WATER RESOURCES RESEARCH, VOL. 29, NO. 8, PAGES 2515–2532, AUGUST 1993

A Multicomponent Decomposition of Spatial Rainfall Fields 1. Segregation of Large- and Small-Scale Features Using Wavelet Transforms

PRAVEEN KUMAR[1] AND EFI FOUFOULA-GEORGIOU

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

Issues of scaling characteristics in spatial rainfall have attracted increasing attention over the last decade. Several methods based on simple and multiscaling and multifractal ideas have been proposed and parameter estimation techniques developed for the hypothesized models. Simulations based on these models have realistic resemblance to "generic rainfall fields." In this research we analyze rainfall data for scaling characteristics without an a priori assumed model. We look at the behavior of rainfall fluctuations obtained at several scales, via orthogonal wavelet transform of the data, to infer the precise nature of scaling exhibited by spatial rainfall. The essential idea behind the analysis is to segregate large-scale (long wavelength) features from small-scale features and study each of them independently. The hypothesis is set forward that rainfall might exhibit scaling in small-scale fluctuations, if at all, and at large scale this behavior will break down to accommodate the effects of external factors affecting the particular rain-producing mechanism. The validity of this hypothesis is examined. In the first of these papers we develop the methodology for the segregation of large- and small-scale features and apply it to a severe spring time midlatitude squall line storm. The second paper (Kumar and Foufoula-Georgiou, this issue) develops a framework for testing the presence and studying the nature of self-similarity in the fluctuations.

1. INTRODUCTION

A characteristic feature of precipitation is its extreme variability over time scales of minutes to years and over space scales of a few to thousands of square kilometers. One of the major challenges of hydrologists, meteorologists, and climatologists is to measure, model, and predict the nature of this variability over different scales. Recent research (e.g., *Lovejoy and Schertzer* [1990], *Gupta and Waymire* [1990], and references therein) has indicated the exciting possibility that rainfall may exhibit scaling-multiscaling characteristics. The presence of such a hidden structure in the highly irregular patterns of rainfall at different spatial scales promises improved understanding of the precipitation process and new approaches to efficient modeling, measurement, and prediction.

Early on, empirical study of contours of rain intensities [*Lovejoy*, 1982] and probability distribution functions of rain rates [*Lovejoy and Mandelbrot*, 1985] suggested scaling in rainfall. Later, it was argued by *Kedem and Chiu* [1987] that, since rainfall is an intermittent positive process giving rise to a mixed distribution with an "atom at zero," it could not be self-similar or simple scaling at least to the extent that a single parameter would not be sufficient to characterize this process. *Lovejoy and Schertzer* [1989] argued that although the rainfall intensities suffer this limitation, the rainfall fluctuations do not and they could be modeled as a self-similar process. They, however, did not empirically study the rainfall fluctuations for scaling characteristics but rather, using an analogy from turbulence which is also an intermittent process, concentrated on further developing their multifractal models of rain [*Schertzer and Lovejoy*, 1987]. *Gupta and Waymire* [1990] studied the moments of marginal distribution function of rainfall process conditioned on being positive and reported deviations from simple scaling. They proceeded with the development of a multiscaling theory of rain.

In this research the hypothesis is set forward that rainfall can be decomposed in a large-scale component representing the mean behavior of the process, and small-scale fluctuations which exhibit self-similarity. The motivation for our hypothesis is based on physical arguments and empirical evidence. It has been observed that the spectrum of spatial rainfall deviates from power law behavior at very low frequencies (see Figure 1), indicating that it is the fluctuation process (deviations from a large-scale mean component) that may exhibit self-similarity, if at all. Notice here that we do not refer to a possible break in the spectrum caused from the lack of data at large scales but rather to a break due to the physical nature of the process which itself has limited extent. The very low frequencies in the spatial rain-intensity spectrum represent the morphological organization due to the large-scale forcing specific to that rain producing mechanism (for example, effects of a front on a squall line). When this effect is subtracted, the deviations which result from the microscopic effects may obey some universality condition like self-similarity. It might also be that we can attribute low-frequency components to a deterministic process that should be eliminated before any stochastic consideration is taken into account. In all of the previous research, such a consideration of segregating large-scale from small-scale behavior, has not been taken into account for the purpose of analysis and inference of rainfall process. The reason for that seems to be that although theoretical attempts have been made to study rainfall fluctuations [e.g., *Waymire*, 1985] there seems to be no consistent method of obtaining fluctuations in two or more dimensions. Our research is an effort

[1]Now at Hydrologic Sciences Branch, Laboratory for Hydrospheric Processes, NASA Goddard Space Flight Center, Greenbelt, Maryland.

Paper number 93WR00548.
0043-1397/93/93WR-00548$05.00

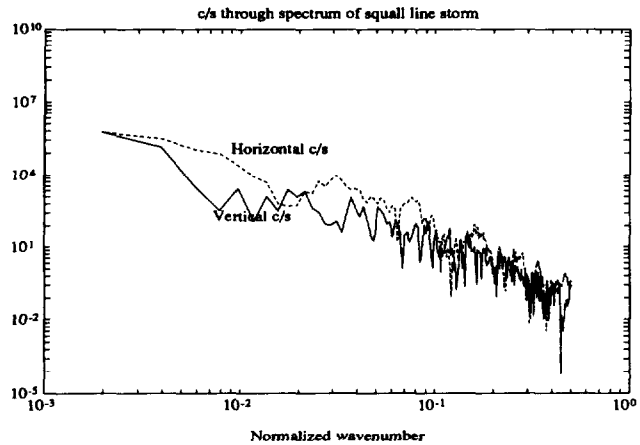

Fig. 1. Horizontal and vertical cross sections through the two-dimensional spectrum of the temporally integrated (10 min) spatial rainfall field shown in Figure 10.

toward this direction and develops a theoretical framework for segregating rainfall in large- and small-scale features and studying small-scale features (fluctuations) for self-similarity.

In addition to segregation of features of different scales, other distinct advantages of our proposed method of analysis are as follows:

1. It can incorporate anisotropy and inhomogeneity. For example, the analysis, interpretation and modeling under the framework of multifractals is based on the assumption of homogeneity and isotropy. *Schertzer and Lovejoy* [1987] developed a methodology to incorporate anisotropy to some extent but the problem of inhomogeneity has not been addressed successfully.

2. It offers a natural way of coupling the hierarchical preferred organization known to exist in rainfall fields with self-similar features of rainfall fluctuations. Prior to evidence of scaling in rainfall, attempts were made to model it using cluster point process models [see *Waymire et al.*, 1984] designed to preserve the second-order properties and mimic the hierarchical imbedding and clustering of high-intensity areas within lower-intensity areas identified as large and small mesoscale areas, and cells [e.g., *Orlanski*, 1975]. The preferred organization, clustering, and hierarchical imbedding of rainfall patterns at a few discrete scales seems, at least at first, to be inconsistent with the current scale invariant modeling frameworks mainly for two reasons. Since, scale invariance implies absence of any preferred scales we can either have clustering and imbedding at all scales or none at all. Second, most models are based on the assumption of homogeneity and isotropy, and consequently they may not be able to account for any preferential organization that rainfall may exhibit. No progress has been made so far to unify these two approaches, although research is currently underway to identify conditions under which multiplicative cascade processes could give rise to desired clustering structures [*Gupta and Waymire*, 1993]. We offer an alternate approach to this problem.

3. It offers a way of dealing with processes which do not possess point values. Rainfall is not a continuous process in either space or time. If one recalls the falling of raindrops on a car's windshield, one would recognize that at any instant of time not the entire windshield is covered with rain, but over a period of time it does get wet everywhere. We distinguish such a process by saying that it does not have point values; i.e., values of the process at every point are not defined, although the integral of the process at certain spatial and temporal scales is still meaningful. Thus both spatial and temporal scales of integration are important for the characterization of the process. However, even for large scales of integration, areas of no rainfall amidst rainy areas continue to exist. We call such a process intermittent. Intermittency gives rise to its own distinctive characteristics. For example, the mean and the standard deviations of the process conditioned on being nonzero change with the scale of description making "scale" an essential parameter in the description of the process; i.e., we have to specify the scale of analysis whenever discussing these properties. In addition to intermittency, the absence of point values makes certain operations like derivatives meaningless for the rainfall field. One has to rely on equivalent integral transforms, so as to have a certain amount of "smoothing," to accomplish the objective.

4. It is ideal for studying the evolutionary behavior of the storm. Our main interest is in analyzing and eventually modeling particular storms, e.g., a squall line or a winter storm, and not a "generic storm" derived from "averaging" over many realizations often obtained by invoking the assumption of stationarity over time and homogeneity over space. Thus being able to account for anisotropy and nonhomogeneity and identify how they characterize the features of the rainfall process under study as it evolves in time is an essential element of our analysis.

This paper deals with developing the methodology and relevant mathematical results for the segregation of large- and small-scale features. It is structured as follows. Section 2 outlines the proposed methodology and broad framework of analysis and provides at an intuitive level the motivation and justification for choosing wavelet transforms for the decomposition of large- and small-scale features. Section 3 presents a brief review of wavelet transforms focusing only on the mathematical results essential for our analysis. In section 4, the theoretical framework of analyzing rainfall fields through wavelet transforms is presented, and some new results related to wavelet transforms of stochastic processes are discussed. Section 5 illustrates the results of the wavelet transform applied to actual rainfall fields for the segregation of large- and small-scale features. The companion paper [*Kumar and Foufoula-Georgiou*, this issue] develops a framework for testing the presence and studying the nature of scaling exhibited by rainfall fluctuations.

2. Methodology

In this research we assume that the observed rain rate process $\mathcal{R}(\mathbf{t})$, described on the two-dimensional space $\mathbf{t} = (t_1, t_2) \in \mathbf{R}^2$, is a composition of two independent processes $X(\mathbf{t})$ and $I(\mathbf{t})$, and is given as $\mathcal{R}(\mathbf{t}) = X(\mathbf{t})I(\mathbf{t})$, where $I(\mathbf{t})$ is a 0–1 valued intermittency random field and $X(\mathbf{t})$ is a positive "inner variability" field (this terminology is due to *Barancourt et al.* [1992]). The essential idea of our framework of analysis of rainfall for identifying self-similarity is to break up the nonhomogeneous inner variability field $X(\mathbf{t})$ as

$$X(\mathbf{t}) = \bar{X}(\mathbf{t}) + X'(\mathbf{t}) \tag{1}$$

where $\bar{X}(\mathbf{t})$ is some large (mean or trend) process and $X'(\mathbf{t})$ is a fluctuation process such that $\bar{X}(\mathbf{t})$ and $X'(\mathbf{t})$ are uncorre-

lated. In the above decomposition the statistics of $X(\mathbf{t})$ are broken up in a very tractable manner. For example, the mean and the variances (expectations taken over ensembles) are additive, i.e., $m(\mathbf{t}) = \bar{m}(\mathbf{t}) + m'(\mathbf{t})$ and $\sigma^2(\mathbf{t}) = \bar{\sigma}^2(\mathbf{t}) + \sigma'^2(\mathbf{t})$, where $\bar{m}(\mathbf{t})$ and $\bar{\sigma}^2(\mathbf{t})$ are mean and variances of $\bar{X}(\mathbf{t})$ and correspondingly for $m'(\mathbf{t})$ and $\sigma'^2(\mathbf{t})$. The "correlation" function (or "noncentral" covariance function) of $X(\mathbf{t})$ can be written in terms of the correlation function of the component processes as

$$R(\mathbf{t}, \mathbf{s}) = \bar{R}(\mathbf{t}, \mathbf{s}) + R'(\mathbf{t}, \mathbf{s}) \qquad \mathbf{t}, \mathbf{s} \in \mathbf{R}^2 \qquad (2)$$

where $R(\mathbf{t}, \mathbf{s}) = E[X(\mathbf{t})X(\mathbf{s})]$, $\bar{R}(\mathbf{t}, \mathbf{s}) = E[\bar{X}(\mathbf{t})\bar{X}(\mathbf{s})]$, and $R'(\mathbf{t}, \mathbf{s}) = E[X'(\mathbf{t})X'(\mathbf{s})]$.

The decomposition of the spatial rainfall into mean and fluctuations is achieved using wavelet transforms. Wavelet decompositions, like Fourier decompositions, are series expansions of a function using orthonormal bases. They possess the following properties (among others) that are attractive for our research: (1) scale-space or multiresolution transformation of processes for multiscale study to identify and extract properties of self-similarity, (2) multirate filtering using quadrature mirror filter pairs for segregation of large- and small-scale features without a priori information on the size of these features, (3) "time-frequency" localization for the study of nonhomogeneous/nonstationary processes through localized fluctuations, and (4) spatially oriented frequency channels to extract anisotropic behavior.

"Scale-space" transformation, i.e., transformation of a given process at some scale to obtain a family of processes indexed by a scale parameter, is desired if need exists for the study, representation, and segregation of features of different characteristic scales. For example, in image processing large-scale features set the context of the image like house or street, while the small-scale features give the details like the window or car, etc. An analogous situation also exists for physical processes like rainfall. The large-scale morphological organization is governed by factors such as topography, wind direction and speed, and thermodynamic conditions like temperature gradient, etc. However, the small-scale features are less likely to be dependent on such effects. The scale-space transformation is achieved via recursively smoothing the process with a smoothing function, called scale function, of varying scale parameter. Thus large-scale features, i.e., features that continue to persist for large changes of scale, will be seen with increasing scale but information about the details or small-scale features is lost. When no a priori information about the size of the features of interest is available, it is difficult to decide on the right scale of analysis and one needs to look at the process at several scales. The representation of the process at any scale (or resolution) is most efficiently obtained by projection onto the space defined by that scale (or resolution). In addition, *Koenderink* [1984] argued that in order to conform to the intuitive notion of no preferred scale analysis, one needs to sample the scale parameter uniformly on the logarithmic scale. The multiresolution framework discussed here encompasses both these requirements. The representation at different scales is obtained by orthogonal projections, and by construction, the scale parameter is uniformly sampled on the logarithmic scale.

