Horizontal structure of marine boundary layer clouds from centimeter to kilometer scales

Anthony B. Davis, 1 Alexander Marshak, 2 4 H. Gerber, 3 and Warren J. Wiscombe 4

Abstract. Horizontal transects of cloud liquid water content (LWC) measured at unprecedented 4-cm resolution are statistically analyzed scale-by-scale. The data were collected with a Particulate Volume Monitor (PVM) probe during the winter Southern Ocean Cloud EXperiment (SOCEX) on July 26, 1993, in a broken-stratocumulus/towering-cumulus cloud complex. Two scaling regimes are found in the sense that two distinct power laws, $k^{-\beta}$, are needed to represent the wavenumber spectrum $E(k)$ over the full range of scales $r \approx 1/k$. Detailed numerical simulations show that the scale break at 2–5 m is not traceable to the normal variability of LWC in the PVM’s instantaneous sampling volume (1.25 cm$^3$) driven by Poissonian fluctuations of droplet number and size. The two regimes therefore differ physically. The non-Poissonian character of the small-scale LWC variability is consistent with a similar finding by Baker [1992] for droplet number concentration obtained from Forward Scattering Spectrometer Probe (FSSP) data: at scales of a few centimeters, spatial droplet distributions do not always follow a uniform Poisson law. With $\beta = 0.9 \pm 0.1$, the small-scale ($8–12 \text{ cm} \leq r \leq 2–5 \text{ m}$) regime is stationary: jumps in LWC are highly variable in size and rapidly cancel each other, leading to short-range correlations. By contrast, the large-scale ($5 \text{ m} \leq r \leq 2 \text{ km}$) variability with $\beta = 1.6 \pm 0.1$ is nonstationary: jumps are generally quite small, conveying a degree of pixel-to-pixel continuity and thus building up long range correlations in the low-pass filtered signal. The large-scale structure of the complex SOCEX cloud system proves to be multifractal, meaning that large jumps do occur on an intermittent basis, that is, on a sparse fractal subset of space. Low-order, hence more robust, multifractal properties of the SOCEX clouds are remarkably similar to those of their First ISCCP Regional Experiment (FIRE) and Atlantic Stratocumulus Transition Experiment (ASTEX) counterparts, and also to those of passive scalars in fully developed turbulence. This is indicative of a remarkable similarity in the microphysical and macrophysical processes that determine cloud structure in the marine boundary layer at very remote locales, especially since the particular SOCEX cloud system investigated here was rather atypical. Interesting differences are also found: in the scaling ranges on the one hand, and in higher order moments on the other hand. Finally, we discuss cloud radiative effects of the large and small-scale variabilities.

1. Introduction and Overview

Internal variability of stratiform clouds has a first-order effect on the Earth’s radiation budget [Cahalan, 1994] and on the remote sensing of cloud properties [Davis et al., 1997a]. This variability affects scales ranging from the size of a high-resolution satellite footprint to that of a general circulation model (GCM) grid box. To better understand cloud microphysics and cloud-radiation interactions, vast amounts of data on the distribution of atmospheric liquid water from a variety of sources have been collected during the last decade. In situ measurements of cloud liquid water content (LWC) have been obtained during many intensive field programs among which we will take a particular interest in (1) the First ISCCP Regional Experiment (FIRE) in June–July 1987 [Cox et al., 1987], (2) the Atlantic Stratocumulus Transition Experiment (ASTEX) in summer 1992 [Albrecht et al., 1995], and (3) the Southern Ocean Cloud Experiment in July 1993 for its winter phase (SOCEX 1) [Boers et al., 1996]. In Table 1 we list selected characteristics of these three experiments. Another, less direct, source of information on cloud structure is column integrated LWC, or liquid water path (LWP), retrieved from passive microwave radiometry during FIRE, ASTEX, and SOCEX. LWP is also recorded continuously at the Atmospheric Radiation Measurement (ARM) Southern Great Plains facility and elsewhere. All these data contain information on spatial and/or temporal correlations in cloudiness, enabling scale-by-scale analyses over a range from a few centimeters to hundreds of kilometers.

Cahalan and Snider [1989] found excellent scaling in wavenumber spectra in day-long time series of 1-min averages of LWP recorded during FIRE. Using LWP data from ASTEX at a similar resolution, Gallmer et al. [1995] find scale-invariance in wavelet coefficients. Wiscombe et al.’s [1995] analysis of LWP data from the ARM archive reveals multifractal be-
Table 1. Comparison of IWC Measurements and Basic Variability Parameters From Three Field Programs

