A New Look at Rainfall Fluctuations and Scaling Properties of Spatial Rainfall Using Orthogonal Wavelets

PRAVEEN KUMAR AND EFI FOUFOULA-GEORGIOU

St. Anthony Falls Hydraulic Laboratory, Department of Civil and Mineral Engineering, University of Minnesota, Minneapolis, Minnesota

(Manuscript received 24 June 1991, in final form 7 April 1992)

ABSTRACT

It has been observed that the finite-dimensional distribution functions of rainfall cannot obey simple scaling laws due to rainfall intermittency (mixed distribution with an atom at zero) and the probability of rainfall being an increasing function of area. Although rainfall fluctuations do not suffer these limitations, it is interesting to note that very few attempts have been made to study them in terms of their self-similarity characteristics. This is due to the lack of unambiguous definition of fluctuations in multidimensions. This paper shows that wavelet transforms offer a convenient and consistent method for the decomposition of inhomogeneous and anisotropic rainfall fields in two dimensions and that the components of this decomposition can be looked at as fluctuations of the rainfall field. It is also shown that under some mild assumptions, the component fields can be treated as homogeneous and thus are amenable to second-order analysis, which can provide useful insight into the nature of the process. The fact that wavelet transforms are a space-scale method also provides a convenient tool to study scaling characteristics of the process. Orthogonal wavelets are used, and these properties are investigated for a squall-line storm to study the presence of self-similarity.

1. Introduction

Rainfall is produced by atmospheric processes that are highly nonlinear and operate and interact at a range of scales. Despite such nonlinearity and interscale dependence, the rainfall patterns are well organized, exhibiting hierarchical structure and clustering, although they are highly variable at different scales and from storm to storm. One of the main challenges to hydrologists, meteorologists, and climatologists is to measure, model, and predict the nature of this variability exhibited by rainfall at different scales. Recent research (e.g., Lovejoy and Schertzer 1990; Gupta and Waymire 1990) has indicated the exciting possibility that rainfall may exhibit scaling–multiscaling characteristics; that is, relatively simple transformations may exist to characterize the behavior of rainfall at some scale if the behavior at another scale is known. The presence of such a hidden structure in an otherwise highly irregular and variable, although organized, pattern bears great significance for the measurement and modeling of the rainfall process.

In characterizing rainfall as a self-similar process, several nuances arise due to its peculiar characteristics such as intermittency, positivity, nonhomogeneity, and anisotropy. Kedem and Chiu (1987) argue against rainfall being self-similar (simple scaling) based on two critical observations: 1) the probability of rain as a function of area is in general an increasing function of area; and 2) the distribution of rain rate is of mixed type having an atom at zero. In order to overcome these restrictions, Lovejoy and Schertzer (1989) suggested the use of fluctuations of rainfall as the underlying process for identifying self-similarity since that process does not have the restrictions identified by Kedem and Chiu. Gupta and Waymire (1990) used conditional probabilities, that is, rainfall process conditioned on being nonzero, in order to eliminate the problem of atom at zero. Lovejoy and Schertzer (1985) also introduced the concept of generalized scale invariance to model certain types of anisotropy, but both Lovejoy and Schertzer (1985) and Gupta and Waymire (1990) assume homogeneity of the rainfall fields.

It is interesting that only a few attempts have been made to study fluctuations of the rainfall process. The reason seems simple: there has been no consistent method up to now to define fluctuations in two or more dimensions. One such attempt is described by Waymire (1985), where fluctuations are obtained by certain limiting operations on the rainfall field. This paper is another attempt in this direction: we show that wavelet multiresolution analysis recently introduced by Meyer (1988) and Mallat (1989b) provides such a framework

Corresponding author address: Efi Foufoula-Georgiou, St. Anthony Falls Hydraulic Laboratory, University of Minnesota, Dept. of Civil and Mineral Engineering, Mississippi River at Third Avenue SE, Minneapolis, MN 55414-2196.

of analysis. In addition, it provides a method to account for nonhomogeneity and anisotropy and also to study the presence of self-similarity characteristics.

The study of self-similarity calls for studying characteristics of the process at different scales. Usually, however, sampling of the process is done at one particular scale, and in order to infer the behavior of the process at a different scale, a transformation of the available dataset to a dataset at other scales is needed. The question then arises as to how to best transform a process to a new scale and how to measure the goodness of the transformation. Wavelet multiresolution analysis provides a solution to both these problems where the transformation is done via a convolution operation and the goodness is in the least-squares sense. The usual method of averaging adjacent pixels to traverse across scales is shown to be a special case of this more generalized framework.

When we transform a process across scales by averaging, that is, low-pass filtering, we lose the information contained in higher-frequency components. As one of the major results of this paper, we show that this information that is lost in the averaging process is significant for identifying self-similarity. In fact, we show that it is this information that is important for the identification of self-similarity for a particular class of processes (processes with self-similar and stationary increments). Multiresolution analysis again plays a significant role in that it allows us to capture this information for analysis. We show that this information, which is the wavelet transform of the process, can be looked at as "fluctuations" of the process. Although it is difficult to define fluctuations in two dimensions, wavelet coefficients, which essentially provide the same information, are easily extendable to two dimensions and are amenable to analysis for identifying self-similarity. Moreover, as we discuss later, they also contain directional information that allows us to understand anisotropic behavior.

In summary, in this paper we show how wavelet multiresolution analysis is an elegant framework for the identification of self-similarity by virtue of (a) being a framework for interscale analysis, (b) providing a method to study nonhomogeneous anisotropic processes, and (c) providing an unambiguous method for defining fluctuations in two dimensions. We also emphasize the role of fluctuations in the study of self-similar processes. The paper is structured as follows. In section 2, we briefly introduce the concepts and relevant terminology of self-similar processes. In section 3, we review multiresolution wavelet analysis focusing only on the theoretical background essential for our study. For complete treatment of wavelet transforms, the reader is referred to the original works of Mallat (1989b), Meyer (1988), Daubechies (1988), and references therein. In section 4, we develop our own results related to wavelet multiresolution analysis of random fields. Section 5 contains the results of multiresolution wavelet analysis of a squall-line storm. We conclude the paper with some remarks.

2. Review of self-similar stochastic processes

a. Simple scaling

A real-valued stochastic process $X(t)$, $t \in \mathbf{R}$ is said to be stochastically self-similar if there exists a constant C_λ such that the finite-dimensional distribution of $X(t)$ satisfies the equation

$$P[C_\lambda^{-1}X(\lambda t_1) < x_1, \cdots, C_\lambda^{-1}X(\lambda t_n) < x_n] = P[X(t_1) < x_1, \cdots, X(t_n) < x_n], \quad (1)$$

which is equivalently expressed as $\{C_\lambda^{-1}X(\lambda t)\} \overset{d}{=} \{X(t)\}$. It can be shown that since C_λ satisfies $C_{\lambda_1\lambda_2} = C_{\lambda_1}C_{\lambda_2}$, it follows, as the homogeneous solution of this equation, that $C_\lambda = \lambda^H$, $\lambda > 0$, $H \in \mathbf{R}$ [see Lamperti's (1962) theorem 1 for a rigorous derivation]. Therefore, for a self-similar process, (1) can be equivalently written as

$$\{X(\lambda t)\} \overset{d}{=} \{\lambda^H X(t)\} \quad \lambda > 0, \quad H \in \mathbf{R}. \quad (2)$$

Lamperti (1962) introduced a class of self-similar processes called semistable processes wherein the process $X(t)$ has to satisfy additional constraints of being defined only for a nonnegative parameter t and satisfying the continuity condition $\lim_{h\to 0} \Pr(|X(t+h) - X(t)| > \epsilon) = 0$ for every $t \geq 0$ and $\epsilon > 0$. He called such semistable processes *proper* if they have a nondegenerate distribution for every $t > 0$. Under these constraints of nonnegative t, continuity in probability (or equivalently in the mean-square sense), and being proper, $X(t)$ is necessarily a nonstationary process [Lamperti 1962, his Eq. (7)]. Such processes will be abbreviated as $H - ss$. A typical example of such a process is the standard Brownian motion.

Mandelbrot and Van Ness (1968, definition 3.2) introduced a wider class of self-similar processes as a class of processes $\{X(t), -\infty < t < \infty\}$ whose increments satisfy the equation

$$\{X(t_0 + \lambda\tau) - X(t_0)\} \overset{d}{=} \{\lambda^H[X(t_0 + \tau) - X(t_0)]\} \quad (3)$$

for any $t_0 \in \mathbf{R}$ and $\lambda > 0$, that is, a process whose increments are stationary and self-similar. Such a process, which has self-similar and stationary increments, will be abbreviated as $H - sssi$. The increment process itself, being self-similar, has an important bearing in our analysis as is discussed later. In this construct of self-similar processes, no condition of continuity is imposed on the process, and this definition admits stationary processes. It can be further shown that (see Mandelbrot and Van Ness 1968, proposition 3.7) if $X(t)$ has self-similar stationary increments and is mean-square continuous, then $0 \leq H < 1$. Fractional Brownian motion $B_H(t)$ is a typical example of a process that

gives rise to self-similar stationary increments called (discrete) fractional Gaussian noise (fGn).

Self-similar processes that are scaling in the sense of (1) or (3) are usually termed *strict-sense simple scaling.* Simple scaling is indicator of the fact that there is only one scaling exponent H for the process. Strict refers to the context of scaling in the sense of distribution as against wide-sense simple scaling, in which case scaling of the covariance function is meant. More precisely, a process is wide-sense simple scaling if it has zero mean and its covariance function $R_X(t, s)$ satisfies

$$R_X(\lambda t, \lambda s) = \lambda^{2H} R_X(t, s). \quad (4)$$

Evidently, if the process is Gaussian and wide-sense simple scaling, then it is also strict-sense simple scaling.

