

QUANTIFYING ORGANIZATION OF ATMOSPHERIC TURBULENT EDDY MOTION USING NONLINEAR TIME SERIES ANALYSIS

KAREN H. WESSON^{1,2*}, GABRIEL G. KATUL¹ and MARIO SIQUEIRA^{1,3}

¹*School of the Environment, Duke University, Box 90328, Durham, NC 27708-0328, U.S.A.;* ²*The Cadmus Group, Inc., 6330 Quadrangle Drive, Suite 180, Chapel Hill, NC 27517, U.S.A.;*

³*Department of Civil and Environmental Engineering, Duke University, Durham, NC 27708, U.S.A.*

(Received in final form 19 March 2002)

Abstract. Using three methods from nonlinear dynamics, we contrast the level of organization in the vertical wind velocity (w) time series collected in the atmospheric surface layer (ASL) and the canopy sublayer (CSL) for a wide range of atmospheric stability (ξ) conditions. The nonlinear methods applied include a modified Shannon entropy, wavelet thresholding, and mutual information content. Time series measurements collected over a pine forest, a hardwood forest, a grass-covered forest clearing, and a bare soil, desert surface were used for this purpose. The results from applying all three nonlinear time series measures suggest that w in the CSL is more organized than that in the ASL, and that as the flows in both layers evolve from near-neutral to near-convective conditions, the level of organization increases. Furthermore, we found that the degree of organization in w associated with changes in ξ is more significant than the transition from CSL to ASL.

Keywords: Canopy turbulence, Mutual information content, Nonlinear time series analysis, Organized motion, Shannon entropy, Wavelet thresholding.

1. Introduction

Concerns of global warming have generated interest in understanding the exchange rates of mass and energy between the terrestrial biosphere and atmosphere. Around the world, programs such as Euroflux and Ameriflux have been established to provide reliable and continuous measurements of carbon and energy exchange between the vegetation and the atmosphere (Kaiser, 1998). It has been understood that the turbulent exchange of mass and momentum from, and within, canopies is dominated by large, coherent eddy structures (Finnigan, 2000). However, the links between these eddy structures and a quantitative description of canopy dynamics are still a subject of active research (Finnigan and Shaw, 2000). This paper explores the use of several novel techniques from the field of nonlinear dynamics that appear to provide promising tools for the analysis of atmospheric turbulence. These techniques provide “scalar measures” that can be related to the system’s complexity.

It is convenient to decompose the region of the atmospheric boundary layer (ABL), within and above the canopy, into two layers based on the distinct flow

* E-mail: kwesson@cadmusgroup.com



characteristics that exist in each (Finnigan, 2000). The region extending from the ground to 2–3 times the canopy height is labeled the canopy sublayer (CSL), and above this domain, stretches the atmospheric surface layer (ASL). Within the ASL, the flow statistics are typically one-dimensional and well described by Monin and Obukhov surface-layer similarity theory. In sharp contrast, the flow statistics in the CSL are typically non-homogeneous and three-dimensional, and its canonical structure may not resemble a rough-wall boundary layer (Raupach et al., 1989, 1996). While the ASL provides a simpler environment for understanding turbulent transport, taller vegetation such as forests necessitate detailed understanding of the structure of turbulence in the CSL, as measurements at 2–3 times the canopy height become prohibitive. For example, recent CO₂ flux monitoring initiatives above forested ecosystems positioned their sensors well within the CSL in all the forests (Kaiser, 1998).

In this paper, we seek to quantify key attributes about the degree of organization of complex eddy motion in the ASL and CSL using nonlinear dynamics methods and to understand how these attributes are affected by roughness and atmospheric stability. We focus on three techniques: namely a modified, Shannon entropy; wavelet thresholding; and mutual information content, and we apply these measures to the vertical wind velocity (w) time series. Data are analyzed from six different experiments performed in the ASL and CSL, above several different vegetation types, and for a variety of atmospheric stability conditions.

2. Methods of Analysis

A brief review of the methods is considered next. These methods are chosen because they are sensitive to distinct measures of organization of a given time series. Hence, agreement among them provides the necessary confidence that what is realized is related to organization in turbulence rather than an artifact of the definition of organization or its estimation.

2.1. SHANNON ENTROPY

One topological measure of the structure of flows is Shannon, or ‘information’, entropy (Wijesekera and Dillon, 1997), which is defined as

$$I_S = - \sum_i p_i \ln p_i, \quad (1)$$

where p_i ($i = 1, 2, \dots, M$) is a discrete probability distribution that is non-negative, additive, and normalized such that,

$$p_i \geq 0, \quad (2a)$$

$$\sum_i p_i = 1, \tag{2b}$$

$$p_{i \cup j \cup \dots} = p_i + p_j + \dots, \tag{2c}$$

(Shannon, 1948). To allow the Shannon entropy to characterize the shape of a spectrum, we represent the spectrum of a measured scalar quantity as a sequence of M positive numbers ψ_i ($i = 1, \dots, M$) with $p_i = \psi_i (\sum_i \psi_i)^{-1}$ as suggested by Wijesekera and Dillon (1997). Since the components of a discrete power spectrum are positive, additive, and have uniform intrinsic measure, the probability axioms in Equation (2) are satisfied (Skilling and Gull, 1985).

Using this definition of the Shannon entropy provides a technique that gives information about the order or disorder of the flow. For example, when flow is dominated by organized structures, most of the energy of the spectrum resides at a peaked, small wavenumber band. This will produce a very small Shannon entropy value. If the organization of the flow decreases, the energy shifts to smaller scales and the shape of the spectrum becomes less peaked and more flattened. In this case, the complexity of the flow increases and the Shannon entropy, correspondingly, becomes larger. The largest possible value of the Shannon entropy is obtained when the signal becomes characterized by a flat, white noise spectrum where

$$\psi_i = 1, \tag{3a}$$

$$p_i = M^{-1}, \tag{3b}$$

$$I_S = \ln(M). \tag{3c}$$

Recently, a study in which Shannon entropy was applied to turbulent overturns in the ocean demonstrated the usefulness of entropy in quantifying organization within turbulent mixing (Wijesekera and Dillon, 1997). In this study, Wijesekera and Dillon (1997) analyzed the Shannon entropy of a laboratory jet and found that the value of the Shannon entropy increased with time as the flow of the laboratory jet evolved from a simple, laminar to a complex, turbulent flow. The study concluded that the more intense the organization of the flow, the lower the value of the Shannon entropy. These results and the simplicity of the Shannon entropy make it a promising tool in the analysis of complex eddy motion in the ASL and CSL.