<table>
<thead>
<tr>
<th></th>
<th>SOCEX</th>
<th>ASTEX</th>
<th>FIRE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Where?</td>
<td>Tasmania</td>
<td>Azores</td>
<td>Southern California</td>
</tr>
<tr>
<td>Cloud type?</td>
<td>broken Sc, embedded towering Cu</td>
<td>marine Sc-to-Cu transition</td>
<td>marine Sc</td>
</tr>
<tr>
<td>IWC Probe (type)</td>
<td>PVM (optical)</td>
<td>PVM (optical)</td>
<td>King (hot wire)</td>
</tr>
<tr>
<td>Frequency</td>
<td>2 kHz</td>
<td>10 Hz</td>
<td>20 Hz</td>
</tr>
<tr>
<td>Aircraft velocity</td>
<td>80 m/s</td>
<td>80 m/s</td>
<td>5 m</td>
</tr>
<tr>
<td>Resolution</td>
<td>4 cu</td>
<td>8 m</td>
<td>5 m</td>
</tr>
<tr>
<td>Points/segment</td>
<td>60,000–650,000</td>
<td>16,348</td>
<td>8,192–65,536</td>
</tr>
<tr>
<td>Longest segment</td>
<td>26 km</td>
<td>131 km</td>
<td>228 km</td>
</tr>
<tr>
<td>Number of segments</td>
<td>6 (see Table 2)</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Scaling range</td>
<td>8–16 cm&lt;sup&gt;*&lt;/sup&gt;</td>
<td>=60 m</td>
<td>20–40 m</td>
</tr>
<tr>
<td>To</td>
<td>2–5 m&lt;sup turno&lt;/sup&gt;</td>
<td>=60 km</td>
<td>=20 km</td>
</tr>
<tr>
<td>Exponents</td>
<td>N/A&lt;sup&gt;*&lt;/sup&gt;</td>
<td>0.31&lt;sup turno&lt;/sup&gt;</td>
<td>0.28</td>
</tr>
<tr>
<td>Spectral (β)</td>
<td>0.92±0.1&lt;sup&gt;1,1&lt;/sup&gt;</td>
<td>1.43±0.08</td>
<td>1.36±0.06</td>
</tr>
<tr>
<td>Roughness (H&lt;sub&gt;i&lt;/sub&gt;)</td>
<td>N/A&lt;sup&gt;*&lt;/sup&gt;</td>
<td>0.29</td>
<td>0.28</td>
</tr>
<tr>
<td>Sparseness (C&lt;sub&gt;i&lt;/sub&gt;)</td>
<td>N/A&lt;sup&gt;*&lt;/sup&gt;</td>
<td>0.12&lt;sup turno&lt;/sup&gt;</td>
<td>0.10</td>
</tr>
</tbody>
</table>

For graphical illustration of LWC data sets observed during FIRE, see Davis et al. [1996a] or Marshak et al. [1997a], and Davis et al. [1994a] for ASTEX. The first 10 rows contain instrumental specifications and ancillary information. Note the 10<sup>2</sup> ratio between the smallest pixel and the longest flight. The last five rows are devoted to three fundamental exponents and the associated range of scales: further discussion in sections 4–6. N/A, not applicable. *Small scales. †Large scales.

Behavior in a scaling regime spanning 3 decades, comparable in extent to those of in situ LWC data sets. (At this point, “multifractal” simply means that scaling exponents for different statistical moments are not linearly related, as predicted by a straightforward dimensional analysis; a more useful definition for data processing is given later.)

In situ measurements of LWC have been collected and analyzed on a continuous scale-by-scale basis in many ways. To the best of our knowledge, this activity started with King et al.’s [1981] spectral analyses of data from a hot-wire probe; they found scaling (power-law behavior) in the wave-number spectrum of LWC from ~3 m to over 30 m scales. Duroure and Guillermot [1990], as well as Malinowski and Zawadzki [1993], find power-law distributions of (almost) droplet-free air in Forward Scattering Spectrometer Probe (FSSP) data at ~10 m resolution; these laws were found to hold over at least 2 decades in scale for both droplet concentration and LWC. Also using FSSP data at 10-m resolution (or better), Korolev and Mozin [1993] investigate the variability of droplet concentration and size distribution over a broad range of scales (10 m to 10 km), finding strong positive and negative fluctuations in number density at scales that rarely exceed 10–20 m; they interpret these observations in terms of turbulent mixing processes and entrainment instabilities. Recently, Davis et al.'s [1994a] studies of ASTEX data at 8-m resolution and Marshak et al.’s [1997a] for FIRE at 5-m resolution show that internal marine stratocumulus (Sc) structure is multifractal over at least 3 decades in scale, from tens of meters to tens of kilometers.

At much finer scales, of the order of a few centimeters, Baker [1992] has found, at least in some areas, surprisingly strong variability (gradients) in droplet concentration. His finding largely confirmed by Brenguier [1993] and Malinowski et al. [1994] using different instrumentation and/or analysis methodology. Is this small-scale variability just what we expect by extrapolating the large-scale properties to such fine scales? If not, scale-by-scale statistical analyses of LWC data should reveal a transition between two physically distinct regimes, providing the spatial resolution is sufficient. We show here that this is indeed the case.

