UNIVERSITÀ DEGLI STUDI DI BERGAMO
DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE
E METODI MATEMATICI°

QUADERNI DEL DIPARTIMENTO

Department of Information Technology and Mathematical Methods

Working Paper

Series "*Mathematics and Statistics*"

n. 9/MS – 2008

2-D Wavelet-Based Spectra with Application in Analysis of Geophysical Images

by

Orietta Nicolis, Claudio Garutti, Brani Vidakovic

° Viale Marconi. 5, I – 24044 Dalmine (BG), ITALY, Tel. +39-035-2052339; Fax. +39-035-562779

[§] L'accesso alle *Series* è approvato dal Comitato di Redazione. I *Working Papers* della Collana dei Quaderni del Dipartimento di Ingegneria dell'Informazione e Metodi Matematici costituiscono un servizio atto a fornire la tempestiva divulgazione dei risultati dell'attività di ricerca, siano essi in forma provvisoria o definitiva.

2-D Wavelet-Based Spectra with Application in Analysis of Geophysical Images

ORIETTA NICOLIS

Dept. of Information Technology and Mathematical Methods

V.le Marconi 5, 24044 Dalmine BG I, Italy

CLAUDIO GARUTTI

Dept. of Information Engineering

Via Gradenigo 6/A, 35131 Padova, Italy

BRANI VIDAKOVIC

Georgia Institute of Technology

Atlanta, GA 30332-0205, USA

Abstract

We propose a wavelet-based spectral method for estimating the (directional) Hurst parameter in isotropic and anisotropic non-stationary fractional Gaussian fields. The method can be applied to self-similar images and, in general, to d-dimensional data that scale. In the application part, we consider denoising of 2-D fractional Brownian fields and the classification of the clouds/temperature

satellite images. In the first application, we use Bayesian inference calibrated by information from the wavelet-spectral domain to separate the signal, in this case the 2-D Brownian field, and the noise. For the classification of geophysical images we first estimate directional Hurst exponents and use them as an input to selected machine learning algorithms.

KEY WORDS: Scaling, Wavelets, Self-similarity, 2-D wavelet spectra.

1 Introduction

It is more the rule than the exception that high frequency data collected in real-life experiments, scale in a regular fashion. This scaling is manifested as regular "propagation of energy" when observations are inspected at different scales/frequencies, and this regularity is often described as ubiquitous or omnipresent. Examples are numerous: high frequency bio-responses, atmospheric data, stock market and exchange rates fluctrations, internet data, etc. In many scenarios involving analysis of such data, standard statistical modeling techniques are simply not applicable.

The methodology used to analyze scaling is based on analysis of autocovariances. The covariance dynamic in the domain of original data corresponds to the "energy" scaling in the frequency or scale domains. The term "energy" is an informal name for the squared coefficients in frequency/scale representations of signals and images. The standardly accepted measure of regular scaling is the Hurst exponent which can also be connected with measures of long memory, dimension, and fractality in signals and images.

Many strategies for assessing the Hurst exponent exist. This assessment can be done in the domain of the original measurements or in some transformed domain (usually scale/frequency domains such as Fourier or wavelet). From the statistical point of view the Hurst exponent can be viewed as an informative summary of data and in many cases standard statistical techniques are applied not on the data directly, but on their scaling exponents (classification, regression, statistical design).

The literature on the topic is vast (Beran, 1994; Chan and Wood, 2000; Constantine and Hall, 1994; *et al.*, 1995; Mandelbrot and Van Ness, 1968; Pesquet-Popescu, 1999a,b; Pipiras, 2004, 2005; Taqqu *et al.*, 1997). Most of the published research concerns the scaling in one-dimensional data. The 1-D theoretical models are well understood, the estimation and simulation methodology is conceptually and computationally straightforward, and univariate high frequency signals that scale are abundant. The definition of scaling in higher dimensions is more complex since multiple formulations are possible. If the scaling is spatial, various directions and choices of neighborhoods are possible, leading to the definition of various anisotropies. Also, the computational complexity of estimation and simulation methods is higher. However, in geophysical, medical, and other applications the scaling analysis of images and higher-dimensional objects is critically important.

The contribution of the paper is in formal definition of d-dimensional wavelet spectra, assessment of some of its properties, and in providing real-life applica-

tions for two dimensional case.

The definition of d-dimensional spectra is conceptually straightforward but computationally challenging due to lack of d-dimensional wavelet transforms in standard software. Parra *et al.* (2003) consider 2-D spectra based only on a hierarchy of diagonal detail spaces, while Heneghan *et al.* (1996) consider general definition based on continuous wavelets.

The idea behind the definition of d-dimensional wavelet spectra is the following: since the tensor-product wavelet multiresolution analysis of d dimensional data comprises of $2^d - 1$ detail spaces, with each space containing the hierarchy of subspaces with nested dyadic resolutions, it is quite natural to assess the energy scaling in each hierarchy. This leads to $2^d - 1$ concurrent power spectra describing a single d-dimensional data set. For example, multiresolution analysis of images leads to three detail spaces described as "horizontal", "vertical" or "diagonal," depending on the selection of the decomposing 2-D wavelet, or equivalently, the order of applications of high- and low-pass wavelet filters on the rows and columns of 2-D objects. Each of the three directional detail spaces contains a nested hierarchy of submatrices corresponding to image details at different scales and each leads to a distinctive power spectrum.

In this work we provide a simulation study to show that the 2-D wavelet spectra gives a consistent estimator of H. In the isotropic case, it can be considered an alternative method to the estimation procedures proposed by other authors (Istas and Lang, 1997; Chan and Wood, 2000; Zhu and Stein, 2002). Moreover, in the

anisotropic case, we show that it can be proposed as a simple and fast new method for estimating the directional Hurst exponents. We then provide two applications in which the 2-D wavelet spectra is instrumental. In the first application we use the 2-D spectra to estimate the Hurst exponent in the signal part of a noisy image and use this information in deconvolution of noise. In the second application we use the 2-D spectra parameters as discriminatory descriptors in classification of geophysical images. These type of images are influenced by "background conditions" (such as clouds, wind, temperature level, humidity, etc.) which can be well summarized by directional Hurst exponents.

The paper is organized as follows. In Section 2 we discuss the theoretical background necessary for describing self-similarity and scaling in d dimensions. In Section 3 we define the wavelet based directional spectra and discuss some of its properties. Section 4 provides simulational and comparative study in which a noisy fractional Brownian field is filtered using properly calibrated Bayes rules. In this section a real-life application of the proposed methodology is discussed: classification of satellite images with respect to time of their acquisition. In Section 5 we provide conclusions and delineate some possible directions for future research.

2 Background

In this section we review the most popular statistical model for data that regularly scale, the fractional Brownian motion. In one dimension, this process is a unique

Gaussian self-similar process with stationary increments. We also briefly discuss multidimensional wavelet domains since the spectrum will be defined there.