The covariance function of a process $X(t)$, $t \geq 0$, which has self-similar stationary increments, is given by [see Yaglom 1987, his Eq. (4.264a)]

$$R_X(t, s) = \frac{C}{2}(|t|^m + |s|^m - |t - s|^m) \quad t, s \geq 0, \quad m > 0, \quad (5)$$

where C is a constant. Gaussian processes with the above covariance function are the fractional Brownian motions (fBm) with $m = 2H$. The properties of such covariance functions are discussed in further detail in Ossiander and Waymire (1989). For fBm, $C = E[B_H^2(1)] = -[\Gamma(2 - 2H)\cos\pi H]/\pi H(2H - 1)$ (see Barton and Poor 1988).

b. *Multiscaling*

Most of the processes in nature are more complex than can be described by a stochastic process obeying strict-sense simple scaling laws. It therefore becomes necessary to describe such processes in terms of more than one parameter while still keeping the objective of parsimonious representation in mind. This gives rise to the notion of multiscaling processes discussed below.

From (2), we expect that moments of a simple-scaling process (if they exist) will satisfy

$$E[X^h(\lambda)] = \lambda^{hH} E[X^h(1)], \quad (6)$$

or equivalently

$$\log m_h(\lambda) = s(h)\log\lambda + \log m_h(1), \quad (7)$$

where $m_h(\lambda) = E[X^h(\lambda)]$ and $s(h) = hH$. It is often evidenced from empirical observations, however, that $s(h)$ is a nonlinear function of h. For example, Gupta and Waymire (1990) observed that $s(h)$ is a concave function of h for a variety of processes. They plotted moments of different orders of rainfall intensity averaged at various scales versus scale length λ and observed that $s(h)$ is a concave function. They found similar observations for growth of slopes in rivers with respect to basin area as a measure of scale parameter. Based on these observations and certain mathematical derivations, they concluded that the concavity of $s(h)$ is an indicator of increasing variability with decreasing scale, and based on this they derived a representation for self-similar characteristics as multiplicative cascades. The terminology *multiplicative cascade* is due to the fact that, since the spatial variability increases with a decrease in scale, the spatial variability of the process can be seen as a consequence of the cascading down of some large-scale flux to successively smaller scales.

In order to accommodate such a representation, Gupta and Waymire (1990) modified the scale function C_λ to satisfy the functional equation

$$C_{\lambda_1\lambda_2} \overset{d}{=} C_{\lambda_1} C_{\lambda_2} \quad (8)$$

in the sense of distribution by taking C_λ to be a random function of λ. Based on such a modification, they concluded that C_λ is of the form

$$C_\lambda \overset{d}{=} \exp[\mu \log\lambda + Z_{\log(1/\lambda)}] \quad \lambda \leq 1. \quad (9)$$

Hence, a continuous cascade stochastic process $X(t)$ satisfies the relationship

$$X(\lambda) \overset{d}{=} X(1)\exp[\mu \log\lambda + Z_{\log(1/\lambda)}], \quad (10)$$

with Z_t being a stochastic process with stationary increments, not necessarily independent, starting at zero; that is, $Z_0 \equiv 0$. They provide an example of such a process derived from the Brownian motion process.

It is interesting that similar multiplicative processes have independently been studied as multifractal processes quite extensively over the last decade from a different point of view (see Schertzer and Lovejoy 1987). There is an equivalence between these two representations, namely, stochastic multiscaling processes and multifractal processes, in that $s(h)$ is the Legendre transform of the codimension function $c(\gamma)$ of the multifractal process (see Schertzer and Lovejoy 1987, p. 9700). This representation is, however, not relevant to our analysis and therefore will not be discussed further. The reader is referred to Lovejoy and Schertzer (1990) and references therein for more details in this rich area of research.

3. Introduction to wavelet transforms

Wavelet transforms are integral transforms using integration kernels called wavelets. These transforms enable the study of nonstationary processes (signals) in that they enable both localization in time and frequency. Hence, locations of rapid changes in a signal can be detected easily. This is unlike Fourier transforms, which are localized in frequency but not in time; that is, they contain information about the frequencies of the signal over all times instead of showing how the frequencies vary in time. Wavelets have been introduced by Grossman and Morlet (1984) as functions whose translations and dilations could be used for expansion of functions in $L^2(\mathbf{R})$; that is, any function

in $L^2(\mathbf{R})$ can be characterized from its decomposition on the wavelet family $\{\sqrt{s}\psi(s(t-u))\}_{(s,u)\in\mathbf{R}^2}$ for some function ψ. In contrast, it is interesting and we emphasize that Fourier transforms are functions whose modulations provide a basis for expansion of functions in $L^2(\mathbf{R})$. An important class of wavelet transforms was found by Meyer (1988). He proved that there exists some discrete wavelet $\psi(t)$ such that $\{\sqrt{2^m}\psi(2^m(t-2^{-m}n))\}_{(m,n)\in\mathbf{Z}^2}$ is an orthonormal basis of $L^2(\mathbf{R})$. These wavelets are discrete in the sense of using a discrete lattice of translates and dilates rather than the entire plane of possibilities. These discrete orthonormal wavelet bases provide an important new tool in functional analysis as they yield simple orthonormal bases for $L^2(\mathbf{R})$ whose elements have good localization properties in both the spatial and frequency domains. Meyer (1988) and Mallat (1989b) developed the concept of *multiresolution analysis* from these orthonormal bases.

Wavelet decompositions have found many applications in sound analysis (Risset 1989), image processing (Mallat 1989c), robotics and vision (Mallat 1989a), and fluid dynamics and turbulence (Everson et al. 1990; Liandrat and Moret-Bailly 1990), and in the study of fractals (Holschneider 1988), among others (e.g., see Combes et al. 1987). The discrete orthonormal wavelets have found applications in theoretical mathematics (Perrier 1989) and signal processing (Mallat 1989a).

a. Continuous wavelet transforms

Let $L^2(\mathbf{R})$ denote the vector space of complex-valued, square integrable functions $f(t)$. By (f, g) we denote the L^2 inner product of $f, g \in L^2(\mathbf{R})$ given by $(f, g) = \int_{-\infty}^{\infty} f(t)\bar{g}(t)dt$, where $\bar{g}(t)$ is the complex conjugate of $g(t)$. The L^2 norm of $f(t)$ is then $\|f\| = (f, f)^{1/2} = [\int_{-\infty}^{\infty} |f(t)|^2 dt]^{1/2}$. The Fourier transform of $f(t) \in L^2(\mathbf{R})$ is obtained as $\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt$, and the convolution of two functions $f, g \in L^2(\mathbf{R})$ is given by $(f*g)(t) = \int_{-\infty}^{\infty} f(u) \times g(t-u)du$. By $l^2(\mathbf{Z})$, we denote the vector space of square summable sequences, that is, $l^2(\mathbf{Z}) = \{\epsilon_i : \sum_{i=-\infty}^{\infty} |\epsilon_i|^2 < \infty, \epsilon_i \in \mathbf{R}, i \in \mathbf{Z}\}$.

In general, the wavelet transform of a function $f(t)$ is defined as the L^2 inner product

$$Wf(s, u) = (f, \psi_s(\cdot - u)) \tag{11}$$

$$= \sqrt{s}\int_{-\infty}^{\infty} f(t)\psi(s(t-u))dt, \tag{12}$$

where $\psi_s(t) = \sqrt{s}\psi(st)$. In order for the wavelet transform to be invertible, it should satisfy (see Grossman and Morlet 1984) $C_\psi = \int_0^\infty |\hat{\psi}(\omega)|^2/\omega d\omega < \infty$, which translates to the requirement that $\psi(t)$ should have a zero mean, that is,

$$\int_{-\infty}^{\infty} \psi(t)dt = 0. \tag{13}$$

The inverse formula is then given by $f(t) = C_\psi^{-1} \times \int_{-\infty}^{\infty}\int_0^\infty Wf(s, u)\psi_s(t-u)dsdu$.

The function $\psi(t)$ can be interpreted as the impulse response to a bandpass filter. Indeed, $(f, \psi(\cdot - u))$ can be equivalently written as $(f*\tilde{\psi})(u)$, where $\tilde{\psi}(t) = \psi(-t)$. The wavelet transform $Wf(s, u) = (f, \psi_s(\cdot - u))$ can therefore be viewed as filtering of f with a bandpass filter $\tilde{\psi}(t)$. The Fourier transform of $\psi_s(t) = \sqrt{s}\psi(st)$ is given by

$$\hat{\psi}_s(\omega) = \frac{1}{\sqrt{s}}\hat{\psi}\left(\frac{\omega}{s}\right). \tag{14}$$

Hence, by dilating the function, the center of passing band $s\omega_0$ of $\hat{\psi}_s(\omega)$, where ω_0 is given by the solution of $\int_0^\infty (\omega - \omega_0)|\hat{\psi}(\omega)|^2 d\omega = 0$, decreases. The root-mean-square bandwidth $s\sigma_\omega$ of $\psi_s(t)$ around $s\omega_0$, where σ_ω is given by $\sigma_\omega^2 = \int_0^\infty (\omega - \omega_0)^2|\hat{\psi}(\omega)|^2 d\omega$, also decreases. Hence, by changing the value of s, we can extract the behavior of the function in different frequency bands. Usually ψ is chosen to be a function of compact support, that is, it is zero everywhere outside a finite interval, and it thus gives localization in time. Also, it should be noted that if $\psi(t)$ is real, then $|\hat{\psi}(\omega)| = |\hat{\psi}(-\omega)|$. It can be shown (see Mallat 1989c) that these frequency bands have a constant size on the logarithmic scale. Thus, when scale s is small, the resolution is coarse in the spatial domain and fine in the frequency domain. If the scale s increases, the resolution increases in the spatial domain and decreases in the frequency domain (see Fig. 1).

b. Wavelet multiresolution analysis

Meyer (1988) and Mallat (1989b) developed a rigorous mathematical framework for the concept of multiresolution decomposition. As defined by Mallat (1989b), a multiresolution approximation of $L^2(\mathbf{R})$ is a sequence $(V_m)_{m\in\mathbf{Z}}$ of closed subspaces of $L^2(\mathbf{R})$ such that the following hold:

M1: $V_m \subset V_{m+1}\ \forall m \in \mathbf{Z}$,
M2: $\cup_{m=-\infty}^{\infty} V_m$ is dense in $L^2(\mathbf{R})$ and $\cap_{m=-\infty}^{\infty} V_m = \{0\}$,
M3: $f(t) \in V_m$ iff $f(2t) \in V_{m+1}\ \forall m \in \mathbf{Z}$,
M4: $f(t) \in V_m$ implies $f(t - [k/2^m]) \in V_m\ \forall k \in \mathbf{Z}$,
M5: there exists an isomorphism $\mathbf{I}$ from V_0 onto $l^2(\mathbf{Z})$, which commutes with the action of $\mathbf{Z}$; this property is best explained by the cummutative diagram

$$\begin{array}{ccc} f(t) \in V_0 & \xrightarrow{\mathbf{I}} & \epsilon_n \in l^2(\mathbf{Z}) \\ \downarrow z & & \downarrow z \\ f(t-k) \in V_0 & \xrightarrow{\mathbf{I}} & \epsilon_{n-k} \in l^2(\mathbf{Z}), \end{array}$$

which essentially depicts the translation invariance property (for an integer translate) of the transform.

