Wavelet Variance Analysis for Random Fields

Debashis Mondal and Donald B. Percival

Abstract

There has been considerable recent interest in using wavelets to analyze time series and images that can be regarded as realizations of certain one- and two-dimensional stochastic processes. Wavelets give rise to the concept of the wavelet variance (or wavelet power spectrum), which decomposes the variance of a stochastic process on a scale-by-scale basis. The wavelet variance has been applied to a variety of time series, and a statistical theory for estimators of this variance has been developed. While there have been applications of the wavelet variance in the two-dimensional context (in particular in works by Unser, 1995, on wavelet-based texture analysis for images and by Lark and Webster, 2004, on analysis of soil properties), a formal statistical theory for such analysis has been lacking. In this paper, we develop the statistical theory by generalizing and extending some of the approaches developed for time series, thus leading to a large sample theory for estimators of two-dimensional wavelet variances. We apply our theory to simulated data from Gaussian random fields with exponential covariances and from fractional Brownian surfaces. We also use our methodology to analyze images of four types of clouds observed over the south-east Pacific ocean.

Index Terms

analysis of variance, cloud data, Daubechies filters, fractional Brownian surface, semi-variogram.

I. Introduction

Two-dimensional versions of the wavelet transform have been used extensively in the past to address problems such as image compression and segmentation, edge detection, deconvolution and denoising; see, for example, [12]. The application of statistical methodology to images has largely been devoted to

D. Mondal is with the Department of Statistics and the College, University of Chicago, Chicago, IL, 60637 USA (e-mail: mondal@galton.uchicago.edu).

D. B. Percival is with the Applied Physics Laboratory, University of Washington, Seattle, WA 98195 USA (e-mail: dbp@apl.washington.edu).

Manuscript received November 23, 2009

extending the pioneering work of Donoho, Johnstone and co-workers [7], [8] on wavelet shrinkage for one-dimensional signals to the two-dimensional case; see, for example, [9]. A statistical theory for the wavelet variance (or wavelet spectrum) of time series has been developed in the literature [14], but to date little work has been done on extending this theory to the two-dimensional case, which would facilitate the study of the second-order properties of random fields. Despite this lack of a theoretical basis, the two-dimensional wavelet variance has already been used in the literature. In pioneering work appearing in this journal, Unser [22] used a localized version of the sample wavelet variance and covariance to extract second-order features related to textures for the purposes of classification and segmentation. Lark and Webster [11] used a two-dimensional version of the wavelet variance to analyze soil thickness and related variables. These works do not consider issues related to sampling variability. In this paper, we develop a basic statistical theory for the two-dimensional wavelet variance that – to some degree – mimics the theory already developed for the one-dimensional case. A key contribution of the paper (Theorem 3 in Section VI) gives the large-sample distribution for estimators of two-dimensional wavelet-based variances, thus providing a firm basis for statistical inference for the work of Unser [22], Lark and Webster [11] and related applications.

The remainder of this paper is organized as follows. We review the basic theory for the wavelet transform of a random field in Section II, after which we go over the theory for the two-dimensional version of the wavelet variance in Section III. We note here that the wavelet variance offers a decomposition of the process variance for stationary random fields. In Section IV we present examples of the theoretical wavelet variance for two classes of Gaussian random fields, namely, a stationary field with an exponential covariance struture and an intriniscally stationary field known in the literature as a fractional Brownian surface. We show through examples how the parameters that specify these fields manifest themselves in plots of wavelet-based variances. As we note in Section V, the sample two-dimensional wavelet variance for any image decomposes the sample variance for the image on a scale-by-scale basis, a result that parallels the wavelet-based decompositon of the process variance for stationary random fields. In Section VI we present the key theory that allows us to make large-sample inferences from estimators of the two-dimensional wavelet variance. We demonstrate that this asymptotic theory is applicable even for random fields of modest size (e.g., 32×32) in Section VII. We devote Section VIII to an analysis of some cloud data off the Chilean coast over the south-east Pacific ocean. We conclude the paper with a summary and discussion (Section IX).

II. Wavelet Transform of a Random Field

Let $\{X_{u,v}, u, v = 0, \pm 1, \pm 2 \ldots\}$ be a random field on the two-dimensional integer lattice $\mathcal{Z}^2$. We are interested in its two-dimensional wavelet transform. For $j = 1, 2, \ldots$, let $\{h_{j,l}, l = 0, 1, \ldots L_j - 1\}$ be the jth level Daubechies wavelet filter normalized such that $\sum_l h_{j,l}^2 = 1/2^j$, and let $\{g_{j,l}\}$ be the corresponding scaling filter [6], [15]. We use a tensor product of the one-dimensional wavelet filters to define the *two-dimensional wavelet filters*. This is in agreement with the existing literature in image processing (see, for example, [11], [12], [22]), but we note that an alternative approach is to use a genuine two-dimensional wavelet filter (i.e., one that is not based directly on applying a one-dimensional transform in two directions). A *wavelet transform* consists of filtering the random field by scaling or wavelet filters along the rows and the columns of the two-dimensional lattice process. Accordingly we define

$$W_{j,j',u,v} = \sum_{l=0}^{L_j-1} \sum_{l'=0}^{L_{j'}-1} h_{j,l} h_{j',l'} X_{u-l,v-l'}, \tag{1}$$

$$U_{j,j',u,v} = \sum_{l=0}^{L_j-1} \sum_{l'=0}^{L_{j'}-1} g_{j,l} h_{j',l'} X_{u-l,v-l'}, \tag{2}$$

$$V_{j,j',u,v} = \sum_{l=0}^{L_j-1} \sum_{l'=0}^{L_{j'}-1} h_{j,l} g_{j',l'} X_{u-l,v-l'} \tag{3}$$

and

$$Z_{j,j',u,v} = \sum_{l=0}^{L_j-1} \sum_{l'=0}^{L_{j'}-1} g_{j,l} g_{j',l'} X_{u-l,v-l'}, \tag{4}$$

which we refer to as, respectively, the *wavelet-wavelet*, *scaling-wavelet*, *wavelet-scaling* and *scaling-scaling* coefficient processes. By definition the wavelet-wavelet process $\{W_{j,j',u,v}\}$ is associated with changes at the scale $\tau_j \equiv 2^{j-1}$ along the x direction of the random field and $\tau_{j'}$ along its y direction. On the other hand the scaling-wavelet process $\{U_{j,j',u,v}\}$ is associated with changes in scale $\tau_{j'}$ of the row averages corresponding to scale $2\tau_j$, and $\{V_{j,j',u,v}\}$ corresponds to changes at the scale τ_j of the column averages associated with the scale $2\tau_{j'}$. The final process $\{Z_{j,j',u,v}\}$ corresponds to averages of the random field at the scales $2\tau_j$ and $2\tau_{j'}$ in the x and y directions respectively. As in the one dimensional case, a pyramid algorithm can be used for fast computations of all these coefficients.

III. Basic Theory for Wavelet Variance

The wavelet transform provides an analysis of variance for a stationary random field and a related decomposition for fields that are nonstationary but are intrinsically stationary of order d (see [3], [20] for background on the latter). For both cases we need to define the wavelet variance in two dimensions.

We set aside the scaling-scaling coefficients because they correspond to the scaling coefficients in time series analysis and are usually the least interesting part of a variance decomposition. Note that, if the underlying random field $\{X_{u,v}\}$ is stationary (a sufficient but not necessary condition) with spectral density function (SDF) $S_X(f, f')$, then the coefficients $\{W_{j,j',u,v}\}$, $\{U_{j,j',u,v}\}$ and $\{V_{j,j',u,v}\}$ are also realizations of stationary random fields. First we define a *wavelet-wavelet variance* by

$$\nu^2_{X,h,h}(\tau_j, \tau_{j'}) = \text{var}\,(W_{j,j',u,v}).$$

If $\{X_{u,v}\}$ is intrinsically stationary of order d, then $S_X(f, f')$ has a pole of order d at the origin, and $\nu^2_{X,h,h}(\tau_j, \tau_{j'})$ is well-defined and finite if $L \geq d$. When $\{X_{u,v}\}$ is stationary with autocovariance sequence (ACVS)

$$s_{X,\tau,\tau'} = \text{cov}(X_{u,v}, X_{u+\tau,v+\tau'}),$$

we can rewrite the above as

$$\nu^2_{X,h,h}(\tau_j, \tau_{j'}) = \sum_{l,l'=0}^{L_j-1} \sum_{k,k'=0}^{L_{j'}-1} h_{j,l} h_{j,l'} h_{j',k} h_{j',k'} s_{X,l-l',k-k'}. \tag{5}$$

When $\{X_{u,v}\}$ is intrinsically stationary of order 1, the increment $X_{u,v} - X_{0,0}$ has finite variance, in which case the above equation can be replaced by one involving the ACVS for $X_{u,v} - X_{u-1,v-1}$ and the cumulative sum of $h_{j,l}$ [4]. Alternatively, let

$$\gamma_{X,k,k'} = \tfrac{1}{2}\text{var}\,(X_{0,0} - X_{k,k'})$$

denote the semi-variogram of $\{X_{u,v}\}$. Then the wavelet-wavelet variance can be expressed as

$$\nu^2_{X,h,h}(\tau_j, \tau'_j) = -\sum_{l,l'=0}^{L_j-1} \sum_{k,k'=0}^{L_{j'}-1} h_{j,l} h_{j,l'} h_{j',k} h_{j',k'} \gamma_{X,l-l',k-k'}. \tag{6}$$

The above equation also holds when $\{X_{u,v}\}$ is stationary.