2.1 Fractional Brownian motion and extensions

The *fractional Brownian motion* (fBm) is one of the most popular models for modeling self-similar phenomena. It is a Gaussian, zero mean, non-stationary stochastic process, originally proposed by Mandelbrot and Van Ness (1968). This process is indexed by self-similarity parameter H, also known as the Hurst exponent. In a one-dimensional case the fBm process, denoted by $\{B_H(t), t \in \mathbb{R}\}$, is characterized by the following correlation function

$$R_{B_H}(t,s) = E\{B_H(t)B_H(s)\} = \frac{\sigma_H^2}{2}\left[|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right], \qquad (1)$$

where $\sigma_H^2 = \Gamma(1-2H)\frac{\cos(\pi H)}{\pi H}$ and $0 < H < 1$. As it can be seen from (1), the fBm is a non stationary process ($R_{B_H}(t,s)$ is a function of $|t-s|$), but it has stationary increments. In addition, the fBm is a *self-similar process*, that is, for all $a > 0$ it satisfies $B_H(at) \stackrel{d}{=} a^H B_H(t)$, where $\stackrel{d}{=}$ denotes the equality in distribution. Because of its nonstationarity, the spectrum in the strict sense does not exist. From the correlation function (1) and the definition of generalized power spectrum (see Reed *et al.*, 1995) we obtain the "power spectrum" of $B_H(t)$ as

$$S_{B_H}(\omega) = |\omega|^{-2H-1}. \qquad (2)$$

These definitions can be extended to any dimension. Unlike the 1-D case the generalization of fBm to higher dimensions is not unique. A simple generalization

to a 2-D surface is the *fractional Brownian field* (fBf). The fBf is a Gaussian, zero mean, random field $B_H(\mathbf{u})$, where $\mathbf{u}$ denotes the position in a selected domain, usually $[0,1] \times [0,1]$. Then, the autocorrelation function is

$$R_{B_H}(\mathbf{u},\mathbf{v}) = E\left[B_H(\mathbf{u})B_H(\mathbf{v})\right] = \frac{\sigma_H^2}{2}\left(\|\mathbf{u}\|^{2H} + \|\mathbf{v}\|^{2H} + \|\mathbf{u}-\mathbf{v}\|^{2H}\right), \quad (3)$$

where $0 < H < 1$, the variance σ_H^2 is

$$\sigma_H^2 = \frac{2^{-(1+2H)}\Gamma(1-H)}{\pi H \Gamma(1+H)} \quad (4)$$

and $\|\cdot\|$ is the usual Euclidean norm in $\mathbb{R}^2$. The increments of a fBf represent stationary, zero mean Gaussian random fields: the variance of such increments depends only on the distance $\|\mathbf{h}\|$ so that

$$E\left[B_H(\mathbf{u}+\mathbf{h}) - B_H(\mathbf{u})\right]^2 = \sigma_H^2 \|\mathbf{h}\|^{2H},$$

where σ_H^2 is given in (4). As in the one-dimensional case, the power spectrum of a fBf is obtained from the correlation function (3) and the 2-D Generalized Power Spectrum can be defined as

$$S_{B_H}(\omega) = \|\boldsymbol{\omega}\|^{-2H-2}. \quad (5)$$

The extension to a d-dimensional case is straightforward (see Reed *et al.* 1995). For a $d-$dimensional fractional Brownian motion the correlation function is given by Eq. (3) with **u, v** in $\mathbb{R}^d$ and

$$\sigma_H^2 = \frac{2^{-1-d-2H}\Gamma(1-H)}{\pi^{\frac{d}{2}} H \Gamma(\frac{d}{2}+H)}, \quad (6)$$

In this case, the power spectrum is

$$S_{B_H}(\omega) = \|\boldsymbol{\omega}\|^{-2H-d} \tag{7}$$

Many generalizations have been proposed to define anisotropy in Gaussian random fields. Bonami and Estrade (2003) defined an anisotropic fractional Brownian field, with stationary increments, by considering a spectral density of the form

$$S(\boldsymbol{\omega}) = \|\boldsymbol{\omega}\|^{-2H(\boldsymbol{\omega})-2},$$

where $H(\boldsymbol{\omega}) \in (0,1)$ is an even function which depends on the direction $\frac{\boldsymbol{\omega}}{|\boldsymbol{\omega}|}$ of $\mathbb{R}^2$. Popescu and Vehel (2002) introduced anisotropy by linear spatial transforms of an isotropic fractional field. Related generalizations are Kamont (1996), Wu and Xiao (2005), Peltier and Lévy-Véhel (1996), and Benassi *et al.*(1997).

2.2 Wavelets

Wavelets are the building blocks of wavelet transforms the same way that the functions e^{inx} are the building blocks of the ordinary Fourier transform. But in contrast to sines and cosines, wavelets can be supported on an arbitrarily small closed interval. Wavelet domains provide adaptive locality which is manifested as a trade off between the scale and time localizations subject to Heisenberg's law constraints. Basics on wavelets can be found in many texts, monographs, and papers at many different levels of exposition. The interested reader should consult monographs by Daubechies (1992), Vidakovic (1999), among others.

In two or higher dimensions wavelets provide an appropriate tool for analyzing self-similar objects and in particular, fractional Gaussian fields. The *energy preservation* in orthogonal wavelet analysis allows for defining wavelet spectra in a manner similar to that in the Fourier domains. Operationally, the traditional, tensor-product, 2-D wavelet transforms are constructed through the translations and the dyadic scaling of a product of univariate wavelets and scaling functions,

$$\begin{aligned} \phi(u_x, u_y) &= \phi(u_x) \cdot \phi(u_y) \\ \psi^h(u_x, u_y) &= \phi(u_x) \cdot \psi(u_y) \\ \psi^v(u_x, u_y) &= \psi(u_x) \cdot \phi(u_y) \\ \psi^d(u_x, u_y) &= \psi(u_x) \cdot \psi(u_y), \end{aligned} \tag{8}$$

which are known as *separable* 2-D wavelets. The symbols h, v, d in (8) stand for horizontal, vertical and diagonal directions, respectively, and relate to the ability of atoms in (8) to describe features along these three directions. Any function $f \in L_2(\mathbb{R}^2)$ can be represented as

$$f(\mathbf{u}) = \sum_{\mathbf{k}} c_{j_0,\mathbf{k}} \phi_{j_0,\mathbf{k}}(\mathbf{u}) + \sum_{j \geq j_0} \sum_{\mathbf{k}} \sum_{i} d^i_{j,\mathbf{k}} \psi^i_{j,\mathbf{k}}(\mathbf{u}) \tag{9}$$

where $\mathbf{u} = (u_x, u_y) \in \mathbb{R}^2$, $i \in \{h, v, d\}$, $\mathbf{k} = (k_1, k_2) \in Z^2$, and

$$\begin{aligned} \phi_{j,\mathbf{k}}(\mathbf{u}) &= 2^j \phi(2^j u_x - k_1, 2^j u_y - k_2) \\ \psi^i_{j,\mathbf{k}}(\mathbf{u}) &= 2^j \psi^i(2^j u_x - k_1, 2^j u_y - k_2). \end{aligned}$$

for $i = h, v, d$. The decomposition in (9) can be extended to an arbitrary function

$f \in L_2(\mathbb{R}^d)$,

$$\begin{aligned} f(\mathbf{u}) &= \sum_{\mathbf{k}} c_{j_0,\mathbf{k}} \phi_{j_0,\mathbf{k}}(\mathbf{u}) \\ &\quad + \sum_{j \geq j_0} \sum_{\mathbf{k}} \sum_{i=1}^{2^d-1} d^i_{j,\mathbf{k}} \psi^i_{j,\mathbf{k}}(\mathbf{u}), \end{aligned} \tag{10}$$

where $\mathbf{k} = (k_1, \ldots, k_d) \in Z^d$, $\mathbf{u} = (u_1, \ldots, u_d) \in \mathbb{R}^d$, and

$$\begin{aligned} \phi_{j;\mathbf{k}}(\mathbf{u}) &= 2^{jd/2} \prod_{i=1}^{d} \phi(2^j u_i - k_i) \\ \psi^l_{j;\mathbf{k}}(\mathbf{u}) &= 2^{jd/2} \prod_{i=1}^{d} \xi(2^j u_i - k_i) \\ &\quad \text{with } \xi = \phi \text{ or } \psi, \text{ but not all } \xi = \phi. \end{aligned}$$

The index l corresponds to one of $2^d - 1$ possible directions. The d-dimensional wavelet spectra will be defined using the wavelet coefficients in (10), namely $2^d - 1$ nested detail spaces with coefficients $d^i_{j,\mathbf{k}}$, along the scale index j.