The approximation of a function $f(t) \in L^2(\mathbf{R})$ at a resolution m (i.e., we have 2^m data points per unit length) is given by the orthogonal projection of $f(t)$

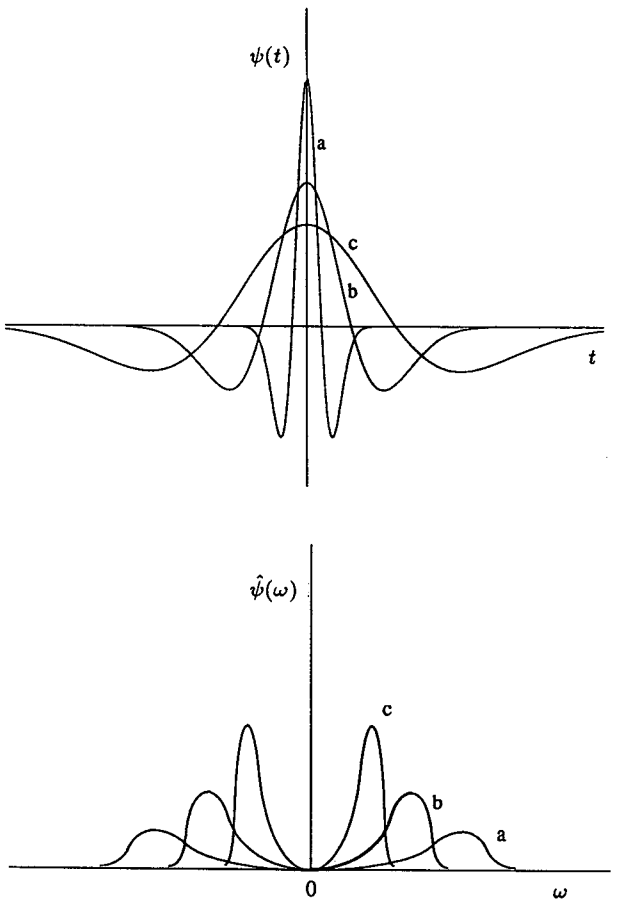

FIG. 1. This figure illustrates the effect of dilation on a "generic" wavelet and the corresponding change on its Fourier transform $|\hat{\psi}(\omega)|$. When the wavelet dilates, its Fourier transform contracts and vice versa.

on V_m. Let P_m represent this projection operator, that is,

$$f(t) \in L^2(\mathbf{R}) \Rightarrow P_m f \in V_m \subset L^2(\mathbf{R}). \quad (15)$$

Since P_m represents an orthogonal projection, it implies that

$$\|f(t) - P_m f(t)\| = \inf \|f(t) - g(t)\| \quad \forall g(t) \in V_m \quad (16)$$

and also that $P_m \circ P_m = P_m$, where $\circ$ denotes the composite operator. As a consequence of property M1, $P_{m+1}f$ contains all the necessary information required to compute $P_m f$. Also, $P_m f$ is the best approximation of $f(t)$ in the least-squares sense.

It can be shown that for a given multiresolution approximation $(V_m)_{m\in\mathbf{Z}}$, there exists a unique function $\phi(t) \in L^2(\mathbf{R})$, called a *scaling function,* such that if $\phi_m(t) = 2^m\phi(2^m t)$, that is, $\int \phi_m(t)dt = 1$ for all $m \in \mathbf{Z}$, then $[2^{-m/2}\phi_m(t - 2^{-m}n)]_{n\in\mathbf{Z}}$ is an orthonormal basis of V_m, that is, V_m = linear span $\{[2^{-m/2}\phi_m(t - 2^{-m}n)]_{n\in\mathbf{Z}}\}$.

Let O_m denote the orthogonal complement of V_m in V_{m+1}, that is,

$$V_{m+1} = V_m \oplus O_m, \quad (17)$$

where $\oplus$ denotes the direct sum. Then the difference in information between $P_{m+1}f$ and $P_m f$ can be characterized by finding the orthogonal projection of f onto O_m. Let Q_m denote this projection operator, that is, $f \in L^2(\mathbf{R}) \Rightarrow Q_m f \in O_m$, or equivalently $O_m = Q_m L^2(\mathbf{R})$. Note that $Q_m = I_{m+1} - P_m$, where I_m is the identity operator on V_m. The function $Q_m f$ is called the *detail function* at resolution m.

It can be shown that there exists a function $\psi(t) \in L^2(\mathbf{R})$, depending upon $\phi(t)$, such that if $\psi_m(t) = 2^m\psi(2^m t)$, then $[2^{-m/2}\psi_m(t - 2^{-m}n)]_{n\in\mathbf{Z}}$ is an orthonormal basis of O_m, that is, O_m = linear span $\{[2^{-m/2}\psi_m(t - 2^{-m}n)]_{n\in\mathbf{Z}}\}$. By recursively using (17), it can be shown that the translates and dilates of ψ, that is, $[2^{-m/2}\psi_m(t - 2^{-m}n)]_{(m,n)\in\mathbf{Z}^2}$, are an orthonormal basis of $L^2(\mathbf{R})$. The function $\psi(t)$ is called an *orthogonal wavelet,* and (f, ψ_{mn}) is called the *orthogonal wavelet transform.* Also, $\psi(t)$ satisfies the admissibility condition of (13). It follows immediately that O_m are orthogonal spaces that sum to $L^2(\mathbf{R})$, that is, $L^2(\mathbf{R}) = \oplus_{m\in\mathbf{Z}} O_m$.

Using the basis functions ϕ and ψ, as described above for the orthogonal transforms,

$$P_m f(t) = 2^{-m} \sum_{n=-\infty}^{\infty} (f, \phi_{mn})\phi_{mn}(t) \quad (18)$$

and

$$Q_m f(t) = 2^{-m} \sum_{n=-\infty}^{\infty} (f, \psi_{mn})\psi_{mn}(t) \quad (19)$$

can be obtained by orthogonal projections, where $\phi_{mn} = 2^m\phi(2^m t - n)$ and $\psi_{mn} = 2^m\psi(2^m t - n)$. Therefore, the set of inner products $P_m^d f = \{(f, \phi_{mn})_{n\in\mathbf{Z}}\}$ gives the discrete approximation of $f(t)$ [or sampled $f(t)$] at resolution m and $Q_m^d f = \{(f, \psi_{mn})_{n\in\mathbf{Z}}\}$ gives the discrete detail approximation of $f(t)$ (or difference in information between resolutions m and $m - 1$). As described previously, the function $\psi(t)$ can be interpreted as impulse response to bandpass filter. Similarly, $\phi(t)$ can be interpreted as impulse response to low-pass filter.

A simple example of a multiresolution approximation, which has the Haar function as the basis for $L^2(\mathbf{R})$ (see Mallat 1989b), is now presented. Let V_m be the space of piecewise constant functions, that is,

$$V_m = \{g \in L^2(\mathbf{R}) : g \text{ is constant on } [2^{-m}n, 2^{-m}(n + 1)), \quad \forall n \in Z\}. \quad (20)$$

It is known that the vector space of piecewise constant functions is dense in $L^2(\mathbf{R})$. The projection of f onto the interval $[2^{-m}n, 2^{-m}(n + 1))$ is given by

$$P_m f|_{[2^{-m}n,\, 2^{-m}(n+1))} = 2^m \int_{2^{-m}n}^{2^{-m}(n+1)} f(t)dt. \quad (21)$$

As m increases, the successive $P_m f(t)$ correspond to finer and finer scales. The scaling function can therefore be chosen as the characteristic function of [0, 1), that is,

$$\phi(t) = \begin{cases} 1, & 0 \leq x < 1 \\ 0, & \text{otherwise.} \end{cases} \quad (22)$$

The corresponding wavelet $\psi(t)$ is the Haar function (see Fig. 2) given by

$$\psi(t) = \begin{cases} 1, & 0 \leq x < 1/2 \\ -1, & 1/2 \leq x < 1 \\ 0, & \text{otherwise,} \end{cases} \quad (23)$$

thereby giving $Q_m f$ onto the interval $[2^{-m}n, 2^{-m}(n + 1))$ as

$$Q_m f|_{[2^{-m}n,\, 2^{-m}(n+1))} = 2^m \left[\int_{2^{-m}n}^{2^{-m}(n+1/2)} f(t)dt - \int_{2^{-m}(n+1/2)}^{2^{-m}(n+1)} f(t)dt \right]. \quad (24)$$

This is a crude example of multiresolution approximation since the functions in V_m are not smooth. Nevertheless, it serves as a good example for our purpose of characterizing self-similarity in multiresolution framework as described in the next section. The reader will recognize that $\phi(t)$ is the usual function used in transforming data from one scale to another by averaging adjacent pixels. This, however, is a special case of a more generalized framework outlined here. Daubechies (1988) has described a method for constructing wavelets $\psi(t)$, with corresponding scaling functions $\phi(t)$, with compact support having an arbitrary high order of smoothness.

