.29	42.79	46.80
d4	-.000549612	35.91	.14	1473.63	33.74	38.30
d5	-.001045508	55.99	.22	505.57	51.30	61.40
d6	-.003244198	28.33	.11	366.81	25.05	32.35

Table 3. Series  $X_t$ , MRA outcomes for the time frame  $Z_2$

Crystals \ Parameters	Lyapunov exponent	$\sigma^2$	$\frac{\sigma^2(d_j)}{\sigma_x^2}$	$\delta_j$	$CI_{low}$	$CI_{high}$
d1	.0002995352	32.92	.11	16640.04	32.43	33.41
d2	.0004067459	40.80	.14	9568.94	39.95	41.68
d3	.0003309755	41.50	.14	5167.68	41.50	44.07
d4	-.0004087144	39.02	.14	3313.77	39.02	42.48
d5	.0002110077	75.04	.27	1428.65	75.04	84.62
d6	.00087144	55.49	.20	852.20	55.49	65.79

Table 4. Series  $X_t$ , MRA outcomes for the time frame  $Z_3$

Crystals \ Parameters	Lyapunov exponent	$\sigma^2$	$\frac{\sigma^2(d_j)}{\sigma_x^2}$	$\delta_j$	$CI_{low}$	$CI_{high}$
d1	.0001502846	8.38	.066	9611.07	8.22	8.55
d2	-.001889886	11.18	.088	5641.95	10.87	11.50
d3	-.001456725	18.31	.14	2631.98	17.60	19.05
d4	.003439542	20.30	.16	1718.19	19.87	22.23
d5	.01866641	38.90	.31	685.46	36.00	42.20
d6	.01247388	29.27	.23	471.07	26.24	32.88

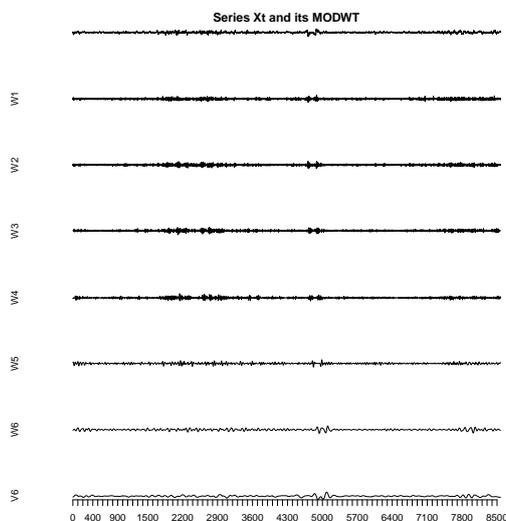


Figure 1. Time series  $X_t$  and its  $MODWT$  coefficient sequence  $d_{j,t}; j = 1, \dots, 6$

## References

- Aittokallio, T., Gyllenberg, M., Hietarinta, J., Kuusela, T., & Multamaäki, T. (1999). Improving the false nearest neighbors method with graphical analysis. *Physical Review E* 60(1), 416. <http://dx.doi.org/10.1103/PhysRevE.60.416>
- Brock, W. A., Brock, A., & Malliaris, A. (1989). Differential equations, stability and chaos in dynamic economics. No. 90A16 BROd
- Chadwick, A. (1998). Archaeology at the edge of chaos: further towards reflexive excavation methodologies. *Assemblage* 3, 97-117.
- Deyle, E. R., & Sugihara, G. (2011). Generalized theorems for nonlinear state space reconstruction. *PLoS One* 6(3), e18295. <http://dx.doi.org/10.1371/journal.pone.0018295>
- Eve, R. A., Horsfall, S., & Lee, M. E. (1997). *Chaos, complexity, and sociology: Myths, models, and theories*. Sage.
- Freeman, W. J. (1992). Tutorial on neurobiology: from single neurons to brain chaos. *International journal of bifurcation and chaos* 2(03), 451-482. <http://dx.doi.org/10.1142/S0218127492000653>
- Guastello, S. J. (2013). *Chaos, catastrophe, and human affairs: Applications of nonlinear dynamics to work, organizations, and social evolution*. Psychology Press.
- Harman, S., Fenwick, A., & Williams, C. (2006). Chaotic signals in radar? In: *2006 European Radar Conference* (pp. 49-52). IEEE. <http://dx.doi.org/10.1109/EURAD.2006.280270>
- Lamarque, C. H. , & Malasoma, J. M. (1996). Analysis of nonlinear oscillations by wavelet transform: Lyapunov exponents. *Nonlinear Dynamics* 9(4), 333-347. <http://dx.doi.org/10.1007/BF01833360>
- Liu, Z., Zhu, X., Hu, W., & Jiang, F. (2007). Principles of chaotic signal radar. *International Journal of Bifurcation and Chaos* 17(05), 1735-1739. <http://dx.doi.org/10.1142/S0218127407018038>
- Merry, R. , & Steinbuch, M. (2005). Wavelet theory and applications. Literature Study, Eindhoven University of Technology, Department of Mechanical Engineering, Control Systems Technology Group.
- Noakes, L. (1991). The takens embedding theorem. *International Journal of Bifurcation and Chaos* 1(04), 867-872. <http://dx.doi.org/10.1142/S0218127491000634>
- Olsen, N., Manda, M., Sabaka, T. J., & Tøffner-Clausen, L. (2009). Chaos-2a geomagnetic field model derived from one decade of continuous satellite data. *Geophysical Journal International* 179(3), 1477-1487. <http://dx.doi.org/10.1111/j.1365-246X.2009.04386.x>
- Percival, D. B. (2002). *Wavelets*. Encyclopedia of environmetrics.
- Percival, D. B., & Walden, A. T. (2006). *Wavelet methods for time series analysis*, vol. 4. Cambridge university press.
- Peters, E. E. (1994). *Fractal market analysis: applying chaos theory to investment and economics*, vol. 24. John Wiley & Sons.
- Provenzale, A., Smith, L. A., Vio, R., & Murante, G. (1992). Distinguishing between low-dimensional dynamics and randomness in measured time series. *Physica D: nonlinear phenomena* 58(1), 31-49. [http://dx.doi.org/10.1016/0167-2789\(92\)90100-2](http://dx.doi.org/10.1016/0167-2789(92)90100-2)
- Puu, T. (2013). *Attractors, bifurcations, & chaos: Nonlinear phenomena in economics*. Springer Science & Business Media.
- Sauer, T., & Yorke, J. A. (1993). How many delay coordinates do you need? *International Journal of Bifurcation and Chaos* 3(03), 737-744. <http://dx.doi.org/10.1142/S0218127493000647>

## Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).