3 Wavelet Spectra of fBf

Time-frequency or time-scale transforms are indispensable tools in analysis of the signals and images. The spectra defined by such representations describe distribution of energies in the signal/image along a range of frequencies or scales. Various definitions of spectra exist, depending on the signal representation. Wavelets and wavelet based spectra have been instrumental in analysis of self-similarity (Flandrin, 1989, 1992; Doukhan *et al.*, 2003; Wornell, 1995). Orthogonal discrete

wavelets are "energy preserving," and as such, are natural for defining the power spectrum.

Suppose that 1-D signal y of length n has wavelet decomposition $d = Wy = (c_{j_0}, d_{j_0}, d_{j_0+1}, \ldots, d_j)$, where j_0 is a fixed level smaller than $j = \log_2 n - 1$, c_{j_0} are scaling coefficients, and d_j's are levels of detail coefficients. The wavelet spectra is defined as

$$S(j) = \log_2 \left(\overline{d_j^2} \right)$$

where $\overline{d_j^2}$ is an average of squared components in vector of detail coefficients at level j.

In the 2-D case three different hierarchies constitute detail spaces and the natural definition of wavelet spectra involves power spectra corresponding to the three hierarchies. Since the detail hierarchies are characterized by their direction (horizontal, vertical and diagonal), this spectra will be sensitive in assessing the energy content and dissipation along the angles of 0, $\pi/2$, and $\pi/4$.

Consider a fBf process $B_H(\mathbf{u})$, the standard model for self-similar isotropic random fields. For this process the wavelet coefficients are given by

$$d^i_{j,\mathbf{k}} = 2^j \int B_H(\mathbf{u}) \psi^i(2^j \mathbf{u} - \mathbf{k}) d\mathbf{u}, \tag{11}$$

where the integral is taken over $\mathbb{R}^2$ and $i = h, v$ or d. The detail coefficients are random variables with zero mean and variance

$$E\left[\left| d^i_{j,\mathbf{k}} \right|^2 \right] = 2^{2j} \iint \psi^i \left(2^j \mathbf{u} - \mathbf{k} \right) \psi^i \left(2^j \mathbf{v} - \mathbf{k} \right) E\left[B_H(\mathbf{u}) B_H(\mathbf{v}) \right] d\mathbf{u} d\mathbf{v}, \tag{12}$$

(Heneghan *et al.*, 1996). From the definition of ψ^h and ψ^v in (8), the integrand in (12) is symmetric and the variances of the wavelet coefficients coincide for these two directions, that is

$$E\left[\left|d^h_{j,\mathbf{k}}\right|^2\right] = E\left[\left|d^v_{j,\mathbf{k}}\right|^2\right]. \tag{13}$$

These two variances differ from the variance of the wavelet coefficients from the diagonal hierarchy. From (12) one can derive

$$E\left[\left|d^i_{j,\mathbf{k}}\right|^2\right] = \frac{\sigma_H^2}{2} V_{\psi^i} 2^{-(2H+2)j}, \tag{14}$$

where

$$V_{\psi^i} = \iint \psi^i(\mathbf{p}+\mathbf{q}) \cdot \psi^i(\mathbf{q}) \left|\mathbf{p}\right|^{2H} d\mathbf{p} d\mathbf{q} \tag{15}$$

depends only on wavelets ψ^i and exponent H, but not on the scale j. The derivation of this result is deferred to Appendix.

An application of the logarithm to both sides of Eq. (14) leads to the following equation

$$\log_2 E\left[\left|d^i_{j,\mathbf{k}}\right|^2\right] = -(2H+2)j + C_i, \tag{16}$$

where

$$C_i = \log_2 \frac{\sigma_H^2}{2} V_{\psi^i}(H). \tag{17}$$

The Hurst coefficient of a fBf be estimated from the slope in the linear equations given in (16). The empirical counterpart of (16) is a regression defined on pairs

$$\left(j, \log_2 \overline{\left|d^i_{j,\mathbf{k}}\right|^2}\right), \quad i = h, v, d, \tag{18}$$

where $\overline{\left|d^i_{j,\mathbf{k}}\right|^2}$ is an empirical counterpart of $E\left[\left|d^i_{j,\mathbf{k}}\right|^2\right]$. The sample mean in (17) can be replaced by sample median or any other location estimation to produce more robust estimators of the spectra. Also the regression should be weighted since the variances in the detail levels are not equal. Veich and Abry (1999) discuss the bias of estimators in (18) and the method to correct it. When the field is isortopic, the Hurst exponent H can be estimated as an average of corresponding directional estimates.

An anisotropic generalization is straightforward; the parameter H in the above equations (11-18) can depend on the direction i and may be replaced by H_i. Given the wavelet ψ, the intercept C_i in (16) is uniquely determined by H, and initial energy, σ^2_H. Thus, if H and C_i vary independently, a novel, wavelet-specific class of anisotropic self-similar random fields can be defined. A few examples are provided for the isotropic and anisotropic cases.

Example 1. Figure 1(a) depicts a simulated isotropic fractional Brownian field with $H = 0.3$. Its 2-D wavelet spectra based on the *Symmlet 4* filter, shown in Figure 1(b), demonstrates the estimation process is consistent. The resulting estimates are $\hat{H}_h = 0.295$, $\hat{H}_v = 0.298$, and $\hat{H}_d = 0.299$ for the horizontal, vertical and diagonal directions, respectively, which are close to the original simulated value of $H = 0.3$. In order to select the best basis and test performance of the wavelet-based estimator, we simulated 1000 fractional Brownian fields with various H and for each field estimated the Hurst parameter in each of the three directions. The averaged wavelet-based estimator was compared with Quadratic

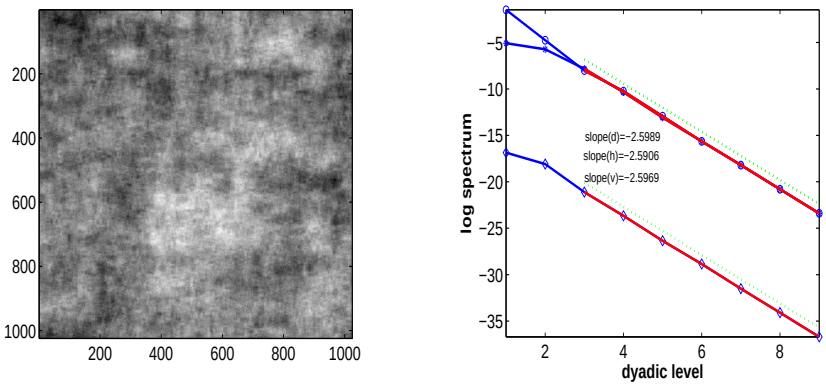
Variations (QV) estimator introduced by Istas and Lang (1997). As is typical for many wavelet based procedures, the choice of basis is important but not decisive for the results and performance of the estimation algorithm. We comprehensively explored Daubechies, Symmlet and Coiflet families for a range of parameter values (vanishing moments) and the differences found were not significant (the Haar basis being an exception). We adopted short filters with different smoothness and symmetry properties such as the Daubechies 4, Symmlet 4 and Coiflet 1 (in the notation of Donoho's Wavelab). For comparative purposes we used Symmlet 4 since this filter provides a good compromise of smoothness, locality and near-symmetry. In Table 1 we provide the summary of this experiment. We found that the averaged wavelet-based estimates are close to those obtained by the QV method for isotropic fields simulated with $H = 0.4$ and with $H = 0.6$. These two exponents are selected to represent antipersistency and long memory. Moreover, the wavelet based estimator is more robust when the data are contaminated by noise, even at a low level. For the estimation procedure of the wavelet spectra, we used the Ordinary Least Square (OLS) estimator. We also implemented weighted least squares (WLS) in the spirit of Veitch and Abry (1999). For calculating regression weights we resampled detail spaces to obtain a surrogate sample of logarithms of average level energies. These are further utilized to obtain proper weights via bootstrap variances. The difference between the weighted and the ordinary least squares regression was found to be minimal which was a consequence of typically large sample sizes in image processing. The WLS also substantially