2.2. WAVELET THRESHOLDING METHODS

Since large-scale eddy motion is known to be intermittent and inherently non-periodic (Farge et al., 1992), the use of Fourier transforms may mask critical information about time-frequency relationships and how they evolve with stability and morphology. Wavelet transforms allow a signal to be unfolded into both time and frequency thereby making the wavelet transforms desirable for such application. For atmospheric turbulence measurements, discrete wavelet transforms are

known to disbalance the number of coefficients containing energy and the dimension of the time series (Vidakovic, 1995; Katul and Vidakovic, 1996, 1998). Given this, and based on the idea that turbulence is composed of two dynamical parts, a ‘high-dimensional’ passive part and a ‘low-dimensional’ active part associated with organized motion, wavelet transforms can be used to reduce the degrees of freedom needed to describe key dynamical features relevant to organized eddies (Wickerhauser et al., 1994; Farge et al., 1992; Zubair et al., 1992; Katul and Vidakovic, 1998). The term ‘active’ here implies large activity or high variability and should not be confused with the terms ‘inactive’ and ‘active’ turbulence of Townsend (1976). Farge et al. (1992) demonstrated that the active part of turbulence is associated with organized structures while the passive part is composed of slaved modes produced by the interactions among the coherent structures. Since organized eddy motions contribute significantly to the variability in velocity and temperature, Katul and Vidakovic (1998) demonstrated that Lorentz thresholding can be applied to extract the low-dimensional energy containing events. The Lorentz curve is a general measure of energy disbalance of a given data vector whether be it in time, frequency, or wavelet domains. A sample Lorentz curve in the wavelet domain is shown in Figure 1. To allow a simple comparison between two signals using the information above, Katul and Vidakovic (1998) proposed a global thresholding criterion, referred to as the Lorentz threshold, in which the proportion of retained coefficients, L_o , at the point of tangency of the Lorentz curve and the ideally balanced signal be used as a statistic. An example of this optimal proportion L_o is shown in Figure 1. The smaller the value of L_o , the smaller the number of coefficients needed to describe most of the energy in the signal, and therefore, the more organized the signal. And vice versa – the larger the value of L_o , the greater the number of coefficients needed to reconstruct the signal, and the more equally the energy of the signal is spread out among the coefficients. Hence, such a measure can also be used to describe the degree of organization for turbulent time series measurements.

2.3. MUTUAL INFORMATION CONTENT

Mutual information, like Shannon entropy, is a commonly studied property of dynamical systems (Williams, 1997). In chaos theory, mutual information is often used to (i) estimate the optimum embedding dimension for attractor reconstruction (Fraser, 1989), (ii) reveal chaos by indicating whether the Kolmogorov–Sinai entropy is greater than or equal to zero (Fraser, 1989; Palus, 1993), (iii) estimate the optimum lag to use in attractor reconstruction (Fraser, 1989; Pineda and Sommerer, 1993), and (iv) estimate how far into the future we can make reliable predictions. Mutual information is a jointed-entropy than can be derived from the Shannon entropy, and is defined by

$$I_M(s, q) = \sum_{i,j} p_{ij}(s, q) \ln p_{ij}(s, q) - \sum_i p_i(s) \ln p_i(s) - \sum_j p_j(q) \ln p_j(q), \quad (4)$$

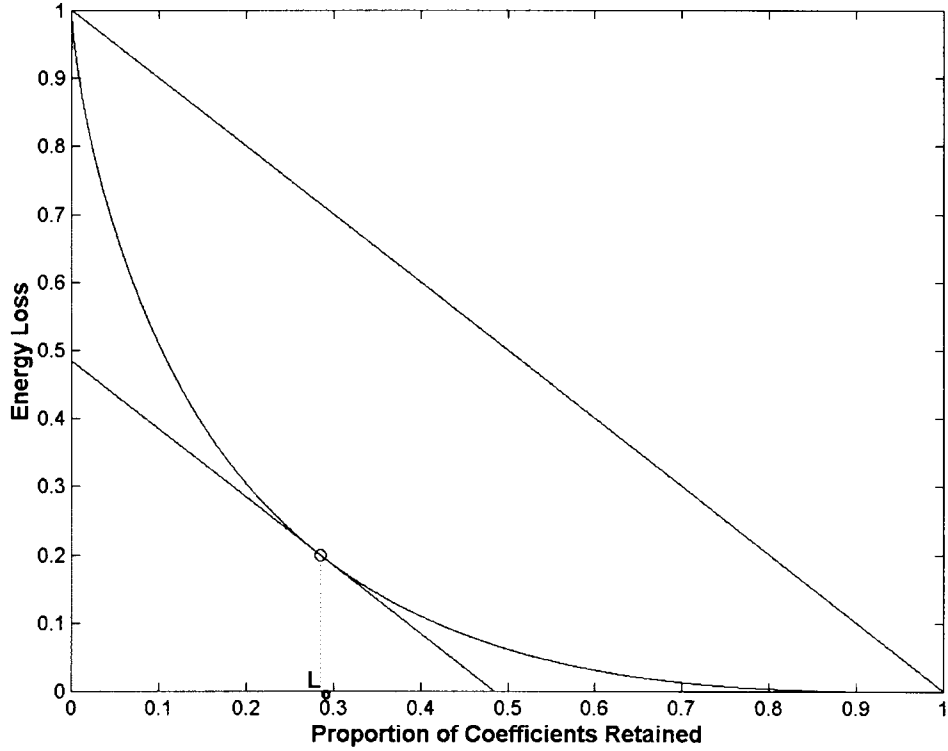


Figure 1. Lorentz curve in the wavelet domain. The diagonal line corresponds to an ideally balanced signal where every coefficient describes the same amount of energy. The convexity of the Lorentz curve in relation to the ideally balance signal is directly proportional to the energy disbalance of the signal.

where $p_{ij}(s, q)$ is the probability that s is in the i th bin and q is in the j th bin, $p_i(s)$ is the probability that s is equal to the i th bin value, and $p_j(q)$ is the probability that q is equal to the j th bin value. In measuring how dependent the values of $x(t + \tau)$ and $x(t)$ are for a measured variable x and time delay τ , we can make the assignment

$$[s, q] = [x(t), x(t + \tau)], \tag{5}$$

and the equation for time-delayed mutual information becomes

$$I_M(\tau) = \sum_{i,j} p_{ij}(\tau) \ln p_{ij}(\tau) - 2 \sum_i p_i \ln p_i, \tag{6}$$

where $p_{ij}(\tau)$ is the probability that $x(t)$ is in the i th bin and $x(t + \tau)$ is in the j th bin and p_i is the probability that $x(t)$ is equal to the i th bin value (Fraser and Swinney, 1986).