This paper focuses primarily on data obtained during SOCEX (phase 1) from the PVM–100A probe [Gerber et al., 1994] carried by CSIRO’s Fokker F–27 aircraft flying off the coast of Tasmania. This instrument samples LWC and Particle Surface Area (PSA, not used in this study) at 2 kHz, that is, 4 cm resolution for a nominal air speed of 80 m/s. The important characteristics of the SOCEX database are entered in the first 10 rows of Table 1 and compared to their FIRE and ASTEX counterparts. Table 2 describes in more detail the six data sets used in the present study. They were all collected during the July 26, 1993, flight which was not considered by Boers et al. [1996] primarily because of the complexity of the cloud cover. The cloudiness was indeed rather atypical that day: boundary layer over 2 km deep, no uniform cloud layer per se yet considerable drizzle/shower activity; cloud top was generally less than 1.5 km but strong convective cores and outflow at the top were “mushrooming” through stretches of broken Sc (R. Boers, private communication, 1997). We selected the six data sets for being contiguous in time and “in cloud” at least 60% of the time, as defined by LWC ≥ 0.02 g/m<sup>2</sup>; overall, about 2/3 of the data described in Table 2 are in-cloud, with mean 0.290 g/m<sup>2</sup> and standard deviation 0.167 g/m<sup>2</sup>. In section 2, spectral analysis is used to separate the two scaling regimes in the SOCEX data. On the basis of this separation of scales at 2–5 m, we visualize separately the large- and small-scale variabilities in section 3 and briefly discuss the physical processes that dominate in each regime. The simplest hypothesis is that the scale break is caused by the natural Poisson noise that contaminates the PVM’s LWC channel due to its finite sampling volume; this eventuality is ruled out (in appendix A) by detailed numerical simulations of PVM operation in a variable LWC environment. Section 4 calls on qth-order structure functions to revisit the scale break, investi-
Table 2. Characteristics of the High-Resolution LWC Data Sets From SOCEX

| Segment | Fig. 2 Start, | End, | \( N_{\text{tot}} \) (pts) | \( \log_{10} N_{\text{tot}} \) | \( L, \) km | \( L^*, \) km | \( N_{\text{real}} \) (pts) | \( f_{\text{in}} \) % | Mean \( \mu \) g/m³ | s. d. \( \sigma \) g/m³ | max \( \mu \) g/m³ | max \( \sigma \) g/m³ | \( \beta_D \) | \( \beta_C \) |
|---------|---------------|------|-----------------|-----------------|-------------|-------------|-----------------|------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------|-----------|
| 1 c     | 1017          | 1018 | 180,000         | 17.5            | 7.2         | 5.1         | 5               | 1,406       | 0.213           | 0.128           | 0.54            | 1.23            | 1.59            | 1.11      |           |
| 2 d     | 1026          | 1027 | 150,000         | 17.2            | 6.0         | 5.1         | 4               | 1,171       | 0.387           | 0.208           | 0.73            | 1.15            | 1.66            | 1.07      |           |
| 3 e     | 1030          | 1031 | 120,000         | 16.9            | 6.0         | 5.1         | 3               | 937         | 0.248           | 0.125           | 0.52            | 0.63            | 1.59            | 1.29      |           |
| 4 b     | 1047          | 1049 | 190,000         | 15.5            | 5.6         | 5.2         | 5               | 1,484       | 0.425           | 0.225           | 1.03            | 1.23            | 1.59            | 1.24      | 0.94      |
| 5 a     | 1109          | 1117 | 650,000         | 19.3            | 26.0        | 20.5        | 19              | 5,078       | 0.201           | 0.132           | 0.58            | 1.27            | 1.53            | 0.81      |           |
| 6 f     | 1151          | 1152 | 60,000          | 15.9            | 2.4         | 1.3         | 1               | 468         | 0.661           | 0.174 (<)       | 1.22            | 0.58            | 1.16 (> )      | 1.04      | 1.74      |
| Total   |               |      | 1,350,000       | 59.0            | 39.8        | 37          | 10,544          |             | 0.290           | 0.167           |                |                |                |           |

In-cloud average

Measurements were made on July 26, 1993, off the coast of Tasmania (Australia). A nominal air speed of 80 m/s was used for time-to-space conversion under Taylor's (1953) hypothesis, the sampling rate being 2 kHz. The integer part of \( \log_{10} N_{\text{tot}} \) indicates the number of points in the various double-log plots (note that there is one less datum for wavenumber spectra). \( L^* \) denotes the largest scale corresponding to an integer power-of-two number of pixels, and \( N_{\text{real}} \) is the number of independent segments of length \( \min(L_{o}, L_{1}, \ldots, L_{n}) \). The parameter \( f_{\text{in}} \) is the percentage of “in-cloud” data-points, as defined by LWC > 0.02 g/m³. The number of “large-scale” points on a 5 x 17-m grid for each segment is \( N_{s} = \left[N_{\text{tot}}/25\right] \). Mean \( \mu \) and standard deviation \( \sigma \) refer to in-cloud data only; differences between 4-cm and 5-m resolution are in the third decimal, except for \( \sigma \). Maximum values, of course, depend strongly on the resolution. The last two columns give the large- and small-scale spectral exponents, as defined by the regressions shown in the right hand side of Figure 2; in Table 1, we entered exponents for the average spectrum, not the average of the exponents, and we note that segments 4–6 clearly dominate this average.