	$H = 0.4$		$H = 0.6$	
	snr$= \infty$	snr$= 20$	snr$= \infty$	snr$= 20$
D4	0.3920 (0.042)	0.3828 (0.325)	0.6027 (0.035)	0.5775 (0.340)
S4	0.3968 (0.042)	0.3838 (0.326)	0.6005 (0.033)	0.5766 (0.359)
C1	0.3865 (0.042)	0.3766 (0.041)	0.5917 (0.037)	0.5715 (0.045)
Haar	0.3508 (0.041)	0.3427 (0.040)	0.5554 (0.035)	0.5327 (0.045)
QV	0.3886 (0.016)	0.3365 (0.027)	0.5886 (0.015)	0.2663 (0.086)

Table 1: Means and standard deviations (in brackets) of the estimated Hurst exponents, by the wavelet-based estimators (D4, S4, C1 and Haar) and the QV estimator, evaluated on 1000 simulated random fields with $H = 0.4$ and $H = 0.6$ and length $n = 256 \times 256$, with and without noise in each case.

increased computational complexity. For example, in the case of $n = 10$ simulated images 512×512 with $H = 0.6$ the OLS gave $\hat{H} = 0.6238$ with a standard deviation of 0.04 while the WLS gave $\hat{H} = 0.6232$ with similar standard deviation.

Example 2. In the second example we simulated an anisotropic Gaussian random field using Daubechies 4 wavelet by controlling the scaling of variances in detail spaces. An example of a simulated field is given in Figure 2(a). In particular, we considered scaling equivalent to Hurst parameters equal to $H_h = 0.3$, $H_v = 0.8$, and $H_d = 0.5$, for the horizontal, vertical, and diagonal direction, respectively. The 2-D wavelet-spectra assessed by a wavelet different than generating (Daub 8)



(a) (b)

Figure 1: (a) Isotropic fractional Brownian field with $H = 0.3$; (b) The wavelet spectra of the field in (a) estimated by Symmlet 4.

gave the following estimates: $\hat{H}_h = 0.297$, $\hat{H}_v = 0.820$, and $\hat{H}_d = 0.511$ (Figure 2(b)) which are very close to those utilized in the simulation of the field. The goal of this exercise was to produce a specific directional anisotropy and to check that 2-D spectra consistently estimated the scaling when basis was changed.

In order to demonstrate the behavior of the intercepts C_i for each direction, we have simulated $N = 200$ isotropic fractional Brownian fields on a regular grid (512×512) with parameters H ranging from $H = 0.1$ to $H = 0.9$. Figure 3 plots the average difference of the intercepts for the horizontal and diagonal directions, $\overline{C}_h - \overline{C}_d$. It is evident that the intercept is affected by the value of H: for higher H the estimated difference $\overline{C}_h - \overline{C}_d$ is larger. The message of this analysis is the following: even for isotropic random fields the amount of energy attributed to different directions differs. Note that $C_i = \log_2(\sigma_H^2 V_{\psi^i})$, where σ_H^2 is the variance

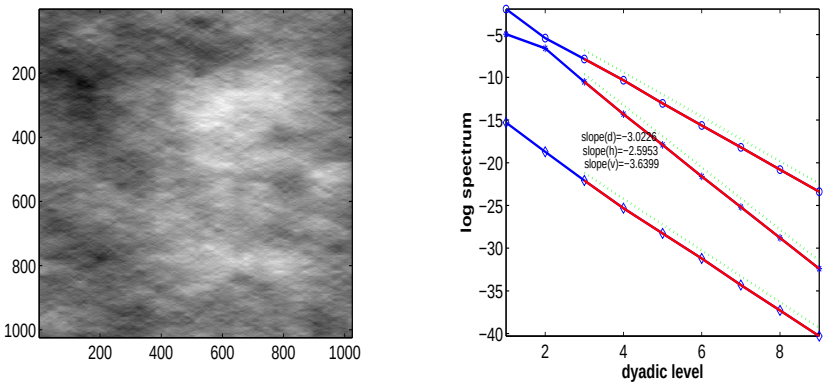


Figure 2: (a) Gaussian random field with $H_h = 0.3$, $H_v = 0.8$, and $H_d = 0.5$ simulated by Daubechies 4; (b) The 2-D wavelet spectra of the field in (a) estimated by Symmlet 4.

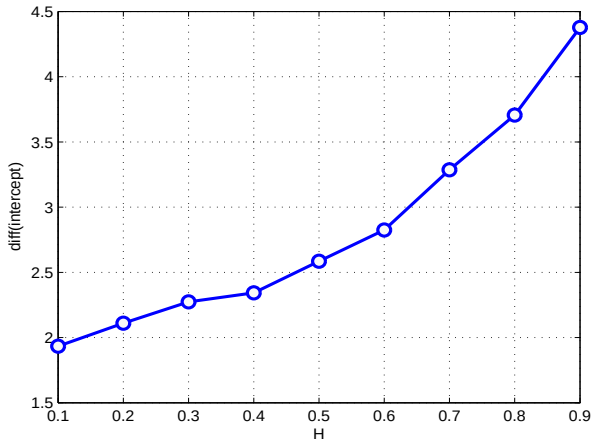


Figure 3: Average differences of the intercepts for horizontal and diagonal directions, $\overline{C}_h - \overline{C}_d$, for different values of parameter H.

of fBf and V_{ψ^i} is given in (15), and numerical evaluation of an intercept C_i is possible but involved due to non-existence of finite form for ψ^i. Evaluation of C_i's is critical for the simulation of random fields using 2-D wavelets, since the C_i's specify how the total energy should be distributed to the directional subspaces.

4 Applications

In this section we provide two applications in which 2-D wavelet spectra is instrumental. The first application concerns a denoising task in which the signal image exhibits scaling and the noise is white. The separation of the signal image and noise is done by Bayesian wavelet filtering calibrated by the properties of the 2-D spectra of the signal-part. The second application involves the statistical task of image classification with 2-D spectra parameters as discriminatory descriptors. In the spirit of reproducible research all MATLAB© codes utilized in these applications are available from the authors on request.

4.1 Bayesian filtering guided by spectral information

In this application we demonstrate how 2-D wavelet spectra can be utilized in filtering noisy images in which the "signal" part scales.