Calculation of the time-delayed mutual information is often used to determine the optimum embedding dimension for attractor reconstruction and is followed by the computation of the correlation dimension, as described by Grassberger and Procaccia (1983). The correlation dimension d_c indicates how complex the system is and gives an indication of the actual number of nonlinear ordinary differential equations needed to describe the behaviour of the system. To find d_c from an observed time series $x(t)$ of length N , the vectors $[x(t), x(t + \tau), \dots, x(t + n\tau - \tau)]$ of length N are constructed for a time lag τ and an embedding dimension n . A direct count is made of the number $m(r)$ of pairs of vectors whose difference is less than r in magnitude, for varying r . If, as n is increased, $m(r)$ varies as $r^{d(n)}$, then $d(n)$ is assumed to be the correlation dimension of the reconstructed attractor. If $d(n)$ approaches a limit d_c as n is increased, one assumes that d_c is the estimated correlation dimension of the attractor of the original system (Lorenz, 1991).

If the correlation dimension is fractal then the attractor is known as a 'strange attractor'. Where there is a strange attractor, there is order and structure. Strange attractors offer the potential for obtaining information about the predictability of the system directly from the data rather than indirectly from modelling (Pool, 1989). Because of this enticing possibility, hunting for order in weather patterns and fluid flow by looking for strange attractors has become one of the most controversial chaos-inspired techniques among scientists.

Several studies have estimated correlation dimensions for atmospheric boundary-layer turbulence from time series measurements (e.g., Tsonis and Elsner, 1988; Jaramillo and Puente, 1993; Xin et al., 2001; Gallego et al., 2001), finding values between 3 and 9. Because of the complexity of the atmosphere, many researchers, including the originators of the method, doubt the appropriateness of such small correlation dimensions (e.g., Procaccia, 1988; Pool, 1989; Lorenz, 1991). As Procaccia (1988) stated in reference to the study of Tsonis and Elsner (1988), the computation of small correlation dimensions can stem from one of two reasons: (i) There is something fundamental in the dynamics of the atmosphere that we do not understand; or (ii) the method of analysis does not apply, or is outside, the range of validity. One criticism that the method of analysis has raised is the use of data sets that are too small for the computation of the correlation dimension (e.g., Procaccia, 1988; Pool, 1989; Lorenz, 1991). Others have also pointed out that noise in the data, poor resolution, and the repetition of one number many times in the data set can lead to incorrect estimation of the correlation dimension (see Sivakumar, 2000 for review).

The application of chaos theory is limited by the fundamental assumption that the time series used to analyze chaos are infinite in resolution and size. In practice, however, infinite time series do not exist and one must use a limited time series sampled at a discrete number of points. Assuming a time series of high resolution, the minimum number of points, N_{\min} , needed to accurately construct the attractor has been a matter of some debate (Sivakumar, 2000). Numerous recommendations have been made including $N_{\min} = 10^d$ (Procaccia, 1988) and $N_{\min} = 42^d$ (Smith,

1988), where d is the dimension of the attractor. In estimating the dimension of the attractor using wind velocity and other measured meteorological variables, Tsonis and Elsner (1988) analyzed 3,960 measured data points, Jaramillo and Puento (1993) used data sets of 18,000 points, Xin et al. (2001) analyzed data sets of 28,000 and 36,000 data points, and Gallego et al. (2001) used data sets of 32,000 data points. The number of data points used in these studies is small compared to the value of N_{\min} recommended by Procaccia (1988) and Smith (1988). For a correlation dimension of 9, it would be recommended to have a minimum of 10^9 (Procaccia, 1988) or 4.067×10^{14} measured data points (Smith, 1988). It has been shown that small data sets will result in systematic underestimation of the correlation dimension (Procaccia, 1988; Havstad and Ehlers, 1989; Lorenz, 1991). For example, Lorenz (1991) demonstrated that using a data set of only 4,000 points to analyze a known mathematical system of 21 coupled variables leads to gross underestimation of the correlation dimension when the variables are ‘weakly’ coupled. In the atmosphere, much of the coupling is weak and indirect.

Some studies have suggested that one does not need to place such strict restrictions on N_{\min} . For example, Nerenberg and Essex (1990) suggested that the estimate of N_{\min} by Smith (1988) was too extreme and that a more appropriate estimate is $N_{\min} \sim 10^{2+0.4d_c}$. Though data sets smaller than those recommended by Smith (1988) might be useful in the computation of d_c , one must remember that d_c is defined as $r \rightarrow 0$. When N is too small, the lowest value of r , for which $m(r)$ is statistically meaningful, may not be particularly close to zero (Ruelle, 1990). In this way, when researchers use data sets that are too small, they are assuming that if N were made large enough, the decrease of $m(r)$ with r , when r is very small, is the same as when r is fairly large (Lorenz, 1991). This implicitly assumes that the fine structure of the attractor resembles that of the coarse structure, and though this has been shown to be true in some simple systems, it does not necessarily hold for more complex systems (Lorenz, 1991).

Procaccia (1988) pointed out that the low values estimated for the correlation dimension of measured meteorological variables could be attributed to the fact that deterministic strange attractors are dense with unstable orbits. When the number of data points is small, one could be analyzing one of these unstable orbits and not the whole attractor. Lorenz (1991) echoed this and suggested that the atmosphere should be viewed loosely as a coupled set of lower-dimensional subsets, and that the studies estimating low values of the correlation dimension for atmospheric turbulence, as they were performed, attempted to measure the dimension of a subsystem.

We believe that the arguments against the existence of a small correlation dimension of the atmosphere are reasonable and convincing. It is for this reason that we only focus on the time-delayed mutual information values without progressing to the estimation of the correlation dimension. Given the current method, the physical restraints of collecting ‘infinite’ data sets, and the high dimensionality of the atmosphere, we do not believe that computation of the correlation dimension in

TABLE I

Summary of site characteristics, the vegetation heights h , the measurement heights z_m from the forest floor or soil surface, the surface roughness lengths z_o , and measured meteorological variables for the nine experiments.

Surface cover	Date	h (m)	z_m (m)	z_o (m)	Sampling frequency (H)	Sampling period (min)
Pine Trees	Jan. 4–8, 2000	14	14.5	1.40	5	27.3
	May 16–20, 2000					
	Aug. 22–26, 2000					
	Oct. 20–23, 2000					
Hardwood Trees	June 15–July 11, July 23, 1996	35	35	3.50	10	27.3
	July 12–16, 1995					
Grass	Sept. 18–21, 1996	1	5.2	0.10	56	19.5
Grass	May 14–25, 2001	1	3.75	0.10	5/10	13.6
Grass	July 7–27, 2001	0.3/0.05	3.0	0.03–0.005	10	27.3
Sand	June–July 1993	N/A	2–3.5	0.00013	56	9.75

this paper would be of benefit to the scientific community. Instead, we compute the mutual information content, which is more accurate to the objectives of this study.