Maximum pixel values correspond to instrumental saturation (only at 718 of the 840,000 points in segments 4–5 are involved).

2. Spectral Evidence for the Scale Break at 2–5 m

Consider a signal \( \varphi(x) \) where \( x \) represents time or space coordinate. To study correlations, or “structures,” in the data, we use statistical properties of \( \varphi(x) \) that depend parametrically on scale, denoted \( r \). Alternatively, one can use wavenumber \( k = 1/r \). If the statistics of \( \varphi(x) \) follow power laws in \( r \) or \( k \) over some reasonably large range, we say that \( \varphi(x) \) is (statistically) “scale-invariant.” If there are several ranges of scales over which different power laws prevail, we call them “scaling regimes,” and the special scales that separate two scaling regimes define “scale breaks.” Baring instrumental artifacts, the occurrence of a scale break suggests that different physical processes dominate on either side.

Assume that \( \varphi(x), 0 \leq x \leq L \), is sampled on a grid of constant \( t \ll L \). This random process certainly has a scale break if its wavenumber spectrum \( E(k) \) follows power laws with spectral exponents \( \beta_{D} \neq \beta_{C}, \) for two separate scaling regimes \( \left\{ r_{D} \right\} \) and \( \left\{ r_{C} \right\} \) with a scale break somewhere in the interval \( \left(r_{D}, r_{C} \right) \). Explicitly,

\[
E(k) \sim \begin{cases} 
    k^{-\beta_{D}}, & (r_{D} \leq r \leq 1/k \leq r_{D}^{*} \leq L) \\
    k^{-\beta_{C}}, & (1/k \leq r \leq r_{C} \leq 1/k \leq r_{C}^{*}) 
\end{cases},
\]

\( r_{D} < r_{D}^{*} \leq r \leq r_{C}^{*} < r_{C} < r \).

Figure 1 shows two wavenumber spectra. The longest data set (5 in Table 2), containing 650,000 points, gives us an unprecedented range of scales in LWC analysis: 8 cm (Nyquist scale) to 21 km (largest power-of-2 number of pixels). We
clearly see two scaling regimes: from 8 cm, to 2–5 m and from there to 5–10 km. The ensemble average spectrum is also shown, yielding

\[
\begin{align*}
\beta_\lambda &= 1.6 \pm 0.1 \quad (r_\lambda = 5 \text{ m} \leq r \leq 1.3 \text{ km} \sim r_\psi), \\
\beta_c &= 0.9 \pm 0.1 \quad (r_c = 8 \text{ cm} \leq r \leq 2 \text{ m} \sim r_c).
\end{align*}
\]

(2)

in equation (1) since the break occurs approximately at the 12th wavenumber-octave bin \((r = 1/k)\) is between 2.5 m and 3.1) starting at \(k = 1/L_s^\lambda (L_s^\lambda = 21 \text{ km})\). The statistical ensemble is defined here by 37 independent intervals of length 1.3 km that can be extracted from the database (compare Table 2). This choice of scale for spectral averaging corresponds to the minimum of the largest integer power-of-two of number of pixels (which is determined by segment 6). As indicated by the data from the longest segment (5), the large-scale scaling goes beyond this operational limit; further on we estimate it to be at least 3 decades. In contrast, it is hardly justified to talk about scaling when the range is hardly more than a single decade [Hamburger et al., 1996; Matcak et al., 1997]; the small-scale exponent is therefore to be interpreted as a clear indication of the break. For the make up of the average spectrum in Figure 1, refer to the segment-wise averages shown in Figure 2.

Note that \(\beta_\lambda < 1\) and \(1 < \beta_c < 3\), indicating that LWC fluctuations are (broad-sense statistically) stationary at small scales and nonstationarity, but with stationary increments, for large ones. The practical utility of this classification is discussed at length by Davis et al. [1996a]. Important properties used further on are, on the one hand, that spatial averages are only meaningful in stationary regimes for LWC (§3.1) or related quantities (§5.1) and, on the other hand, that cumulating incremental statistics by spatial averaging is fully justified (§4.1).

As seen in Table 1, the large-scale spectral exponent \(\beta_c\) is in reasonable agreement with other spectral analyses of LWC. Since \(\beta_c < 1\), the small scales show far more variability than expected by extrapolating spectrally the large-scale behavior (small \(k\) to small scales (large \(k\)). This is consistent with Gerber’s [1991] conclusion that surfaces formed between mixing cloud elements have a significant effect on droplet distributions. Baker [1992] shows compelling evidence of strong centimeter-scale variability in droplet concentration using a customized statistical analysis of the arrival times of cloud droplets in the laser beam of an FSSP. Brenguier [1993] comes to similar conclusions using data from a “fast” FSSP equipped with new electronics. Malinowski et al. [1994] also confirmed the main points in Baker’s study by applying fractal methods to the same data.