The scale-based decomposition of the process variance can be summarized in the following theorem, which generalizes an analogous result for time series [15].

Theorem 1: Let $\{X_{u,v}\}$ be a stationary random field with SDF $S_X(f, f')$. Then

$$\text{var}\,(X_{u,v}) = \sum_{j=1}^{\infty} \sum_{j'=1}^{\infty} \nu^2_{X,h,h}(\tau_j, \tau_{j'}) \tag{7}$$

(the proof is in Appendix A).

Since Daubechies wavelet filters are approximate band-pass filters, the decomposition given by Theorem 1 corresponds to a partitioning of the frequency domain. Figure 1 depicts this partition. For simplicity we only look at f and f' frequencies ranging from 0 to $1/2$.

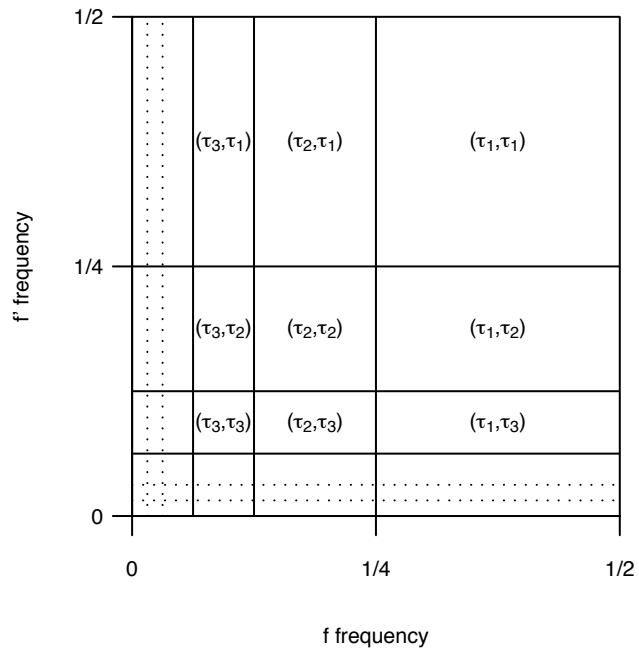

Fig. 1. Frequency partition of a random field.

Next we turn to the scaling-wavelet and wavelet-scaling coefficient processes. We define a *scaling-wavelet variance* as the variance of the scaling-wavelet coefficient process:

$$\nu^2_{X,g,h}(\tau_j, \tau_{j'}) = \text{var}\,(U_{j,j',u,v}).$$

If $\{X_{u,v}\}$ is intrinsically stationary of order d, then $S_X(f, f')$ has a pole of order d at the origin, and $\nu^2_{X,g,h}(\tau_j, \tau_{j'})$ is finite if $L \geq 2d$. Depending on the choice of wavelet filter, the scaling-wavelet variance might not be finite for certain processes, even though the wavelet-wavelet variance is. When $\{X_{u,v}\}$ is stationary, we can rewrite the above as

$$\nu^2_{X,g,h}(\tau_j, \tau_{j'}) = \sum_{l,l'=0}^{L_j-1} \sum_{k,k'=0}^{L_{j'}-1} g_{j,l} g_{j,l'} h_{j',k} h_{j',k'} s_{X,l-l',k-k'}. \tag{8}$$

When $\{X_{u,v}\}$ is intrinsically stationary of order 1, the above equation can be replaced by

$$\nu^2_{X,g,h}(\tau_j, \tau_{j'}) = -\sum_{l,l'=0}^{L_j-1} \sum_{k,k'=0}^{L_{j'}-1} g_{j,l} g_{j,l'} h_{j',k} h_{j',k'} \gamma_{X,l-l',k-k'}. \tag{9}$$

In a similar manner, the *wavelet-scaling variance* is given by

$$\nu^2_{X,h,g}(\tau_j, \tau_{j'}) = \text{var}\,(U_{j,j',u,v})$$

and satisfies

$$\nu^2_{X,h,g}(\tau_j, \tau_{j'}) = \sum_{l,l'=0}^{L_j-1} \sum_{k,k'=0}^{L_{j'}-1} h_{j,l} h_{j,l'} g_{j',k} g_{j',k'} s_{X,l-l',k-k'}$$

when the ACVS exists and

$$\nu^2_{X,h,g}(\tau_j, \tau_{j'}) = -\sum_{l,l'=0}^{L_j-1} \sum_{k,k'=0}^{L_{j'}-1} h_{j,l} h_{j,l'} g_{j',k} g_{j',k'} \gamma_{X,l-l',k-k'}$$

when the semi-variogram exists.

Although the scaling-wavelet and wavelet-scaling variances do not appear in equation (7), they have a close connection to wavelet-wavelet variances and are important factors in an alternative variance decomposition. The properties of the Daubechies wavelet filters imply that both

$$\nu^2_{X,g,h}(\tau_j, \tau_{j'}) = \sum_{k=j+1}^{\infty} \nu^2_{X,h,h}(\tau_k, \tau_{j'}), \quad j \geq j',$$

and

$$\nu^2_{X,h,g}(\tau_j, \tau_{j'}) = \sum_{k=j'+1}^{\infty} \nu^2_{X,h,h}(\tau_j, \tau_k), \quad j \leq j',$$

allowing us to rewrite equation (7) as

$$\text{var}\,(X_{u,v}) = \sum_{j=1}^{\infty} \nu^2_{X,h,h}(\tau_j, \tau_j) + \sum_{j=1}^{\infty} \nu^2_{X,g,h}(\tau_j, \tau_j) + \sum_{j=1}^{\infty} \nu^2_{X,h,g}(\tau_j, \tau_j). \tag{10}$$

Note that the above can be interpreted as a *tri-diagonal representation* and is a simplification of (7). As illustrated in Figure 2, the main diagonal involves the wavelet-wavelet variances, while the first diagonal above (below) this involves the scaling-wavelet (wavelet-scaling) variances. In practice, this tri-diagonal representation enables both faster computations and the use of graphical displays already developed for time series.

The variance decompositions given by (7) and (10) provide scale-based analyses of variance that are useful alternatives to the one given by SDFs for stationary and intrinsic random fields. The decompositions can serve as a diagnostic tool for checking inhomogeneity and are also useful for estimating space-varying SDFs for locally stationary and certain intrinsically stationary random fields (for related work in the frequency domain see [10], [16]). They can be used to investigate directionality effects as well. In particular, if $\{X_{u,v}\}$ is isotropic on $\mathcal{Z}^2$, then

$$\nu^2_{X,h,h}(\tau_j, \tau_{j'}) = \nu^2_{X,h,h}(\tau_{j'}, \tau_j),$$

and

$$\nu^2_{X,g,h}(\tau_j, \tau_{j'}) = \nu^2_{X,h,g}(\tau_{j'}, \tau_j)$$

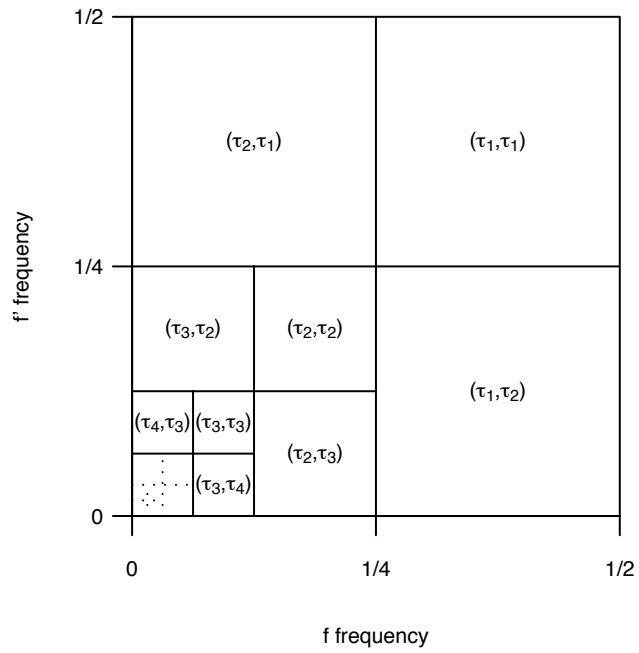

Fig. 2. Tri-diagonal frequency partition of a random field.

for all (j, j'). In practice, given a realization of a random field, we can estimate these quantities and apply large sample results to test if there is significant departure from equality in the above equations. Any departure will provide evidence against the rotational symmetry of the process

IV. Examples of Theoretical Wavelet Variances

Here we look at the theoretical wavelet variance for two classes of Gaussian random fields. The first class is stationary and is based on exponential covariances. The second class – fractional Brownian surfaces – is first order intrinsic. Both classes are simple in that their members depend on just two parameters.