3. Experimental Setup

Eddy correlation measurements used in this study were taken at four different sites in six different experiments. Some statistics, such as the dates of the experiments, the vegetation height h , the measurement height z_m , the surface roughness length z_o , and the sampling period and frequency, are listed in Table I. Maximum, minimum, and mean values of the mean longitudinal wind velocity $\langle u \rangle$; the variance of w , σ_w^2 ; and mean atmospheric temperature, $\langle T \rangle$ for each site are listed in Table II. Typical values of other meteorological variables were calculated and listed in Table II. These variables include the friction velocity,

$$u_* = \sqrt{-\langle u'w' \rangle}; \quad (7)$$

the Obukhov length,

$$L = -\frac{u_*^3 \langle T \rangle}{\kappa g \langle w'T' \rangle}; \quad (8)$$

TABLE II

Range of values and mean values of the six major, measured meteorological variables used in the nonlinear time series analysis.

Site description		z_m/L	$\langle u \rangle$ (m s ⁻¹)	u_* (m s ⁻¹)	σ_w^2 (m ² s ⁻²)	H (W m ⁻²)	$\langle T \rangle$ (°C)
Pine forest	Max.	0.0	2.92	1.03	1.06	282.67	32.09
	Min.	-28.83	0.18	0.01	0.02	0	2.70
	Mean	-0.89	1.02	0.34	0.17	78.67	18.71
Hardwood forest	Max.	~0.0	3.79	1.04	1.67	330.06	30.88
	Min.	-11.67	0.39	0.09	0.05	0.74	2.16
	Mean	-0.45	1.67	0.49	0.32	125.70	15.83
Grass-covered surface (1995)	Max.	-0.01	5.47	0.46	0.23	374.26	35.37
	Min.	-10.11	0.72	0.07	0.02	0.37	25.16
	Mean	-0.43	2.14	0.24	0.12	140.72	28.47
Grass-covered surface (1996)	Max.	0.0	4.25	0.59	0.49	159.06	28.24
	Min.	-10.28	0.54	0.08	0.02	3.69	17.82
	Mean	-0.45	2.28	0.28	0.15	63.65	21.74
Grass-covered surface (2001)	Max.	0.0	3.39	0.50	0.32	290.54	33.31
	Min.	-2.97	0.26	0.02	0.01	0.0	8.97
	Mean	-0.49	1.57	0.18	0.09	64.62	23.10
Dry Owens	Max.	-0.04	7.52	0.38	0.24	288.08	44.42
Lake Bed	Min.	-1.64	1.71	0.06	0.01	0	10.35
	Mean	-0.36	4.34	0.22	0.09	71.53	20.57

and the sensible heat flux,

$$H = \rho c_p \langle w'T' \rangle, \quad (9)$$

where $\langle x \rangle$ and x' symbolize the time averaging and the deviation from the mean of a measured variable x , u is the longitudinal wind velocity, T is atmospheric temperature, κ ($= 0.4$) is the von Karman constant, g is the acceleration due to gravity, ρ is the density of air, and c_p is the specific heat capacity of air. Atmospheric stability is denoted by $\xi = z/L$, where $z = z_m - d$ and d is the displacement height. Only unstable and near-neutral stability conditions, where $L < 0$, are used in the analysis and shown in Table II.

3.1. PINE FOREST SITE

The pine forest is located in the Blackwood Division of the Duke Forest (35°58' N, 79°08' W, elevation 163 m) in Durham, North Carolina, U.S.A. The site is a uniformly aged Loblolly pine forest that extends 1000 m in the north-south direction

and 600 m in the east–west direction. Though the stand was grown from pine seedlings planted in 1983 following a clear cutting and burning, the pine forest has a significant hardwood understorey. The eddy correlation system used is composed of a Campbell Scientific Krypton hygrometer co-located with a triaxial sonic anemometer and the gas inlet for a LICOR 6262 CO₂/H₂O infrared gas analyzer at $z_m/h \approx 1.0$. The experiment took place during several different months in the year 2000. Further details of the site description, the understorey species, and the measurement apparatus are given in Katul et al. (1999), Katul and Albertson (1999), and Lai et al. (2000 a, b).

3.2. HARDWOOD FOREST SITE

The hardwood forest is also located in the Blackwood Division of the Duke Forest and is composed of unevenly aged, mature second growth, deciduous hardwood trees approximately 35 m in height. The tower at this site is a 40-m walk-up tower on which the eddy correlation system, composed of the triaxial sonic anemometer and the Campbell Scientific Krypton hygrometer, is mounted at $z_m/h = 1.0$. The area surrounding the tower, known as Meadow Flats, is fairly level for about 500 m in all directions. More details concerning the site location and set-up and the stand's species composition can be found in Katul et al. (1997) and Conklin (1994), respectively.

3.3. GRASS-COVERED CLEARING SITE

The site is a 480 m by 305 m grass covered-clearing located in the Blackwood Division of the Duke Forest. For the 1995 and 1996 experiments, the eddy correlation system was mounted on a mast located 250 m from the north-end and 160 m from the west-end of the Loblolly pine forest. For the 2001 experiment, the mast was removed and replaced by a 6-m walk-up tower; the eddy correlation system was mounted on the tower. During the 2001 experiment, the grass was mowed on July 5, 2001 such that the height of the grass was reduced from 0.3 m to approximately 0.05 m. Information on the height at which the instruments were mounted, the vegetation height, and the sampling frequency and period for each of the three experiments is listed in Table I. Further details on the 1995 and 1996 experiments can be found in Szilagyi et al. (1996) and Katul et al. (1998).

3.4. DRY OWENS LAKE BED

The field experiment was carried out in late June to early July in 1993 at the dry Owens Lake bed in Owens Valley, California, U.S.A. Dry Owens Lake bed is part of a larger basin and is surrounded by the Sierra Nevada to the west and the White and Inyo Mountains to the east. The dry lake bed is covered with a crusted sand surface with substantial amounts of evaporated salt. Meteorological measurements were made using a Campbell Scientific eddy correlation system with z_m ranging

from 2.0 m to 3.5 m. More information on this experiment can be found in Katul (1994), Katul et al. (1995), and Albertson et al. (1995).

4. Analysis and Results

We apply the Shannon entropy, Lorentz thresholding, and mutual information measures to w for the six experiments and analyze the results for unstable stability conditions where $z/L < 0$. The experiments were grouped into two categories based on whether measurements were taken in the ASL or CSL. Measurements at the pine forest and hardwood forest sites were taken at approximately $z_m/h = 1$, and therefore, were made in the CSL. At the grass-covered clearing and the dry Owens Lake bed sites, measurements were taken at $z_m > 3h$ such that these experiments were performed in the ASL.