When we smooth a process we loose structure or detail. This is because details or small-scale features have a limited range of resolution in which they can be identified. If a need exists to analyze the details that are lost during smoothing, one has to find a way to preserve this information. This information is essential in the study of self-similar processes because a wide range of their spectrum contains information that is significant to the process. The multiresolution framework discussed here is again suited for this purpose. This is achieved by constructing a band pass filter, called wavelet, complementary to the smoothing filter (scale function), i.e., it grasps all the information lost during smoothing (the filters in the pair are called quadrature mirror filters [see *Vaidyanathan*, 1987]). Like the smoothing filters this is implemented recursively at different scales, and this multirate filtering using a quadrature mirror filter pair enables one to segregate large- and small-scale features at any scale.

The wavelets, under the multiresolution framework, are orthogonal to all their translates and dyadic dilates (i.e., dilates by powers of two) thereby allowing us to cover the entire domain at several different scales. In this decomposition, the coefficients used in the expansion, i.e., the inner product between the wavelets and the function under study, are called wavelet coefficients and have a very specific meaning. As will be shown later, they represent the discretization (or sampling) of the fluctuation process at the given scale. Scale is an explicit parameter in these decompositions allowing us to (1) make direct influence about the behavior of the process at different scales and (2) isolate features of different characteristic sizes. Hence wavelet multiresolution decomposition allows us to simultaneously address the problem of segregating large-scale behavior from small-scale behavior and overcome the problem that rainfall does not have point values by defining all operations through integral transforms.

Another useful property of wavelets is that of time-frequency localization. So far, Fourier transforms have been used to study processes and their fluctuations at different scales. However, in Fourier decomposition, the field is assumed, a priori, to be homogeneous. Also, we would like to go beyond the restriction of "global" description of the process as is the case in Fourier decompositions. Since rainfall exhibits "violent" behavior (i.e., the probability distribution of the rain intensities is thick tailed and consequently very large values can be obtained with high probability), we would like a definition of fluctuation that is "local," i.e., captures the violent behavior locally as large value of the fluctuation, rather than being an average of the behavior over the entire domain. Some windowing techniques [see *Gabor*, 1946] have been marginally successful in overcoming the limitations of Fourier transform. However, wavelets achieved a dramatic success in this respect due to two characteristics fundamentally different from Fourier transform: (1) compact support of basis functions and (2) basis functions that are obtained through dilations and modulations of a basic function. Whereas the basis functions sin nt and cos nt in the Fourier transforms have infinite support, the wavelets have a compact support; i.e., they are zero everywhere outside the domain of finite size. This enables the localization in time or space. Also, the basis functions in Fourier analysis are constructed by the modulation of a single function, i.e., sin t or cos t, whereas the wavelet basis are dilates and translates of a "mother wavelet" which (as we will see in the following section) enable localization in frequency such that the size of the support is proportional to the "size of the feature" it represents. We have small support for high-frequency features and large

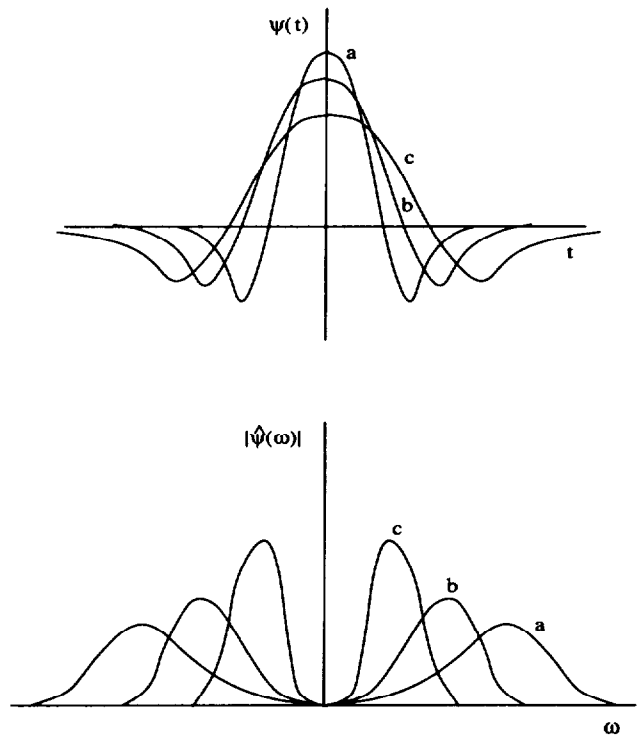

Fig. 2. The effect of dilation on a "generic" wavelet and the corresponding change on its Fourier transform $|\hat{\psi}(\omega)|$. When the wavelet dilates, its Fourier transform contracts and vice versa.

support for low frequency or large wavelength features. This property enables one to zoom into the irregularities of a function and characterize them locally. Furthermore, using wavelets, fluctuations at different scales can also be obtained due to the multiscale transform properties.

In addition to the above properties, two-dimensional wavelet transforms enable the decomposition of a process into spatially oriented frequency channels, i.e., features with dominant frequencies in different directions are extracted as separate components. This property is exploited to study the anisotropic behavior of rainfall fluctuations. As is shown in this paper, using wavelets we obtain further decomposition of the fluctuation field into three components as

$$X'(\mathbf{t}) = X'_1(\mathbf{t}) + X'_2(\mathbf{t}) + X'_3(\mathbf{t}). \tag{3}$$

The three components X'_1, X'_2, and X'_3 represent three different aspects of the process, namely, the vertical, horizontal, and diagonal high correlations. Therefore the fields X'_1, X'_2, and X'_3 can be seen as capturing the anisotropic behavior of the process. Each of the fields X'_1, X'_2, and X'_3 are studied for self-similarity in a manner made precise in paper 2 [see *Kumar and Foufoula-Georgiou*, this issue]. Here it is sufficient to remark that the condition of positivity and atom at zero discussed by *Kedem and Chiu* [1987] no longer applies for the field $X'(\mathbf{t})$ and, therefore it can be self-similar. As we will see later, the fluctuations of the rainfall obtained in this manner are responsible for a significant contribution to the total variance (up to 55%).

3. Review of Wavelet Transforms

In this section we briefly review the theory of wavelet transforms in order to introduce the notation and provide a brief exposition so as to make the paper self-contained. A comprehensive treatment of this theory can be found in the work by *Daubechies* [1992] and references therein.

3.1. *Continuous Wavelet Transforms*

Let $L^2(\mathbf{R})$ denote the vector space of complex valued, square integrable functions $f(t)$, where t denotes a physical coordinate of space or time depending upon the context. 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 $(f * g)(t) = \int_{-\infty}^{\infty} f(u)g(t-u)\, du$. By $l^2(\mathbf{Z})$, we denote the vector space of square summable sequences, i.e., $l^2(\mathbf{Z}) = \{\varepsilon_i: \sum_{i=-\infty}^{\infty} |\varepsilon_i|^2 < \infty, i \in \mathbf{Z}, \varepsilon \in \mathbf{R}\}$, where $\mathbf{Z}$ is the set of integers.

3.1.1. *Definition.* The wavelet transform of a function $f(t)$ is defined as the integral transform

$$Wf(\lambda, u) = \int_{-\infty}^{\infty} f(t)\psi_\lambda(t-u)\, dt \tag{4}$$

where λ is a scale parameter, u is a location parameter, and the function $\psi_\lambda(t) \equiv \sqrt{\lambda}\psi(\lambda t)$ is called a wavelet. Changing the value of λ has the effect of dilating or contracting the function $\psi(t)$ (see Figure 2 (top)) and changing u has the effect of analyzing the function $f(t)$ around the point u. The choice of the wavelet $\psi(t)$ is neither unique nor arbitrary. The function $\psi(t)$ is chosen so that it has a compact support, or has a sufficiently fast decay, to obtain localization in space. Invertibility requirement of the wavelet transform dictates that $\psi(t)$ should at least have a zero mean, i.e., $\int_{-\infty}^{\infty} \psi(t)\, dt = 0$, although higher-order moments may also be zero; i.e.,

$$\int_{-\infty}^{\infty} t^k\psi(t)\, dt = 0 \qquad k = 0, \cdots, N-1. \tag{5}$$

The inverse relation is then given by

$$f(t) = \frac{1}{C_\psi}\int_{-\infty}^{\infty}\int_{0}^{\infty} Wf(\lambda, u)\psi_\lambda(t-u)\, d\lambda\, du \tag{6}$$

where

$$C_\psi = \int_0^{\infty} \frac{|\hat{\psi}(\omega)|^2}{\omega}\, d\omega < \infty. \tag{7}$$

A typical example of a wavelet is obtained from the second derivative of the Gaussian (also called the Mexican hat wavelet because of its shape) given by $\psi(t) = (2/\sqrt{3})\pi^{-1/4}(1-t^2)e^{-t^2/2}$. This wavelet has found application in edge detection [see *Mallat*, 1989*b*].

3.1.2. *Time-frequency localization.* The function $\psi_\lambda(t)$ can be interpreted as the impulse response of a band pass filter. Indeed, $Wf(\lambda, u)$ can be equivalently written as $(f * \bar{\psi}_\lambda)(u)$, where $\bar{\psi}_\lambda(t) = \psi_\lambda(-t)$. The wavelet transform $Wf(\lambda, u)$ can therefore be viewed as filtering of $f(t)$ with a band pass filter $\bar{\psi}_\lambda(t)$ (notice that $\bar{\psi}(0) = 0$ due to (5)). The Fourier transform of $\psi_\lambda(t)$ is given by

$$\hat{\psi}_\lambda(\omega) = \frac{1}{\sqrt{\lambda}} \hat{\psi}\left(\frac{\omega}{\lambda}\right) \tag{8}$$

Hence by dilating the function ($\lambda < 1$), we see that the center of passing band λw_0 of $\bar{\psi}_\lambda(\omega)$, where ω_0 is given as

$$\omega_0 = \frac{\int_0^\infty \omega |\hat{\psi}(\omega)|^2 \, d\omega}{\int_0^\infty |\hat{\psi}(\omega)|^2 \, d\omega} \tag{9}$$

decreases (see Figure 2 (bottom)). Consequently, longer wavelength features will become dominant in the corresponding wavelet transform. The root-mean-square bandwidth $\lambda\sigma_\omega$ of $\bar{\psi}_\lambda(\omega)$ around $\lambda\omega_0$ where σ_ω is given by

$$\sigma_\omega^2 = \int_0^\infty (\omega - \omega_0)^2 |\hat{\psi}(\omega)|^2 \, d\omega \tag{10}$$

also decreases. Hence by changing the value of λ we can extract the behavior of the function in different frequency bands. It can be shown [see *Mallat*, 1989*b*] that these frequency bands have a constant size on the logarithmic scale. When scale λ is small, the resolution is coarse in the spatial domain and fine in the frequency domain. If the scale λ increases, the resolution increases in the spatial domain and decreases in the frequency domain, i.e., when the wavelet dilates its Fourier transform contracts and vice versa. This enables localization in frequency. Thus time-frequency localization is achieved by the properties of compact support and dilation. When support is small we capture high-frequency components and vice versa and this information is localized in time and space.

3.2. *Discrete Wavelet Transforms and Multiresolution Representation*

3.2.1. *Multiresolution framework.* For the implementation of the wavelet transform on a sampled function $f(t)$, i.e., a sequence of numbers $\{c_n\}_{n\in\mathbf{Z}}$, the scale and location parameters can be discretized giving rise to discrete wavelet transforms. Note that the wavelet itself is continuous but the location and scale parameters are discrete. The scales can be selected as some sequence $\{\alpha^m\}_{m\in\mathbf{Z}}$, where α is an elementary dilation step. The location parameter is discretized so that its sampling rate is uniform and proportional to α^m, i.e., u is chosen as $k\beta/\alpha^m$, where α^m/β is the sampling rate of u at the scale α^m. The particular choices of $\alpha = 2$ and $\beta = 1$ can be used to construct orthonormal wavelets for $L^2(\mathbf{R})$. This is accomplished elegantly in the wavelet multiresolution analysis by constructing a sequence of closed subspaces $\{V_m\}_{m\in\mathbf{Z}}$ of $L^2(\mathbf{R})$. These subspaces characterize the behavior of a function at different resolutions. For example, V_m describes functions at 2^m samples per unit length. The subspaces satisfy five important properties (referred to hereafter as M1 through M5) which are listed in Appendix A. The representation of functions in these subspaces V_m are obtained through an orthogonal projection by constructing an orthonormal basis for these subspaces. The orthogonal projection of a function onto the subspace V_m corresponds to its approximation at resolution m, i.e., approximation using 2^m sample points per unit length (see the following discussion).