Suppose the observed image $\mathbf{y}$ is a convolution of an unknown "true" image $\mathbf{s}$, exhibiting scaling, and a random noise $\mathbf{e}$,

$$\mathbf{y} = \mathbf{s} + \mathbf{e}. \tag{19}$$

It is assumed that the random noise is a matrix of i.i.d. zero mean Gaussians with standard deviation σ_e and that the "true" image is well modeled by an isotropic fractional Brownian field, with its parameter H not known in advance. In the wavelet domain, expression (19) becomes $d^i_{j\mathbf{k}} = \theta^i_{j\mathbf{k}} + \varepsilon^i_{j\mathbf{k}}$, where $d^i_{j\mathbf{k}}, \theta^i_{j\mathbf{k}}$, and $\varepsilon^i_{j\mathbf{k}}$ are the $(j, \mathbf{k})$ coordinates in the traditional scale/shift wavelet-enumeration of transformed images $\mathbf{y}, \mathbf{s}$ and $\mathbf{e}$, respectively. This model preservation is a consequence of the linearity and orthogonality of wavelet transforms. In the exposition that follows, we omit the double index $j, \mathbf{k}$ and the direction i, and work with a "typical" wavelet coefficient, d. The conditional distribution of d given θ and σ^2, $[d|\theta, \sigma^2]$, is $\mathcal{N}(\theta, \sigma^2)$. We utilize Bayesian Adaptive Multiscale Shrinkage (BAMS), a technique proposed in Vidakovic and Ruggeri (2001) to statistically estimate wavelet coefficients, corresponding to fBf, using a shrinkage rule in a Bayesian framework.

In BAMS, σ^2 and θ are assumed to be independent random variables. The variance σ^2 is given exponential $\mathcal{E}(\mu)$ prior, while θ is given a mixture prior, as standardly done. The mixture prior consists of a point mass at zero (representing the "parsimony" part) and a double exponential distribution (representing the "spread" part) mixed in proportion $(1 - \varepsilon) : \varepsilon$,

$$(1 - \varepsilon)\delta_0 + \varepsilon \mathcal{DE}(0, \tau).$$

The resulting Bayes rule is given by:

$$\delta^*(d) = \frac{(1 - \epsilon)\ m(d)\ \delta(d)}{(1 - \epsilon)\ m(d) + \epsilon\ \mathcal{DE}\left(0, \frac{1}{\sqrt{2\mu}}\right)}, \tag{20}$$

and this is the shrinkage rule to be utilized. In (20),

$$m(d) = \frac{\tau e^{-|d|/\tau} - \frac{1}{\sqrt{2\mu}} e^{-\sqrt{2\mu}|d|}}{2\tau^2 - 1/\mu}$$

is the prior-predictive distribution of d, and

$$\delta(d) = \frac{\tau(\tau^2 - 1/(2\mu))de^{-|d|/\tau} + \tau^2(e^{-|d|\sqrt{2\mu}} - e^{-|d|/\tau})/\mu}{(\tau^2 - 1/(2\mu))(\tau e^{-|d|/\tau} - (1/\sqrt{2\mu})e^{-|d|\sqrt{2\mu}})}$$

is the Bayes rule only for the spread part of prior, $\mathcal{DE}(0, \tau)$.

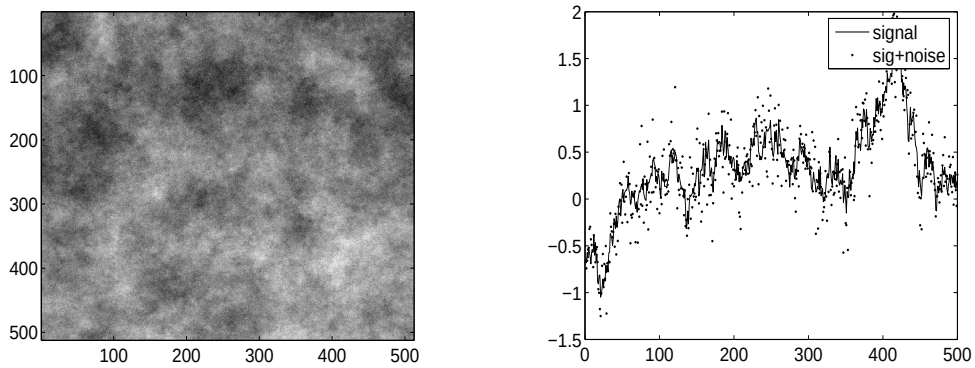
The Bayes rule (20) depends on the hyper-parameters ε, τ, and μ. The elicitation of the hyper-parameters is critical for good performance of Bayesian filtering. Default choices for the hyperparameters are not suitable in function estimation, since observations vary tremendously and to accommodate for possibility of widely different images and signal-to-noise ratios, a degree of informativeness and/or data dependence needs to be exploited. The hyper-parameters have been set using Empirical Bayes (EB) arguments, as in Vidakovic and Ruggeri (2001) or Katul *et al.* (2006). The rule in (20) is close to a thresholding rule: it heavily shrinks small-magnitude arguments while the large arguments are only slightly shrunk.

After the rule in (20) is calibrated, the separation of $\mathbf{s}$ and $\mathbf{e}$ is performed as follows: each wavelet coefficient d is split as

$$d = \delta^*(d) + (d - \delta^*(d)) = \hat{\theta} + \hat{\varepsilon},$$

with $\delta^*(d)$ and $(d - \delta^*(d))$ estimating signal and noise contributions, respectively. All $\hat{\theta} = \delta^*(d)$ form a matrix $\widehat{\boldsymbol{\theta}}_{j,\mathbf{k}}$ which back-transformed to the original domain gives the estimator of denoised image.

This process of filtering is illustrated in Figures 4 and 5, where we consider a fBf with Hurst exponent $H = 1/3$ with an additive Gaussian noise with signal-to-noise ratio equal to $snr = 2$. Fig. 4 (a) shows a simulated fBf. In order to emphasize the effect of the noise on the image, we show in panel (b) of Figure 4 the 100th row of the simulated image with and without noise.



(a) (b)

Figure 4: a) Simulated fBf with $H = 1/3$; (b) Signal of $100-$th and signal with noise.

Figure 5(a) shows the wavelet power spectra for the noisy and denoised images based on the *Symmlet 4* wavelet filter. The estimates for H, in each direction, inferred from the slopes in Eq. (16), are $\hat{H}_h = 0.3204$, $\hat{H}_v = 0.3161$, and $\hat{H}_d = 0.2739$. Note the flattening of directional spectra (solid lines) of the noisy image. That means that noise which is affecting all scales and all coefficients, has sufficient energy to leave its signature only on the finest few levels of detail (high dyadic resolution) where its relative energy compared to that of the signal image

