Available online at www.sciencedirect.com
ScienceDirect
Physica A 379 (2007) 503–512
PHYSICA A
www.elsevier.com/locate/physa

Wavelet entropy of stochastic processes

L. Zunino[a,b,c,*], D.G. Pérez[d], M. Garavaglia[a,c], O.A. Rosso[e]

[a] Centro de Investigaciones Ópticas (CIOp), CC. 124 Correo Central, 1900 La Plata, Argentina
[b] Departamento de Ciencias Básicas, Facultad de Ingeniería, Universidad Nacional de La Plata (UNLP), 1900 La Plata, Argentina
[c] Departamento de Física, Facultad de Ciencias Exactas, Universidad Nacional de La Plata (UNLP), 1900 La Plata, Argentina
[d] Instituto de Física, Pontificia Universidad Católica de Valparaíso (PUCV), 23-40025 Valparaíso, Chile
[e] Facultad de Ciencias Exactas y Naturales, Instituto de Cálculo, Universidad de Buenos Aires (UBA), Pabellón II, Ciudad Universitaria, 1428 Ciudad de Buenos Aires, Argentina

Received 10 January 2006; received in revised form 30 October 2006
Available online 28 February 2007

Abstract

We compare two different definitions for the wavelet entropy associated to stochastic processes. The first one, the normalized total wavelet entropy (NTWS) family [S. Blanco, A. Figliola, R.Q. Quiroga, O.A. Rosso, E. Serrano, Time–frequency analysis of electroencephalogram series, III. Wavelet packets and information cost function, Phys. Rev. E 57 (1998) 932–940; O.A. Rosso, S. Blanco, J. Yordanova, V. Kolev, A. Figliola, M. Schürmann, E. Başar, Wavelet entropy: a new tool for analysis of short duration brain electrical signals, J. Neurosci. Method 105 (2001) 65–75] and a second introduced by Tavares and Lucena [Physica A 357(1) (2005) 71–78]. In order to understand their advantages and disadvantages, exact results obtained for fractional Gaussian noise ($-1<\alpha<1$) and fractional Brownian motion ($1<\alpha<3$) are assessed. We find out that the NTWS family performs better as a characterization method for these stochastic processes.

Keywords: Wavelet analysis; Wavelet entropy; Fractional Brownian motion; Fractional Gaussian noise; α-parameter

1. Introduction

The advantages of projecting an arbitrary continuous stochastic process in a discrete wavelet space are widely known. The wavelet time–frequency representation does not make any assumptions about signal stationarity and is capable of detecting dynamic changes due to its localization properties. Unlike the harmonic base functions of the Fourier analysis, which are precisely localized in frequency but infinitely extend in time, wavelets are well localized in both time and frequency. Moreover, the computational time is significantly shorter since the algorithm involves the use of fast wavelet transform in a multi-resolution framework. Finally, contaminating noise contributions can be easily eliminated when they are concentrated in

*Corresponding author. Centro de Investigaciones Ópticas (CIOp), CC. 124 Correo Central, 1900 La Plata, Argentina.
Tel.: +54 2214714341; fax: +54 2214717872.
E-mail addresses: lucianoz@ciop.unlp.edu.ar (L. Zunino), dario.perez@ucv.cl (D.G. Pérez), garavagliam@ciop.unlp.edu.ar (M. Garavaglia), oarosso@fibertel.com.ar (O.A. Rosso).

0378-4371/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2006.12.057

some frequency bands [1,2]. These important reasons justify the introduction, within this special space, of entropy-based algorithms in order to quantify the degree of order or disorder associated with a multi-frequency signal response. With the entropy estimated via the wavelet transform, the time evolution of frequency patterns can be followed with an optimal time–frequency resolution. Several recent papers have confirmed the effectiveness, relevance and suitability of the wavelet entropy as a quantifier of experimental and synthetic signals. These include applications to the characterization of brain electrical signals (EEG and EP/ERP) and neuronal activity [3–14]), solar physics [15,16], erythrocytes deformation [17], laser propagation throughout turbulent media and other lasers applications [18–20], pseudo-random number generators [21], the quantum-classical limit [22], and fractional Brownian motion [23].

In this paper we focus on two definitions for this quantifier: the normalized total wavelet entropy (NTWS) family introduced by one of us (O.A. Rosso) [3,4], and another definition given recently by Tavares and Lucena [24]. We compare their performances while characterizing two important stochastic processes: the fractional Brownian motion (fBm) and the fractional Gaussian noise (fGn). They have been employed as stochastic models in different and heterogeneous scientific fields, like atmospheric turbulence [18,19], econophysics [25] and coastal dispersion [26]. We will show that the NTWS family gives a better characterization for both of them.

2. Wavelet quantifiers

2.1. Wavelet energies

The wavelet analysis is one of the most useful tools when dealing with data samples. Any signal can be decomposed by using a wavelet dyadic discrete family $\{2^{j/2}\psi(2^j t - k)\}$, with $j, k \in \mathbb{Z}$ (the set of integers)—an *orthonormal* basis for $L^2(\mathbb{R})$ consisting of finite-energy signals—of translations and scaling functions based on a function ψ: the mother wavelet [1,2]. In the following, given a stochastic process $s(t)$ its associated *signal* is assumed to be given by the sampled values $\mathcal{S} = \{s(n), n = 1, \ldots, M\}$. Its wavelet expansion has associated wavelet coefficients given by

$$C_j(k) = \langle \mathcal{S}, 2^{j/2}\psi(2^j t - k) \rangle, \tag{1}$$

with $j = -N, \ldots, -1$, and $N = \log_2 M$. The number of coefficients at each resolution level is $N_j = 2^j M$. Note that this correlation gives information on the signal at scale 2^{-j} and time $2^{-j}k$. The set of wavelet coefficients at level j, $\{C_j(k)\}_k$, is also a stochastic process where k represents the discrete time variable. It provides a direct estimation of local energies at different scales. Inspired by the Fourier analysis we define the energy at resolution level j by

$$\mathcal{E}_j = \sum_k \mathbb{E}\left|C_j(k)\right|^2, \tag{2}$$

where $\mathbb{E}$ stands for the average using some, at first, unknown probability distribution. In the case the set $\{C_j(k)\}_k$ is proved to be a stationary process the previous equation reads

$$\mathcal{E}_j = N_j \mathbb{E}\left|C_j(k)\right|^2. \tag{3}$$

Observe that the energy $\mathcal{E}_j$ is only a function of the resolution level. Also, under the same assumptions, the temporal average energy at level j is given by

$$\widetilde{\mathcal{E}}_j = \frac{1}{N_j}\sum_k \mathbb{E}\left|C_j(k)\right|^2 = \mathbb{E}\left|C_j(k)\right|^2, \tag{4}$$

where we have used Eq. (3) to arrive to the last step in this equation. Since we are using dyadic discrete wavelets the number of coefficients decreases over the low frequency bands (at resolution level j the number is halved with respect to the previous one $j+1$); thus, the latter energy definition reinforces the contribution of these low frequency bands.