A. Gaussian Random Fields with Exponential Covariance

Let $X_{u,v}$ be a Gaussian stationary random field with exponential covariances

$$s_{X,k,k'} = \phi^{d(k,k')}, \quad d^2(k,k') = k^2 + k'^2,$$

where $0 < \phi < 1$. As a stochastic process on the Euclidean plane, this random field finds wide use in various geostatistical applications (see, e.g., [5]). Its SDF is proportional to $(\phi^2 + |\omega|^2)^{-3/2}$, with

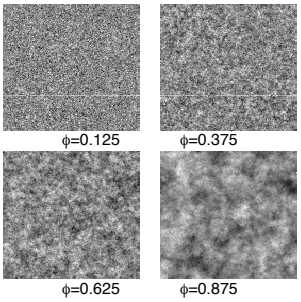

Fig. 3. Plot of simulated Gaussian random fields with exponential covariances for four values of ϕ.

$|\omega|^2 = \omega_1^2 + \omega_2^2$ and $0 < \{\omega_1, \omega_2\} < \infty$. When restricted to the integer lattice, the SDF takes a slightly different form:

$$S_X(f, f') = \left(\phi^2 + \sin^2(\tfrac{f}{2}) + \sin^2(\tfrac{f'}{2})\right)^{-\frac{3}{2}}.$$

Nearby observations from a realization of this random field are similar, but widely separated observations are almost independent because the correlation decreases exponentially. Here the parameter ϕ determines the practical range of dependence between two observations. In particular, at a lag distance $d(k, k') = 3\phi$, the correlation decreases by 95% from its value at the origin. Figure 3 provides four realizations of random fields with four different values of ϕ on 128×128 grids. As ϕ increases, there is substantial growth in dependence, and the realizations become much smoother in appearance.

We apply equations (5) and (8) to compute Haar wavelet-wavelet and scaling-wavelet variances for the exponential covariance model. Figures 4 and 5 plot these values for diagonal scales indexed by $j = 1, \ldots, 7$ with ϕ ranging from 0.125 to 0.875. The effect of ϕ is visible only through the small scale variability, whereas, at the larger scales, the variances fall off exponentially (linear on the log scale) and have a pattern characteristic of a white noise random field.

B. Fractional Brownian Surfaces

Fractional Brownian surfaces (FBSs) are a class of Gaussian random fields that can represent many physical processes; see, for example, [13] and [23]. FBSs produce realizations that are continuous but not differentiable at any point. An FBS is a Gaussian random field with constant mean and semi-variogram structure given by

$$\gamma_{X,u,v} = 2\kappa(u^2 + v^2)^\alpha,$$

where $\alpha \in (0, 1)$ [17]. The parameter κ is a scale parameter, and α is a smoothness (or persistence) parameter that controls the roughness of the surface. An α close to one (zero) corresponds to a smooth

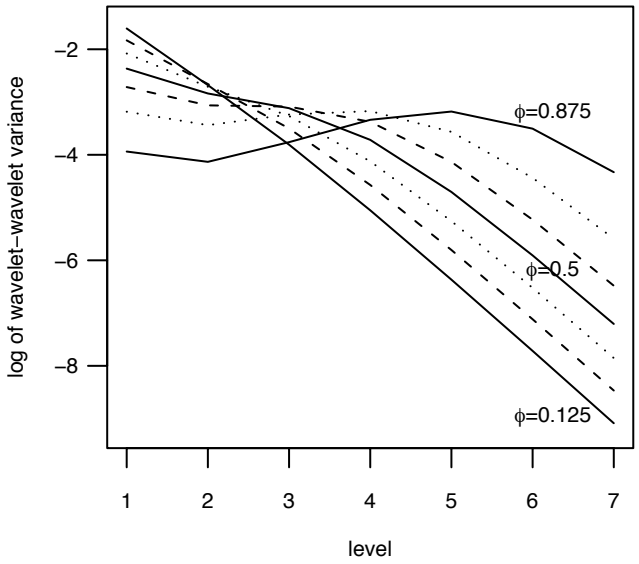

Fig. 4. Plot of log Haar wavelet-wavelet variance versus level j for exponential covariance model with values of ϕ ranging from 0.125 to 0.875 in steps of 0.125.

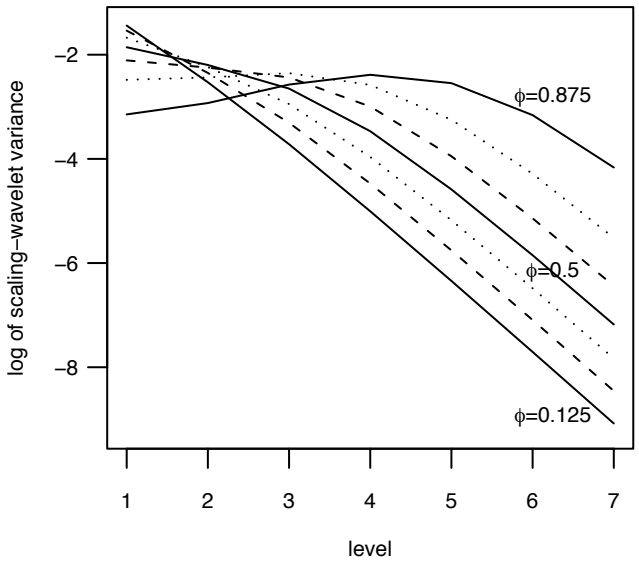

Fig. 5. Plot of log Haar scaling-wavelet variance versus level j for exponential covariance model with values of ϕ ranging from 0.125 to 0.875 in steps of 0.125.

(very rough) surface. Figure 6 shows images of simulated FBSs for four values of α generated using the R package RandomFields [19]. As α increases, the image gets smoother. Moreover, α is linearly related to the fractal dimension D of an FBS, namely, $D = 3 - \alpha$ [5].

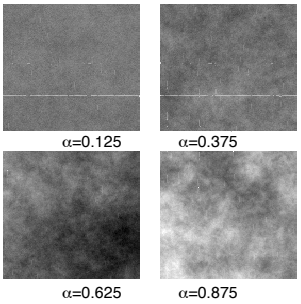

Fig. 6. Plot of simulated fractional Brownian surfaces for four values of α.

There has been considerable interest in investigating the statistical properties of estimators of α using the empirical semi-variogram of the random field observed on a rectangular array; see, for example, [25] and the references therein. The wavelet variance is in a sense a generalization of the semi-variogram and hence provides another way of determining α. Near zero frequencies, the FBS obeys a generalized power spectral density approximately of the form

$$S_X(f_1, f_2) \propto |f_1^2 + f_2^2|^{-\beta},$$

where the spectral decay is characterized by the spectral exponent β, and $\beta = 2 + 2\alpha$. In the one-dimensional case, known as fractional Brownian motion, it is well known that the log of the wavelet variance satisfies a linear relationship with the logarithm of the scale. Thus an estimate of α is readily available by looking at the slope of $\log_2(\nu_X^2(\tau_j))$ versus $\log_2(\tau_j)$. This is also the case for FBSs.

Theorem 2: Let $\{X_{u,v}\}$ be a fractional Brownian surface. Then, for large j and j', we have

$$\log_2\left(\nu^2_{X,h,h}(\tau_j, \tau_{j'})\right) \approx C - j - j' - (1+\alpha)\log_2\left(2^{-2j'} + 2^{-2j}\right)$$

and

$$\log_2\left(\nu^2_{X,g,h}(\tau_j, \tau_{j'})\right) \approx C - j - j' - (1+\alpha)\log_2\left(9 \cdot 2^{-2j'} + 2^{-2j}\right),$$

where C in the above equations stands for a generic constant that does not depend on j or j' (the proof is in Appendix B).

When $j' = j$, we have

$$\log_2\left(\nu^2_{X,h,h}(\tau_j, \tau_j)\right) \approx C + 2\alpha j$$

and

$$\log_2\left(\nu^2_{X,g,h}(\tau_j, \tau_j)\right) \approx C + 2\alpha j$$

also. Using equations (6) and (9), we can compute Haar wavelet-wavelet and scaling-wavelet variances exactly for fixed values of κ and α to assess the efficacy of these approximations. Figure 7 plots the exact wavelet-wavelet variance at levels $j = 1, \ldots, 7$ for values of α ranging from 0.125 to 0.875, while Figure 8 does the same for the scaling-wavelet variance. The approximations are very good except for small levels j with small values of α, and the approximation gets better for all α as the value of j increases. We can thus use the wavelet-wavelet and scaling-wavelet variances to assess the scale exponent α of an FBS.