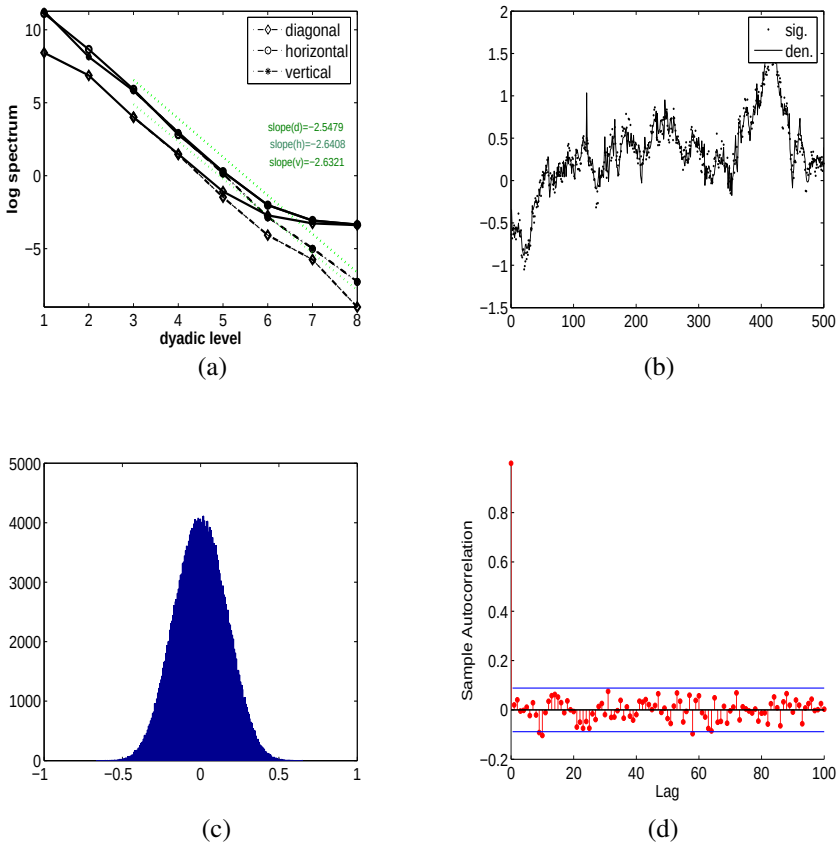


Figure 5: (a) 2-D wavelet-based spectra. Black solid lines depict the directional spectra of the noisy images while the gray lines are "straightened" and correspond to denoised image; (b) Signal of 100th row and denoised image; (c) Histogram of all residuals; and (d) Autocorrelation function of 100th row in the residual matrix.

is high.

The described filtering procedure based on Bayesian rule was able to recover the simulated image since the estimator of H is close to the original and, at the same time, the estimator of the noise is close to a matrix of i.i.d. Gaussians. This is evident from Figure 5. Figure 5 (b) shows the performance of the denoising procedure by comparing the 100th row of the original image to the 100th row reconstructed image after applying the Bayes rule. The marginal distribution of magnitudes of all residuals, depicted in panel (c), suggests that the components have zero-mean and bell-shaped distribution consistent with the originally simulated noise. In order to show that the residuals are not "colored", we again selected 100th row (out of 512) from the estimated noise matrix. In addition to their marginal Gaussianity, the autocorrelation of components in the selected vector is consistent with their "whitenes", i.e., no autocorrelations at nonzero lags are significant. The autocorrelations for the first 100 lags are shown in panel (d).

4.2 Classification of Cloud/Temperature Maps

In this application we illustrate how the wavelet-based estimators of directional Hurst exponents can be utilized in classification of satellite images. The emphasis here is on discrimination abilities of the Hurst summaries, and not on a solution of a realistic environmental problem. It is straightforward to implement the described analysis in various scientific areas in which 2-D data are instrumental: medical imaging, geoscience, industrial applications, etc.

The source of the data is EUMETSAT (`http://www.eumetsat.int`). EUMETSAT is an intergovernmental organization created through an international convention signed by 17 European Member States. EUMETSAT's Meteosat system is intended primarily to support the National Meteorological Services (NMS) of Member States. The NMS in turn distributes the image data to other end-users, notably through the provision of forecasts on television several times a day. In addition to the provision of images of the Earth and its atmosphere every half an hour in three spectral channels (Visible, Infrared and Water Vapour), a range of processed meteorological parameters is produced.

The satellite receives that part of the sun radiation which is reflected by the earth surface or by cloudiness. It is a so-called *window channel* which means that radiation is not significantly absorbed by the gases in the troposphere. The satellite receives radiation which is emitted by the earth and the clouds because of their temperature. Infra Red (IR) images via window channel (Wavelength 3.9-13.4 microns (μ)) are useful for day and night cloud-mapping and determination of surface temperature. A range of grey shades in the IR channel represent different temperatures of the radiating surface which can either be the earth surface or the cloud tops.

Our data set contains 160 IR satellite images of the Gulf of Guinea (West cost of Africa and South Atlantic Ocean). The images are taken at 3.9 μ IR band for 40 consecutive days (11/1/2006 - 12/10/2006), and subsequently divided into 4 groups according to the hour of their acquisition: (i) night (0:12am), (ii) morning

(6:12am), (iii) noon (12:12pm), and (iv) evening (6:12pm). A typical observation (6:12am, 11/1/06, IR 3.9 μ) is shown in Figure 6.

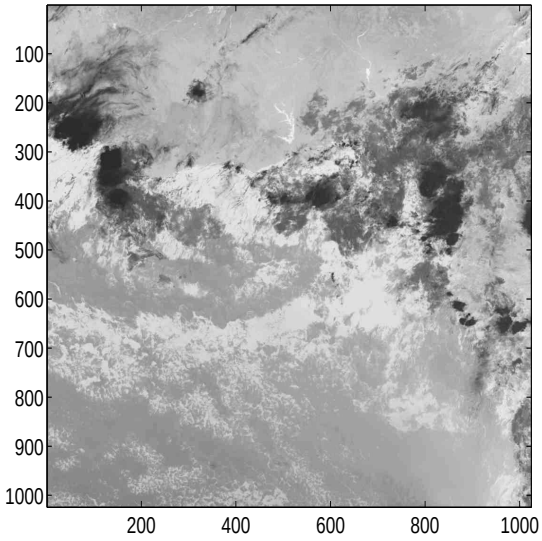


Figure 6: Satelite IR image with wavelength 3.9 μ, taken on November 11, 2006 at 06:12am (morning group).

There are factors, different than geography (terrain), possibly influencing the scaling in the satellite image. These "background conditions" such as clouds, wind, temperature level, humidity, etc., are influenced by the time of day and exhibit no regular behavior. We base discrimination on global scaling properties of the observed images since the scaling is influenced by the background conditions.

In order to assess the efficacy of Hurst exponents in separation of images to groups with different scaling characteristics, the following experiment was performed. We randomly selected a portion of the data to fit the classification model

Testing Proportion	Lin.	Quad.	Lin. SVM	Quad. SVM	Poly. SVM	RBF SVM
50%	0.072	0.081	0.085	0.088	0.084	0.086
30%	0.069	0.077	0.086	0.086	0.079	0.084

Table 2: Misclassification errors for selected classical classifiers Linear, Quadratic and SVM methods (Lin. SVM, Quad. SVM, Poly. SVM, and RBF SVM) for 50% and 30% of data used for testing.

and used the remainder of the data to assess the model. Two scenarios are considered: the first utilizes 50% of the data and the second 70% of the data as training sets. The random selection of training data was repeated 10000 times, so the reported prediction errors are averaged over 10000 runs.

Since several exogenous variables (such as temperature, wind, humidity, pressure) at noon differ from the same variables at the other times of the day, we considered only two groups for the classification purpose: the "noon" group and "others" group. For classifying the images we considered several classification procedures: linear, quadratic and SVM (Support Vector Machines) (see Hastie *et al.*, 2001). Table 2 provides the results obtained with linear and quadratic classifiers and with SVM with linear, quadratic, polynomial (with degree $d = 3$), and radial basis (with scaling parameter $c = 2$) kernel functions. Figure 7 depicts classification of "noon" and "others" images by the radial basis SVM (with $c = 2$).