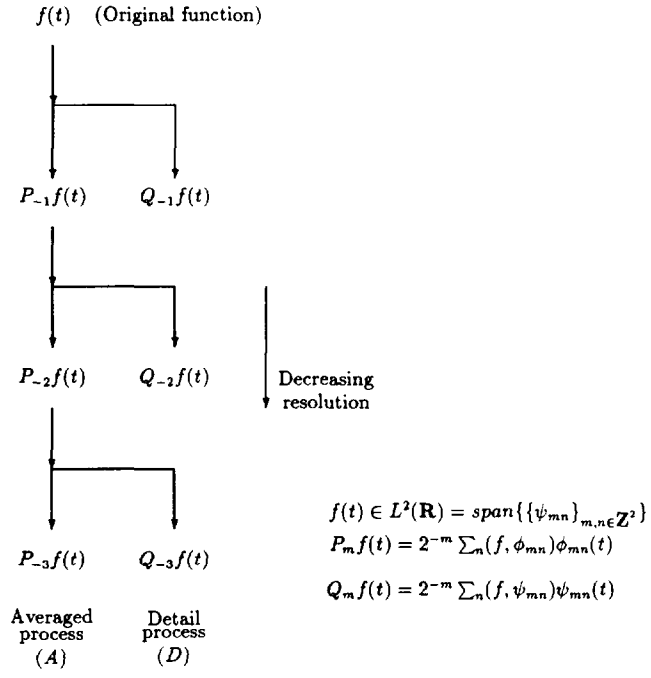

Fig. 3. Schematic showing one-dimensional multiresolution decomposition.

Therefore by successively traversing through the projections of $f(t)$ on the spaces V_m we obtain multiscale representation of the function $f(t)$.

Given the nested structure of V_m (see property M1, appendix A), it is possible to construct a function $\phi(t)$ in V_0 such that $\{\phi(t - n)\}_{n\in\mathbf{Z}}$ is an orthonormal basis of V_0. The function $\phi(t)$ is called a scale function and satisfies $\int \phi(t)\, dt = 1$. Let O_0 be the orthogonal complement of V_0 in V_1, i.e.,

$$V_1 = V_0 \oplus O_0 \tag{11}$$

It is possible to find a function $\psi(t)$, based on $\phi(t)$, such that $\{\psi(t - n)\}_{n\in\mathbf{Z}}$ is an orthonormal basis of O_0. The function $\phi(t)$ is orthogonal to its integer translates, and $\psi(t)$ is orthogonal to its integer translates and dyadic dilates; i.e., $\phi(t) \perp \phi(t - n)\ \forall n$ and $\psi(t) \perp \psi(2^m t - n)\ \forall m, n$. The function $\phi(t)$ is such that if its integer translates constitute an orthogonal basis of V_m then the integer translates of $\phi(2t)$ form an orthogonal basis of V_{m+1}. Using the recursive definition of (11) along with property M1 and the orthogonality of $\psi(t)$ with its integer translates and dyadic dilates, it can be shown that the dilates on the dyadic sequence and translates on integers of $\psi(t)$, i.e., $\{2^{m/2}\psi(2^m t - n)\}_{m,n\in\mathbf{Z}^2}$, form an orthonormal basis of $L^2(\mathbf{R})$ (see Figure 3), or equivalently, $L^2(\mathbf{R}) = \oplus_{m\in\mathbf{Z}} O_m$. The function $\psi(t)$ is called an orthogonal wavelet. The space O_m also satisfies the scaling property M3 [see *Daubechies*, 1992, equation (5.1.12)], i.e.,

$$f(t) \in O_m \text{ if } f(2t) \in O_{m+1} \qquad \forall\, m \in \mathbf{Z}. \tag{12}$$

An example of scale function is

$$\begin{aligned} \phi(t) &= 1 \qquad 0 \le t < 1 \\ \phi(t) &= 0 \qquad \text{otherwise,} \end{aligned} \tag{13}$$

and the corresponding wavelet is the Haar function (see Figure 4*a*) given as

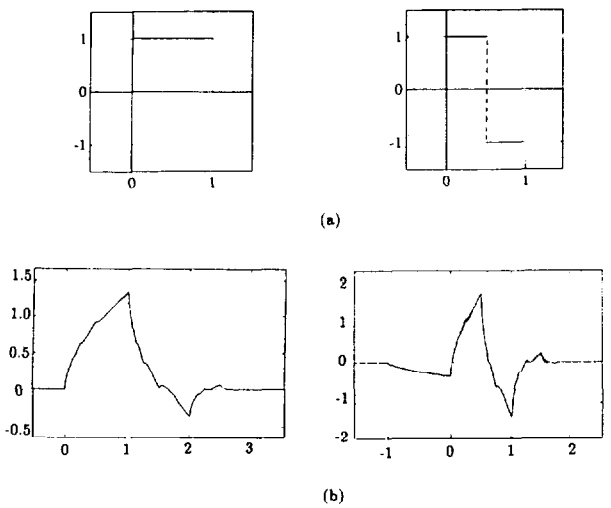

Fig. 4. Typical examples of scale functions ϕ and corresponding wavelets ψ. The wavelets are called (*a*) Haar and (*b*) D4 (after Daubechies) wavelets and they have one and two vanishing moments, respectively.

$$\psi(t) = 1 \qquad 0 \le t < \tfrac{1}{2}$$

$$\psi(t) = -1 \qquad \tfrac{1}{2} \le t < 1 \tag{14}$$

$$\psi(t) = 0 \qquad \text{otherwise.}$$

Another example of scale function and the corresponding wavelet which has two vanishing moments is given in Figure 4*b*. For the general case of constructing the orthonormal bases of scale functions and wavelets, see *Daubechies* [1992].

3.2.2. *Multiresolution approximation.* The approximation of a function $f(t) \in L^2(\mathbf{R})$ at a resolution m, i.e., 2^m sample points per unit length, is given by the orthogonal projection of $f(t)$ on V_m. Let P_m represent this projection operator, i.e., $f(t) \in L^2(\mathbf{R}) \Rightarrow P_m f(t) \in V_m \subset L^2(\mathbf{R})$. Since P_m represents an orthogonal projection it implies that $\|f(t) - P_m f(t)\| = \inf \|f(t) - g(t)\| \ \forall g(t) \in V_m$. This means that the approximation of the function $f(t)$ at resolution m is optimal in the least squares sense. As a consequence of property M1, we understand that $P_{m+1} f(t)$ contains all the necessary information required to compute $P_m f(t)$. Note that the integral transform (4) can be also looked at as an inner product in $L^2(\mathbf{R})$. Using the basis functions $\phi(t)$ and $\psi(t)$ as described above, we can obtain

$$P_m f(t) = 2^{-m} \sum_{n=-\infty}^{\infty} (f(t), \phi_{mn}(t))\phi_{mn}(t) \tag{15}$$

where $\phi_{mn}(t) = 2^m \phi(2^m t - n)$ (notice that $\int \phi_{mn}(t)\, dt = 1$). Let $Q_m f(t)$ represent the orthogonal projection of $f(t)$ onto O_m. Then we can obtain

$$Q_m f(t) = 2^{-m} \sum_{n=-\infty}^{\infty} (f(t), \psi_{mn}(t))\psi_{mn}(t) \tag{16}$$

where $\psi_{mn}(t) = 2^m \psi(2^m t - n)$.

In the interpretation of multiresolution approximation, the projections of a function $f(t)$ on the subspaces V_m are viewed as successive approximations of $f(t)$ at finer and finer resolutions as m increases. The values of the samples of the function at resolution m are then exactly the inner products of $f(t)$ with the scale function $\phi_{mn}(t)$ for various values of n. The wavelet coefficients are used to express the additional details needed to go from one resolution to the next finer resolution level. Therefore the set of inner products $P_m^d f = \{(f(t), \phi_{mn}(t))\}_{n \in \mathbf{Z}}$ gives the discrete approximation of $f(t)$ (or sampled $f(t)$) at resolution m), and the wavelet coefficients $Q_m^d f = \{(f(t), \psi_{mn}(t))\}_{n \in \mathbf{Z}}$ give the discrete detail approximation of $f(t)$ (or difference in information between functions at different resolutions). In other words, we need to add the information contained in $Q_m^d f$ to $P_m^d f$ to go from resolution level m to the next higher resolution level $m + 1$. For this reason, $Q_m f(t)$ is also referred to as the detail function. Due to orthogonality of the decomposition the number of values in $P_m^d f$ and $Q_m^d f$ are each half of that in $P_{m+1}^d f$. Thus there is no increase in the data size.

It was indicated that the integral transform (4) can be also looked upon as convolution of $f(t)$ with the function $\bar{\psi}(t)$, and the wavelet transform can be looked upon as a filtering operation. The inner product with the scale function therefore corresponds to low pass filtering and that with the wavelet to band pass filtering. It is therefore evident that the construction of a function from a higher resolution to a lower resolution is accomplished through low pass filtering and details lost in this process are kept (as wavelet coefficients) through band pass filtering.

3.3. *Two-Dimensional Multiresolution Representation*

For two-dimensional multiresolution approximation we consider the function $f(t_1, t_2) \in L^2(\mathbf{R}^2)$. A multiresolution approximation of $L^2(\mathbf{R}^2)$ is a sequence of subspaces which satisfy the two-dimensional extension of properties M1 through M5 enumerated in Appendix A for one-dimensional multiresolution approximation. We denote such a sequence of subspaces of $L^2(\mathbf{R}^2)$ by $(V_m)_{m \in \mathbf{Z}}$. The approximation of the function $f(t_1, t_2)$ at the resolution m, i.e., 2^{2m} samples per unit area, is the orthogonal projection on the vector space V_m.

A two-dimensional multiresolution approximation is called separable if each vector space V_m can be decomposed as a tensor product of two identical subspaces V_m^1 of $L^2(\mathbf{R})$; i.e., 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 a separable multiresolution approximation of $L^2(\mathbf{R}^2)$,

$$V_m = V_m^1 \otimes V_m^1 \tag{17}$$

where $\otimes$ represents a tensor product. It therefore follows (by expanding V_{m+1} as in (17) and using property M1) that the orthogonal complement O_m of V_m in V_{m+1} consists of the direct sum of three subspaces, i.e.,

$$O_m = (V_m^1 \otimes O_m^1) \oplus (O_m^1 \otimes V_m^1) \oplus (O_m^1 \otimes O_m^1). \tag{18}$$

The orthonormal basis for V_m is given by

$$(2^m \Phi(2^m t_1 - n, 2^m t_2 - k))_{(n,k) \in \mathbf{Z}^2} = (2^m \phi_m(2^m t_1 - n)\phi_m(2^m t_2 - k))_{(n,k) \in \mathbf{Z}^2}. \tag{19}$$

Analogous to the one-dimensional case, the detail function at the resolution m is equal to the orthogonal projection of the

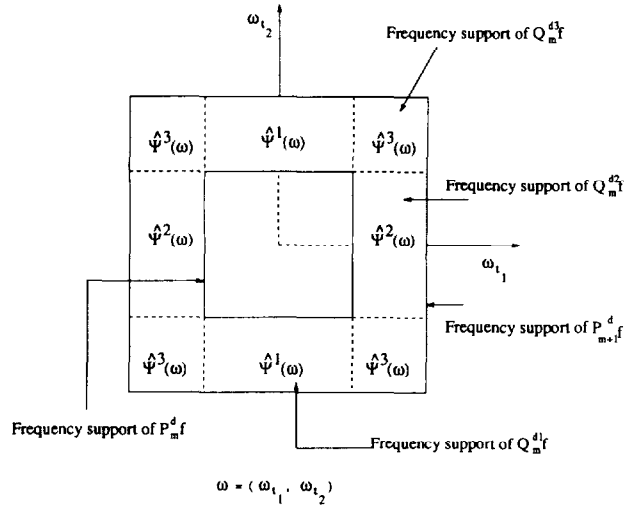

Fig. 5. Frequency support of wavelets in two-dimensional multiresolution decomposition.

function onto the space O_m which is the orthogonal complement of V_m in V_{m+1}. An orthonormal basis for O_m can be built based on theorem 4 in the work by *Mallat* [1989*a*, p. 683], who shows that if $\psi(t_1)$ is the one-dimensional wavelet associated with the scaling function $\phi(t_1)$, then, the three "wavelets" $\Psi^1(t_1, t_2) = \phi(t_1)\psi(t_2)$, $\Psi^2(t_1, t_2) = \psi(t_1)\phi(t_2)$, and $\Psi^3(t_1, t_2) = \psi(t_1)\psi(t_2)$ are such that

$$\{(\Psi^1_{mnk}, \Psi^2_{mnk}, \Psi^3_{mnk})_{(n,k)\in \mathbf{Z}^2}\}$$

is an orthonormal basis for O_m.

The discrete approximation of the function $f(t_1, t_2)$ at a resolution m is obtained through the inner products

$$P_m^d f = \{(f, \Phi_{mnk})_{(n,k)\in \mathbf{Z}^2}\} = \{(f, \phi_{mn}\phi_{mk})_{(n,k)\in \mathbf{Z}^2}\} \quad (20)$$

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

$$\begin{aligned} Q_m^{d1}f &= \{(f, \Psi^1_{mnk})_{(n,k)\in \mathbf{Z}^2}\}, \\ Q_m^{d2}f &= \{(f, \Psi^2_{mnk})_{(n,k)\in \mathbf{Z}^2}\}, \\ Q_m^{d3}f &= \{(f, \Psi^3_{mnk})_{(n,k)\in \mathbf{Z}^2}\} \end{aligned} \quad (21)$$

The decomposition of O_m into the sum of three subspaces (see (18)) gives the behavior of spatially oriented frequency channels. Assume that we have a discrete process at some resolution $m + 1$ whose frequency domain is shown in Figure 5 as the domain of $P_{m+1}^d f$. When the same process is reduced to resolution m, its frequency domain shrinks to that of $P_m^d f$. The information lost can be divided into three components as shown in Figure 5: vertical high frequencies (high horizontal correlation), horizontal high frequencies (high vertical correlation), and high frequencies in both direction (high vertical and horizontal correlations, for example, features like corners). These components are captured as $Q_m^{d1}f$, $Q_m^{d2}f$, and $Q_m^{d3}f$, respectively. We will use this property to characterize the directional behavior of rainfall. Although wavelets with more than three frequency channels can be constructed [see *Daubechies*, 1992], they will not be considered in this research.