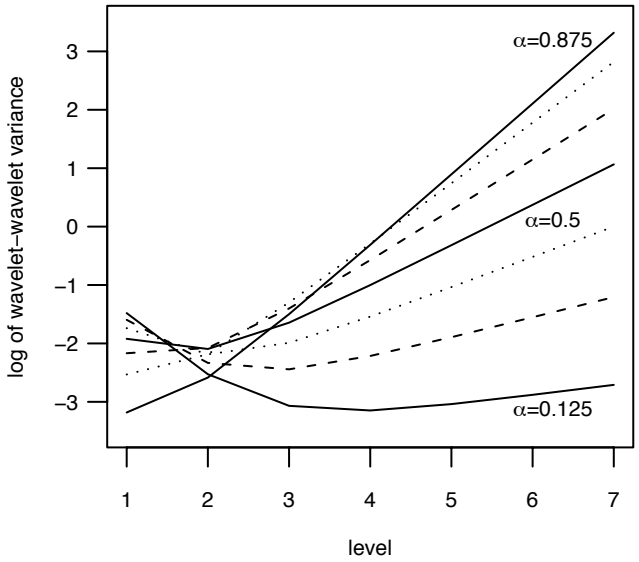

Fig. 7. Plot of log Haar wavelet-wavelet variance versus level j for fractional Brownian surfaces with values of α ranging from 0.125 to 0.875 in steps of 0.125.

V. Decomposition of Sample Variance

Given a realization of a random field $X_{u,v}$ on a finite array $\{(u,v) : u = 0, \ldots, N-1,\ v = 0, \ldots, M-1\}$, let us now consider the decomposition of its sample variance. Define

$$\tilde{W}_{j,j',u,v} = \sum_{l=0}^{L_j-1} \sum_{l'=0}^{L_{j'}-1} h_{j,l} h_{j',l'} X_{u-l \bmod N, v-l' \bmod M},$$

which is similar in form to $W_{j,j',u,v}$ as defined in (1), but involves circular filtering to avoid using elements of the random field outside of the finite array. If $L_j - 1 \leq u \leq N-1$ and $L_{j'} - 1 \leq u \leq M-1$, then $\tilde{W}_{j,j',u,v} = W_{j,j',u,v}$, but equality need not hold otherwise. Define $\tilde{U}_{j,j',u,v}$, $\tilde{V}_{j,j',u,v}$ and $\tilde{Z}_{j,j',u,v}$ analogously by using $X_{u-l \bmod N, v-l' \bmod M}$ as a replacment for $X_{u-l,v-l'}$ in (2), (3) and (4). Let $\tilde{W}_{j,j'}$

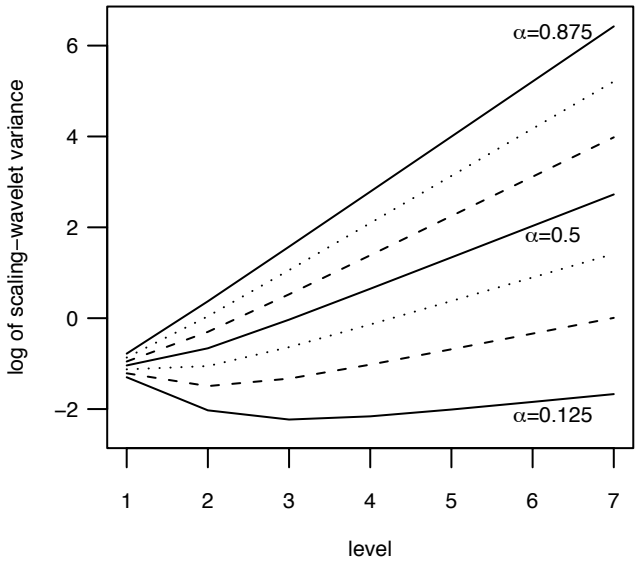

Fig. 8. Plot of log Haar scaling-wavelet variance versus level j for fractional Brownian surfaces with values of α ranging from 0.125 to 0.875 in steps of 0.125.

denote an $N \times M$ array whose (u, v)th element is $\tilde{W}_{j,j',u,v}$, and define $\tilde{U}_{j,j'}$, $\tilde{V}_{j,j'}$ and $\tilde{Z}_{j,j'}$ in a similar way. Consider $\tilde{W}_{1,1}$, $\tilde{U}_{1,1}$, $\tilde{V}_{1,1}$ and $\tilde{Z}_{1,1}$, which are obtained by circularly filtering $X_{u,v}$ with unit level Daubechies wavelet and scaling filters in appropriate orders. Letting X denote the $N \times M$ array with elements $X_{u,v}$, it follows that

$$\|X\|^2 = \|\tilde{W}_{1,1}\|^2 + \|\tilde{U}_{1,1}\|^2 + \|\tilde{V}_{1,1}\|^2 + \|\tilde{Z}_{1,1}\|^2,$$

where $\|.\|^2$ denotes sum of squares of all the elements. Proof of the above is analogous to one given in [15] for the time series case and makes use of the fact that the unit level Daubechies wavelet and scaling filters are orthogonal to each other and to their even shifts. Because the level j filters $h_{j,l}$ and $g_{j,l}$ can be formed from the level $j-1$ filter $g_{j-1,l}$ using the unit level wavelet and scaling filters, we can derive $\tilde{W}_{j,j}$, $\tilde{U}_{j,j}$, $\tilde{V}_{j,j}$ and $\tilde{Z}_{j,j}$ by circularly filtering $\tilde{Z}_{j-1,j-1}$, thus leading to the decomposition

$$\|\tilde{Z}_{j-1,j-1}\|^2 = \|\tilde{W}_{j,j}\|^2 + \|\tilde{U}_{j,j}\|^2 + \|\tilde{V}_{j,j}\|^2 + \|\tilde{Z}_{j,j}\|^2,$$

$j = 2, 3 \ldots$. On the other hand, we can circularly filter the rows of $\tilde{U}_{j,j'}$ to obtain $\tilde{W}_{j+1,j'}$ and $\tilde{U}_{j+1,j'}$, which gives

$$\|\tilde{U}_{j,j'}\|^2 = \|\tilde{W}_{j+1,j'}\|^2 + \|\tilde{U}_{j+1,j'}\|^2,$$

$j \geq j', j, j' = 1, 2 \ldots$. A similar argument leads to

$$\|\tilde{V}_{j,j'}\|^2 = \|\tilde{W}_{j,j'+1}\|^2 + \|\tilde{V}_{j,j'+1}\|^2,$$

for $j \leq j', j, j' = 1, 2 \ldots$. Collecting these decompositions together, we find that $\|X\|^2$ is equal to

$$\sum_{j=1,j'=1}^{J,J'} \|\tilde{W}_{j,j'}\|^2 + \sum_{j=1}^{J} \|\tilde{V}_{j,J'}\|^2 + \sum_{j'=1}^{J'} \|\tilde{U}_{J,j'}\|^2 + \|\tilde{Z}_{J,J'}\|^2,$$

where J, J' are usually chosen such that $2^J \leq N$ and $2^{J'} \leq M$. Thus we can decompose the sample variance $\hat{\sigma}_X^2 = N^{-1}M^{-1}\|X\|^2 - \bar{X}^2$ as

$$\hat{\sigma}_X^2 = \frac{1}{NM}\left(\sum_{j,j'} \|\tilde{W}_{j,j'}\|^2 + \sum_{j} \|\tilde{V}_{j,J}\|^2 + \sum_{j'} \|\tilde{U}_{J,j'}\|^2 + \|\tilde{Z}_{J,J'}\|^2 - \bar{X}^2\right). \tag{11}$$

so that $\|\tilde{W}_{j,j'}\|^2/(NM)$, $\|\tilde{U}_{J,j'}\|^2/(NM)$, $\|\tilde{V}_{j,J}\|^2/(NM)$ are the contributions to the sample variance due to the respective scale pairs $(\tau_j, \tau_{j'})$, $(\tau_J, \tau_{j'})$, $(\tau_j, \tau_{J'})$. In addition, the sample mean of $\tilde{Z}_{J,J'}$ is equal to $\bar{X}$ and hence that the sample variance of $\tilde{Z}_{J,J'}$ is given by $\|\tilde{Z}_{J,J'}\|^2/(NM) - \bar{X}^2$. This decomposition is useful for defining a discrete power spectrum for random fields.

As we have already noted in (10), there is an alternative form of the variance decomposition involving only the diagonal terms. This corresponds to an alternative analysis of variance given by

$$\hat{\sigma}_X^2 = \frac{1}{NM}\left(\sum_{j=1}^{J_0}\left(\|\tilde{W}_{j,j}\|^2 + \|\tilde{U}_{j,j}\|^2 + \|\tilde{V}_{j,j}\|^2\right) + \|\tilde{Z}_{J_0,J_0}\|^2 - \bar{X}^2\right). \tag{12}$$

VI. Estimation and Large Sample Inference for Wavelet Variance

In this section, we consider statistical estimation and inference for wavelet variances of a random field. We therefore assume that we have a realization of a stationary or an intrinsically stationary random field $\{X_{u,v}\}$ on a finite array $\{(u, v) : u = 0, \ldots, N-1,\ v = 0, \ldots, M-1\}$. We can then estimate the wavelet-wavelet variance by the unbiased estimator

$$\hat{\nu}^2_{X,h,h}(\tau_j, \tau_{j'}) = \frac{1}{N_j M_{j'}} \sum_{u=L_j-1}^{N-1} \sum_{v=L_{j'}-1}^{M-1} W^2_{j,j',u,v}, \tag{13}$$

where $N_j = N - L_j + 1$ and $M_j = M - L_j + 1$. Similarly we estimate the scaling-wavelet and wavelet-scaling variance as follows:

$$\hat{\nu}^2_{X,g,h}(\tau_j, \tau_{j'}) = \frac{1}{N_j M_{j'}} \sum_{u=L_j-1}^{N-1} \sum_{v=L_{j'}-1}^{M-1} U^2_{j,j',u,v} \tag{14}$$

and

$$\hat{\nu}^2_{X,h,g}(\tau_j, \tau_{j'}) = \frac{1}{N_j M_{j'}} \sum_{u=L_j-1}^{N-1} \sum_{v=L_{j'}-1}^{M-1} V^2_{j,j',u,v}. \tag{15}$$

The following theorem gives the large sample properties of estimators (13), (14) and (15) when the process $\{X_{u,v}\}$ is Gaussian.