3. Dynamical Nature of LWC Variability

3.1. Large Scales

To visualize the large-scale behavior of IWC, we need a coarse-grained representation: each point corresponds to a spatial average of the raw (4-cm) data at a scale close to where the break occurs, in the range \(r_\psi = 2 \text{ m} \sim r_\lambda = 5 \text{ m}\). Setting the new grid constant at

\[
\ell_s = 128 \times \ell - 5.12 \text{ m},
\]

we have plotted

\[
\bar{\varphi}(x) = \frac{1}{\ell_s} \int_0^{\ell_s} \varphi(x')dx', \quad 0 \leq x < L - \ell_s.
\]

(3b)

Figure 2 shows this field for all six flight segments in Table 2, from the one with the largest outer scale \((L_o = 26 \text{ km})\) to that with the smallest \((L_o = L_2 = 10)\). Segment 5 is not only the longest data set but also the most variable in the sense of the \(\sigma/\mu\) ratio; Figure 3a shows the newly observed small-scale fluctuations for a typical subsegment. The converse can be said about segment 6 where LWC is higher and somewhat more constant, at least according to the \(\sigma/\mu\) ratio, this is actually a rather questionable statement in view of the subset of small-scale fluctuations extracted from this segment in Figure 3e.

The LWC traces in Figure 2 are visually similar to those seen in Davis et al. [1994a] and Gerber [1996] who analyzed LWC data from the PVM-100A probe at 8 m resolution during ASTEX. That measurement program targeted Sc-to-Cu transition in the marine boundary layer, a complex situation closer to the present one than when only marine Sc are present (as, e.g., during PIRTe). However, similar spectral exponents are found in all three cases (compare Table 1). These exponents, in the range 1.4–1.7, are not far from the “5/3” prediction for the density fluctuations of velocity [Kolmogorov, 1941] and passive scalars [Obukhov, 1949; Corrsin, 1951] in fully developed 3-D turbulence. In 2-D turbulence, the prediction is also 5/3 but for somewhat different reasons [Kraichnan, 1967]. We conclude that, one way or another, this regime is dominated by inertial range turbulent dynamics. For turbulent velocity however, the inertial sub-range goes down to the much smaller (mm–cm) Kolmogorov dissipation scales. It is worth mentioning that identifying water, in any phase, with a passive tracer in atmospheric dynamics offers limited insight into cloud structure. Indeed, by definition, clouds are in near-saturation conditions and latent/sensible heat production affects buoyancy, hence dynamical forcing. So cloud microphysical processes interact with the circulation at some, if not all, scales [Mouang et al., 1996; and references therein]; the scale break in Figure 1 is a tangible manifestation of this poorly understood interaction. Further on we will qualify the similarity between LWC in boundary layer clouds and passive scalars in turbulence, and identify the source of their differences.

3.2. Small Scales

Typical small-scale fluctuations are illustrated in Figure 3. Each panel contains 12,500 points, covering scales from 4 cm to 0.5 km. The obvious difference between Figures 3a–3b and 3d–3f is that upspikes tend to occur when LWC is low on average and, if it is high, deviations from the mean seem have a
Figure 3. Sampler of LWC data at 4-cm resolution, showing small scale variability. The various $x$-values can be used to find in Figure 2 the positions where these 0.5-km details (12,500 points) were extracted. The white curves consist of 97 points reproduced from Figure 2. Low LWC values on average (note dominance of up-spikes): (a) 5, Figure 2a; (b) 4, Figure 2b. Intermediate-average LWC: (c) 3, Figure 2e (note conspicuous lack of spikes); (d) 4, Figure 2b. High-average LWC: (e) 6, Figure 2f; (f) 4, Figure 2b (note saturation, fortunately, an infrequent occurrence). The right hand side (Figures 3b, 3d, and 3f) were taken from the same segment 4.

random sign. (Note that, throughout this paper, we use the word "spike" without any insinuation that it describes some kind of artificial glitch in the instrumentation.) A priori all this small-scale activity could simply be Poisson noise resulting from the natural fluctuation of the number of droplets present in the PVM's sampling volume.

To quantify the contribution of Poisson noise to the small-scale variability, we simulated numerically each step of the PVM-100A's operation when collecting cloud LWC data. These simulations are described in full detail in appendix A. We conclude that Poisson noise alone is insufficient to explain the scale break at 2-5 m. Specifically, the excess variance is contained in the strong up-spikes in Figures 3a–3b (compare Figures A2a and A2f) and the largest jumps in Figures 3d–3f (compare Figures A2a’ and A2f’). The scale break uncovered in Figure 1 is therefore likely to be physical. This interpretation is reinforced by the fact that an absence of spikes (Figure 3c) is associated with a lesser scale break (Figure 2e).
and Table 2) in flight segment 3. These strong events are therefore responsible for the scale break even though they are at most a few 4-em pixels wide. Indeed, isolated $\delta$ functions (even at finite resolution) excite variance (equally) at all scales larger than their width, at least in the Fourier representation used here. Furthermore, the spikes appear to come in clusters of variable extension.