For an efficient algorithm to implement orthogonal wavelet transforms to discrete data, see *Mallat* [1989*a*].

4. Wavelet Analysis of Rainfall Fields: Theory

4.1. *Basic Components of Analysis*

For the purpose of this study we consider rainfall intensities (obtained using meteorological radar) averaged over small intervals of time on a spatial grid at ground level. We therefore have a temporally integrated two-dimensional process which we call a frame. A sequence of frames with temporal averaging over nonoverlapping windows and covering the entire time domain characterizes the behavior of rainfall in time. Note that the actual storm is a three-dimensional process that evolves over time. Here we do not consider the three-dimensional storm structure but the "derived rain intensity field" at the ground level at some temporal integration scale. Inference about the evolution of the rain intensity field can be made by studying the changes in the parameters of interest, i.e., parameters found to characterize the structure of the rainfield, across a sequence of frames [see *Kumar and Foufoula-Georgiou*, this issue].

The process of arriving at the derived rain intensity field at the ground level from observations made using meteorological radar is quite complex involving vertical integration of cloud reflectivities, applying corrections for ground clutter, etc., converting them to rain intensities through nonlinear relationships, and calibrating with rain gage observations if available. Within the scope of this research we will assume that the average rainfall over an area A during a time period T is obtained as

$$\mathcal{R}(\mathbf{t}) = \frac{1}{|A|}\int_A ds \frac{1}{T}\int_0^T r(\mathbf{s}, \eta)\, d\eta \quad (22)$$

$$\mathbf{s} = (s_1, s_2) \in \mathbf{R}^2, \qquad \eta \in \mathbf{R}^+$$

$$\mathcal{R}(\mathbf{t}) \equiv \frac{1}{|A|}\int_A \bar{r}(\mathbf{s})\, ds \quad (23)$$

where $r(\mathbf{t}, \eta)$ is some underlying rainfall field defined over continuous space and time parameters $\mathbf{t}$ and η, respectively. That is, we will assume rain intensities obtained at the grid points at an instant of time to be the output of a linear filter applied to the underlying rainfall field. The field $\mathcal{R}(\mathbf{t})$ corresponds to the actual observations obtained using the radar. For several advantages, to be described below, we will assume that the temporally integrated rain intensity of a frame, at a grid point given by the index n, $k \in \mathbf{Z}$, where the grid has a resolution of 2^{2m} samples per unit area, is obtained as

$$\mathcal{R}_m(n, k) = \int \Phi_{mnk}(\mathbf{t})\bar{r}(\mathbf{t})\, d\mathbf{t} \qquad m, n, k \in \mathbf{Z} \quad (24)$$

i.e., the spatial filer is a scale function $\Phi_{mnk} = 2^m\Phi(2^m t_1 - n, 2^m t_2 - k)$ corresponding to a multiresolution framework (see section 3.3).

In several situations, average behavior provides only one facet of the nature of the process. Information about fluctuations, i.e., deviations from the mean, are as essential. In

analogy with (24) we define a fluctuation field, at resolution m, by

$$\mathcal{R}'_m(n, k) = \int \varphi_{mnk}(\mathbf{t})\bar{r}(\mathbf{t})\, d\mathbf{t} \tag{25}$$

where $\varphi_{mnk}(\mathbf{t})$ is any spatial filter (not necessarily a wavelet) such that $\varphi_{mnk}(\mathbf{t}) = 2^m\varphi(2^m t_1 - n, 2^m t_2 - k)$, it has a compact support and satisfies

$$\int \varphi_{mnk}(\mathbf{t})\, d\mathbf{t} = 0. \tag{26}$$

By virtue of the property (26), $\mathcal{R}'_m(n, k)$ gives the value of the deviation from the local mean at resolution m or scale $1/2^m$. To see this consider $E[\mathcal{R}'_m(n, k)] = \int \varphi_{mnk}(\mathbf{t})E[\bar{r}(\mathbf{t})]\, d\mathbf{t}$. Now assume $E[\bar{r}(\mathbf{t})]$ is some constant within the support of $\varphi_{mnk}(\mathbf{t})$. Then by (26) the above integral is zero. Hence if at resolution m, the nonhomogeneous mean can be approximated by a piecewise constant function then the fluctuation can be obtained easily by the above procedure. The more general case, where the nonhomogeneous mean can be approximated by a polynomial of degree N, is discussed in Appendix B. The advantage of such a methodology is quite evident: in the region where $\bar{r}(\mathbf{t})$ has a large variation from the local mean, $\mathcal{R}'$ will be large and vice versa and this information is localized due to the compact support of φ. The only restriction placed on $\bar{r}(\mathbf{t})$ is that it be a measurable function.

These characterizations enable us to overcome the problem of "absence of point values" in rainfall, since the averages and fluctuations are both defined through integral transforms. Notice also that the description of scale is built into the transform, which is a very desirable feature for our purpose of multiscale study. If we have the rainfall process sampled at some resolution m_0, then under the multiresolution framework (property M1, Appendix A), fluctuations corresponding to all resolutions $m < m_0$ can be obtained by choosing the wavelets as the kernel function. This choice also has certain other desirable features which are detailed in the following subsections. The choices of Ψ^1, Ψ^2, or Ψ^3 extract different behavior of the process and this is also discussed in the following subsections. Specifically, in what follows we develop the following properties of wavelet analysis of stochastic processes.

1. We introduce the notion of discretization of a stochastic process using integration kernels that are orthogonal to their integer translates. We discuss that this discretization has no redundancy due to the linear independence of the integration kernel, and that the discretization is optimal in a sense to be defined. An attractive feature of this discretization is that these properties continue to hold when we change the scale of observation (on a dyadic scale). Both the scale functions and wavelets provide integration kernels that give attractive discretization of the process, albeit characterizing different properties. These properties are essential to justify the inference drawn about the underlying process from discrete samples of the process.

2. The transformation of a given process using the scale function provides a description of the process at different scales. Alternatively, it can be also viewed as the discretization of the mean of the process with the mean obtained using the scale function. This property enables us to characterize the large-scale features of the process obtained using the scale function at some small resolution.

3. The transformation of a process using wavelets with N vanishing moments is equivalent to removing a stochastic trend of polynomial order $N - 1$. By using this we show that the wavelet coefficients may be regarded as the fluctuations of the process at the given scale. This property forms the basis of treating the wavelet coefficients as (discretizations of) fluctuations at various scales which characterize the small-scale behavior of the process.

4. Given the interpretation of averages and fluctuations to the processes obtained using the scale functions and wavelets, we show that the two can be regarded as containing complementary information that are "practically uncorrelated." In practical situations where nonhomogeneity in a process can be attributed to the mean, the fluctuations (i.e., wavelet coefficients) may be assumed homogeneous. This gives us dual advantage: (1) it allows us to address the issue of nonhomogeneity by segregating it from the original process and (2) fluctuations, under the assumption of homogeneity, are easily amenable to statistical analysis.

The reader not interested in the mathematical developments of the above properties can skip the rest of this section and go directly to section 5.

4.2. *Wavelet Analysis of Stochastic Processes*

The objective of this section is to illustrate an optimal method (in a sense to be discussed below) of obtaining values of the process $X_\varphi(t) = \int X(t)\varphi(t)\, dt$, $t \in \mathbf{R}$ at discrete arguments t_n, i.e., obtain an optimal discretization of the process $X_\varphi(t)$. Without loss of generality, let us denote t_n by n. To minimize the redundancy of information contained in the discrete samples we consider discretizations of $X_\varphi(t)$ by considering the values of $X_{\varphi 0n}(t)$ for $\varphi_{0n}(t)$ chosen orthogonal to all of its translates $\varphi_{0j}(t)$ on the set of integers $n, j \in \mathbf{Z}$, where $\varphi_{mn} = 2^m\varphi(2^m t - n)$. In view of the multiresolution wavelet analysis framework, obvious candidates for $\varphi_{mn}(t)$ are either the scale functions $\phi_{mn}(t)$ or the wavelets $\psi_{mn}(t)$, both orthogonal to their translates. In the next subsection we study the properties of the discretizations obtained by choosing the scale function $\phi_{mn}(t)$ as the integration kernel. In section 4.2.2 the discretization obtained using $\psi_{mn}(t)$ is discussed. The two-dimensional extensions are given in the subsections 4.3.1 and 4.3.2.

4.2.1. *Optimal multiscale discretization of a process.* From the theory of multiresolution decomposition it can be seen that the discrete values $\{(X, \phi_{mn})\}_{n\in\mathbf{Z}}$ obtained as $(X, \phi_{mn}) = \int X(t)\phi_{mn}(t)\, dt$ constitute an optimal discretization of $X(t)$ at the resolution m in the least squares sense if the realization of $X(t)$ is regarded as a function in $L^2(\mathbf{R})$. However, in a probabilistic framework, the multiresolution framework provides additional, significant, nontrivial extensions as can be seen from the following three properties.

P1: There is no redundancy in the discretization at any resolution since the integration kernels are linearly independent of their translates. This property is essentially a statement about the nondegeneracy of the finite dimensional joint distribution function of the random variables $\mathbf{X} = \{(X, \phi_{mn_1}), \cdots, (X, \phi_{mn_\nu})\}$. Since $\{\phi_{mn}\}_{n\in\mathbf{Z}}$ are linearly independent, the dimension of the space spanned by $\{\phi_{mn_i}\}_{i=1,\nu}$ is ν. The above property implies that there exists no sub-

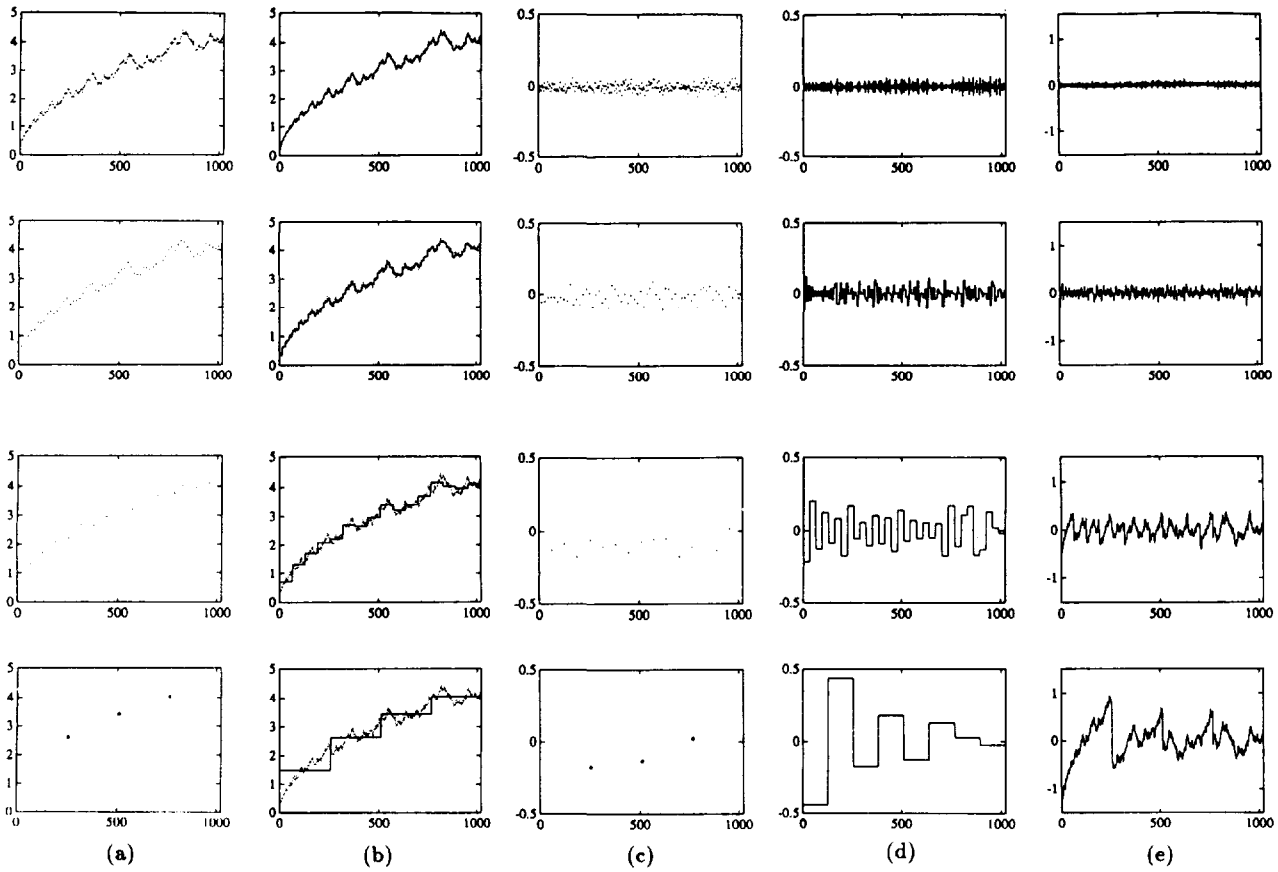

Fig. 6. (*a*) Discrete approximation of the Weierstrass-Mandelbrot cosine curve at decreasing resolutions using the scale function shown in Figure 4*a*. (*b*) Continuous approximation obtained by interpolating with the scale function, i.e., $P_m X(t) = \sum_n (X, \phi_{mn})\phi_{mn}(t)$. (*c*) Discrete wavelet coefficients using the Haar wavelet. (*d*) Continuous approximation obtained by interpolating with the wavelets, i.e., $Q_m X(t) = \sum_n (X, \psi_{mn})\psi_{mn}(t)$. (*e*) Mean removed process obtained via $\sum_{m \geq m_0} \sum_n (X, \psi_{mn})\psi_{mn}(t)$ (or equivalently $X(t) - \sum_n (X, \phi_{m_0 n})\phi_{m_0 n}(t)$).

space of dimension $\nu' < \nu$ such that $\int_{\mathbf{R}^{\nu'}} p_m(\mathbf{x}')\, d\mathbf{x}' = \int_{\mathbf{R}^{\nu'}} p_m(\mathbf{x})\, d\mathbf{x}$, where $\mathbf{x}' \in \mathbf{R}^{\nu'}$; i.e., the joint distribution function of the vector $\mathbf{X} \in \mathbf{R}^{\nu}$ is nondegenerate at scale m. It is noted that for this property to hold the condition of linear independence of $\{\phi_{mn}\}$ is sufficient and orthogonality is not needed. This property will be useful in characterizing a process in general L^P (Banach) spaces as wavelets form a linearly independent basis for those spaces but the concept of orthogonality does not exist.