As evidenced from Table 2 the standard linear classifier outperforms all others. This is a consequence of almost perfect linear separation between the "noon"

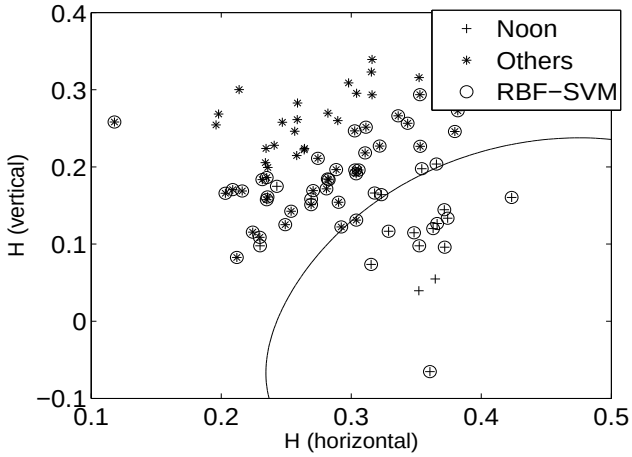


Figure 7: Training set (80 images) and non linear boundaries obtained by radial bases SVM ($c = 2$). The circles are the support vectors.

vs "others". We also performed classification experiments in which harder-to-separate cases are considered. In these cases, the SVMs were distinctly superior to the linear and quadratic classifiers. For example, in the "evening" vs "others" case the error rate was about 15% and in the "morning" vs "others" case about 25%. This latter rate was affected by scaling similarities between "morning" and "midnight" images belonging to different classes.

Figure 8 shows a linear classifier based on 50% of the data and the testing set consisting of the remaining 50%. Only a few observations are misclassified. On a horizontal-vertical Hurst plane, the *asterisks* correspond to the "noon" group while the *pluses* correspond to "others." The misclassified cases are circled. This example demonstrates that the 2-D spectra can capture information well on di-

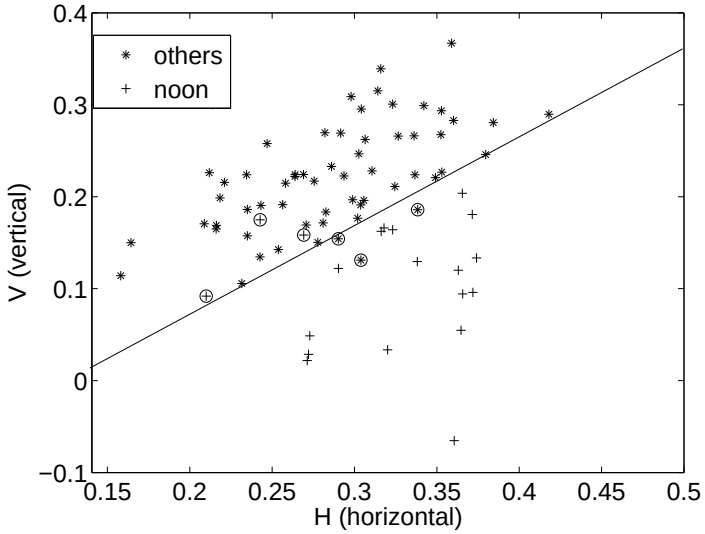


Figure 8: Testing set (based on 50% of data) and linear classifier. The circled observations are misclassified.

rectional anisotropies and produce discriminatory summaries. In particular, the horizontal and vertical directions captured most of the significant differences between the two groups of satellite images.

5 Conclusions

We have shown that 2-D wavelet-based spectra, evaluated through the sample moments of wavelet coefficients, can be used for estimating the Hurst parameter vector in a variety of self-similar random fields. Examples of standard isotropic fractional Brownian fields as well as anisotropic non-stationary Gaussian fields

are provided in the context of estimation of their directional Hurst parameters.

The methodology involving statistical models in the wavelet spectral domain has been developed and applied in denoising of composite images in which the "signal part" is self-similar. This is done by considering 2-D spectra and empirically calibrating a Bayesian shrinkage rule which preserves regular scaling in the estimator of the signal image and assures marginal normality and independence of the residuals. It is interesting that the signal image is a random field itself and this application is in fact a challenging deconvolution of two random fields.

We also utilized the 2-D wavelet spectra to classify geophysical images. In particular, we classified clouds/temperature map images to their corresponding groups by a linear discriminator fed by the vector of directional Hurst exponents. The Hurst descriptors have shown to be discriminatory, leading to a classifier with an excellent percentage of correct predictions (up to 93.1%). Further extension of this methodology to other classes of anisotropic processes is under investigation.

Acknowledgments: This work was done while the first and second authors were visiting the Georgia Institute of Technology. The work of O. Nicolis has been partially supported by Italian Research grants by MIUR (PRIN2004). Support for C. Garutti was provided in part by *Fondazione Ing. Aldo Gini* The work of B. Vidakovic was partially supported by the NSF grant DMS 0004131 at Georgia Tech. The authors are grateful to EUMETSAT, and in particular to Mr Bob Barrett for help about satellite images used in the application part.

Appendix

The tensor product wavelet transform are standard way of generating $d-$dimensional multiresolution analysis. The atomic function in (8) is such that $\{\psi^i_{j,k_1,k_2} = 2^{\frac{j}{2}}\psi^i\left(2^j u_x - k_1, 2^j u_y - k_2\right)\}$ for each direction i. For a fBm process, $B_{\,H}(u_x, u_y)$, the wavelet coefficients are

$$d^i_{j,k_1,k_2} = 2^j \iint B_{\,H}(u_x, u_y)\psi^h(2^j u_x - k_1, 2^j u_y - k_2)du_x du_y \tag{21}$$

By setting $\mathbf{k} = (k_1, k_2)$ and $\mathbf{u} = (u_x, u_y)$, Eq. (21) can be written as

$$d^i_{j,\mathbf{k}} = 2^j \iint B_{\,H}(\mathbf{u})\psi^i(2^j\mathbf{u} - \mathbf{k})d\mathbf{u}. \tag{22}$$

The variance of the detail coefficients $d^i_{j,\mathbf{k}}$ is obtained in a similar way to the continuous wavelet approach described in (Heneghan *et al.*, 1996), is

$$E\left[\left|d^i_{j,\mathbf{k}}\right|^2\right] = 2^{2j} \iint \psi^i\left(2^j\mathbf{u} - \mathbf{k}\right)\psi^i\left(2^j\mathbf{v} - \mathbf{k}\right) E\left[B_H(\mathbf{u})B_H(\mathbf{v})\right] d\mathbf{u}\, d\mathbf{v} \tag{23}$$

Taking into the account the definition of the covariance function for a fBm in 2-D, we have

$$\begin{aligned} E\left[\left|d^i_{j,\mathbf{k}}\right|^2\right] &= \frac{\sigma_H^2}{2}2^{2j} \iint \psi^i\left(2^j\mathbf{u} - \mathbf{k}\right)\psi^i\left(2^j\mathbf{v} - \mathbf{k}\right)|\mathbf{u}|^{2H} d\mathbf{u}d\mathbf{v} \\ &\quad + \iint \psi^i\left(2^j\mathbf{u} - \mathbf{k}\right)\psi^i\left(2^j\mathbf{v} - \mathbf{k}\right)|\mathbf{v}|^{2H} d\mathbf{u}d\mathbf{v} \\ &\quad + \int \psi^i\left(2^j\mathbf{u} - \mathbf{k}\right)\psi^i\left(2^j\mathbf{v} - \mathbf{k}\right)|\mathbf{u} - \mathbf{v}|^{2H} d\mathbf{u}d\mathbf{v}. \end{aligned} \tag{24}$$