At present, we have no definitive hydrodynamical or thermodynamical explanation to advance for the anomalous spikes that raise above the Poissonian noise in the sampling volume; however, we note that these events unfold on the same scale as the deviations from spatial Poissonian behavior detected by Baker [1992] and Brenguier [1992] in records of droplet arrival times in an FSSP’s laser beam. In further agreement with the analyses of these authors, the strong features (and associated non-Poissonian behavior) is not observed everywhere: see Figure 3c. A common cause is therefore very likely for the FSSP-based results and for our findings. Both Shaw et al.’s [1998] direct numerical simulations of preferential droplet concentration by small-scale turbulence and Lu and Hallett’s [1998] new conceptual model for turbulence-microphysics interaction are likely to bear on our statistical findings. In modeling rain measurements, non-Poissonian models have also proven useful for drop count statistics [Kosinski and Jameson, 1997] and radar reflectivities [Lovejoy et al., 1996]. Discussion of our data and spectral analyses in more cloud-microphysical detail will be pursued elsewhere (II. Gerber et al., Spectral density of cloud liquid water content at high frequencies, submitted to Journal of the Atmospheric Sciences, 1999).

4. Structure-Function Analysis

Until now, we have used only wavenumber spectra to establish that cloud structure exhibits a scale break a few meters. This marks the transition from a large-scale turbulence-dominated regime to a poorly understood small-scale regime where the variability exceeds the natural Poissonian fluctuations in the sampling volume. However, we are unable to distinguish Gaussian-like signals from more strongly variable ones, scaling or not, with spectral analysis alone; see Davis et al. [1996b], Marshak et al. [1997a], and Appendix B. By “Gaussian-like,” we mean statistics determined largely by low-order moments, typically first-order (means) and/or second-order (variances and autocorrelations). To overcome this intrinsic limitation of spectral analysis, which is second-order, we turn to more sophisticated scale-by-scale analyses based on a broader choice of moments, starting with qth-order structure functions.

4.1. Background and Definitions

The earliest application of higher-order structure-function analysis was to turbulent velocity signals [Monin and Yaglom, 1975, and references therein]. In this approach, we consider qth-order statistical moments of the absolute increments across scale r:

$$\langle \Delta \phi(r) \rangle = \langle [\phi(x+r) - \phi(x)]^q \rangle .$$

In a scaling range where $1 < \beta < 3$ (nonstationary process with stationary increments), the statistical properties of $\Delta \phi(r;x)$ are independent of position x, so $\langle \Delta \phi(r;x)^q \rangle \equiv \langle \Delta \phi(r)^q \rangle$. Because of the scale invariance, we anticipate

$$\langle \Delta \phi(r)^q \rangle - \zeta(q), q \geq 0,$$

where $\zeta(\cdot)$ denotes, in theory, an ensemble average. In practice, we use spatial averaging as a substitute and/or complement for its ensemble counterpart (incremental stationarity justifies this). As long as the proportionality factors in equation (5), commonly called “prefactors,” depend only weakly on q, then $\zeta(q)$ is necessarily concave (i.e., $\zeta''(q) \leq 0$) [Frisch and Parisi, 1985]. Furthermore, due to the normalization, we have $\zeta(0) = 0$.

The (global) Hölder exponent, also known as (Hurst’s) roughness exponent,

$$H_1 = \zeta(1) < (0,1),$$

characterizes the “smoothness” of the signal. Indeed, we will have typically $\Delta \phi(r;x)/r^{\beta + 1}$, so the signal’s derivative (obtained when $r \to 0$) exists only in the limit $H_1 \to 1$: otherwise, the signal is singular. The opposite limit, $H_1 \to 0$, leads to stationary ($\beta < 1$) random processes. Indeed, scale-invariant processes that are stationary per se, statistically invariant under translation in x, have not only scale-invariant but scale-independent increments: $\zeta(q) \equiv 0$ in equation (5) and, in particular, $H_1 = 0$ in equation (6). This results ultimately from the inherent discontinuity or “jumpsiness” of stationary scale-invariant processes associated with short-range decorrelation [Davis et al., 1994b, 1996a].

In short, $H_1$ quantifies nonstationarity whereas the remainder of the $\zeta(q)$ essentially qualify the prevailing nonstationarity. Is it “monofractal,” with $\zeta(q) = H_1 q$, in Mandelbrot’s [1977] well-known “fractional” Brownian motion (fBM)? Or is it “multifractal,” with a nonlinear $\zeta(q)$ as in the “bounded cascade” model [Cahalan et al., 1994; Marshak et al., 1994] where $\zeta(q) = \min(H(q),1)$? Davis et al. [1994b, 1996a,b] discuss this analysis in more detail, providing numerous illustrations with data and models.