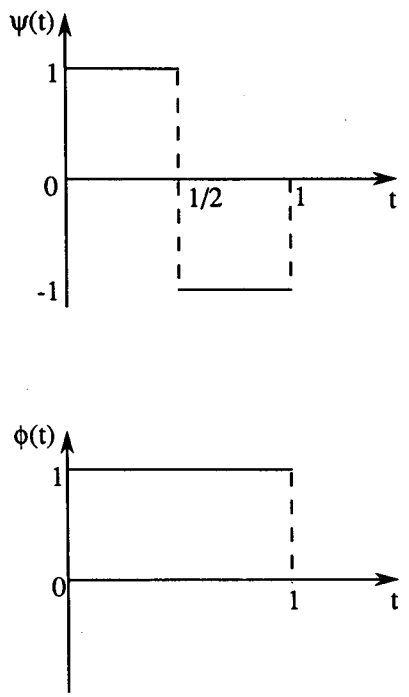

FIG. 2. Haar wavelet and the corresponding scale function. This pair provides the simplest example of wavelet multiresolution framework.

If we have a function $f(t)$, sampled at any resolution, say m, giving rise to a data sequence $\{c_n^m\}_{m\in\mathbf{Z}}$, then we assume that $P_m f = 2^{-m} \sum_n c_n^m \phi_{mn}$; that is, $c_n^m = \langle f, \phi_{mn} \rangle$. Mallat (1989a) has described an elegant discrete convolution algorithm to convert the data sequence $\{c_n^m\}$ at resolution m to a data sequence $\{c_n^l\}$ at any other resolution l, $l < m$. The detail information, that is, the information lost in going from the higher to the lower resolution, is captured in the sequence $\{d_n^l\}_{n\in\mathbf{Z}}$, where $\{d_n^l\} = \{\langle f, \psi_{ln} \rangle\}$ and ψ is the wavelet corresponding to the chosen ϕ. The transformation equations are

$$c_k^{m-1} = \sum_n h(n - 2k) c_n^m \quad (25)$$

and

$$d_k^{m-1} = \sum_n g(n - 2k) c_n^m, \quad (26)$$

where $h(n)$ and $g(n)$ are obtained as $\langle \phi_{00}, \phi_{1n} \rangle$ and $\langle \phi_{00}, \psi_{1n} \rangle$ [for a more complete description on choosing $h(n)$ and $g(n)$, see Daubechies (1988)]. This thus gives us a method to traverse across scales and obtain the data sequences $\{c_n^l\}$ at various scales l while capturing the lost information due to averaging as the sequences $\{d_n^l\}$ at these scales.

c. *Two-dimensional multiresolution representation*

The multiresolution approximation can be easily extended to two dimensions. The function under consideration is now $f(t_1, t_2) \in L^2(\mathbf{R}^2)$. A multiresolution approximation of $L^2(\mathbf{R}^2)$ is a sequence of subspaces that satisfy the straightforward two-dimensional extension of properties M1–M5 enumerated in the definition of one-dimensional multiresolution approximation. Such a sequence of subspaces of $L^2(\mathbf{R}^2)$ is denoted by $(V_m)_{m\in\mathbf{Z}}$. The approximation of the function $f(t_1, t_2)$ at the resolution m is the orthogonal projection on the vector space V_m.

A two-dimensional multiresolution approximation is called separable if the representation is computed by filtering the signal with a low-pass filter of the form $\Phi(t_1, t_2) = \phi(t_1)\phi(t_2)$. For separable multiresolution approximation of $L^2(\mathbf{R}^2)$, each vector space V_m can be decomposed as a tensor product of two identical subspaces V_m^1 of $L^2(\mathbf{R})$, that is, $V_m = V_m^1 \otimes V_m^1$. It therefore follows that

$$V_{m+1} = V_{m+1}^1 \otimes V_{m+1}^1 \quad (27)$$

$$= (V_m^1 \oplus O_m^1) \otimes (V_m^1 \oplus O_m^1), \quad (28)$$

which can be rewritten as

$$V_{m+1} = (V_m^1 \otimes V_m^1) \oplus (V_m^1 \otimes O_m^1) \oplus (O_m^1 \otimes V_m^1) \oplus (O_m^1 \otimes O_m^1). \quad (29)$$

Therefore, the orthogonal complement O_m of V_m in V_{m+1} is $(V_m^1 \otimes O_m^1) \oplus (O_m^1 \otimes V_m^1) \oplus (O_m^1 \otimes O_m^1)$.

The sequence of vector spaces $(V_m)_{m\in\mathbf{Z}}$ forms a multiresolution approximation of $L^2(\mathbf{R}^2)$ if and only if $(V_m^1)_{m\in\mathbf{Z}}$ is a multiresolution approximation of $L^2(\mathbf{R})$. Analogous to the one-dimensional case, the detail function at the resolution m is equal to the orthogonal projection of the function onto the space O_m, which is the orthogonal complement of V_m in V_{m+1}. It can be shown that the orthonormal basis for V_m is then given by

$$[2^{-m}\Phi_m(t_1 - 2^{-m}n, t_2 - 2^{-m}k)]_{(n,k)\in\mathbf{Z}^2} = [2^{-m}\phi_m(t_1 - 2^{-m}n)\phi_m(t_2 - 2^{-m}k)]_{(n,k)\in\mathbf{Z}^2}.$$

The orthonormal basis for each component of O_m, that is, $(V_m^1 \otimes O_m^1)$, $(O_m^1 \otimes V_m^1)$, and $(O_m^1 \otimes O_m^1)$, is given by $\{[2^{-m}\Psi_m^1(t_1 - 2^{-m}n, t_2 - 2^{-m}k)]_{(n,k)\in\mathbf{Z}^2}\}$, $\{[2^{-m}\Psi_m^2(t_1 - 2^{-m}n, t_2 - 2^{-m}k)]_{(n,k)\in\mathbf{Z}^2}\}$, and $\{[2^{-m}\Psi_m^3(t_1 - 2^{-m}n, t_2 - 2^{-m}k)]_{(n,k)\in\mathbf{Z}^2}\}$, respectively, where $\Psi^1(t_1, t_2) = \phi(t_1)\psi(t_2)$, $\Psi^2(t_1, t_2) = \psi(t_1)\phi(t_2)$, and $\Psi^3(t_1, t_2) = \psi(t_1)\psi(t_2)$.

The discrete approximation of the function $f(t_1, t_2)$ at a resolution m is characterized by the inner product

$$P_m^d f = \{(f, \Phi_m(\cdot - 2^{-m}n, \cdot - 2^{-m}k))_{(n,k)\in\mathbf{Z}^2}\} = \{(f, \phi_m(\cdot - 2^{-m}n)\phi_m(\cdot - 2^{-m}k))_{(n,k)\in\mathbf{Z}^2}\}. \quad (30)$$

The discrete detail approximation of the function is obtained by the inner product of $f(t_1, t_2)$, with each of the vectors of the orthonormal basis of O_m. This is thus given by

$$Q_m^{d^1} f = \{(f, \Psi_m^1(\cdot - 2^{-m}n, \cdot - 2^{-m}k))_{(n,k)\in\mathbf{Z}^2}\}$$
$$Q_m^{d^2} f = \{(f, \Psi_m^2(\cdot - 2^{-m}n, \cdot - 2^{-m}k))_{(n,k)\in\mathbf{Z}^2}\}$$
$$Q_m^{d^3} f = \{(f, \Psi_m^3(\cdot - 2^{-m}n, \cdot - 2^{-m}k))_{(n,k)\in\mathbf{Z}^2}\}. \quad (31)$$

The (separable) wavelet decomposition can thus be interpreted as a function decomposition in a set of independent, *spatially oriented* frequency channels. Here $Q_m^{d^1} f$ gives the vertical high frequencies (horizontal edge or horizontal high correlation), $Q_m^{d^2} f$ gives the horizontal high frequencies (vertical edge or vertical high correlation), and $Q_m^{d^3} f$ gives high frequencies in both the directions (the corners or high correlation in both horizontal and vertical direction).

Only the special case of a separable multiresolution approximation that results in the decomposition of a two-dimensional function using three orientations has been indicated. One can, however, build a nonseparable wavelet representation having as many orientation tunings as desired by using nonseparable wavelet orthonormal bases (see Cohen and Daubechies 1991).

4. Identifying self-similarity from wavelet coefficients

This section deals with multiresolution analysis of stochastic processes, and therefore our notation will be switched from function $f(t)$ to stochastic process $X(t)$. It is argued that the wavelet coefficients essentially give the description of the fluctuations of the process, and if the process $X(t)$ is self-similar or has self-similar stationary increments, then the wavelet coefficients are self-similar. The latter property is important in that it guarantees that property of self-similarity is preserved during the wavelet transform. In the next section, radar data are analyzed and the results are interpreted under the hypothesis that wavelet coefficients give the description of the fluctuations.

The previous section has shown that for a given data sequence $\{c_n^m\}$ obtained by sampling the process $X(t)$ at some resolution m, we can consider that c_n^m is obtained as $\int X(t)\phi_{mn}(t)dt$. Then, through a convolution operation, c_n^m can be transformed to another scale $l > m$ and the detail information d_n^l can be obtained. This transformation has been shown to be good in the sense of least squares [see (16) and discussion thereafter].

We now need to address the question: If $X(t)$ is self-similar, that is, $\{X(\lambda t)\} \overset{d}{=} \{\lambda^H X(t)\}$, then is $P_m^d X(t)$ or $Q_m^d X(t)$ also self-similar? The answer, which is affirmative for both, can be argued heuristically by noting that

$$(X, \phi_{mn}) = \int X(t)\phi_{mn}(t)dt \overset{d}{=} \lambda^{-H} \int X_\lambda(t)\phi_{mn}(t)dt \quad (32)$$
$$= \lambda^{-H}(X_\lambda, \phi_{mn}) \quad (33)$$

where $X_\lambda(t) = X(\lambda t)$. A similar argument also shows that for such a process, $\{Q_m^d X_\lambda(t)\} \overset{d}{=} \{\lambda^H Q_m^d X(t)\}$, and therefore both $P_m^d X(t)$ and $Q_m^d X(t)$ are self-similar. The process $X(t)$ is known to be a nonstationary process, and therefore $P_m^d X(t)$ and $Q_m^d X(t)$ are in general not necessarily stationary (unless further conditions are imposed).