4.1. SHANNON ENTROPY

The Shannon entropy, estimated based on the power spectrum, is computed for w . We depart from the traditional definition of entropy and employ the power spectrum for two reasons: (1) The probability density function of w in the CSL and ASL is approximately Gaussian for both layers (Katul et al., 1997); hence, probability-based Shannon entropy measures cannot discern organization differences between CSL and ASL; and (2) we follow the analysis of Wijesekera and Dillon (1997) who demonstrated that usefulness of using the spectrum instead of the probability density function for laboratory and ocean turbulent mixing. For further calculations, we will use the normalized Shannon entropy,

$$S_N = I_S / \ln(M), \quad (10)$$

where $\ln(M)$ is the largest possible value of the Shannon entropy and is produced by a flat, white noise spectrum. S_N is guaranteed to be between 0 and 1. This normalization ensures that entropy differences attributed to different sampling lengths and duration (e.g., Table I) are minimized.

The values of the Shannon entropy for the power spectrum of w were computed for the experiments in the ASL and CSL for unstable and near-neutral conditions and are shown in Figure 2. The solid and dashed lines illustrate the mean behaviour of the Shannon entropy with changes in stability, $\ln(-z/L)$, for experiments in the ASL and the CSL, respectively. The vertical bars represent one standard deviation. As the stability changes from near-neutral conditions ($z/L \approx 0$) to near-convective conditions ($z/L \ll 0$), the Shannon entropy decreases. On average, the value of the Shannon entropy in the ASL is greater than that in the CSL for all unstable conditions. This implies that the flow organization is higher in the CSL for all stability classes. Furthermore, for both ASL and CSL flow types, we find that as the flow approaches near-convective conditions, entropy drops and the level of

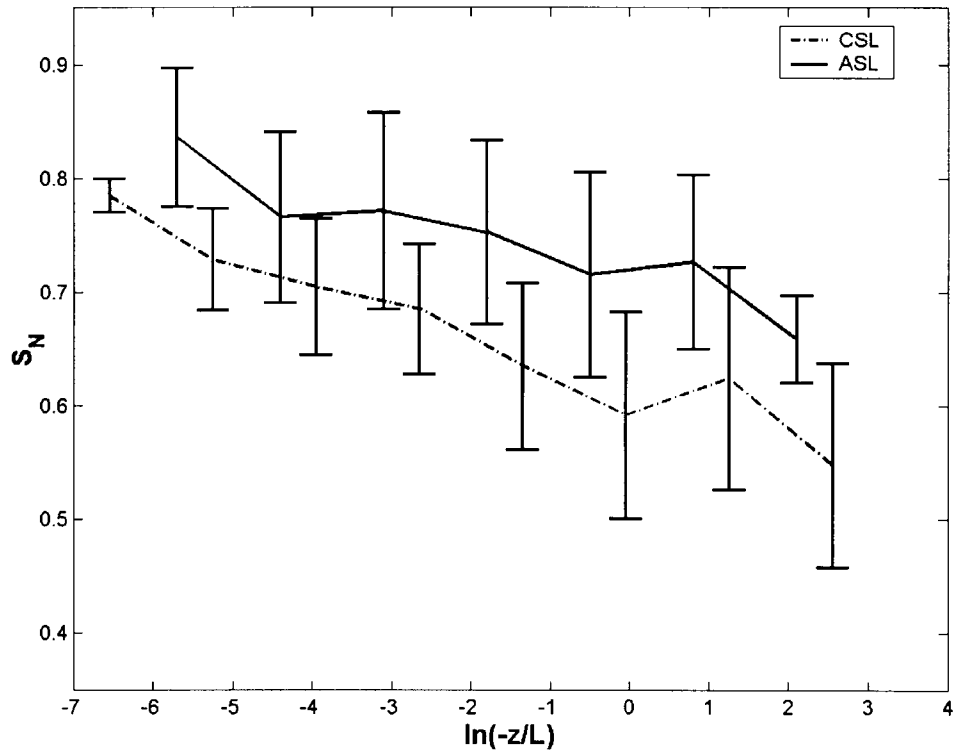


Figure 2. Modified Shannon Entropy is plotted versus stability for unstable and near-neutral conditions, $\ln(-z/L)$. The solid and dashed lines show the ensemble behaviour of the Shannon entropy with changes in stability for experiments in the ASL and the CSL, respectively. The vertical bars represent one standard deviation.

organization increases. When contrasting the effects of stability vis-à-vis the flow type (i.e., CSL versus ASL) on entropy, Figure 2 suggests that organization due to stability is much more critical than flow type. We explore whether the other measures confirm the above two findings.

4.2. WAVELET THRESHOLDING METHODS

The simple, orthogonal Haar wavelet, is used to compute the Lorentz curve of w . This wavelet is chosen because of its simplicity and because it has been found to be the optimal wavelet for identifying energy-containing turbulent eddy motion in the wavelet domain (Katul and Vidakovic, 1996). Furthermore, Katul and Vidakovic (1998) conducted a detailed analysis to quantify how the choice of wavelet biases L_o . They found that variations in L_o due to the choice of the wavelet function are minor.

The percentage of coefficients L_o retained after applying the Lorentz thresholding criteria to w were computed for unstable and near-neutral conditions and shown

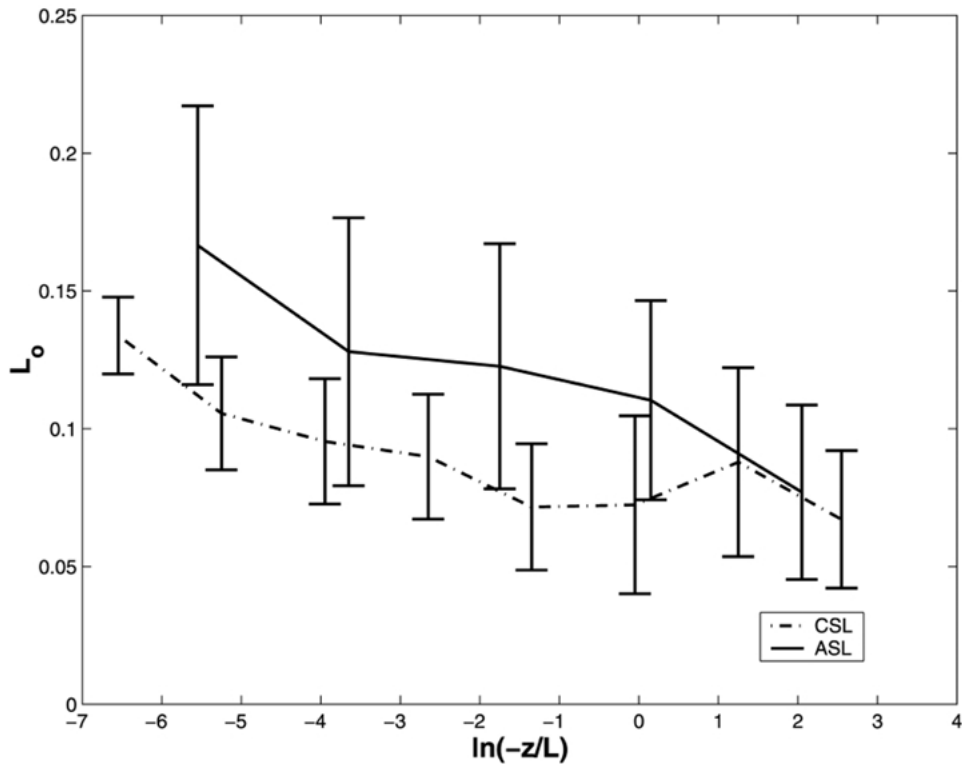


Figure 3. The percentage of coefficients L_o retained after applying the Lorentz thresholding criteria to w for each of the experiments is plotted versus stability for unstable and near-neutral conditions, $\ln(-z/L)$. The solid and dashed lines show the mean behaviour of L_o with changes in stability for experiments in the ASL and the CSL, respectively. The vertical bars represent one standard deviation.