P2: The discretization at any resolution m is "maximal" as the translates of the integration kernel span the complete space V_m. This is a statement about the uniqueness of the joint distribution function. Consider the infinite set $\{X_n = (X, \phi_{mn})\}_{n \in \mathbf{Z}}$ of random variables. The uniqueness condition states that the joint distribution function of any finite ν number of random variables $\mathbf{X} = \{X_n\}_{n=1,\nu}$ does not change if any additional information about another random variable $X_{n'}$, $n' \notin [1, \nu]$ becomes available. In other words, the equality

$$p(x_1, \cdots, x_\nu) = \int p(x_1, \cdots, x_\nu, x_{n'})\, dx_{n'}$$

holds. This property stems essentially from the property of the completeness of the space V_m. This property is trivially true for finite sets $\{x_n\}$ but not necessarily for infinite sets.

P3: The above properties hold for all dyadic scales (resolutions), thus providing elegant discretizations at a hierarchy of scales. This property makes multiresolution framework very attractive for multiscale studies.

The scale functions ϕ_{mn} provide a very large class of functions that can be used for the discretization of the process $X(t)$. A special case which is often used in practice is the discretization obtained by choosing $\varphi_{0n}(t)$ as the indicator function of intervals that arise from the partitioning of the real line using the set of integers, i.e.,

$$\begin{aligned} \varphi_{0n}(t) &= 1 \qquad n \leq t < n+1 \\ \varphi_{0n}(t) &= 0 \qquad \text{otherwise.} \end{aligned} \tag{27}$$

Clearly, this is a particular case of the multiresolution wavelet transform (see (13)) corresponding to the scale function of the Haar wavelet.

To illustrate the approximation of a function at various resolutions, we applied the multiresolution transform on the Weierstrass-Mandelbrot cosine fractal function (see Appendix C for a brief description). Its discrete approximation, using the two scale functions shown in Figure 4 at decreasing resolutions are shown in Figures 6*a* and 7*a*, respectively. Figures 6*b* and 7*b* show the continuous approximations obtained by interpolating with the corresponding scale functions. It is observed that the scale function corresponding to D4 wavelet (see Figure 4*b*) gives a better approximation of

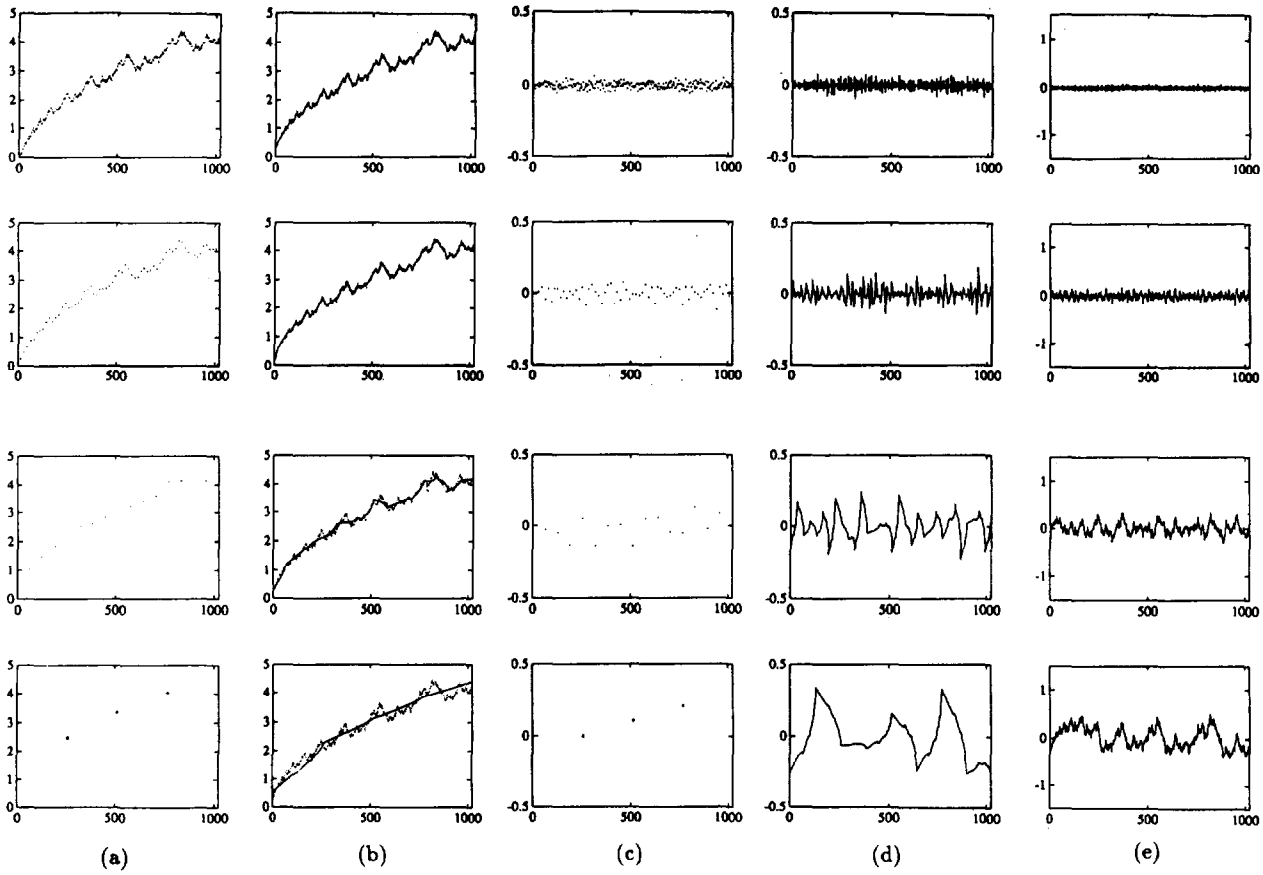

Fig. 7. Same as Figure 6 with D4 wavelet and the corresponding scale function (see Figure 4*b*).

the process than the scale function corresponding to the Haar wavelet and this is particularly noticeable at lower resolutions.

It is easy to show by change of variables in $(X, \phi_{mn}) = 2^m \int X(t)\phi(2^m t - n)\, dt$ that this is the same as $\int X(2^{-m}t)\phi(t - n)\, dt$, i.e., resolution and scale are inversely related. In the first case we look at long-scale behavior of $X(t)$ by spreading out $\phi(t)$ $(m < 0)$, which is equivalent to studying a contracted version of $X(t)$, i.e., $X(2^{-m}t)$ $(2^{-m} > 1)$, through a window of constant size and vice versa. This interpretation has the notion of map scale where large scale indicates global view and small scale indicates detailed view [see also *Rioul and Vetterli*, 1991]. Notice the structure of the transformation where scaling of $X(t)$ is accomplished by equivalently scaling the scale function $\phi(t)$. Therefore the choice of $\phi_{mn}(t)$ as an averaging function proves very useful for multiscale analysis where discretization of $X(t)$ (which may not have point values) at different scales can be obtained by changing m. Alternatively, given the discretization of $X(t)$ at a particular scale, its discretization at different scales can be obtained by an efficient algorithm [see *Mallat*, 1989*a*].

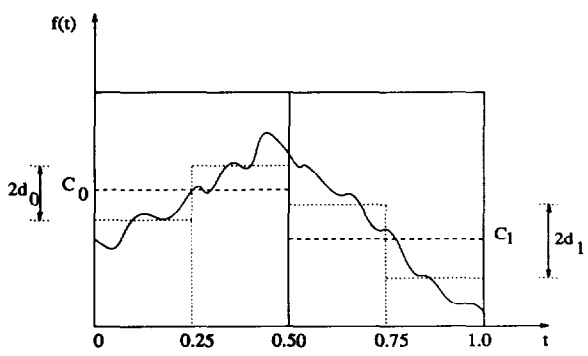

Fig. 8. Schematic showing the concept of fluctuations using the Haar wavelet for discretization with two samples per unit length. The mean of the function in each interval [0, 0.5) and [0.5, 1) is c_0 and c_1 (obtained as the scaling coefficients). Each of these intervals are further divided into two intervals and the difference between the averages over these smaller intervals, as is shown, gives an estimate of the fluctuation (here $2d_0$ where d_0 is the wavelet coefficient). It is easy to see that at the next lower resolution the wavelet coefficient will be $(c_0 - c_1)/2$.

4.2.2. *Wavelet coefficients as fluctuations.* While the representation of the process from one resolution to a lower one is obtained via the scale function, the information lost during this transformation is preserved as the sequence (wavelet coefficients) $Q_m^d X \equiv \{(X, \psi_{mn})\}_{n \in \mathbf{Z}}$. Alternatively, these coefficients can be also viewed as discretization of $X_\varphi(t)$ for $\varphi(t)$ satisfying equation

$$\int t^k \varphi(t)\, dt = 0 \qquad k = 0, \cdots, N - 1. \tag{28}$$

By choosing the wavelet $\psi_{mn}(t)$ for the integration kernel $\varphi_{mn}(t)$ we see that the discretization we obtain satisfies properties P1, P2, and P3 with the space under consideration being O_m. The wavelet coefficients can be regarded as

discretization of the fluctuations of $X(t)$ in the sense that these values give the deviation of the process from its local mean. This concept is schematically shown in Figure 8 for the Haar wavelet. The local mean is obtained using the corresponding scale functions $\phi_{mn}(t)$ as the integration kernel.

Figures 6*c* and 7*c* show the wavelet coefficients at different resolutions obtained from the Weierstrass-Mandelbrot cosine fractal function using the Haar and D4 wavelets. Figures 6*d* and 7*d* show the continuous approximations obtained by the interpolation using the corresponding wavelets. The interpretation of wavelet coefficients as fluctuations and the scaling coefficient as mean is now apparent by comparison of Figures 6*b* and 7*b* with Figures 6*d* and 7*d*, respectively. The interpretation of the wavelet coefficients as discretization of fluctuations leads to interesting developments for using wavelet transforms in the study of stochastic processes with stationary increments of order N as discussed in Appendix B.

So far we have focussed on the behavior, interpretation, and significance of the averaged process and the wavelet coefficients independently. However, at any scale, they contain complementary information. Whereas the wavelet coefficients represent the process with the (stochastic) trend removed, the averaged process gives the approximation (at the given scale) of the mean (see Figures 6*b* and 7*b*). We therefore have an approximation at the given scale 2^{-m} (or resolution m) of the form

$$X(t) \approx \bar{X}_m(t) + X'_m(t) \tag{29}$$

where $\bar{X}_m(t)$ is the approximation of the mean (moving average component) at scale 2^{-m}; i.e., $\bar{X}_m(t) = \sum_n (X, \phi_{mn})\phi_{mn}(t)$ and $X'_m(t)$ is the corresponding fluctuation given by $\sum_n (X, \psi_{mn})\psi_{mn}(t)$. Equality in the above equation could be obtained by summing up the fluctuation processes at all scales greater than 2^{-m_0}, where 2^{-m_0} is the largest scale of approximation, i.e.,

$$X(t) = \bar{X}_{m_0}(t) + \sum_{m \geq m_0} X'_m(t) \tag{30}$$

where $\sum_{m \geq m_0} X'_m(t)$ is obtained by $\sum_{m \geq m_0} \sum_n (X, \psi_{mn})\psi_{mn}(t)$ (or equivalently as $X(t) - \sum_n (X, \phi_{m_0 n})\phi_{m_0 n}(t)$). The approximation $\sum_{m \geq m_0} \sum_n (X, \psi_{mn})\psi_{mn}(t)$ for the Weierstrass-Mandelbrot function using the Haar and D4 wavelets, for decreasing resolutions m_0, are shown in Figures 6*e* and 7*e*, respectively. That is, when we add Figure 6*b* and Figure 6*e* we get the original signal.