Theorem 3: Let $\{C_{j,j',u,v}\}$ stand for either the wavelet-wavelet $\{W_{j,j',u,v}\}$, scaling-wavelet $\{U_{j,j',u,v}\}$ or wavelet-scaling $\{V_{j,j',u,v}\}$ coefficients. Suppose that $\{C_{j,j',u,v}\}$ is a mean zero Gaussian stationary random field satisfying the squared integrability condition

$$\sigma^2_{C,j,j'} = 2\sum_{t,t'} \mathrm{cov}^2(C_{j,j',0,0}, C_{j,j',t,t'}) < \infty.$$

Then, as $\min(N_j, M_{j'}) \to \infty$, the corresponding coefficient variance $\hat{\nu}^2_{X,\cdot,\cdot}(\tau_j, \tau_{j'})$ is asymptotically normal with mean $\nu^2_{X,\cdot,\cdot}(\tau_j, \tau_{j'})$ and large sample variance $2\sigma^2_{C,j,j'}/(N_j M_{j'})$.

Proof of Theorem 3. Take $Y_{u,v} = C_{j,j',u,v}$ and $P(Y_{u,v}) = Y^2_{u,v} - E(Y^2_{u,v})$ so that P is the second order Hermite polynomial. Note that conditions of Theorem 2 of [2] are trivially satisfied, and hence the central limit theorem follows. ■

Theorem 3 enables us to construct approximate confidence intervals for the purpose of statistical inference. As an example, let us consider the wavelet-wavelet variance $\nu^2_{X,h,h}(\tau_j, \tau_{j'})$. By Theorem 3, the true $\nu^2_{X,h,h}(\tau_j, \tau_{j'})$ lies within

$$\hat{\nu}^2_{X,h,h}(\tau_j, \tau_{j'}) \pm \Phi^{-1}(1-p)\left(\frac{\sigma^2_{W,j,j'}}{N_j M_{j'}}\right)^{\frac{1}{2}} \tag{16}$$

asymptotically with probability $(1-2p)$, where Φ is the cumulative distribution function of a standard Gaussian random variable. However $\sigma^2_{W,h,h}$ is generally unknown, and we need to estimate it from the observed data. Since

$$\sigma^2_{W,j,j'} = 2\sum_{t}\sum_{t'} \mathrm{cov}^2(W_{j,j',0,0}, W_{j,j',t,t'}),$$

we can estimate $\mathrm{cov}(W_{j,j',t,t'}, W_{j,j',0,0})$ by its sample covariance

$$\hat{s}_{W,j,j'}(t,t') = \frac{1}{N_j M_{j'}} \sum_{u=0}^{N-|t|-1} \sum_{v=0}^{M-|t'|-1} W_{j,j',u,v} W_{j,j',t+u,t'+v}.$$

Using the argument given in p. 312 of [15], we can estimate $\sigma^2_{W,j,j'}$ by

$$\hat{\sigma}^2_{W,j,j'} = \tfrac{1}{2} \sum_{t=-(N_j-1)}^{N_j-1} \sum_{t'=-(M_j-1)}^{M_{j'}-1} \hat{s}^2_{W,j,j'}(t,t'). \tag{17}$$

Moreover the argument in p. 537 of [15] shows that $\hat{\sigma}^2_{W,j,j'}$ is asymptotically an unbiased estimator of $\sigma^2_{W,j,j'}$. Finally we obtain an approximate confidence interval for $\nu^2_{X,h,h}(\tau_j, \tau_{j'})$ by replacing $\sigma^2_{W,j,j'}$ in (16) with $\hat{\sigma}^2_{W,j,j'}$. The construction of large sample confidence intervals for wavelet-scaling and scaling-wavelet variances is similar to that for the wavelet-wavelet variance.

An undesirable consequence of the above approach is that the lower confidence limit constructed from (16) can be negative. Since the estimator $\hat{\nu}^2_{X,h,h}(\tau_j, \tau_{j'})$ is quadratic in Gaussian-distributed wavelet-wavelet coefficients, its exact distribution is amenable to approximation by a scaled χ^2 distribution. A moment matching scheme yields the approximation

$$\hat{\nu}^2_{X,h,h}(\tau_j, \tau_{j'}) \sim \nu^2_{X,h,h}(\tau_j, \tau_{j'}) \eta^{-1}_{j,j'} \chi^2_{\eta_{j,j'}}, \tag{18}$$

where the effective degrees of freedom $\eta_{j,j'}$ can be estimated by

$$\hat{\eta}_{j,j'} = \frac{N_j M_{j'} \hat{\nu}^4_{X,h,h}(\tau_j, \tau_{j'})}{\hat{\sigma}^2_{W,j,j'}}.$$

Confidence intervals based upon this χ^2 approximation are guaranteed to have a nonnegative lower limit and are asymptotically equivalent to ones based upon the Gaussian distribution.

VII. Monte Carlo Experiments

Here we report on a simulation study to assess the approximations described in section VI. We consider two examples. In the first, we generate 1000 realizations of a Gaussian random field from an exponential covariance model with $\phi = 0.875$, each on a 32×32 grid. For each realization (indexed by r), we use the Haar wavelet transform to obtain estimates of the wavelet-wavelet variance $\hat{\nu}^2_{X,r,h,h}(\tau_j, \tau_j)$ for four diagonal scales indexed by $j = 1, 2, 3, 4$. We compute their Monte Carlo averages and standard deviations. In addition, for each realization, we obtain estimates of the standard deviation (SD) of $\hat{\nu}^2_{X,r,h,h}(\tau_j, \tau_j)$ using both $\hat{\sigma}_{W,r,j,j'}$ from (17) and the χ^2 approximation from (18). The latter is in fact equal to $\tilde{\sigma}_{W,r,j,j} = \sqrt{(2\hat{\nu}^4_{X,r,h,h}(\tau_j, \tau_j)/\eta_{r,j,j}}$, where $\eta_{r,j,j}$ denotes the effective degrees of freedom.

Table I summarizes the study for the first example. The Monte Carlo averages of the wavelet-wavelet variance estimates (row 2) match very well with the true variances (row 1). Similarly, the Monte Carlo averages of SDs obtained from (17) and (18) (rows 4 and 5) match well with the Monte Carlo SDs of the wavelet-wavelet variance estimates (row 3) for scales indexed by $j = 1, 2$, but not for $j = 3, 4$. There is significant drop in the effective degrees of freedom for these two scales (row 6), which is undoubtedly what causes the observed downward bias (an image larger than 32×32 would be needed to correct this bias). Consequently, the confidence intervals described in section VI would approximately achieve the desired coverage probabilities for $j = 1, 2$, but would results in an undercoverage for $j = 3, 4$. The χ^2 approximation (row 5) performs better than the normal approximation (row 4) at the highest scale ($j = 4$). (A simulation study using scaling-wavelet coefficients yielded similar results.)

For the second example, we replicate the first, but now using 1000 realizations of FBSs on 32×32 grids with $\alpha = 0.0875$. Table II summarizes this second study, which leads to conclusions similar to

TABLE I

SUMMARY OF MONTE CARLO STUDY FROM EXPONENTIAL COVARIANCE MODEL

j	1	2	3	4
$\nu^2_{X,h,h}(\tau_j,\tau_j)$	0.0195	0.0160	0.0233	0.0355
mean of $\hat{\nu}^2_{X,r,h,h}(\tau_j,\tau_j)$	0.0195	0.0160	0.0232	0.0350
SD of $\hat{\nu}^2_{X,r,h,h}(\tau_j,\tau_j)$	0.0012	0.0015	0.0050	0.0176
mean of $\hat{\sigma}_{W,r,j,j}$	0.0012	0.0016	0.0038	0.0075
mean of $\tilde{\sigma}_{W,r,j,j}$	0.0012	0.0014	0.0037	0.0106
mean of $\eta_{r,j,j}$	491.8	278.7	83.4	24.1

TABLE II

SUMMARY OF MONTE CARLO STUDY USING FRACTIONAL BROWNIAN SURFACES

j	1	2	3	4
$\nu^2_{X,h,h}(\tau_j,\tau_j)$	0.0415	0.0754	0.2234	0.7317
mean of $\hat{\nu}^2_{X,r,h,h}(\tau_j,\tau_j)$	0.0417	0.0759	0.2239	0.7324
SD of $\hat{\nu}^2_{X,r,h,h}(\tau_j,\tau_j)$	0.0023	0.0080	0.0533	0.4225
mean of $\hat{\sigma}_{W,r,j,j}$	0.0032	0.0085	0.0400	0.1735
mean of $\tilde{\sigma}_{W,r,j,j}$	0.0025	0.0071	0.0388	0.2470
mean of $\eta_{r,j,j}$	552.3	233.8	70.4	20.5

those stated for the first example. Our methodology is thus capable of handling fields with both short- and long-range dependence.