A slightly different situation arises if the process $X(t)$ has self-similar stationary increments, that is, $\{X(t_0 + \lambda\tau) - X(t_0)\} \overset{d}{=} \{\lambda^H[X(t_0 + \tau) - X(t_0)]\}$. As shown later for such processes, $\{Q_m^d X(\lambda t)\} \overset{d}{=} \{\lambda^H Q_m^d X(t)\}$ [recall $Q_m^d X$ is the discrete detail approximation of $X(t)$]. Such a result, however, cannot necessarily be proven for $P_m^d X(t)$ unless additional constraints like

$X(t_0) = 0$ are imposed on the process. The proof of the above result is as follows. We know that

$$(X_\lambda, \psi_{mn}) = \int \psi_{mn}(t) X(\lambda t) dt. \quad (34)$$

Since $\int \psi_{mn}(t)dt = 0$ [from (13)], it implies $\int \psi_{mn}(t)X(t_0)dt = 0$ [$X(t_0)$ being a random constant]. Subtracting this from the rhs of the previous equation gives

$$(X_\lambda, \psi_{mn}) = \int \psi_{mn}(t)[X(\lambda t) - X(t_0)]dt \quad (35)$$

$$\stackrel{d}{=} \lambda^H \int \psi_{mn}(t)[X(t) - X(t_0)]dt \quad (36)$$

$$= \lambda^H \int \psi_{mn}(t) X(t) dt \quad (37)$$

$$= \lambda^H (X, \psi_{mn}). \quad (38)$$

By considering all n's, $\{Q_m^d X_\lambda(t)\} \stackrel{d}{=} \{\lambda^H Q_m^d X(t)\}$.

Flandrin (1989) has shown that the wavelet transform of a process $X(t)$, which has self-similar stationary increments having covariance function of the form in (5), is stationary at any fixed scale. It is also easy to verify that $Q_m^d X(t)$ is a zero-mean process. Combining these results, that is, self-similarity of $Q_m^d X(t)$ and its stationarity, shows that the wavelet coefficients $Q_m^d X(t)$ are amenable to analysis for the purpose of identifying self-similarity.

Having shown that the wavelet coefficients are significant in the analysis of self-similar processes, it is now argued that the wavelet transforms can be looked upon as fluctuations of the process $X(t)$. A stochastic process $X(t)$ can be decomposed into two components $X_T(t)$ and $[X(t) - X_T(t)]$, where

$$X_T(t) = \frac{1}{T} \int_{t-T/2}^{t+T/2} X(u) du \quad (39)$$

is the "local average" process and $[X(t) - X_T(t)]$ is the "fluctuation" process. The process can then be written as $X(t) \approx X_T(t) + [X(t) - X_T(t)]$. These two components are practically uncorrelated that is, $\text{cov}[X_T(t), X(t) - X_T(t)] \approx 0$ (the covariance decays as $1/T$ for increasing T). The variances of these components are given by $\text{var}[X_T(t)] = \sigma^2\gamma(T)$ and $\text{var}[X(t) - X_T(t)] \approx \sigma^2[1 - \gamma(T)]$, where σ^2 is the variance of the process $X(t)$, and $\gamma(T)$ is the variance function (see Vanmarcke 1984, p. 228).

We claim that a discrete approximation of $X_T(t)$ and $[X(t) - X_T(t)]$ can be obtained by decomposing the process $X(t)$ onto V_m and O_m, respectively, where T is chosen such that $m = -\log_2 T$, $m \in \mathbf{Z}$. Moreover, $X_T(t)$ can be obtained exactly by choosing the scaling function ϕ_m as $1/T$ times the characteristic function of $[-T/2, T/2]$.

A heuristic argument toward the support of this can be given by noting that the fluctuation process is given by

$$Y_T(t) = X(t) - X_T(t) \quad (40)$$

$$= X(t) - \frac{1}{T} \int_{t-T/2}^{t+T/2} X(u) du \quad (41)$$

$$= X(t) - \int_{-\infty}^{\infty} \varphi(t - u) X(u) du \quad (42)$$

$$= X(t) - \int_{-\infty}^{\infty} \tilde{\varphi}(u - t) X(u) du, \quad (43)$$

where $\varphi(t)$ is $1/T$ times the characteristic function of the interval $[-T/2, T/2]$ and $\tilde{\varphi}(t) = \varphi(-t)$. Taking the Fourier transforms gives

$$\hat{Y}_T(\omega) = \hat{X}(\omega) - \hat{\varphi}(\omega)\hat{X}(\omega)$$

$$= \hat{X}(\omega)[1 - \hat{\varphi}(\omega)], \quad (44)$$

where $[1 - \hat{\varphi}(\omega)]$ represents a high-pass filter. If, however, we regard the very high frequencies (which will be absent for any function sampled at any finite resolution) as (uncorrelated) noise, we can approximate $[1 - \hat{\varphi}(\omega)]$ by a bandpass filter. We can choose $\hat{\psi}(\omega)$, corresponding to the Haar wavelet, as such a filter. The scaling function ϕ, of the multiresolution analysis, is then the same as φ. Equation (44) can be written as $\hat{Y}_T(\omega) = \hat{X}(\omega)\hat{\psi}(\omega)$, which gives $Y_T(t) = (X * \tilde{\psi})(t)$. By using the dilates of ψ, say $\psi_s(t)$, we obtain $Y_T(t)$ for different T's. We therefore have $X_T(u) \approx (X(t), \phi_s(t - u))$ and $Y_T(u) \approx (X(t), \psi_s(t - u))$. In this case where s and u are discrete, $s = 2^m$ and $u = n/2^m$, giving rise to orthogonal ϕ and ψ in the multiresolution framework, we obtain the sampling of the functions $X_T(t)$ and $Y_T(t)$ at resolution m.

The above result implies that we can segregate (approximately) the stochastic process $X(t)$ into two components $X_T(t)$ and $[X(t) - X_T(t)]$, that is, $X(t) \approx X_T(t) + [X(t) - X_T(t)]$, whose realizations belong to orthogonal subspaces V_m and O_m, if the realization of the original process is assumed to belong to $L^2(\mathbf{R})$, where m is such that $2^{-m} = T$. Since the averaging process is essentially a low-pass filtering, it can be improved by selecting more regular scaling functions ϕ (see Daubechies 1988). It is easier, however, to understand the concept with the choice of the Haar wavelet. It is easy to see, for example, that for a discrete dataset, the wavelet coefficients are essentially the increments of the process.

The extension of the decomposition process to two dimensions is simple. Define the component processes as

$$X(\mathbf{t}) \approx X_D(\mathbf{t}) + [X(\mathbf{t}) - X_D(\mathbf{t})], \quad (45)$$

where $X_D(\mathbf{t})$ is defined analogous to (39) but for an integration region of size $D = T_1 T_2$. The variances of the component processes are given by $\text{var}[X_D(\mathbf{t})] = \sigma^2\gamma(D)$ and $\text{var}[X(\mathbf{t}) - X_D(\mathbf{t})] \approx \sigma^2[1 - \gamma(D)]$, and $X_D(\mathbf{t})$ and $[X(\mathbf{t}) - X_D(\mathbf{t})]$ are practically uncorrelated, that is, $\text{cov}[X_D(t), X(t) - X_D(t)] \approx 0$.

A very interesting situation arises when this two-dimensional case is analyzed in the multiresolution framework. From the theory of two-dimensional wavelet transforms (see previous section), the function $f(t_1, t_2)$ can be decomposed into four orthogonal components under the separable multiresolution analysis. It is shown below that for a realization of a stochastic process, these components have important interpretations. In particular, the first component $P_m^d X$ of the process $X(t_1, t_2)$ can be interpreted as the sampling of the average process $X_D(t_1, t_2)$ over an integration domain $D = T_1 T_2$, that is, sampling of

$$X_D(t_1, t_2) = \frac{1}{T_1 T_2} \int_{t_1 - T_1/2}^{t_1 + T_1/2} \int_{t_2 - T_2/2}^{t_2 + T_2/2} X(u_1, u_2) du_1 du_2.$$

The second component $Q_m^{d_1} X$ can be interpreted as the sampling of the "fluctuations of the marginal local average process" in the t_1 direction, that is, sampling of the fluctuations of

$$X_{T_1}(t_2; t_1) = \frac{1}{T_1} \int_{t_1 - T_1/2}^{t_1 + T_1/2} X(u_1, t_2) du_1.$$

In an analogous way, the third component $Q_m^{d_2} X$ can be interpreted as the sampling of the "fluctuations of the marginal local average process" in the t_2 direction, that is, sampling of the fluctuations of

$$X_{T_2}(t_1; t_2) = \frac{1}{T_2} \int_{t_2 - T_2/2}^{t_2 + T_2/2} X(t_1, u_2) du_2.$$

The fourth component does not lend itself to such simple interpretations, but it essentially represents the fluctuations along the diagonals.