In a practical situation we may assume the mean and fluctuations to be uncorrelated. This is also justified since the two components belong to orthogonal subspaces and although the scale function and the wavelets are not orthogonal in the frequency domain, the overlap region is very small, and hence they may be considered approximately uncorrelated. Several situations arise when the nonstationarity in a process can be attributed to the mean and the fluctuations can be assumed stationary. A particular example is a process with stationary Nth order increments. For all such situations when this physical assumption is justified, the wavelet coefficients can be also assumed stationary.

4.3. *Two-Dimensional Analysis*

4.3.1. *Physical interpretation of wavelet transforms.* In the context of signal processing the wavelet decomposition can thus be interpreted as a function decomposition in a set of independent, spatially oriented frequency channels. $Q_m^{d1}f$ gives the vertical high frequencies (horizontal edges), $Q_m^{d2}f$ gives the horizontal high frequencies (vertical edges), and $Q_m^{d3}f$ gives high frequencies in both directions (the corners). However, when dealing with physical processes in a stochastic framework we need a more "tangible" interpretation of the wavelet coefficients. In what follows, we attempt to provide such an interpretation. Consider the discretization of a stochastic process on a uniform grid as shown in Figure 9.

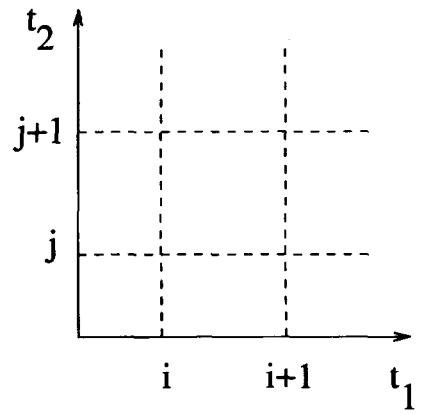

Fig. 9. Schematic showing discretization on a two-dimensional grid.

The mean of the process at scale λ can be obtained as $\bar{X} = (X_{i,j} + X_{i+1,j} + X_{i,j+1} + X_{i+1,j+1})/4$. However, there is no unique way to define increments and usually in practice a scheme such as the one below is adopted:

$$\Delta X_{t_1} = \frac{1}{2}\left[\frac{X_{i,j} + X_{i,j+1}}{2} - \frac{X_{i+1,j} + X_{i+1,j+1}}{2}\right] \tag{31}$$

$$\Delta X_{t_2} = \frac{1}{2}\left[\frac{X_{i,j} + X_{i+1,j}}{2} - \frac{X_{i,j+1} + X_{i+1,j+1}}{2}\right] \tag{32}$$

$$\Delta X_{t_1 t_2} = \frac{1}{2}\left[\frac{X_{i,j} - X_{i+1,j}}{2} - \frac{X_{i,j+1} - X_{i+1,j+1}}{2}\right] \tag{33}$$

They may be looked as analogous to $\partial X/\partial t_1$, $\partial X/\partial t_2$, and $\partial^2 X/\partial t_1 \partial t_2$, respectively.

The choice of Haar wavelet as the analysing wavelet, in the discrete case, gives precisely the above three components (up to a constant) for inner products with Ψ^2, Ψ^1, and Ψ^3, respectively. In the general case we call the above three components as "discretization of fluctuations of the marginal local average process" in t_1 and t_2 directions and "discretization of fluctuations of the marginal fluctuation process." This can be seen by looking at

$$(X, \Psi^1_{mnk}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} X(t_1, t_2) \cdot \Psi^1_m(t_1 - 2^{-m}n, t_2 - 2^{-m}k)\, dt_1\, dt_2 \tag{34}$$

$$(X, \Psi^1_{mnk}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} X(t_1, t_2) \cdot \phi_m(t_1 - 2^{-m}n)\psi_m(t_2 - 2^{-m}n)\, dt_1\, dt_2 \tag{35}$$

$$(X, \Psi^1_{mnk}) = \int_{-\infty}^{\infty} \psi_m(t_2 - 2^{-m}k) \cdot \left(\int_{-\infty}^{\infty} X(t_1, t_2)\phi_m(t_1 - 2^{-m}n)\, dt_1 \right) dt_2 \quad (36)$$

The component inside the parenthesis, with a proper choice of m and n, is the discretization of the marginal local average process in the t_1 direction. The component (X, Ψ^2_{mnk}), when treated analogously, gives fluctuations of the marginal local average process in the t_2 direction. The interpretation for the third component follows analogously.

4.3.2. *Removal of nonhomogeneity in the mean.* Consider the expected value of (X, φ) given by

$$E\left[\int X(t_1, t_2)\varphi(t_1, t_2)\, dt_1\, dt_2\right] = \int m(t_1, t_2)\varphi(t_1, t_2)\, dt_1\, dt_2 \quad (37)$$

$$E\left[\int X(t_1, t_2)\varphi(t_1, t_2)\, dt_1\, dt_2\right] = \int (a_{00} + a_{10}t_1 + a_{02}t_2 + a_{11}t_1t_2 + \cdots)\varphi(t_1, t_2)\, dt_1\, dt_2 \quad (38)$$

where $\sum_p \sum_q a_{pq}t_1^p t_2^q$ is the Taylor series expansion of $m(t_1, t_2)$ at scale 2^{-m}. We can choose Ψ^1, Ψ^2, or Ψ^3 for φ. If the two-dimensional wavelets are constructed using one-dimensional wavelets with N vanishing moments then it is easy to verify that using Ψ^2 and Ψ^1 will eliminate a trend of polynomial order N in the t_1 and t_2 directions, respectively, i.e., trends of the form $m(t_1, t_2) = \sum_{p \le N} a_{p0}t_1^p$ and $m(t_1, t_2) = \sum_{q \le N} a_{0q}t_2^q$, and Ψ^3 will eliminate a polynomial trend of the form $m(t_1, t_2) = \sum \sum_{p+q \le N} a_{pq}t_1^p t_2^q$.

5. Representation of Rainfall as a Composite Process

5.1. *Component Processes Using Wavelets*

In general, a nonhomogeneous process can be made tractable by decomposing it into simpler component processes in a variety of ways [see *Vanmarcke*, 1983, p. 224]. We choose the decomposition (1) for its physical significance of capturing large- and small-scale behavior of the process. We approximate the mean $\bar{X}(\mathbf{t})$ and fluctuations $X'(\mathbf{t})$ at some resolution m_0 using scale functions and wavelets as $\bar{X}(\mathbf{t}) \approx \bar{X}_{m_0}(\mathbf{t})$ and $X'(\mathbf{t}) = \sum_{m > m_0} X'_m(\mathbf{t})$ where

$$\bar{X}_m(\mathbf{t}) = \sum_{n,k} (X, \Phi_{mnk})\Phi_{mnk}(\mathbf{t}) \quad (39)$$

and the fluctuation field at resolution m is composed of three components

$$X'_m(\mathbf{t}) = \sum_{n,k} (X, \Psi^1_{mnk})\Psi^1_{mnk}(\mathbf{t}) + \sum_{n,k} (X, \Psi^2_{mnk})\Psi^2_{mnk}(\mathbf{t}) + \sum_{n,k} (X, \Psi^3_{mnk})\Psi^3_{mnk}(\mathbf{t}) \quad (40)$$

$$X'_m(\mathbf{t}) \equiv X'_{1,m}(\mathbf{t}) + X'_{2,m}(\mathbf{t}) + X'_{3,m}(\mathbf{t}) \quad (41)$$

where $X'_{i,m}(\mathbf{t})$ is used to denote the component $\{Q^i_m X(\mathbf{t})\}$ of the wavelet decomposition. Therefore

$$X'(\mathbf{t}) = \sum_{m \ge m_0} (X'_{1,m}(\mathbf{t}) + X'_{2,m}(\mathbf{t}) + X'_{3,m}(\mathbf{t})) \quad (42)$$

$$X'(\mathbf{t}) \equiv X'_1(\mathbf{t}) + X'_2(\mathbf{t}) + X'_3(\mathbf{t}). \quad (43)$$

The scale 2^{-m_0} (or resolution m_0) will be determined from statistical considerations, as the largest scale up to which rainfall fluctuations exhibit self-similarity. Notice that if rainfall fluctuations exhibit self-similarity at all scales then m_0 does not exist (or is infinite), i.e., all scales are equally significant in characterizing the process. In the case that there is no self-similarity at all, m_0 will have to be determined from other considerations. For example, if the process has two distinct scales then m_0 can be chosen as a scale in between these distinct scales.

The advantages of the above approximation are enumerated below which also illustrate how the various properties of wavelet transforms like multirate filtering, time-frequency localization, and spatially oriented frequency channels provide useful characterization of the process.

1. The mean field $\bar{X}_m(\mathbf{t})$ represents the large-scale behavior of the process. Because the storm is evolutionary, the mean field cannot be obtained from the ensemble average of the several frames of the storm. Since the mean field in general represents the large-scale behavior, (39) offers a very convenient way of representing this field and also studying the evolution of the mean field of the storm. No assumption about the homogeneity of the mean field is required since the representation (39) is constructed using transformation that is local.

2. The fluctuation field at some resolution m, $X'_m(t)$, itself consists of three components giving the directional information about the storm. Since the wavelet transforms are local (recall that wavelet transforms provide time-frequency localization) and essentially represent the fluctuations of the process, they can be assumed homogeneous provided the nonhomogeneity is only the mean (as was discussed earlier) and the mean of the appropriate polynomial order has been eliminated. This second criterion will essentially govern the choice of the wavelet used. For example, if the mean can be approximated fairly well by piecewise constant functions at the resolution of its representation m_0, then Haar wavelet would be good enough for the study of fluctuations.

Since the fluctuation components are extracted using independent spatially oriented frequency channels they are uncorrelated [see *Yaglom*, 1987, equation (2.205)] and the covariance $R'_m(\mathbf{t}, \mathbf{s})$ of $X'_m(\mathbf{t})$ itself can be written as the sum of the covariances of these three components, i.e.,

$$R'_m(\mathbf{t}, \mathbf{s}) = R'_{1,m}(\mathbf{t}, \mathbf{s}) + R'_{2,m}(\mathbf{t}, \mathbf{s}) + R'_{3,m}(\mathbf{t}, \mathbf{s}) \quad (44)$$

$$R'(\mathbf{t}, \mathbf{s}) = \sum_{m \ge m_0} (R'_{1,m}(\mathbf{t}, \mathbf{s}) + R'_{2,m}(\mathbf{t}, \mathbf{s}) + R'_{3,m}(\mathbf{t}, \mathbf{s})) \quad (45)$$

$$R'(\mathbf{t}, \mathbf{s}) \equiv R'_1(\mathbf{t}, \mathbf{s}) + R'_2(\mathbf{t}, \mathbf{s}) + R'_3(\mathbf{t}, \mathbf{s}) \quad (46)$$

We conjecture that the external factors governing the storm mostly influence the large-scale behavior (or the mean field) and the small-scale behavior (or fluctuation field) is relatively independent of this influence. We perform multiscale analy-

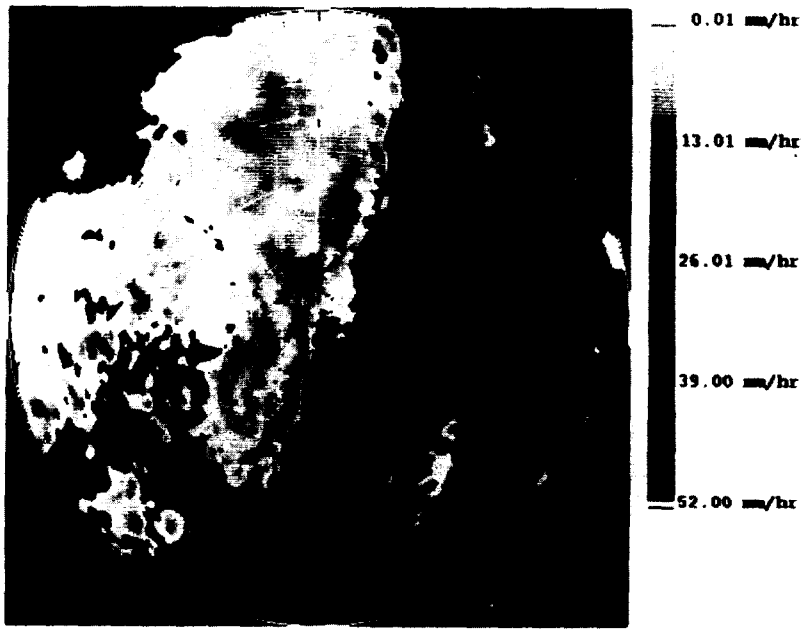

Fig. 10. Squall line storm (at 11:52 A.M.) monitored by National Severe Storms Laboratory, Norman, Oklahoma, on May 27, 1987. The field represents the average rainfall intensity (in millimeters per hour) over 10 min interval. The vertical represents the north-south direction and the radial marks are 25 km apart.

sis [see *Kumar and Foufoula-Georgiou*, this issue] on each of these components to test the presence of scaling and establish its precise nature.

3. By m_0 we will represent the largest scale up to which all fluctuation components X'_{d1}, X'_{d2}, and X'_{d3} show scaling behavior. Although there is no a priori reason to believe that the three components will show scaling up to the same resolution m_0, we will assume that this is true for the purpose of this research (which is further supported by the results of the data analysis). The critical resolution m_0 depends on the storm system under study and may vary from storm to storm and cannot be determined a priori. However, using the multirate filtering capability of wavelets, we find m_0 as the scale at which the fluctuation components cease to show scaling behavior, i.e., m_0 is determined a posteriori from the behavior of the particular storm.