VIII. APPLICATION TO CLOUD DATA

In this application, we focus on south-east Pacific stratocumulus clouds near the Chilean coast. Known as one of world's driest oceanic regions, the stratocumulus clouds here reflect sunlight back into space and thereby check the temperature in the Pacific ocean off of Chile. Consequently, these clouds account for much of the variations associated with the seasonal cycles of the east Pacific ocean. Past studies undertaken by the East Pacific Investigation of Climate (EPIC) have identified the cloud dynamics in this region as complex, and driven partly by atmospheric aerosol produced from industrial activities. The EPIC studies also found pockets of seemingly cloud-free air existing in an otherwise homogenous strato-

cumulus cloud field. These pockets of open cells (POCs) are primarily characterized by low-aerosol air mass and are responsible, in the form of light drizzles, for most of the precipitation that occurs in this region [1], [21], [24].

Figure 9 shows a brightness temperature image from Geo-stationary Operational Environmental Satellites (GOES) Imagery collected on 17 October 2001. The image is of the highest clouds of the entire study region and has been preprocessed using a median filter to weed out outliers and other data aberrations. The red square boxes indicate four cloud regions identified by atmospheric scientists, namely, (1) a POC around longitude -85.2° and latitude -12.5°, (2) uniform stratocumulus clouds at about longitude -78° and latitude -18°, (3) broken clouds near longitude -94° and latitude -16°, and finally (4) a cloud region where a POC is forming. We use these four regions to demonstrate wavelet-based analysis of variance, with a focus on extracting statistical features that might usefully characterize the different clouds in these regions.

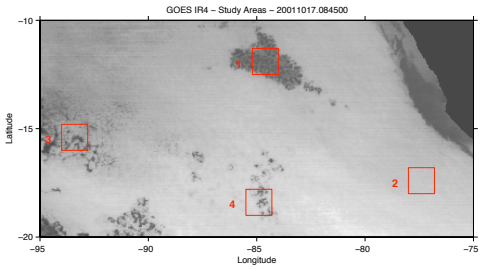

Fig. 9. Image plots of clouds at four different regions

Figure 10 shows Haar wavelet-wavelet variance estimates (13) up to level $j = j' = 4$ in a grey-scale scheme that indicates lower values by darker shades. The wavelet-wavelet variance decreases slowly in region 1 (the POC region) as the level (j, j') increases at each direction. The evident symmetry in the image about (j, j') suggests that the clouds in this region are homogeneous. Moreover, the ratio of the effective degrees of freedom $\hat{\eta}_{j,j'}$ to the actual number of coefficients $N_j M_{j'}$ varies between 0.07 to 0.23. This variability indicates that the strength of dependence between wavelet-wavelet coefficients varies considerably across scales. The ratio attains its maximum at level $(j = 2, j' = 2)$ and minimum at level $(j = 3, j' = 4)$, whereas its value at level $(j = 4, j' = 3)$ is about 0.14. The discrepancy between the values at levels $(j = 3, j' = 4)$ and $(j = 4, j' = 3)$ indicates that, despite the evident symmetry, there might be some differences in behavior at the paired scales. Turning now to region 2 (uniform clouds), the wavelet-wavelet variances are a magnitude smaller than those for the POC region. Here the

variability does not decrease much as the scale increases. The ratio of effective degrees of freedom to the actual number of coefficients varies between 0.05 to 0.19. Interestingly, the strength of dependence between wavelet-wavelet coefficients at the higher scales has the same pattern for regions 1 and 2. In particular, the maximum is attained again at level $(j = 2, j' = 2)$, and the minimum, at $(j = 4, j' = 2)$. For region 3 (broken clouds) the wavelet-wavelet variances are larger than for the POC region. There is now a slight discrepancy between the variances at levels $(j = 4, j' = 2)$ and $(j = 4, j' = 2)$. In addition, the ratio between the effective degrees of freedom and the actual number of wavelet-wavelet coefficients traces an asymmetric pattern. It attains its maximum at level $(j = 2, j' = 1)$ and takes smaller values at levels $(j = 4, j' = 1, 2, 3)$ compared to levels $(j = 1, 2, 3, j' = 4$. For region 4 (forming POC), the wavelet-wavelet variances mostly lie between the estimates for the POC and the broken cloud regions (1 and 3). There is a hint of inhomogeneity as the values at levels $(j = 2, j' = 1)$ and $(j = 1, j' = 2)$ and at levels $(j = 4, j' = 2)$ and $(j = 2, j' = 4)$ differ considerably. The ratio between the effective degrees of freedom and the actual number of wavelet-wavelet coefficients in this region varies between 0.064 to 0.19.

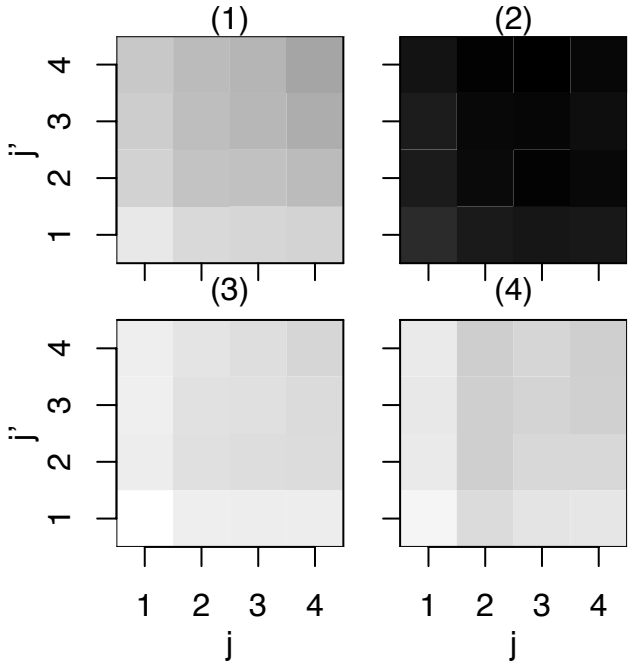

Fig. 10. Wavelet-wavelet variance of cloud image for four regions.

Next we consider an analysis of variance based on the diagonal decomposition (10), which involves use of scaling-wavelet and wavelet-scaling variances in addition to the wavelet-wavelet variances dictated

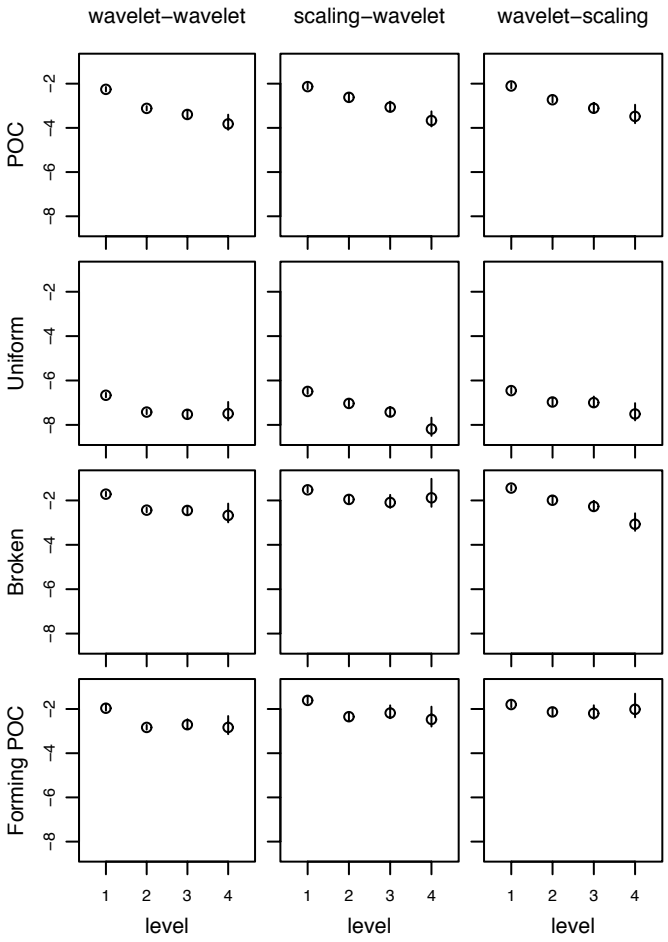

Fig. 11. Wavelet variances of cloud image for, from top to bottom rows, POC region, uniform clouds, broken cloud region and forming POC region.

by decomposition (7). Estimates of the three types of variances – along with 95% confidence intervals – are shown in Figure 11 on a log/log scale, one row for each region. For the POC region (top row), the three variances follow a very similar pattern, although the values of the wavelet-wavelet variances are somewhat smaller than the other two variances. There is no indication of directionality in the POC region since the scaling-wavelet and the wavelet-scaling variances are so similar. For the uniform stratus clouds (second row), all three variances are now much smaller, with somwhat similar values and patterns for all scales. For the broken cloud (third row) and forming plot regions (fourth), the three types of variances do not have the same degree of similarity in their patterns as they do for the POC region. The values of the wavelet-wavelet variances for the broken clouds are somewhat bigger than those for the forming POC region. In addition, there is some evidence of anisotropy in the broken clouds because

the scaling-wavelet and wavelet-scaling patterns are somewhat different. In particular, at scale τ_4, the wavelet-scaling variance is very small and has a much smaller confidence interval in comparison to the scaling-wavelet variance. The wavelet-scaling variances clearly decrease with increasing scale, which is not the case for the scaling-wavelet variances.