5. Analysis of spatial rainfall

This section presents results from a multiresolution wavelet analysis of a radar-depicted squall-line storm. This storm started with a well-organized elliptical embedded structure that later dissipated and formed a convective-type organization. The storm occurred on 27 May 1987 and was monitored by the National Severe Storm Laboratory (NSSL) using a WSR-57 radar, which is a 10-cm wavelength system with a peak power of 305 kW and a beam width of 2.2°. The storm lasted for about 8 h starting at around 1100 LST. The conversion of the cloud reflectivity (in dBZ) to rainfall rates (mm h^{-1}) was done at NSSL in Norman, Oklahoma, using the relationship $Z = 300R^{1.4}$, where R is rainfall rate (mm h^{-1}) and the reflectivity factor in dBZ is related to Z (mm^6 m^3) by the relationship 1 dBZ = 10 log$Z$. The rainfall-intensity values are available at approximately 10-min intervals (and averaged over this interval) for 360 azimuths, with every azimuth containing 115 estimates for a range of 230 km (i.e., data at every 2 km × 1°). Data for each scan over the 360 azimuths will henceforth be referred to as a frame. The precipitation processing system, used to correlate reflectivity and rainfall intensity, taking into account the raingage observations and adjustment for ground clutter, etc., is described by O'Bannon and Ahnert (1986). For the purpose of our analysis, the data were converted to a rectangular grid of 512 × 512 by bilinear interpolation.

The rainfall field was decomposed into four components: the average process A ($P_m X$) and the three detail processes D_1 ($Q_m^{d_1} X$), D_2 ($Q_m^{d_2} X$), and D_3 ($Q_m^{d_3} X$), as was discussed in the previous sections. The decomposition was repeated at different resolutions, which here correspond to grids of 512 × 512 (original field), 256 × 256, 128 × 128, 64 × 64, 32 × 32, and 16 × 16. Figure 3 graphically depicts this decomposition and establishes the terminology used in the discussion that follows. The original rainfall at a specific instant of time (frame 1) is shown in Fig. 4, and the averaged field at grid 32 × 32 and its ACF are shown in Fig. 5. The original field is assumed to be nonhomogeneous in the mean and anisotropic as can be clearly seen from Fig. 4. The same holds true for the average fields at all resolutions. The detail fields (D_1, D_2, and D_3) obtained using wavelet transforms are, however, assumed homogeneous. (See Fig. 6 for all four components of the original field and their ACFs at one resolution corresponding to grid 256 × 256.) An argument in favor of this assumption is that if the rainfall field had self-similar stationary increments, then the wavelet transform (detail processes) constitutes a stationary process [see Flandrin (1989) for a proof]. An implication, as a consequence of this assumption, is that the underlying assumption is made that the nonstationary field is composed of a nonsta-

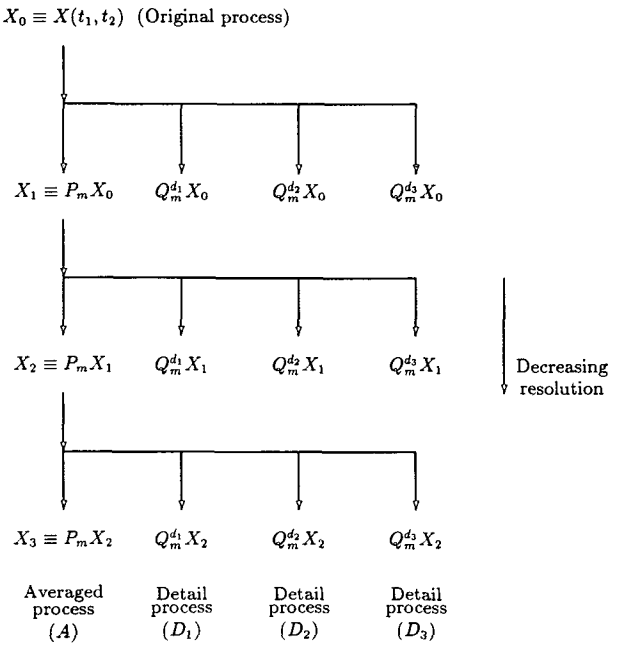

FIG. 3. Schematic of the decomposition of the original rainfall process X_0 into four components (averaged process A and detail processes D_1, D_2, D_3) at decreasing resolutions.

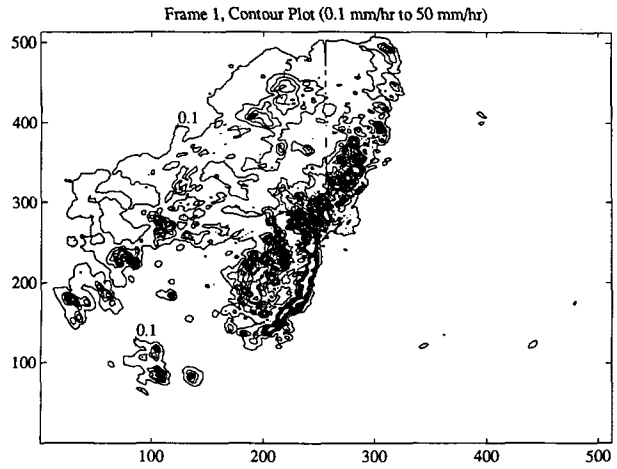

FIG. 4. Original rainfall field at grid 512 × 512 for frame 1 of the squall-line storm. The total areal coverage is 460 km × 460 km. The contours are at intensities between 0.1 and 50 mm h^{-1}.

tionary mean and stationary fluctuations. The stationarity of the detail processes, which may be (as discussed earlier) looked upon as fluctuations of the process, makes them amenable to second-order analysis, which can now be used to check the presence of self-similarity in the rainfall fluctuations. For this, several statistical properties such as the spectral density function (sdf) and spectral bandwidth were estimated for each field at each resolution. Table 1 lists some of these estimates for all four components at three resolutions. These properties, although not properly defined for a nonhomogeneous field, provide some indication of its correlation and persistence characteristics and thus are given in Table 1 for the averaged process A also. The reader is referred to Vanmarcke (1984) for definitions and estimation of these properties and to the Appendix for a brief exposition. It is only noted here that for a process with self-similar stationary increments that are positively correlated ($1/2 < H < 1$), the behavior of these properties is as follows:

1) the sdf behaves as $|\omega|^{1-2H}$, $1/2 < H < 1$, $|\omega| > 0$ and

2) the spectral bandwidth is (see Appendix for definition)

$$\delta \approx \frac{1}{3 - 2H}. \tag{46}$$

It is noted here that for $H = 0.5$, which corresponds to white noise process, the spectral density function is a constant over all frequencies and $\delta = 0.5$.

In computing the above statistical properties, the nonzero mean-removed processes were used. That is, let $X(t_1, t_2)$ be the rainfall field, let $R(t_1, t_2)$ denote the field under analysis (i.e., $P_m^d X$, $Q_m^{d^1} X$, $Q_m^{d^2} X$ or $Q_m^{d^3} X$) as a function of spatial location, and let S denote the domain in this field where $X(t_1, t_2)$ is nonzero. All the estimations, that is, bandwidths, autocorrelation function, and spectral density function, were made on domain S of nonzero values, and the mean ($\bar{R}$) estimated over this domain was subtracted from the actual intensities before the analysis, that is, the field analyzed was $Y(t_1, t_2) = R(t_1, t_2) - \bar{R}$. In order to estimate the bandwidth, the spectrum was obtained by squaring the values obtained from the fast Fourier transform of $Y(t_1, t_2)$ (without any filtering). Since $Y(\cdot, \cdot)$ is a zero mean process, the covariance function was estimated as

$$\hat{B}(\tau_1, \tau_2) = \frac{1}{|S|} \sum_{t_1=1}^{N-|\tau_1|} \sum_{t_2=1}^{N-|\tau_2|} Y(t_1, t_2) Y(t_1 + \tau_1, t_2 + \tau_2)$$

$$\tau_1, \tau_2 = 0, \pm 1, \cdots, \pm(N-1) \tag{47}$$

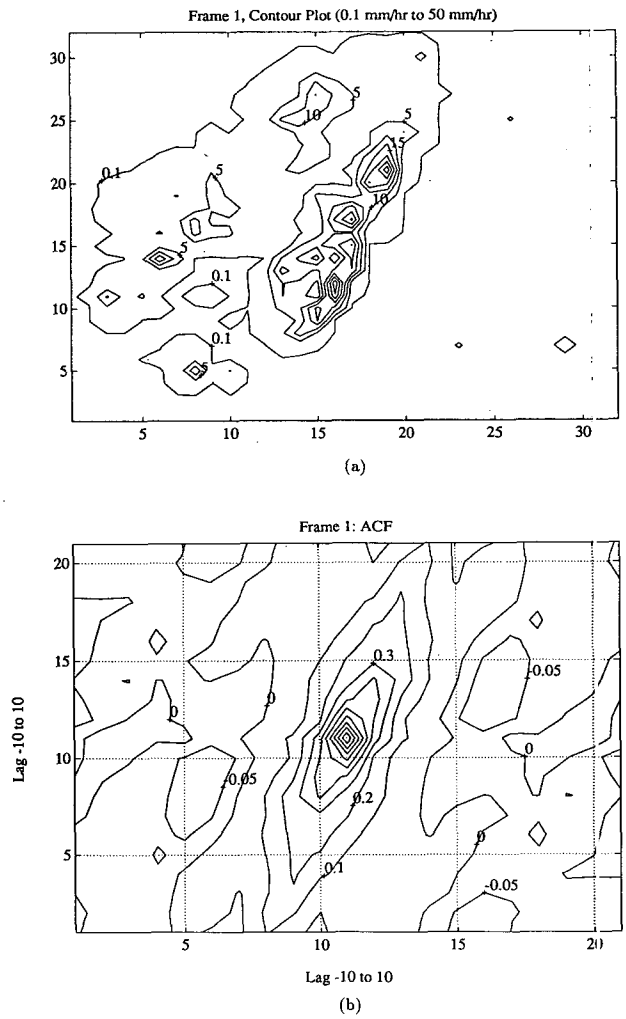

FIG. 5. (a) Averaged field (frame 1) at low resolution (grid 32 × 32; areal coverage 460 km × 460 km) and (b) its autocorrelation function. The contour plots in the rainfall field are at intensities between 0.1 and 50 mm h^{-1}. A unit lag shown in the figure for autocorrelation function corresponds to 14.375 km.