4.2. Results for LWC During SOCEX, Comparison With Turbulence at All Orders

We can now re-visit the 2–5 m scale break with structure functions. Figure 4a is a log-log plot of $\langle \Delta \phi(r)^q \rangle$ versus r for $q = 1, \ldots, 4$ obtained by spatial and ensemble averaging, as for E(k), over the 37 segments of data $2^{15}$ pixels long (the largest power-of-2 number of pixels in shortest flight segment, 6). For all q’s we have two distinct regimes with the same transition around 2–5 m as seen in the Fourier-space representation (Figure 1). The range of scales used in Figure 4a is the same as for the average energy spectrum in Figure 1.

For large scales, we find $\zeta(2) = 0.51 \pm 0.02$ using equation (5) which agrees reasonably well with $\beta_{-1}$ where $\beta_q$ is obtained from equation (7). (See Monin and Yaglom [1975] for proof that $\beta = \zeta(2) + 1$ in regimes where $1 < \beta < 3$.) Furthermore, we find

$$H_1 = 0.31 \pm 0.01$$

in equation (6).

The nonlinearities of $\zeta(q)$ versus q in Figure 4b tells us that the large-scale behavior of LWC during SOCEX is a multifractal nonstationary process from 5 m up to at least 2 km. We note that $\beta = 1.6$ ($\zeta(2) \sim 0.6$) and $H_1 \sim 0.3$ are not far from the Obukhov [1949]–Corrsin [1951] prediction for a passive scalar in fully developed 3-D turbulence: $\beta = 5/3$ (equivalently, $\zeta(2) = 2/3$). and generalized to $\zeta(q) = q/3$ as for velocity...
LWC and scalars such as temperature or an admixture of light particles or molecules are close enough to justify thinking of LWC as passively advected by the turbulence in clouds in a first approximation. However, our data does not support this conceptual model in the details, specifically, with respect to the large jumps that dominate the higher-order moments at all scales.

The small scale regime has been diagnosed spectrally as stationary so, in principle, we should have $\zeta(q) = 0$, that is, structure functions independent of scale $r$. This is not contradicted by residual $r$ dependence observed at small scales in Figure 4a because the weak dependence is compatible with finite-size effects; see Davis et al. [1996b] for details on how to show this. Since structure functions are spuriously scale dependent anyway in a stationary regime, the range of scales involved here, hardly a decade, is insufficient to do a meaningful structure-function analysis. Conversely, Davis et al.'s [op cit.] analysis of finite-size effects certifies that the multifractality of the large-scale behavior is “real” in the following sense: it is stronger than the spurious multifractality likely to be found operationally for a single realization of a monoscaling model over the same finite range of scales.

5. Singularity Analysis

It has been argued that (higher-order) structure-function and singularity analyses are redundant; we differ our discussion of this issue to §6.2 and Appendix B. The above remark for structure-function analysis applies to singularity analysis as well: the range of scales is insufficient in the small-scale regime. We therefore restrict our attention in this and the following sections only to the large-scale variability, as illustrated exhaustively (for July 26, 1993) in Figure 2.

5.1. Background and Definitions

To further study LWC variability in a way that emphasizes the intermittent occurrence of larger-than-average jumps, we compute a nonnegative field, a “measure,” from the $r_p$ scale gradient fields, recalling that $r_p$ is the lower bound of power-law regime for large scales in Figure 1. In the remainder of this section, we will therefore interest ourselves in the statistical properties of

$$r(x) = \bar{\eta}(x + r_p) - \bar{\eta}(x)^m$$

on a scale-by-scale basis, mostly with $m = 1$. On the basis of spectral and structure-function analyses, we take the grid constant as $r_p = 5.12$ m in this study. Sample of the result field, obtained from flight segment 5, is plotted in Figure 5.

We now focus on “coarse-grained” versions of the measure in equation (8), parameterized by the scale $r$ over which the local average is computed:

$$\epsilon(r,x) = \frac{1}{L - r} \int_x^{x + r} \epsilon(x') dx', \quad 0 \leq x < L - r,$$

where $x'$ is sampled on the new 5.12-m grid and $x$ on an even coarser grid of constant $r \geq r_p = 5.12$ m. Being obtained from an increment of $\phi(x)$ in equation (8), the $\epsilon(r,x)$ fields are stationary for any $r$, hence their statistical properties are stationary (a priori independent of $x$) which enables us to use spatial averaging over $x$ as well as ensemble-averaging, that is, over

---

[Frisch, 1995]. In Figure 4b, our $\zeta(q)$ results for LWC are compared with that standard—that is, pre-intermittency corrected—prediction (bold line) and the empirical results of Antonia et al. [1984] (bold symbols). The low-$q$ behaviors of
5.2. Results for LWC During SOCEX, Comparison With Turbulence at Low Order