Let us now concentrate on just the wavelet-wavelet variances (first column of Figure 11). For the POC region, the log of these variances decreases approximately linearly as the log of the scale increases, which is different from the patterns for the other three cloud regions. This enables us to use the wavelet-wavelet variances as a potential criterion for distinguishing the POC region from the other three cloud types. The wavelet-wavelet variances for the uniform stratus clouds are very small in comparison to the other three cloud types, which again gives us a way of distingushing this cloud type from the others. However, the wavelet-wavelet variances for the broken clouds and the forming POC regions are not easy to distinguish because they assume similar values and follow a similar pattern across all four scales.

Finally let us compare the patterns of the wavelet-wavelet variances observed in the first column of Figure 11 with those for exponential covariance models in Figure 4 and FBSs in Figure 7 (recall that both models have a parameter that controls the level of the wavelet-wavelet variances, so only a comparison of the overall patterns is meaningful). The $\phi = 0.75$ and 0.875 exponential covariance models exhibit an upturn between scales τ_3 and τ_4, which does not agree with any of the patterns in the cloud data. The $\phi = 0.125$ and 0.25 models show overall drops from scale τ_1 to τ_4 (> 2 in log space) that are inconsistent with the observed data (< 2); here the fact that the confidence intervals for the wavelet-wavelet variances in Figure 11 are quite small helps convince us that the inconsistency is not just due to sampling variability. Models with $0.375 \leq \phi \leq 0.625$, on the other hand, might be viable for some of the cloud types. Turning now to FBSs, none of the observed patterns exhibit an upturn at the large scales, which would rule out FBSs with $\alpha \geq 0.25$ as viable models; however, the $\alpha = 0.125$ model might be tenable for all but the POC region since it matches the overall declines rather well (about 1.5 in log space). A comparison of the patterns of the scaling-wavelet variances in the second column of Figure 11 with those for exponential covariance models in Figure 5 and FBSs in Figure 8 leads to much the same evaluation of the tenability of the various models.

IX. Summary and Discussion

This paper presents a statistical theory for wavelet-based analysis of variance for random fields on the two-dimensional plane. Unlike in one dimension, the variance decomposition for a random field can take more than one form; in particular one form (7) uses wavelet-wavelet coefficients exclusively,

while another (10) uses wavelet-wavelet, scaling-wavelet and wavelet-scaling coefficients. These variance decompositions are useful for exploratory analysis of the behavior of random fields, as demonstrated through exact computations (Figures 4, 5, 7 and 8); in particular, they are useful for identifying which scales are the dominant contributors to the overall variability. We demonstrated the efficacy of our sampling theory for the various wavelet-based variances via simulations of fields that obey an exponential covariance model and a fractional Brownian surface (Tables I and II). We also applied our methodology to the analysis of cloud data from the south-east Pacific ocean and demonstrated that there is potential for discriminating amongst some cloud types using wavelet-based variances. The cloud data also demonstrates the ability of our theory to provide confidence intervals for the unknown true wavelet-based variances, which is vital for assessing the effect of sampling variability on the variance estimates. We also noted that we can assess the degree of correlation in an image through a comparison of the effective degrees of freedom (needed to construct confidence intervals) with the actual number of wavelet-based coefficients used to estimate the variance.

While we noted in the previous section that some of the patterns of wavelet-based variances for exponential covariance models and FBSs are consistent with those observed in the cloud data, we have not considered how such models can be fit to images with the help of wavelet-based variances. There is thus further scope to investigate how the range parameter ϕ of exponential covariance models or the shape parameter α of FBSs can be estimated based upon some combination of wavelet-wavelet, scaling-wavelet and wavelet-scaling variances. Use of these variances might lead to interesting alternatives to previously proposed methodology for estimating α [25].

Other extensions to our work are possible. One direction would be to look at multiple images on a more formal basis, i.e., above and beyond how we treated the four subimages of cloud data. Development of a multivariate theory would provide a statistical underpining for the pioneering work of Unser [22] on classification of image textures based on wavelet variances. In a multivariate setting involving more than one dependent random field, different decompositions of a wavelet-based variance-covariance matrix could in principle lead to scale-based classification rules, clustering or principal component analysis. Finally, although not considered here, some theory for non-Gaussian random fields can be derived along the lines of Serroukh et al. [18].

APPENDIX A
PROOF OF THEOREM 1

Let $\mathcal{H}_j$ and $\mathcal{G}_j$ be the squared gain functions for, respectively, Daubechies wavelet and scaling filters $\{h_{j,l}\}$ and $\{g_{j,l}\}$. It then follows that, for any $J \geq 1$, we have the identity

$$\sum_{j=1}^{J} \mathcal{H}_j(f) + \mathcal{G}_J(f) = 1. \tag{19}$$

The above allows us to rewrite $S_X(f, f')$ as

$$S_X(f, f') = \sum_{j=1}^{J} \sum_{j'=1}^{J} \mathcal{H}_j(f)\mathcal{H}_{j'}(f')S_X(f, f') + C_J(f, f'), \tag{20}$$

where

$$C_J(f, f') = \Big(\mathcal{G}_J(f) + \mathcal{G}_J(f') - \mathcal{G}_J(f)\mathcal{G}_J(f')\Big)S_X(f, f').$$

Since $\mathcal{H}_j(f)\mathcal{H}_j(f')S_x(f, f')$ is the SDF of the wavelet-wavelet coefficient process $\{W_{j,j',u.v}\}$, integration of both sides of (20) yields

$$\begin{aligned} \text{var}\,(X_{u,v}) &= \int_{-1/2}^{1/2}\int_{-1/2}^{1/2} S_x(f, f')\, df\, df' \\ &= \sum_{j=1}^{J}\sum_{j=1}^{J} \text{var}\,(W_{j,j',u,v}) + c_J, \end{aligned}$$

where

$$c_J = \int_{-1/2}^{1/2}\int_{-1/2}^{1/2} C_J(f, f')\, df\, df'.$$

It then suffices to show that $c_J \to 0$ as $J \to \infty$. When $S_X(f, f')$ is bounded by, say, $c_S = \max S_X(f, f') < \infty$, we use the fact that $\mathcal{G}_J$ integrates to 2^{-J} to obtain the following bounds:

$$\int_{-1/2}^{1/2}\int_{-1/2}^{1/2} \mathcal{G}_J(f)S_X(f, f')\, df\, df' \leq c_S 2^{-J} \tag{21}$$

and

$$\int_{-1/2}^{1/2}\int_{-1/2}^{1/2} \mathcal{G}_J(f)\mathcal{G}_J(f')S_X(f, f')\, df\, df' \leq c_S 2^{-2J}. \tag{22}$$

Clearly, $0 \leq c_J \leq c_S 2^{-J}(2 + 2^{-J})$, and hence $c_J \to 0$ as $J \to \infty$. On the other hand, if $S_X(f, f')$ is unbounded, we use the fact that $\text{var}\,(X_t) < \infty$ to find, for any given $\epsilon > 0$, a constant c_ϵ such that

$$\int_{A_\epsilon} S_X(f, f')\, df\, df' < \epsilon,$$

where $A_\epsilon = \{(f, f') : S_X(f, f') \geq c_\epsilon\}$. We evaluate c_J by dividing the entire frequency range into A_ϵ and its complement $B_\epsilon = \{(f, f') : S_X(f, f') < c_\epsilon\}$. Since $G_J(f)$ is bounded by 1, it follows that integrals of $\mathcal{G}_J(f)S_X(f, f')$ and $\mathcal{G}_J(f)\mathcal{G}_J(f')S_X(f, f')$ over A_ϵ are each less than ϵ. Since $S_X(f, f')$ is

bounded by c_ϵ in the range of frequencies in B_ϵ, we obtain slight modifications of equations (21) and (22), namely,

$$\int_{B_\epsilon} \mathcal{G}_J(f) S_X(f, f')\, df\, df' < c_\epsilon 2^{-J}$$

and

$$\int_{B_\epsilon} \mathcal{G}_J(f)\mathcal{G}_J(f) S_X(f, f')\, df\, df' < c_\epsilon 2^{-2J}.$$

Thus letting J tend to infinity, we obtain c_J to be less than 3ϵ. Since ϵ is an arbitrarily small positive number, this completes the proof. ■