in Figure 3. The solid and dashed lines show the mean variations in the calculated values of L_o with changes in stability for experiments in the ASL and the CSL. The vertical bars represent one standard deviation. Recall that higher L_o implies lower level of organization in the time series. On average, L_o decreases as the stability conditions change from near neutral to very unstable. There also seems to be some pattern of distinction between the value of L_o and the surface roughness conditions; the average value of L_o is larger for experiments in the ASL than for those in the CSL, though the scatter among the results is large. Both of these findings are consistent with the previous entropy calculations, and imply that the flow in the CSL is more organized than the ASL, and that near convective conditions are more organized than near-neutral conditions.

4.3. MUTUAL INFORMATION

We compute the time-delayed mutual information of w for the experiments in the ASL and CSL. To compare each signal, the value of the lag τ_M at the first minimum

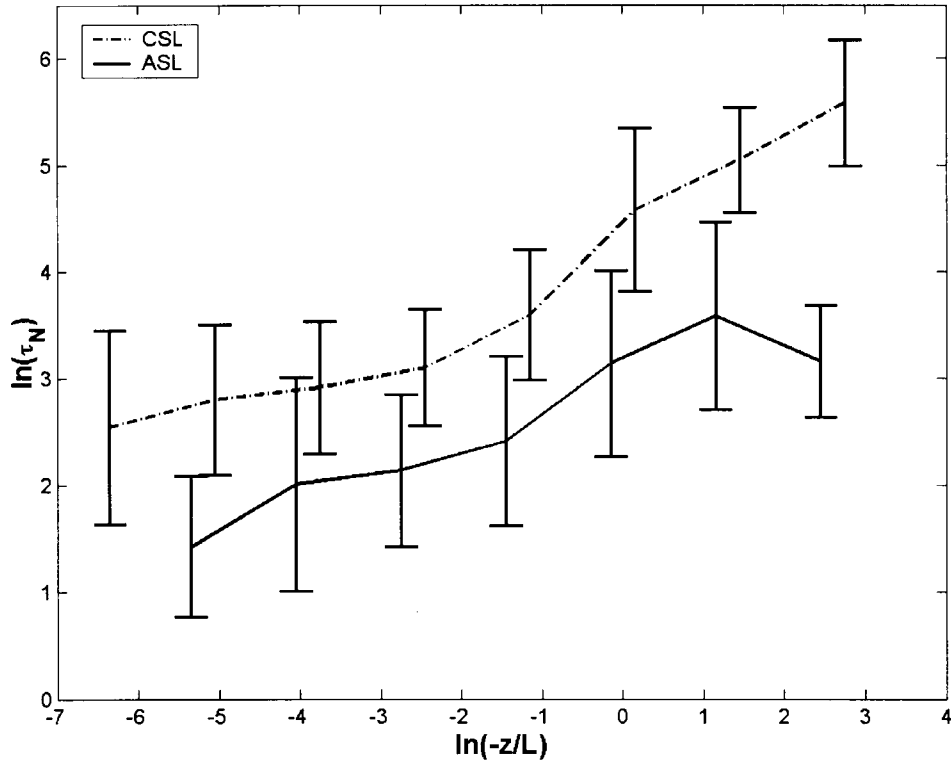


Figure 4. τ_N , the normalized, optimum lag time for the mutual information is plotted versus stability for unstable and near-neutral conditions, $\ln(-z/L)$. The solid and dashed lines show the mean variations in the calculated values of τ_N with changes in stability for experiments in the ASL and the CSL. The vertical bars are for one standard deviation.

in $I_M(\tau)$ is used, as advocated by Fraser and Swinney (1986). At this value, τ_M , the signals of $w(t)$ and $w(t + \tau_M)$ are relatively uncorrelated but still remain similar enough to allow the reconstruction of the attractor. In Figure 4, τ_M is plotted versus stability where τ_m is non-dimensionalized by multiplying by z/u_* such that

$$\tau_N = \tau_M \frac{z}{u_*}. \quad (11)$$

The solid and dashed lines in Figure 4 show the mean variations in the calculated values of τ_N with changes in stability for experiments in the ASL and the CSL. The vertical bars are for one standard deviation. On average, τ_N increases as stability conditions range from near neutral to very unstable. The mutual information is also able to discern the effect of surface roughness. The mean value of τ_N for experiments in the CSL is significantly greater than those in the ASL for all unstable stability conditions. The larger τ_N , the more organized the flow. Again, these results confirm the findings from the entropy calculations: near convectively produced

turbulence is more organized than mechanically produced turbulence, and the CSL is more coherent than the ASL.

The next logical step of this analysis is to compute the correlation dimension. The larger the actual correlation dimension, and therefore the complexity, the larger the amount of data needed. Given the number of points that compose our data sets of measured wind velocity ($<10^5$ measured points, see Table I) and the complex nature of atmospheric turbulence, we have decided not to attempt to estimate the correlation dimension. With such small data sets it is probable that the whole attractor would not be analyzed. Analyses of a subset of the atmospheric attractor will add little insight into the effect of stability and surface roughness on atmospheric turbulence.

5. Conclusions

This analysis presented three simple techniques that contrast the level of organization in the vertical wind velocity time series collected in the ASL and CSL for a wide range of stability conditions. The objective was to separate the degree of organization attributed to stability and that attributed to roughness. Using Shannon entropy, Lorentz wavelet thresholding, and mutual information content, we found that the vertical velocity in the CSL is more organized than in the ASL. To some extent, this conclusion may seem counter-intuitive. The statistics of the flow in the CSL are influenced by canopy morphology and appear to be three-dimensional when compared to the simpler one-dimensional statistics in the ASL. However, if one considers the mode of instability producing the organized eddy motion in the CSL, then this paradox is resolved. For an extensive canopy such as the pine and hardwood forests, the organized eddy motion in the CSL is produced by Kelvin–Helmholtz type instabilities that can be predicted from two-dimensional linear stability analysis (Raupach et al., 1996). Mixing layers are known to have a ‘long-memory’ in the sense that the instability mode remains the energetic mode as more eddy sizes are produced and dissipated. This is not the case for ASL flows, in which the mode of instability originating from interactions between the surface and the fluid is rapidly distorted by an entire population of new eddies (Finnigan, 2000). Hence, the findings from this analysis are consistent with the view that the organized eddies must be originating from a flow instability whose signature remains persistent, such as a mixing layer.