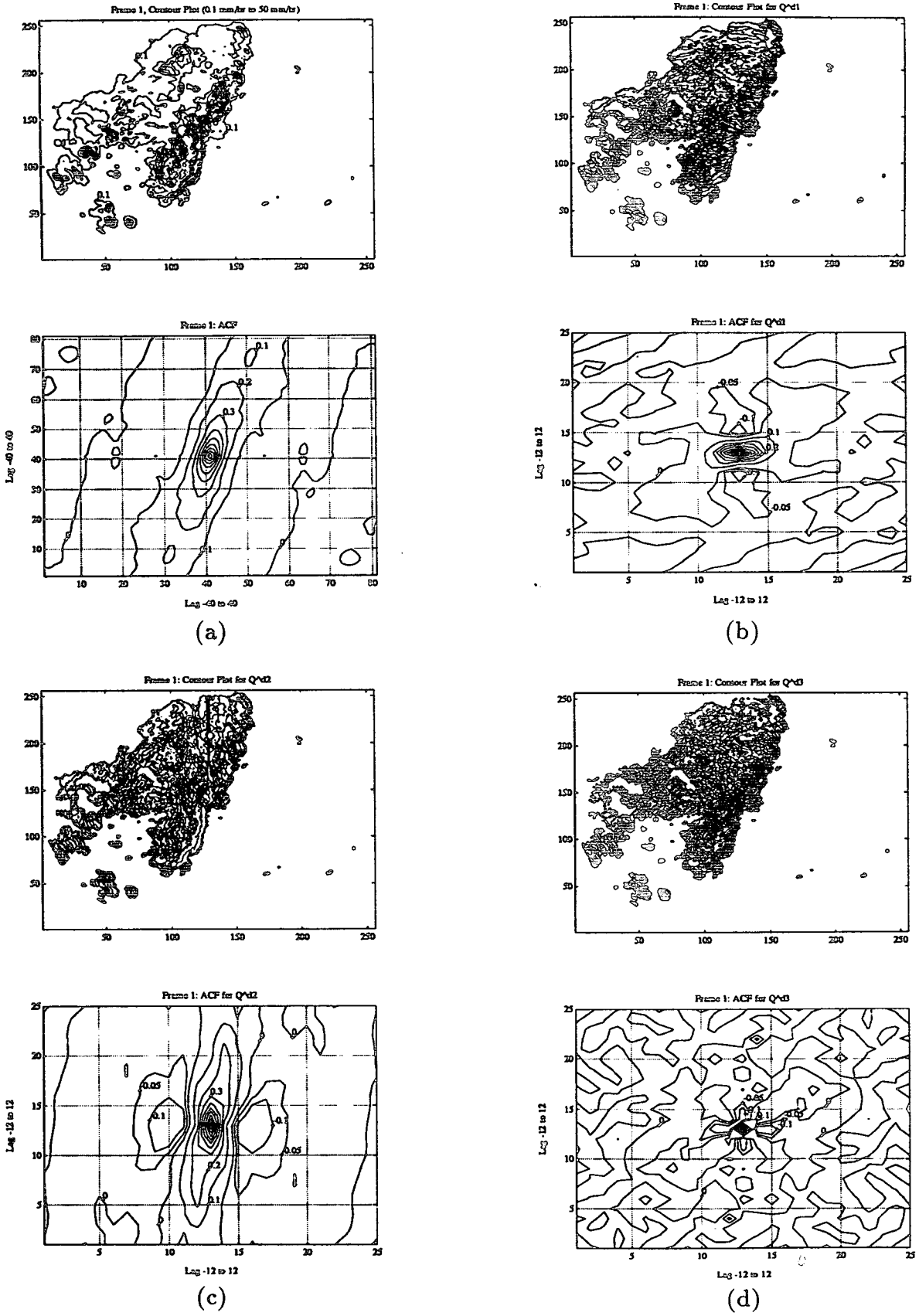

FIG. 6. The four components of the original field at grid 256 × 256 [(a) is $P_m^d X_0$, (b) is $Q_m^{d_1} X_0$, (c) is $Q_m^{d_2} X_0$, and (d) is $Q_m^{d_3} X_0$] and their respective autocorrelation functions. The contours for the averaged process are between 0.1 and 50 mm h^{-1} and for fluctuations between −1.0 and 1.0 mm h^{-1}. A unit lag shown in the figure for the autocorrelation functions corresponds to 1.8 km.

TABLE 1. Second-order statistics of the four components for the squall-line storm.

Component	Grid	$\delta^{(1)}$	$\delta^{(2)}$	Mean (mm h^{-1})	Standard deviation (mm h^{-1})
A	256 × 256	0.84	0.74	4.23	8.05
	128 × 128	0.81	0.69	4.05	7.60
	64 × 64	0.76	0.65	3.81	6.69
	32 × 32	0.74	0.62	3.42	5.43
	16 × 16	0.71	0.58	3.00	4.46
D_1	256 × 256	0.50	0.62	0.0	0.88
	128 × 128	0.45	0.56	0.0	1.25
	64 × 64	0.45	0.54	0.0	1.58
	32 × 32	0.47	0.53	0.0	1.46
	16 × 16	0.53	0.52	0.0	1.44
D_2	256 × 256	0.72	0.49	0.0	1.24
	128 × 128	0.69	0.41	0.0	1.79
	64 × 64	0.69	0.40	0.0	2.66
	32 × 32	0.62	0.46	0.0	2.91
	16 × 16	0.54	0.47	0.0	1.84
D_3	256 × 256	0.41	0.39	0.0	0.34
	128 × 128	0.41	0.41	0.0	0.59
	64 × 64	0.48	0.44	0.0	0.91
	32 × 32	0.50	0.47	0.0	1.24
	16 × 16	0.53	0.56	0.0	1.41

for the $N \times N$ grid. Although this gives a biased estimate, it has minimum mean-square error (see Priestley 1981, section 5.3.3) and is positive semidefinite. A tricky situation arises in the estimation of the moments of the fields generated by the wavelet coefficients due to the intermittency of the original rainfall field. In our analysis, moments were computed using only the nonzero values of the field under consideration. Therefore, these statistics should be looked at as approximate statistics. The estimation of the covariance function and the spectral density function does not pose any serious problem since the contribution of zero values is zero.

Some noteworthy observations of this analysis are as follows.

1) The standard deviation of the storm is much larger than the mean, indicating large variability in the field.

2) The autocorrelation functions of the storm at all resolutions have elliptic patterns (except at very small lags where they are circular) and are well behaved, indicating the direction, organization, and structure of the rainfall field.

3) For the storm, the autocorrelation function decays very sharply initially, that is, for small lags (distances), but decays very slowly for large lags, suggesting two possible regimes of correlation characteristics—the second indicating long-range dependence.

4) The bandwidth values of the marginal sdf's $G(\omega_i)$ indicate that the rainfall is a wide bandwidth process, again indicating long-range spatial dependence.

5) Another observation of great significance is that D_1, D_2, and D_3 are zero-mean processes with standard deviations much smaller than the mean of the original process, suggesting the presence of small (homogeneous) fluctuations "superimposed" on an underlying process. This superposition is not additive because of the interscale dependence. When this observation is combined with observation 1, it indicates that the (nonstationary) mean process itself exhibits large variability. Further research is needed to establish the exact mathematical relationships for this type of interscale dependence.

6) The detail process D_2 ($Q_m^{d^2}X$) exhibits long-range correlation characteristics (and could be modeled as fractional Gaussian noise) and the other two components D_2 and D_3 seem uncorrelated (white noise). These results have significant implications. They lead to the interpretation that if the rainfall has high correlation in a particular direction, the fluctuation field is also highly correlated in that direction and exhibits long-range dependence. For example, the storm analyzed shows high correlation in the vertical direction, and this is reflected in the correlation characteristics of component $Q_m^{d^2}X$ since this component captures horizontal high frequencies, that is, high vertical correlation.

7) It is seen that the components of the wavelet transform of the rainfall field exhibit white-noise characteristics across scales. It is possible to show that such a behavior is indicative of the presence of $1/|\omega|^{\alpha}$, $0 < \alpha < 2$, spectrum in the original process for certain frequency range $\omega_1 < \omega < \omega_2$ (see Wornell 1990). The implication of such a behavior is being explored further (see Kuman and Foufoula-Georgiou 1992a,b).

The above results are from one frame of one storm only, but the results are similar for other frames and are indicative of the power of the proposed methodology to unravel structure in ways never tried before for rainfall fields. Many more data analyses (i.e., for different storms and different frames at various times of evolution of each storm) are needed before conclusions regarding the scaling structure of rainfall can be made. It is clearly evident, however, that the decomposition of the given rainfall field into component processes obtained as the wavelet coefficients offers significant advancement in data-analysis techniques for studying scale dependence. It is also remarked that many more elaborate analyses, such as interscale dependence of the component processes and the dependence of the component processes on the original process (through cross-covariance analysis), can be performed to study scaling characteristics.

6. Concluding remarks

We have shown that the orthogonal wavelet transforms of the rainfall fields provide insight that is otherwise not possible by other methods. They segregate

large-scale features from small-scale features by providing convenient orthogonal decompositions with an inbuilt directionality. The averaged process at very coarse resolutions is an indicator of the large-scale features of the storm and can be studied independently of the small-scale features. We find that rainfall fluctuations, reflected as small-scale features, may exhibit simple scaling characteristics, that is, power-law behavior of the spectral density function, within a certain range of frequencies and that these characteristics are direction dependent, which is a manifestation of the underlying organization and structure of the storm. This observation also corroborates well with the empirical covariance function of the rainfall field itself (treated as homogeneous), where two possible regimes are indicated. The large-scale features that are manifestations of the underlying mechanism particular to the storm will be reflected in the spectral density function for frequencies close to the origin and may not obey power law. The small-scale fluctuations may show long-range dependence and will be reflected in the spectral density function away from the origin that may obey power law. Such a behavior is indicated by the particular storm analyzed. More data analysis is required before these ideas can be firmly established.