Since,

$$\int \psi^i \left(2^j \mathbf{u} - \mathbf{k}\right) d\mathbf{u} = \int \left(2^j \mathbf{v} - \mathbf{k}\right) d\mathbf{v} = 0, \tag{25}$$

the variance in (24) can be simplified as

$$E\left[\left|d^i_{j,\mathbf{k}}\right|^2\right] = \frac{\sigma^2}{2} 2^{2j} \iint \psi^i \left(2^j \mathbf{u} - \mathbf{k}\right) \psi^i \left(2^j \mathbf{v} - \mathbf{k}\right) |\mathbf{u} - \mathbf{v}|^{2H} \, d\mathbf{u} d\mathbf{v} \tag{26}$$

By substituting $\mathbf{p} = 2^j(\mathbf{u} - \mathbf{v})$ and $\mathbf{q} = 2^j \mathbf{v} - \mathbf{k}$, we obtain:

$$\begin{aligned} E\left[\left|d^i_{j,\mathbf{k}}\right|^2\right] &= \frac{\sigma_H^2}{2} 2^{2j} \iint \psi^i \left(\mathbf{p} + \mathbf{q}\right) \psi^i \left(\mathbf{q}\right) 2^{-2jH} \, |\mathbf{p}|^{2H} \, 2^{-4j} d\mathbf{p} d\mathbf{q} \\ &= \frac{\sigma_H^2}{2} 2^{-j(2H+2)} \iint \psi^i \left(\mathbf{p} + \mathbf{q}\right) \psi^i \left(\mathbf{q}\right) |\mathbf{p}|^{2H} \, d\mathbf{p} d\mathbf{q} \\ &= \frac{\sigma_H^2}{2} V_{\psi^i} 2^{-j(2H-2)}, \end{aligned} \tag{27}$$

where V_{ψ^i} denotes

$$\iint \psi^i(\mathbf{p} + \mathbf{q}) \cdot \psi^i(\mathbf{q}) \, |\mathbf{p}|^{2H} \, d\mathbf{p} d\mathbf{q}, \tag{28}$$

an integral depending on wavelet ψ^i and H, but free of the scale j, and σ_H^2 was given in (4). By applying logarithm to both ends in

$$E\left[\left|d^i_{j,\mathbf{k}}\right|^2\right] = \frac{\sigma_H^2}{2} V_{\psi^i} 2^{-(2H+2)j}, \tag{29}$$

the equation

$$\log_2 E\left[\left|d^i_{j,\mathbf{k}}\right|^2\right] = -(2H+2)j + C_i,$$

discussed previously in Section 3 is obtained.

References

BENASSI, A., JAFFARD, S., and ROUX, D. (1997). Elliptic Gaussian random processes. *Rev. Mat. Iberoamericana*, 13, 1, 19-89.

BERAN, J. (1994). *Statistics for Long-Memory Processes*, Volume 61 of Monographs on Statistics and Applied Probability. New York: Chapman & Hall.

BONAMI, A. and ESTRADE, A. (2003). Anisotropic analysis of some Gaussian models, *Journal of Fourier Anal. and Appl.*, 9, Number 3 (2003) 215–236.

CHAN, G. and WOOD, A. T. A. (2000). Increment-based estimators of fractal dimension for two-dimensional surface data, *Statist. Sinica* 10, 343–376.

CONSTANTINE, A. G. and HALL, P. (1994). Characterizing surface smoothness via estimation of effective fractal dimension, *J. Roy. Statist. Soc. Ser. B* 56, 97–113.

DAUBECHIES, I. (1992). Ten lectures on wavelets, CBMS-NSF Regional Conference Series in Applied Mathematics, 61, SIAM, Philadelphia.

DOUKHAN, P., OPPENHEIM, G., and TAQQU, M. S. (2003). *Theory and Applications of Long-Range Dependence*, Birkhauser.

FLANDRIN, P. (1989). On the Spectrum of Fractional Brownian Motions, *IEEE Transaction on Information Theory*, 35, 197–199.

FLANDRIN, P. (1992). Wavelet analysis and synthesis of fractional Brownian motion, *IEEE Transaction on Information Theory*, 38, 910–917.

HENEGHAN, C., LOWEN, S.B., and TEICH, M.C. (1996). Two dimensional fractional Brownian motion: wavelet analysis and synthesis. Image Analysis and Interpretation, *Proceedings of the IEEE Southwest Symposium*, 213–217.

KAMONT, A.(1996). On the fractional anisotropic Wiener field, *Probab. Math. Statist.* 16, 85-98.

KATUL, G. G., RUGGERI, F., and VIDAKOVIC, B. (2006). BAMS filtering and applications to denoising ozone concentration measurements, *Journal of Statistical Planning and Inference*, 136, 2395–2405.

MANDELBROT, B. B. and VAN NESS, J. W. (1968). Fractional Brownian Motions, Fractional Noises and Applications, *SIAM Review* 10(4) 422–437.

PARRA, C., IFTEKHARUDDIN, K., and RENDON, T. (2003). Wavelet based estimation of the fractal dimension in fBm images. Presented in the *1st IEEE EMBS conference on Neural Engineering* , March.

PELTIER, R. and LÉVY-VÉHEL, J. (1995). Multifractional Brownian motion: definition and preliminary results, *Rapport de recherche INRIA*, 2645.

PIPIRAS, V. (2004). On the usefulness of wavelet-based simulation of fractional Brownian motion, Preprint. Available at http://www.stat.unc.edu/faculty/pipiras/.

PIPIRAS, V. (2005). Wavelet-based simulation of fractional Brownian motion revisited. Applied and Computational Harmonic Analysis, 19(1), 49–60.

PESQUET-POPESCU, B. VÉHEL, J. L. (2002) Stochastic Fractal Models for Image Processing, *IEEE Signal Procesing Magazine*, 48-62.

PESQUET-POPESCU, B. (1999). Statistical Properties of the Wavelet Decomposition of Some Non-Gaussian Self-Similar Processes, *invited paper*, *Signal Processing*, 75, 303–322.

PESQUET-POPESCU, B. (1999). Wavelet Packet Analysis of 2D Processes with Stationary Fractional Increments, *IEEE Transactions on Information Theory*, 45(3), 1033–1039.

REED, I. S., LEE, P. C., and TRUONG, T. K. (1995). Spectral representation of fractional Brownian motion in n dimensions and its properties Information Theory, *IEEE Transactions*, 41(5), 1439–1451

TAQQU M. S., TEVEROVSKY V., and WILLINGER, W. (1997). Is network traffic self-similar or multifractal? *Fractals*, 5–63.

VEITCH, D. and ABRY P. (1999). A wavelet based joint estimator for the parameters of LRD, *Special issue on Multiscale Statistical Signal Analysis and its Applications, IEEE Trans. Info. Th.*, 45, 3, 878–897.

VIDAKOVIC B. and RUGGERI F. (2001). BAMS Method: Theory And Simulations, *Sankhya, Series B*, 63, 234–249.

VIDAKOVIC, B. (1999). *Statistical Modeling by Wavelets,* John Wiley & Sons, New York.

WORNELL, G. (1995). *Signal Processing with Fractals: A Wavelet Based Approach*, Prentice Hall.

WU, D. and XIAO, Y. (2005). Geometric properties of the images of fractional Brownian sheets, *J. Fourier Anal. Appl.* To appear.

ZHU, Z. and STEIN, M.L. (2002) Parameter estimation for fractional Brownian surfaces, *Statistica Sinica* 12, 863-883