Finnigan and Shaw (2000) also found results that support this conclusion. Conducting empirical orthogonal function (EOF) analysis, Finnigan and Shaw (2000) compared the contribution of partial sums of the eigenmodes to the vertical distribution of $\langle uw \rangle$ in both the roughness sublayer alone and the lower surface layer plus the roughness sublayer. They found that convergence of the eigenmodes was indeed significantly more rapid in the roughness sublayer than in the surface layer. These results support the findings of this paper and present objective evidence that

the turbulence in the CSL is dominated by more distinct eddy structures than in the ASL (Finnigan, 2000).

Unfortunately, the EOF method is difficult to apply to single-point measurements because the required empirical eigenfunctions are those of two-point velocity covariance tensor and the data must, therefore, be collected in particularly intensive experiments. The novelty of the three techniques proposed in this paper is that, unlike the EOF method, these measures can be applied to time series measurements now collected in many field experiments (e.g., see Baldocchi et al., 2001). These techniques are objective measures of sensitivity to organization and do not require *a priori* knowledge about these organized structures. In addition, the three techniques presented here are able to investigate how stability alters the degree of organization in both the CSL and ASL. We found that as the flows in both layers evolves from near-neutral to unstable conditions, the level of organization in w increases. We note that as the flow approaches pure convective conditions, the traditional definitions of the CSL and ASL as defined in this paper become less applicable. However, for near-neutral conditions, the increase in the degree of organization in w due to stability by far exceeds any difference noted between the CSL and ASL.

The fact that the level of organization did not converge in the CSL and ASL suggests buoyant or thermal plumes alone are not the key mechanism responsible for the organization in the two layers. It is conceivable that with increasing temperature in the CSL and reduced velocity differences between inside the canopy and the air aloft, a thermal mixing layer, influencing vertical velocities, also emerges. A pure thermal mixing is formed when two co-flowing fluids with the same longitudinal velocity but different temperatures are allowed to mix. The resulting instability leading to organized eddy motion is very distinct from the classic convective mixing. These instabilities are also different from the Kelvin-Helmholtz waves in neutral mixing layers.

Acknowledgements

We thank Judd Edburn and the Duke Forest staff for their overall assistance during many of these field experiments. We would also like to thank Henry Greenside for his helpful comments on the application and computation of the mutual information and correlation dimension. Support was provided by the National Science Foundation through their WEAVE program (NSF-EAR) and directorate of Mathematical Statistics (NSF-DMS), the Department of Energy (DOE) through their FACEFACTS project, Terrestrial Carbon Program (TCP), and the National Institute of Global Environmental Change (NIGEC) through the Southeast Regional Center at the University of Alabama, Tuscaloosa. The third author also thanks the 'Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)' of Brazil for their support.

References

- Albertson, J. D., Parlange, M. B., Katul, G. G., Chu, C-R., Stricker, H., and Tyler, S.: 1995, 'Sensible Heat Flux from Arid Regions: A Simple Flux-Variance Method', *Water Resour. Res.* **31**, 969–973.
- Baldocchi, D. D., Falge, E., Gu, L. H., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X. H., Malhi, Y., Meyers, T., Munger, W., Oechel, W. U. K. T. P., Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T., Wilson, K., Wofsy, S.: 2001, 'FLUXNET: A New Tool to Study the Temporal and Spatial Variability of Ecosystem-Scale Carbon Dioxide, Water Vapor, and Energy Flux Densities', *Bull. Amer. Meteorol. Soc.* **82**, 2415–2434.
- Conklin, P.: 1994, 'Turbulent Wind, Temperature, and Pressure in a Mature Hardwood Canopy', Ph.D. Dissertation, School of the Environment, Duke University, Durham, N.C., 105 pp.
- Farge, M., Goirand, E., Meyer, Y., Pascal, F., and Wickerhauser, M. V.: 1992, 'Improved Predictability of Two-Dimensional Turbulent Flows Using Wavelet Packet Compression', *Fluid Dyn. Res.* **10**, 229–250.
- Finnigan, J.: 2000, 'Turbulence inside Plant Canopies', *Annu. Rev. Fluid Mech.* **32**, 519–571.
- Finnigan, J. and Shaw, R. H.: 2000, 'A Wind-Tunnel Study of Airflow in Waving Wheat: An EOF Analysis of the Structure of the Large-Eddy Motion', *Boundary-Layer Meteorol.* **96**, 211–255.
- Fraser, A. M.: 1989, 'Information and Entropy in Strange Attractors', *IEEE Trans. Inform. Theory* **35**, 245–262.
- Fraser, A. M. and Swinney, H. L.: 1986, 'Independent Coordinates for Strange Attractors from Mutual Information', *Phys. Rev. A* **33**, 1134–1140.
- Gallego, M. C., García, J. A., and Cancillo, M. L.: 2001, 'Characterization of Atmospheric Turbulence by Dynamical Systems Techniques', *Boundary-Layer Meteorol.* **100**, 375–392.
- Grassberger, P. and Procaccia, I.: 1983, 'Measuring the Strangeness of Strange Attractors', *Physica D* **9**, 189–208.
- Havstad, J. W. and Ehlers, C. L.: 1989, 'Attractor Dimension of Nonstationary Dynamical Systems from Small Data Sets', *Phys. Rev. A* **39**, 845–853.
- Jaramillo, G. P. and Puente, C. E.: 1993, 'Strange Attractors in Atmospheric Boundary-Layer Turbulence', *Boundary-Layer Meteorol.* **64**, 175–197.
- Kaiser, J.: 1998, 'Climate Change – New Network Aims to Take the World's CO₂ Pulse Source', *Science* **281**, 506–507.
- Katul, G. G.: 1994, 'A Model for Sensible Heat-Flux Probability Density-Function for Near-Neutral and Slightly-Stable Atmospheric Flows', *Boundary-Layer Meteorol.* **71**, 1–20.
- Katul, G. G. and Albertson, J. D.: 1999, 'Modeling CO₂ Sources, Sinks, and Fluxes within a Forest Canopy', *J. Geophys. Res.* **104**, 6081–6091.
- Katul, G. G. and Vidakovic, B.: 1996, 'The Partitioning of the Attached and Detached Eddy Motion in the Atmospheric Surface Layer Using Lorentz Wavelet Filtering', *Boundary-Layer Meteorol.* **77**, 153–172.
- Katul, G. G. and Vidakovic, B.: 1998, 'Identification of Low-Dimensional Energy Containing/Flux Transporting Eddy Motion in the Atmospheric Surface Layer Using Wavelet Thresholding Methods', *J. Atmos. Sci.* **55**, 377–389.
- Katul, G. G., Goltz, S. M., Hsieh, C-I., Cheng, Y., Mowry, F., and Sigmon, J.: 1995, 'Estimation of Surface Heat and Momentum Fluxes Using the Flux-Variance Method above Uniform and Non-Uniform Terrain', *Boundary-Layer Meteorol.* **74**, 237–260.
- Katul, G. G., Hsieh, C-I., Bowling, D., Clark, K., Shurpali, N., Turnipseed, A., Albertson, J., Tu, K., Hollinger, D., Evans, B., Offerle, B., Anderson, D., Ellsworth, D., Vogel, C., and Oren, R.: 1999, 'Spatial Variability of Turbulent Fluxes in the Roughness Sublayer of an Even-Aged Pine Forest', *Boundary-Layer Meteorol.* **93**, 1–28.
- Katul, G. G., Hsieh, C-I., Kuhn, G., and Ellsworth, D.: 1997, 'Turbulent Eddy Motion at the Forest-Atmosphere Interface', *J. Geophys. Res.* **102**(D12), 13409–13421.