Acknowledgments. This research was supported by National Science Foundation Grants BSC-8957469 and EAR-9117866 and by NASA under a graduate student fellowship for Global Change Research. We also thank the Minnesota Supercomputer Institute for providing us with computer time on the Cray-2 supercomputer.

APPENDIX

Second-Order Analysis of Random Fields

In this appendix, only definitions are given for the second-order properties of homogeneous random fields used in our analysis. The reader is referred to Vanmarcke (1984) for more details.

If $S(\omega)$, $\omega \in \mathbf{R}$ denotes the spectral density function (sdf) of the homogeneous random function $X(t)$, then the one-sided sdf $G(\omega)$ for $\omega \geq 0$ is

$$G(\omega) = 2S(\omega), \quad 0 \leq \omega < \infty, \tag{A1}$$

and the kth spectral moment of the stationary random process $X(t)$ is defined s

$$\lambda_k = \int_{-\infty}^{\infty} |\omega|^k S(\omega) d\omega = \int_0^{\infty} \omega^k G(\omega) d\omega, \quad k = 0, 1, 2, \cdots. \tag{A2}$$

Note that

$$\lambda_0 = \int_{-\infty}^{\infty} S(\omega) d\omega = \sigma^2. \tag{A3}$$

The *characteristic frequency* can be defined as

$$\Omega_k = \left(\frac{\lambda_k}{\lambda_0}\right)^{1/k}, \quad k = 1, 2, 3, \cdots. \tag{A4}$$

These frequencies are such that $\Omega_k \leq \Omega_{k+1}$, ($k = 1, 2, \cdots$). Notice that Ω_1 gives the mean frequency of the normalized sdf $G(\omega)$, and Ω_2 is its root-mean-square frequency. Therefore,

$$\Omega_s = (\Omega_2^2 - \Omega_1^2)^{1/2} \tag{A5}$$

gives the frequency standard deviation of the normalized sdf $g(\omega)$. The bandwidth of $G(\omega)$ [or $g(\omega)$] is defined as

$$\delta = \left[\left(1 - \frac{\lambda_1^2}{\lambda_0 \lambda_2}\right)\right]^{1/2} = \left[\left(1 - \frac{\Omega_1^2}{\Omega_2^2}\right)\right]^{1/2} = \frac{\Omega_s}{\Omega_2} \qquad 0 \leq \delta \leq 1. \tag{A6}$$

In the two-dimensional case, the spectral moments can similarly be defined as

$$\lambda_{kl} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \omega_1^k \omega_2^l S(\omega_1, \omega_2) d\omega_1 d\omega_2 \tag{A7}$$

for $k + l$ even; $k, l = 0, 1, 2, \cdots$. If $k + l$ is odd, the fact that $S(\omega_1, \omega_2)$ is symmetrical about $\omega_1 = \omega_2 = 0$ ensures that the double integral vanishes. The kth marginal spectral moment with respect to ω_1 is given by

$$\begin{aligned} \lambda_k^{(1)} \equiv \lambda_{k0} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\omega_1|^k S(\omega_1, \omega_2) d\omega_1 d\omega_2 \\ &= \int_{-\infty}^{\infty} |\omega_1|^k S(\omega_1) d\omega_1 \\ &= \int_0^{\infty} \omega_1^k G(\omega_1) d\omega_1 \quad k = 0, 1, 2, \cdots. \end{aligned} \tag{A8}$$

In the above equation, $S(\omega_1)$ and $G(\omega_1)$ represents the respective marginal sdf, that is, $S(\omega_1) = \int_{-\infty}^{\infty} S(\omega_1, \omega_2) d\omega_2$ and $G(\omega_1) = \int_0^{\infty} G(\omega_1, \omega_2) d\omega_2$. Similarly,

$$\begin{aligned} \lambda_l^{(2)} \equiv \lambda_{0l} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\omega_2|^l S(\omega_1, \omega_2) d\omega_1 d\omega_2 \\ &= \int_{-\infty}^{\infty} |\omega_2|^l S(\omega_2) d\omega_2 \\ &= \int_0^{\infty} \omega_2^l G(\omega_2) d\omega_2 \quad l = 0, 1, 2, \cdots, \end{aligned} \tag{A9}$$

with analogous definitions for $S(\omega_2)$ and $G(\omega_2)$. We therefore have

$$\lambda_{00} = \lambda_0^{(1)} = \lambda_0^{(2)} = \sigma^2. \tag{A10}$$

The bandwidths $\delta^{(i)}$, $i = 1, 2$ of the marginal sdf $G(\omega_i)$ are given by

$$\delta^{(1)} = \left(1 - \frac{\lambda_{11}}{\lambda_{00}\lambda_{20}}\right)^{1/2} \tag{A11}$$

and

$$\delta^{(2)} = \left(1 - \frac{\lambda_{11}}{\lambda_{00}\lambda_{02}}\right)^{1/2}. \tag{A12}$$

REFERENCES

Barton, R. J., and H. V. Poor, 1988: Signal detection in fractional Gaussian noise. *IEEE Trans. Inf. Theor,* **34,** 5, 943–959.

Cohen, A., and I. Daubechies, 1991: *Non-Separable Bidimensional Wavelet Bases.* AT&T Bell Laboratories, 70 pp.

Combes, J. M., A. Grossman, and P. Tchamitchian, Eds., 1987: Wavelets: time-frequency methods and phase-space. *Proc. Int. Conf.,* Marseille, Springer-Verlag, 314 pp.

Daubechies, I., 1988: Orthonormal bases of compactly supported wavelets. *Commun. Pure Appl. Math.,* **41,** 901–996.

Everson, R., L. Sirovich, and K. R. Sreenivasan, 1990: Wavelet analysis of the turbulent jet. *Phys. Lett. A,* **145,** 6, 314–322.

Flandrin, P., 1989: On the spectrum of fractional Brownian motion. *IEEE Trans. Inf. Theory,* **35,** 1, 197–199.

Grossmann, A., and J. Morlet, 1984: Decompostion of hardy functions into square integrable wavelets of constant shape. *SIAM J. Math. Anal.,* **15,** 4, 723–736.

Gupta, V., and E. Waymire, 1990: Multiscaling properties of spatial rainfall and river flow distributions. *J. Geophys. Res.,* **95,** D3, 1999–2009.

Holschneider, M., 1988: On the wavelet transforms of fractal objects. *J. Stat. Phys.,* **50,** 5/6, 963–992.

Kedem, B., and L. S. Chiu, 1987: Are rain rate processes self-similar? *Water Resour. Res.,* **23,** 10, 1816–1818.

Kumar, P., and E. Foufoula-Georgiou, 1992a: A multicomponent decomposition of spatial rainfall fields: 1. Segregation of large and small scale features using wavelet transforms. *Water Resources Res.,* submitted.

——, and ——, 1992b: A multicomponent decomposition of spatial rainfall fields. 2. Self-similarity in fluctuations. *Water Resources Res.,* submitted.

Lamperti, J., 1962: Semi-stable stochastic processes. *Trans. Amer. Math. Soc.,* **104,** 62–78.

Liandrat, J., and F. Moret-Bailly, 1990: The wavelet transform: Some applications to fluid dynamics and turbulence. *Eur. J. Mech., B/Fluids,* **9,** 1, 1–19.

Lovejoy, S., and D. Schertzer, 1985: Generalized scale invariance in the atmosphere and fractal models of rain. *Water Resour. Res.,* **21,** 8, 1233–1250.

——, and ——, 1989: Comment on "Are rain rate processes self-similar?" *Water Resour. Res.,* **25,** 13, 577–579.

——, and ——, 1990: Multifractals, universality classes and satellite and radar measurements of cloud and rain fields. *J. Geophys. Res.,* **95,** D3, 2021–2034.

Mallat, S., 1989a: A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intel.,* **11,** 7, 674–693.

——, 1989b: Multiresolution approximations and wavelet orthonormal bases of $L^2(\mathbf{R})$. *Trans. Amer. Math. Soc.,* **315,** 1, 69–87.

——, 1980c: Multifrequency channel decomposition of images and wavelet models. *IEEE Trans. Acoustics, Speech and Signal Anal.,* **37,** 12, 2091–2110.

Mandelbrot, B., and J. W. Van. Ness, 1968: Fractional brownian motions, fractional noises and applications. *SIAM Rev.* **10,** 4, 422–437.

Meyer, Y., 1988: *Ondelettes et Operateurs.* Hermann.

O'Bannon, T., and P. Ahnert, 1986: A study of the NEXRAD precipitation processing system on a winter-type Oklahoma rainstorm. *Proc. 23d Conf. on Radar Meteor. and the Conf. on Cloud Phys.,* Snowmass, CO, Amer. Meteor. Soc., JP99–JP102.

Ossiander, M., and E. Waymire, 1989: Certain positive-definite kernels. *Proc. Amer. Math. Soc.,* **107,** 2, 487–492.

Perrier, V., 1989: Towards a method for solving partial differential equations using wavelet bases. *Wavelets: Time-Frequency Methods and Phase-Space,* J. M. Combes, A. Grossman, and P. Tchamitchian, Eds., 269–283.

Priestley, M. B., 1981: *Spectral Analysis and Time Series Analysis.* Academic Press, 1890 pp.

Risset, J. C., 1989: The computer, music, and sound models. 102–123.

Schertzer, D., and S. Lovejoy, 1987: Physical modeling of rain and clouds by anisotropic scaling multiplicative processes. *J. Geophys. Res.,* **92,** D8, 9693–9714.

Vanmarcke, E., 1984: *Random Fields: Analysis and Synthesis.* The MIT Press, 382 pp.

Waymire, E., 1985: Scaling limits and self-similarity in precipitation fields. *Water Resour. Res.,* **21,** 8, 1271–1281.

Wornell, G. W., 1990: A Karhunen–Loéve-like expansion for $1/f$ processes via wavelets. *IEEE Trans. Inf. Theory,* **36,** 4, 859–861.

Yaglom, A. M., 1987: *Correlation Theory of Stationary and Related Random Functions, I: Basic Results.* Springer-Verlag, 526 pp.