APPENDIX B
PROOF OF THEOREM 2

The generalized SDF for a continuous parameter two-dimensional FBS is given by

$$S_c(f_1, f_2) = \frac{1}{(f_1^2 + f_2^2)^{\beta/2}}, \quad -\infty < \{f_1, f_2\} < \infty,$$

where $2 < \beta < 4$ [17]. The corresponding SDF for a discrete parameter FBS is

$$S(f_1, f_2) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \frac{1}{([f_1 + k]^2 + [f_2 + l]^2)^{\beta/2}},$$

$-1/2 \le f_1, f_2 \le 1/2$. Now

$$\begin{aligned} \frac{1}{(f_1^2 + f_2^2)^{\beta/2}} &< S(f_1, f_2) \\ &< \frac{1}{(f_1^2 + f_2^2)^{\beta/2}} \\ &\quad +4 \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \frac{1}{([k - 1/2]^2 + [l - 1/2]^2)^{\beta/2}} \\ &\quad +4 \sum_{k=1}^{\infty} \frac{1}{[k - 1/2]^\beta}. \end{aligned}$$

The double and single sums converge to finite values, and hence $S_c(f_1, f_2) \approx 1/(f_1^2 + f_2^2)^{\beta/2}$ at low frequencies. The Daubechies wavelet filter $h_{j,l}$ is approximately a bandpass filter with a passband given by $|f| \in [1/2^{j+1}, 1/2^j]$, so we have

$$\nu^2_{X,h,h}(\tau_j, \tau_{j'}) \approx 4 \int_{1/2^{j+1}}^{1/2^j} \int_{1/2^{j'+1}}^{1/2^{j'}} S(f_1, f_2)\, df_1\, df_2.$$

Hence, at large scales,

$$\nu^2_{X,h,h}(\tau_j, \tau_{j'}) \approx 4 \int_{1/2^{j+1}}^{1/2^j} \int_{1/2^{j'+1}}^{1/2^{j'}} \frac{1}{(f_1^2 + f_2^2)^{\beta/2}}\, df_1\, df_2.$$

The mean value theorem suggests

$$\int_a^b g(x)\,dx \approx (b-a)g\left(\frac{a+b}{2}\right),$$

with the approximation improving as $b-a$ shrinks to zero. Elementary manipulations yield

$$\nu^2_{X,h,h}(\tau_j,\tau_{j'}) \;\approx\; 2^{-j-j'}\left(\frac{4}{3}\right)^{\beta}\left(2^{-2j'}+2^{-2j}\right)^{-\beta/2},$$

from which the stated result follows readily.

For the scaling-wavelet variance,

$$\nu^2_{X,g,h}(\tau_j,\tau_{j'}) \approx 4\int_0^{1/2^{j+1}}\int_{1/2^{j'+1}}^{1/2^{j'}} \frac{1}{(f_1^2+f_2^2)^{\beta/2}}\,df_1\,df_2$$

since the Daubechies scaling filter $g_{j,l}$ is approximately a lowpass filter with a passband given by $[-1/2^{j+1}, 1/2^{j+1}]$. Elementary manipulations yield

$$\nu^2_{X,g,h}(\tau_j,\tau_{j'}) \;\approx\; 2^{-j-j'}4^{\beta}\left(9\cdot 2^{-2j'}+2^{-2j}\right)^{-\beta/2},$$

from which the desired result follows easily. ■

ACKNOWLEDGMENT

The authors would like to thank Peter Guttorp, Chris Bretherton and Charles Cornish for discussions and for help with the application to cloud data.

REFERENCES

[1] C. S. Bretherton, T. Uttal, C. W. Fairall, S. E. Yuter, R. A. Weller, D. Baumgardner, K. Comstock, R. Wood and G. B. Raga, "The Epic 2001 Stratocumulus Study," *Bulletin of the American Meteorological Society*, vol. 85, no. 7, 967–977, July 2004.

[2] P. Breuer and P. Major, "Central limit theorems for nonlinear functionals of Gaussian fields," *J. Multivariate Anal.*, vol. 13, no. 3, 425–441, Sept. 1983.

[3] J.-P. Chilès and P. Delfiner, *Geostatistics: Modeling Spatial Uncertainty*, John Wiley & Sons, New York, 1999.

[4] P. F. Craigmile and D. B. Percival, "Asymptotic decorrelation of between-scale wavelet coefficients," *IEEE Trans. Inform. Theory*, vol. 51, no. 8, 1039–1048, Mar. 2005.

[5] N. Cressie, *Statistics for Spatial Data*, John Wiley, New York, 1993.

[6] I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 61, SIAM, Philadelphia, PA, 1992.

[7] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, 425–455, Aug. 1994.

[8] D. L. Donoho, I. M. Johnstone, G. Kerkyacharian and D. Picard, "Wavelet shrinkage: asymptopia?," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 57, no. 2, 301–369, 1995.

[9] L. Evers and T. J. Heaton, "Locally Adaptive Tree-Based Thresholding," *Journal of Computational and Graphical Statistics*, to be published.

[10] M. Fuentes, "A formal test for nonstationarity of spatial stochastic processes," *J. Multivariate Anal.*, vol. 96, no. 1, 30–54, Sept. 2005.

[11] R. M. Lark and R. Webster, "Analysing soil variation in two dimensions with the discrete wavelet transform," *European Journal of Soil Science*, vol. 55, no. 4, 777–797, Dec. 2004.

[12] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way* (Third Edition), Academic Press Inc., San Diego, CA, 2008.

[13] H. O. Peitgen and D. Saupe, *The Science of Fractal Images*, Springer-Verlag, New York, 1988.

[14] D. B. Percival, "On estimation of the wavelet variance," *Biometrika*, vol. 82, no. 3, 619–631, Sept. 1995.

[15] D. B. Percival and A. T. Walden, *Wavelet Methods for Time Series Analysis*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 4, Cambridge, 2000.

[16] A. Pintore and C. Holmes, "Spatially adaptive non-stationary covariance functions via spatially adaptive spectra," http:\\www.stats.ox.ac.uk~ cholmes\Reports\spectral_tempering.pdf, 2004.

[17] I. S. Reed, P. C. Lee and T. K. Truong, "Spectral representation of fractional Brownian motion in n dimensions and its properties," *IEEE Trans. Info. Theory*, vol. 41, no. 5, 1439–1451, Sept. 1995.

[18] A. Serroukh, A. T. Walden and D. B. Percival, "Statistical properties and uses of the wavelet variance estimator for the scale analysis of time series," *J. Amer. Statist. Assoc.*, vol. 95, no. 449, 184–196, Mar. 2000.

[19] M. Schlather, *RandomFields: Simulation and Analysis of Random Fields*, R package, version 1.2.29, 2006 (code and documentation available at Cran.r-project.org).

[20] M. L. Stein, *Interpolation of Spatial Data*, Springer-Verlag, New York, 1999.

[21] B. Stevens, G. Vali, K. Comstock, R. Wood, M. C. Van Zanten, P. H. Austin, C. S. Bretherton and D. H. Lenschow, "Pockets of Open Cells and Drizzle in Marine Stratocumulus," *Bulletin of the American Meteorological Society*, vol. 86, no. 1, 51–57, Jan. 2005.

[22] M. Unser, "Texure classification and segmentation using wavelet frames," *IEEE Trans. Image Process.*, vol. 4, no. 11, 1549–1560, Nov. 1995.

[23] T. H. Wilson, "Some distinctions between self-similar and self-affine estimates of fractal dimension with case history," *Math. Geol.*, vol. 32, no. 3, 319–335, Apr. 2000.

[24] R. Wood, K. K. Comstock, C. S. Bretherton, C. Cornish, J. Tomlinson, D. R. Collins and C. Fairall, "Open cellular structure in marine stratocumulus cloud sheets," *J. Geophys. Res.*, vol. 113, D12207, doi:10.1029/2007JD009371, 2008.

[25] Z. Zhu and M. L. Stein, "Parameter estimation for fractional Brownian surfaces," *Statist. Sinica*, vol. 12, no. 3, 863–883, July 2002.

Debashis Mondal received the Ph.D. degree from the University of Washington, Seattle, in 2007.

He is an assistant professor with the Department of Statistics, University of Chicago. His research interests include time series analysis, wavelets, spectral analysis, Markov random fields, and geo-statistics.

Donald B. Percival received the B.A. degree in astronomy from the University of Pennsylvania, Philadelphia, in 1968, the M.A. degree in mathematical statistics from George Washington University, Washington, DC, in 1976, and the Ph.D. degree in statistics from the University of Washington, Seatlle, in 1982.

From 1968 to 1978, he was an astronomer with the Time Service Division, U.S. Naval Observatory, Washington, DC, where he worked on the generation of atomic clock time scales and on the analysis of the frequency stability of high-performance oscillators. Since 1982, he has been with the Applied Physics Laboratory and the Department of Statistics, University of Washington, where he currently holds the positions of Principal Mathematician and Professor. His research interests include spectral analysis, wavelets, and the use of statistical methodology in physical sciences. He is the co-author (with A.T. Walden) of the textbooks *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques* (Cambridge, U.K.: Cambridge Univ. Press, 1993) and *Wavelet Methods for Time Series Analysis* (Cambridge, U.K.: Cambridge Univ. Press, 2000). He is an Associate Editor of the Journal of Computational and Graphical Statistics.