- Katul, G. G., Schieldge, J., Hsieh, C.-I., and Vidakovic, B.: 1998, 'Skin Temperature Perturbations Induced by Surface Layer Turbulence above a Grass Surface', *Water Resour. Res.* **34**, 1265–1274.
- Lai, C.-T., Katul, G. G., Ellsworth, D. S., and Oren, R.: 2000a, 'Modeling Vegetation-Atmosphere CO₂ Exchange by a Coupled Eulerian-Lagrangian Approach', *Boundary-Layer Meteorol.* **95**, 91–122.
- Lai, C.-T., Katul, G. G., Oren, R., Ellsworth, D. S., and Schäfer, K.: 2000b, 'Modeling CO₂ and Water Vapor Turbulent Flow Distributions within a Forest Canopy', *J. Geophys. Res.* **105**, 26333–26351.
- Lorenz, E. N.: 1991, 'Dimension of Weather and Climate Attractors', *Nature* **353**, 241–244.
- Nerenberg, M. A. H. and Essex, C.: 1990, 'Correlation Dimension and Systematic Geometric Effects', *Phys. Rev. A* **42**, 7065–7074.
- Palus, M.: 1993, 'Identifying and Quantifying Chaos Using Information-Theoretic Functionals', in A. S. Weigend and N. A. Gershenfeld (eds.), *Time-Series Prediction: Forecasting the Future and Understanding the Past*, Addison-Wesley, Reading, MA, pp. 387–413.
- Pineda, F. J. and Sommerer, J. C.: 1993, 'Estimating Generalized Dimensions and Choosing Time Delays: A Fast Algorithm', in A. S. Weigend and N. A. Gershenfeld (eds.), *Time-Series Prediction: Forecasting the Future and Understanding the Past*, Addison-Wesley, Reading, MA, pp. 367–385.
- Pool, R.: 1989, 'Is Something Strange about the Weather?', *Science* **243**, 1290–1293.
- Procaccia, I.: 1988, 'Complex or Just Complicated?', *Nature* **333**, 498–499.
- Raupach, M. R. and Thom, A. S.: 1981, 'Turbulence in and above Canopies', *Annu. Rev. Fluid Mech.* **13**, 97–129.
- Raupach, M. R., Finnigan, J. J., and Brunet, Y.: 1989, 'Coherent Eddies in Vegetation Canopies', in *Australian Conference on Heat and Mass Transfer*, Christchurch, New Zealand, pp. 75–90.
- Raupach, M. R., Finnigan, J. J., and Brunet, Y.: 1996, 'Coherent Eddies and Turbulence in Vegetation Canopies: The Mixing Layer Analogy', *Boundary-Layer Meteorol.* **78**, 351–382.
- Ruelle, D.: 1990, 'Deterministic Chaos: The Science and the Fiction', *Proc. Roy. Soc. Lond. A* **427**, 241–248.
- Shannon, C. E.: 1948, 'A Mathematical Theory of Communications', *Bell Syst. Tech. J.* **27**, 379–623.
- Sivakumar, B.: 2000, 'Chaos Theory in Hydrology: Important Issues and Interpretations', *J. Hydrol.* **227**, 1–20.
- Skilling, J. and Gull, S. F.: 1985, 'Algorithm and Applications', in C. R. Smyth and W. T. Grandy Jr. (eds.), *Maximum-Entropy and Bayesian Method in Inverse Problems*, D. Reidel, Dordrecht, pp. 83–131.
- Smith, L. A.: 1988, 'Intrinsic Limits on Dimension Calculations', *Phys. Lett. A* **133**, 283–288.
- Szilagyi, J., Katul, G. G., Parlange, M. B., Albertson, J. D., and Cahill, A. T.: 1996, 'The Local Effect of Intermittency on the Inertial Subrange Energy Spectrum of the Atmospheric Surface Layer', *Boundary-Layer Meteorol.* **79**, 35–50.
- Townsend, A. A.: 1976, *The Structure of Turbulent Shear Flow*, Cambridge University Press, Cambridge, 429 pp.
- Tsonis, A. A. and Elsner, J. B.: 1988, 'The Weather Attractor over Very Short Timescales', *Nature* **333**, 545–547.
- Vidakovic, B.: 1995, 'Unbalancing Data with Wavelet Transformations', *Wavelet Applications in Signal and Image Processing III, Proc. SPIE 2569* **2**, 845–857.
- Wickerhauser, M. V., Farge, M., Goirand, E., Wesfried, E., and Cubillo, E.: 1994, 'Efficiency Comparison of Wavelet Packet and Adapted Local Cosine Bases for Compression of a Two-Dimensional Turbulent Flow', in C. K. Chui, L. Montefusco, and L. Puccio (eds.), *Wavelets: Theory, Algorithms, and Applications*, Academic Press, San Diego, pp. 509–531.
- Wijesekera, H. W. and Dillion, T. M.: 1997, 'Shannon Entropy as an Indicator of Age Turbulent Overturns in the Oceanic Thermocline', *J. Geophys. Res.* **102**, 3279–3291.
- Williams, G. P.: 1997, *Chaos Theory Tamed*, Joseph Henry Press, Washington, DC, 499 pp.

- Xin, L., Fei, H., and Gang, L.: 2001, 'Characteristics of Chaotic Attractors', *Boundary-Layer Meteorol.* **99**, 335–345.
- Zubair, L., Sreenivasan, K. R., and Wickerhauser, M. V.: 1992, 'Turbulent Signals and Images Using Wavelet-Packets', in T. Gatski, S. Sarkar, and C. G. Speziale (eds.), *Studies in Turbulence*, Springer-Verlag, New York, 602 pp.