5.2. *Application*

To illustrate the decomposition of rainfall in large- and small-scale component processes, we present the results of multiresolution analysis on a severe squall line storm which occurred over Norman, Oklahoma on May 27, 1987. This storm 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 conversion of the cloud reflectivity (in dbZ) to rainfall rates (in millimeters per hour) was done at NSSL in Norman, Oklahoma using the relationship $Z = 300\mathcal{R}^{1.4}$, where $\mathcal{R}$ is rainfall rate in millimeters per hour 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 for this storm are available at two temporal integration scales, 1 hour and 10 min (for a period of 7 hours beginning with the mature stage of the core of squall line), for 360 azimuths, with every azimuth containing 115 estimates for a range of 230 km (i.e., data at every 2 km by 1 degree). The 10-min integrated rainfall intensities will be used for the illustration of the decomposition procedure.

Figure 10 shows the rainfall intensities of the squall line storm (at 11:52 A.M.). The storm had a broken line formation (see *Bluestein and Jain* [1985] for a classification of squall line storms) and was classified as a severe storm with total precipitation exceeding 25 cm in a period of over 8 hours. The frontal high-intensity core region and the trailing low-intensity anvil region of stratiform precipitation [see *Smull and Houze*, 1985] are clearly evident in this squall line storm.

The wavelet decomposition of the 10-min integrated rainfall intensities at the original resolution will produce four fields: one average field A (corresponding to the scale function Φ) and three fluctuation fields D_1, D_2, and D_3 (corresponding to the wavelets Ψ^1, Ψ^2, and Ψ^3, respectively), each at half the original resolution. For display purposes, at every level of decomposition, the display area of the average field is divided into four quadrants, each containing one of the four fields A, D_1, D_2, and D_3, each at half the previous resolution. The average field at this lower resolution will be displayed in the upper left quadrant of the current average field, and the upper right, lower left, and lower right quadrants are used to display the fields D_1, D_2, and D_3, respectively. Using this scheme for display, Figure 11 shows the wavelet decomposition of the frame depicted in Figure 10 carried on up to four levels. Evidently, the fluctuation field shows large variability in regions of large intensity and vice versa, indicating that the core behaves differently from the anvil region both in mean and fluctuations. That is, the core has large mean and fluctuations, whereas the anvil region has small mean and fluctuations. This storm therefore shows nonhomogeneity not only in the mean but also in the fluctuations. Also, the fluctuation fields D_1, D_2, and D_3 have different characteristics. The fluctuation field D_3 is relatively "mild" as compared to D_1 and D_2.

Figure 12 shows the large-scale features of the storm ob-

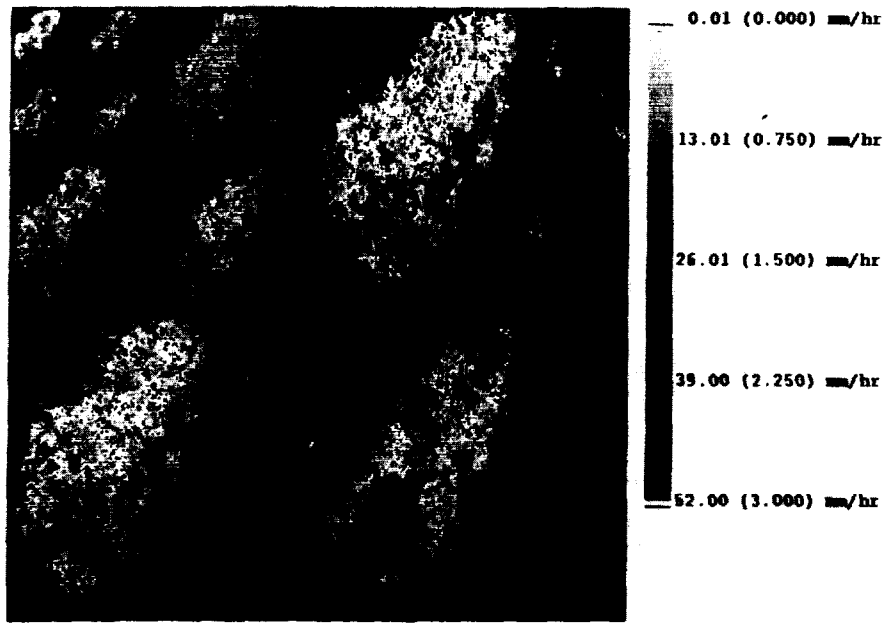

Fig. 11. Three levels of multiresolution decomposition of the squall line storm using Haar wavelet. The absolute values of the wavelet coefficients as they have negative and positive values are displayed. The scales for displaying the magnitude of the wavelet coefficients are indicated in parenthesis along the gray scale bar.

tained by eliminating the details after four levels of decomposition; i.e., it represents $\bar{X}_{m_0}(t)$ at $m_0 = 5$. This evidently captures the morphological organization of the storm and we argue that it should be modeled deterministically, e.g., using mesoscale weather models, to account for the environmental conditions and meteorological forcings responsible for this storm. Figure 13 shows the rainfall field of the storm after the large-scale features depicted in Figure 12 have been eliminated; i.e., it shows the fluctuation field $|X'(t)|$. We argue that this field may exhibit scaling characteristics and that an appropriate way of looking for these characteristics is by studying its components $X'_1(t)$, $X'_2(t)$, and $X'_3(t)$ obtained by decomposing $X'(t)$ using wavelets. This hypothesis is tested in paper 2 [*Kumar and Foufoula-Georgiou*, this issue] and, indeed, it is found that scaling laws hold up to certain scales.

It should be mentioned that a parsimonious statistical parameterization of the fluctuations based on scaling laws (if they are appropriate) is important since fluctuations contain a considerable part of the total energy (variance) of the storm. For example, as is seen from Table 1 for the squall

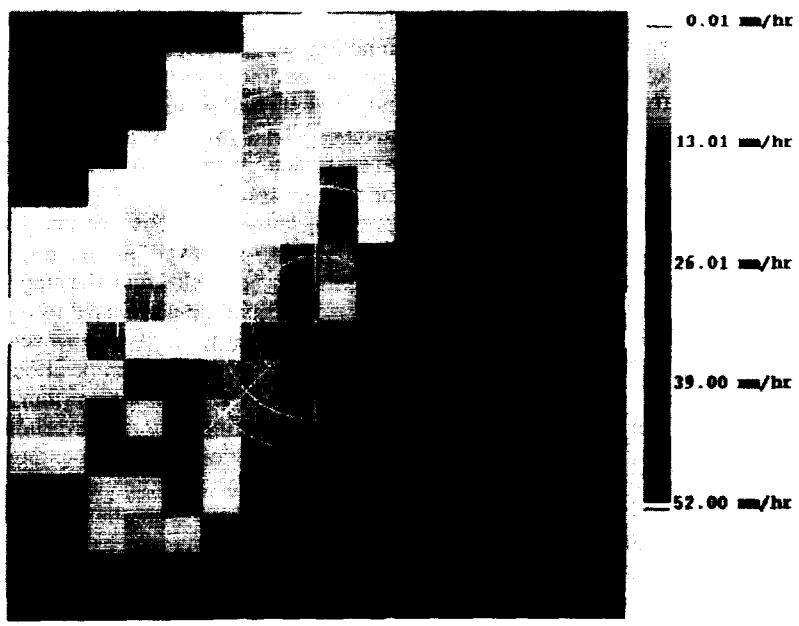

Fig. 12. Large-scale behavior of the squall line storm. This field is obtained by enlarging the average field obtained after five levels of decomposition.

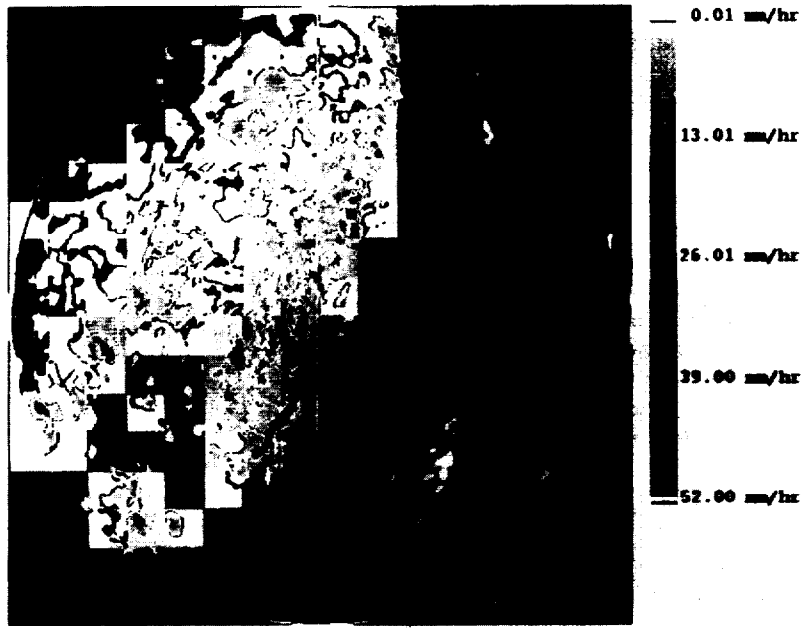

Fig. 13. Small-scale behavior of the squall line storm. This field is obtained by eliminating the large-scale features after five levels of decomposition and reconstructing the small-scale features using the inverse algorithm. Since this field has both positive and negative values, the absolute values are used for display.

line, the fluctuations up to five levels of decomposition account for 55% of the total energy of the storm.

6. Concluding Remarks

Over the last two decades precipitation modeling and analysis has played an increasingly important role in hydrologic research. However, past trends indicate that model development has outpaced the development of data analysis and inference procedures. Most data analysis techniques developed have been geared toward estimating parameters of a hypothesized model. The results presented here are a departure from this approach. We described a versatile data analysis technique, independent of any hypothesized model, that (1) segregates components having characteristic features at different scales through an optimal multiscale transformation without an a priori assumption of a homogeneous field, (2) isolates directional information for the study of anisotropic behavior of fluctuations, and (3) allows one to track the evolutionary behavior of a storm. Since there is no hypothesized model guiding the analysis and inference procedure, the methodology presents a unique way of studying natural phenomena to unravel their statistical structure in scale and space.

The statistical parameterization of rainfall fields at different scales is needed for several hydrologic and atmospheric applications. For example, a problem of considerable current research and practical interest is that of subgrid scale parameterization of hydrologic and land surface processes. This interest is triggered by the fact that the outputs of mesoscale weather prediction models or global circulation models are available only at coarse grid scales, whereas most regional or basin scale hydrologic applications require more detailed information. It is common that processes at unresolved scales are treated statistically either because detailed physical information needed to describe and model these processes is not available or the physics of the process are not well understood at all these small scales or it is computationally inefficient to run physical models with such small grid sizes. Thus a determination of the statistical

TABLE 1. Second-Order Statistics of the Four Components for One Frame (at 11:52 A.M.) of the May 27, 1987, Squall Line Storm Over Norman, Oklahoma, Shown in Figure 10

Grid	Mean	Standard Deviation, mm/h	Energy (Variance)
	Component A		
512 × 512	4.31	8.26	68.23
256 × 256	4.23	8.05	64.80
128 × 128	4.05	7.60	57.76
64 × 64	3.81	6.69	44.76
32 × 32	3.42	5.43	29.48
16 × 16	3.00	4.46	19.89
	Component D_1		
256 × 256	0.0	0.88	0.77
128 × 128	0.0	1.25	1.56
64 × 64	0.0	1.58	2.50
32 × 32	0.0	1.46	2.13
16 × 16	0.0	1.44	2.07
	Component D_2		
256 × 256	0.0	1.24	1.54
128 × 128	0.0	1.79	3.20
64 × 64	0.0	2.66	7.07
32 × 42	0.0	2.91	8.47
16 × 16	0.0	1.84	3.39
	Component D_3		
256 × 256	0.0	0.34	0.12
128 × 128	0.0	0.59	0.35
64 × 64	0.0	0.91	0.83
32 × 32	0.0	1.24	1.54
16 × 16	0.0	1.41	2.0

behavior of the process at unresolved scales and its coupling with the dynamics of the process at the resolved scale is a problem of considerable interest as it will significantly enhance our modeling capability. The methodologies presented in this research can be explored toward development of a framework for coupling the results of large-scale meteorological models with a statistical approach to subgrid scale parameterization. Once the segregation of large- and small-scale features from the data is accomplished using wavelet decomposition and a statistical characterization of the small-scale features (wavelet coefficients) is obtained (as will be discussed in paper 2 [*Kumar and Foufoula-Georgiou*, this issue]), the output of a numerical weather prediction model can be treated as the process at the coarse grid (large-scale features) and simulated fluctuations (small-scale features) can be added to it to get an appropriate description at smaller scales. If scaling laws are found to be present in the fluctuations and thus statistical description of the small-scale features of rainfall can be based on scaling models, then this approach to subgrid scale parameterization is even more appealing as it is very efficient and provides consistent statistical descriptions at any desired scale. Whether fluctuations can be treated as a scaling process, what kind of scaling behavior they exhibit, and how can the scaling parameters be estimated is the subject of paper 2 [*Kumar and Foufoula-Georgiou*, this issue].

As a final comment, it is remarked that the success of both analysis and modeling using the proposed decomposition approach will depend largely on the appropriate choice of the wavelet. As indicated in this paper, proper wavelet selection can be guided by the behavior of the mean at the coarsest scale of description. For the rainfall data sets studied as part of this research, the Haar wavelet is found to be appropriate as piecewise constant function approximates the mean behavior adequately. *Barancourt et al.* [1992] have independently reported that intrinsic random functions of order zero provide an adequate description of the spatial mean behavior of rainfall. The theory presented here is general and provides some guidelines for the appropriate choice of wavelets from the class of orthogonal wavelets if the criterion of selection is removal of polynomial trends. The final choice will of course be governed by the problem at hand.

Appendix A: Properties of Subspaces V_m in the Multiresolution Framework

The subspaces V_m satisfy the following properties: (1) M1, $V_m \subset V_{m+1}\ \forall m \in \mathbf{Z}$; i.e., a space corresponding to some resolution contains all the information about the space at lower resolution; (2) M2, $\cup_{m=-\infty}^{\infty} V_m$ is dense in $L^2(\mathbf{R})$ and $\cap_{m=-\infty}^{\infty} V_m = \{0\}$; i.e., as the resolution increases the approximated function converges to the original function, and as the resolution decreases the approximated function contains less and less information; (3) M3, $f(t) \in V_m$ if and only if $f(2t) \in V_{m+1}\ \forall m \in \mathbf{Z}$; i.e., all spaces are scaled versions of one space (it is this property that leads to the multiresolution framework); (4) M4, $f(t) \in V_m$ implies $f(t - k/2^m) \in V_m\ \forall k \in \mathbf{Z}$; i.e., the space is invariant with respect to "integer translations" of a function; and (5) M5, there exists an isomorphism **I** from V_0 onto $l^2(\mathbf{Z})$ which commutes with the action of **Z**. This property is best explained by the commutative diagram

$$\begin{array}{ccc} f(x) \in V_0 & \xrightarrow{\mathbf{I}} & \{\varepsilon_n\} \in l^2(\mathbf{Z}) \\ \downarrow \mathbf{Z} & & \downarrow \mathbf{Z} \\ f(x-k) \in V_0 & \xrightarrow{\mathbf{I}} & \{\varepsilon_{n-k}\} \in l^2(\mathbf{Z}) \end{array}$$

i.e., translation (by integers) of the approximation of the function is equal to the approximation of the translation of the function (by integers).

Appendix B: Removal of Polynomial Trend of Order N Using Wavelet Transforms

Here it is shown that the wavelet transform of a process with stationary increments of order N, using a wavelet with N vanishing moments (see (5)) denoted herein as ${}_N\psi(t)$ (using the notation of *Daubechies* [1988]) gives rise to a stationary process. We briefly sketch the argument due to *Yaglom* [1958] where, of course, wavelets were not in the context. Consider the Nth order symmetric difference process $\Delta_\tau^{(N)} X(t)$ corresponding to $X(t)$ where

$$\Delta_\tau^{(N)} X(t) = \sum_{k=0}^{N} (-1)^k \binom{N}{k} X(t - k\tau) \tag{47}$$

If $\Delta_\tau^{(N)} X(t)$ is stationary then it can be represented as [see *Yaglom*, 1958, equation (1.28)]

$$\Delta_\tau^{(N)} X(t) = \int_{-\infty}^{\infty} e^{i\omega t}(1 - e^{-i\tau\omega})^N \frac{(1 + i\omega)^N}{(i\omega)^N}\, dZ(\omega) \tag{48}$$

where $Z(\omega)$ is a random process with uncorrelated increments. The process $X(t)$ can be obtained from the above difference equation as

$$X(t) = \int_{-\infty}^{\infty} \left[e^{i\omega t} - \frac{1 + i\omega t + \cdots + \dfrac{(i\omega t)^{N-1}}{(N-1)!}}{1 + \omega^N} \right] \cdot \frac{(1 + i\omega)^N}{(i\omega)^N}\, dZ(\omega) + X_0 + X_1 t + \cdots + X_{N-1} t^{N-1} \tag{49}$$

where $X_0, X_1, \cdots, X_{N-1}$ are arbitrary constants. Taking the inner product of this equation with ${}_N\psi_{mn}$ and integrating we get

$$(X,\ {}_N\psi_{mn}) = \int_{-\infty}^{\infty} {}_N\hat{\psi}_{mn}(\omega) \frac{(1 + i\omega)^N}{(i\omega)^N}\, dZ(\omega) \tag{50}$$

Thus $(X,\ {}_N\psi_{mn})$ is uniquely determined by $\Delta_\tau^{(N)} X(t)$. The mean of the sequence $\{(X,\ {}_N\psi_{mn})\}$ at any scale 2^{-m} is constant and its second-order moments are given by

$$E[(X,\ {}_N\psi_{mn})^2] = \int_{-\infty}^{\infty} |{}_N\hat{\psi}_{mn}(\omega)|^2 \frac{(1 + \omega^2)^N}{\omega^{2N}}\, dF(\omega) \tag{51}$$

$$E[(X,\ {}_N\psi_{mn})(X,\ {}_N\psi_{mn'})] = \int_{-\infty}^{\infty} {}_N\hat{\psi}_{0n}(\omega) \cdot \overline{{}_N\hat{\psi}_{0n'}(\omega)} \frac{(1 + \omega^2)^N}{(\omega)^{2N}}\, dF(\omega) \tag{52}$$

where $E[|dZ(\omega)|^2] = dF(\omega)$, and these moments are invariant under the simultaneous translations of ${}_N\psi_{mn} \rightarrow {}_N\psi_{m(n+k)}$ and ${}_N\psi_{mn'} \rightarrow {}_N\psi_{m(n'+k)}$, $k \in \mathbf{Z}$. We see therefore that the transformed process is second-order stationary implying that orthogonal wavelet transforms using wavelets with N vanishing moments remove a polynomial stochastic trend of order $N - 1$. If the process has stationary increments of order N then we obtain a process that has zero mean and is second-order stationary. Thus wavelet transforms provide an alternative way to study nonstationary stochastic processes. In addition, the above results are valid for all dyadic scales 2^{-m} ($m \in \mathbf{Z}$) which makes wavelets attractive for multiscale study of nonstationary processes.

In practical situations, an assumption of a process having Nth order stationary increments may break down due to physical reasons. For example, (nonstationary) turbulent flows may be approximated by processes with stationary increments of order 1 for time intervals over which the mean characteristics of flow may be regarded as linear [see *Monin and Yaglom*, 1981]. One may therefore have to consider higher-order increments for larger time intervals. Another way to look at such a situation is by considering the equation

$$E[(X, {}_N\psi_{mn})] = \int_{-\infty}^{\infty} E[X(t)]_N\psi_{mn}(t)\, dt$$

$$= \int_a^b m(t)_N\psi_{mn}(t)\, dt \qquad (53)$$

where $[a, b]$ is the region of support for ${}_N\psi_{mn}$. (Notice here that the expectation is an ensemble average). If we consider the truncated Taylor series approximation of $m(t)$, i.e., $m(t) = \sum_{k=0}^{M-1} \alpha_k t^k$, then

$$E[(X, {}_N\psi_{mn})] = 0 \qquad M \leq N \qquad (54a)$$

$$E[(X, {}_N\psi_{mn})] = r_{MN}(m) \equiv \sum_{k=N}^{M} \int_a^b t^k {}_N\psi_{mn}(t)\, dt \qquad (54b)$$

$$M > N$$

Here $r_{MN}(m)$ is the residual component in the mean for the scale 2^{-m} which will depend upon both M and N. By looking at the behavior of $r_{MN}(m)$ for various choices of N and different resolutions m, we can get an estimate for the order of M, i.e., the order of the stochastic trend. Obviously, $r_{MN}(m)$ will increase with scale for a fixed N.

Appendix C: Weierstrass-Mandelbrot Fractal Function

The Weierstrass-Mandelbrot fractal function is a deterministic function given by [see *Berry and Lewis*, 1980]

$$w(t) = \sum_{n=-\infty}^{\infty} \frac{[(1 - e^{i\gamma^n t})e^{i\phi_n}]}{\gamma^{(2-D)n}} \qquad (55)$$

$$1 < D < 2,\ \gamma > 1,\ \phi_n = \text{arbitrary phase}$$

This is a continuous nondifferentiable function. The parameter D gives the fractal dimension of the curve. By choosing $\phi_n = 0$ and taking the real part of $w(t)$ we get the cosine function

$$w_c(t) = \sum_{n=-\infty}^{\infty} \frac{1 - \cos\, \gamma^n t}{\gamma^{(2-D)n}} \qquad (56)$$

The cosine function $w_c(t)$ is nonnegative and has a trend that may be approximated as

$$\bar{w}_c(t) \approx \frac{t^{2-D}\Gamma(D-1)\cos\,(\pi(2-D)/2)}{(2-D)\log\,\gamma} \qquad (57)$$

A plot of $w_c(t)$ for $D = 1.5$ and $\gamma = 1.5$ is shown in Figure 6*a* (top).

Acknowledgments. This research was supported by National Science Foundation grants BSC-8957469 and EAR-9117866 and by NASA grant NAG 5-2108 and a Graduate Student Fellowship for Global Change Research. We also thank the Minnesota Supercomputer Institute for providing us with supercomputer resources and Tim O'Bannon of NEXRAD Operational Support Facility, Norman, Oklahoma for providing us with the rainfall data.

References

Barancourt C., J. D. Creutin, and J. Rivoirard, A method for delineating and estimating rainfall fields, *Water Resour. Res.*, *28*(4), 1133–1144, 1992.

Berry, M. V., and Z. V. Lewis, On the Weierstrass-Mandelbrot fractal function, *Proc. R. Soc. London, Ser. A*, *370*, 459–484, 1980.

Bluestein, H. B., and M. H. Jain, Formation of mesoscale lines of precipitation: Severe squall lines in Oklahoma during the spring, *J. Atmos. Sci.*, *42*(16), 1711–1732, 1985.

Daubechies, I., Orthonormal bases of compactly supported wavelets, *Commun. Pure Appl. Math.*, *16*, 901–996, 1988.

Daubechies, I., *Ten Lectures on Wavelets*, SIAM, Philadelphia, Pa., 1992.

Gabor, D., Theory of communications, *J. Inst. Elec. Eng.*, *93*, 429–457, 1946.

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

Gupta, V., and E. Waymire, A statistical analysis of mesoscale rainfall as a random cascade, *J. Appl. Meteorol.*, *32*(2), 251–267, 1993.

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

Koenderink, J. J., The structure of images, *Biol. Cybern.*, *50*, 363–370, 1984.

Kumar, P., and E. Foufoula-Georgiou, A multicomponent decomposition of spatial rainfall fields, 2, Self-similarity in fluctuations, *Water Resour. Res.*, this issue.

Lovejoy, S., The area perimeter relation for rain and cloud, *Science*, *216*, 185–187, 1982.

Lovejoy, S., and B. B. Mandelbrot, Fractal properties of rain and a fractal model, *Tellus, Ser. A, Dyn. Meteorol. Oceanogr.*, *37*, 209–232, 1985.

Lovejoy, S., and D. Schertzer, Comment on "Are rain rate processes self-similar?", *Water Resour. Res.*, *25*(13), 577–579, 1989.

Lovejoy, S., and D. Schertzer, Multifractals, universality classes and satellite and radar measurements of cloud and rain fields, *J. Geophys. Res.*, *95*(D3), 2021–2034, 1990.

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

Mallat, S., Multifrequency channel decomposition of images and wavelet models, *IEEE Trans. on Acoust. Speech Signal Anal.*, *37*(12), 2091–2110, 1989*b*.

Monin, A. S., and A. M. Yaglom, *Statistical Fluid Mechanics: Mechanics of Turbulence*, MIT Press, Cambridge, Mass., 1981.

Orlanski, I., A rational subdivision of scales for atmospheric processes, *Bull. Am. Meteorol. Soc.*, *52*, 1186–1188, 1975.

Rioul, O., and M. Vetterli, Wavelets and signal processing, *IEEE Signal Process. Mag.*, 14–38, Oct. 1991.

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

Smull, B. F., and R. A. Houze, A midlatitude squall line with a trailing region of stratiform rain: Radar and satellite observations, *Mon. Weather Rev.*, *113*(1), 117–133, 1985.

Vaidyanathan, P. P., Theory and design of M-channel maximally decimated quadrature mirror filters and arbitrary M, having the perfect-reconstruction property, *IEEE Trans. Acoust. Signal Process.*, *ASSP-35*(4), 476–492, 1987.

Vanmarcke, E., *Random Fields: Analysis and Synthesis*, The MIT Press, Cambridge, Mass., 1983.

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

Waymire, E., V. K. Gupta, and I. Rodriguez-Iturbe, A spectral theory of rainfall intensity at the meso-β scale, *Water Resour. Res.*, *20*(10), 1453–1465, 1984.

Yaglom, A. M., Correlation theory of processes with random stationary nth increments, *Am. Math. Soc. Transl., Ser. 2*, *8*, 87–141, 1958.

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

E. Foufoula-Georgiou, St. Anthony Falls Hydraulic Laboratory, Department of Civil and Mineral Engineering, University of Minnesota, Mississippi River at Third Avenue, Southeast, Minneapolis, MN 55414.

P. Kumar, USRA, Hydrologic Sciences Branch, Laboratory for Hydrospheric Processes, Code 974, Building 22, NASA Goddard Space Flight Center, Greenbelt, MD 20771.

(Received July 23, 1992;
revised February 11, 1993;
accepted March 4, 1993.)