We plot \( \log_{10}(\epsilon(r))^q \) versus \( \log_{10} r \) in Figure 6a, using selected integer values of \( q \). We then seek the straight lines that define scaling regimes. Equation (10) can be used for \( \langle \epsilon(r)^q \rangle \) for scales over the full range of scales relevant to this measure, namely, 5 m to 1.3 km; this leads to the function \( K(q) \) plotted in Figure 6b. By numerical differentiation at \( q = 1 \) we find:

\[
C_1 = 0.12 \pm 0.01.
\]

(a)

(b)

![Figure 6. Ensemble-averaged moments of coarse-grained measures. (a) We plot \( \log_{10}(\epsilon(r))^q \) for \( q = 1, \ldots, 4 \) versus \( \log_{10} r \) associated with the large-scale (resolution \( \ell_s = 5.12 \) m) signal in Figure 2. Averaging is over all 37 sub-segments obtained from the six flight segments (compare \( N_s \) and \( N_{\text{real}} \) in Table 2) and regressions apply to scales from 5 m to 21 km. (b) Large-scale exponent functions \( K(q) \); the \( C_1 \) exponent is highlighted.](image)
Prasad et al. [1988] analyzed the intermittency properties of small-scale gradients in the fluctuations of passive scalar concentration in a high Reynolds number flow with a 2-D imaging technique using a fluorescing dye in a water jet; they also used the more standard 1-D technique for temperature fluctuations recorded in an air jet. Following normal usage in turbulence, the authors used the square rather than the absolute value of the small scale gradient to define the measure, that is, $m = 2$ rather than $m = 1$ in equation (8). As expected from the phenomenology, the multifractal properties of temperature and scalar concentration fluctuations are indistinguishable and remarkably universal. According to the "$f(\alpha)$" data in Figure 6 of Prasad et al. [1988] and the well-known Legendre transform relation between $f(\alpha)$ and $(q-1)D_q = (q-1)K(q)$ here, we find $C_1^{(m=2)} = 1-D_1^{(m=2)} = 0.22$ where $q = d\alpha/dx = 1$; following Marshak et al. [1997a], this translates to $C_1^{(m=1)} = 0.11^*$. Comparison of this outcome with equation (12) reinforces our above conclusion about the compatibility of the low-order statistical behavior of LWC and passive scalars.

5.3. An Alternative Estimation of $C_1$

Consider the "refined" similarity formula:

$$\xi(q) = qa - K(q\beta)$$

that links the structure-function exponents to the singularity properties of the small-scale gradients in equation (8). Following Schertzer and Lovejoy [1987], we view $a$ and $b$ as arbitrary constants, thus generalizing the prototypical turbulence case where we know that $a = b = 3$ [Kolmogorov, 1962]. Note how $K(q)$ appears in equation (13) as a (so-called "intermittency") correction term in an otherwise linear model for $\xi(q)$, for turbulence, this linear model is identified with the standard [Kolmogorov, 1941] phenomenology that leads to prediction $\xi(2) = 2/3 (B = 5/3)$, hence $\xi(q) = q^3$.

Equations (11) and (13) yield

$$C_1 = \xi'(b) - bL^*_c(b)$$

which enables us to estimate $C_1$, as well as $H_1$, from structure-function analysis alone. For simplicity, we can take $b = 1$ in equation (14) and, accordingly, $a = 1/H_1$ in equation (13). The data in Figure 4b then yield

$$C_1 - H_1 - \xi'(1) = 0.07.$$  (15)

The significant difference between the estimates of $C_1$ in equations (11)-(12) and (14)-(15) casts doubt on the general applicability of equation (13), bearing in mind that another choice of $b$ in equation (14) could possibly reconcile the two estimation techniques (we return to this issue in the following section). In turbulence, however, where $b$ is a given quantity, there is a similar discrepancy between the intermittency parameter for passive scalars, $\mu_{\text{urb}} = K^{(m=2)}(2)$, estimated by Antonia et al. [1984] from structure-function analysis, $\mu_{\text{urb}} = 0.25$, and estimated by Prasad et al. [1988] from singularity analysis, $\mu_{\text{urb}} = 0.38 \pm 0.08$. (The debate on how the statistics of absolute small (Kolmogorov) scale fluctuations in equation (8) influence those of the inertial-range quantities in equation (4) is still ongoing, e.g., Vainshtein et al., 1994.)

Using arguments similar to those used by Marshak et al. [1997a] for turbulent velocity, we can convert "$m = 2$" results to "$m = 1$" results and then $\mu_{\text{urb}}$ values (at $q = 2$) into $C_1$ values (at $q = 1$). Although uncertainty increases in this process, the outcome is a lower-bound for $C_1^{(m=1)}$ given by $\mu_{\text{urb}}/4$; this yields the following: $C_1^{(m=1)} = 0.06^*$ for structure functions (compare equation (15)); and $C_1^{(m=1)} = 0.10^*$ for singularity analysis (compare equation (12)), in good agreement with the more direct estimate obtained in section 5.2. Again, we conclude (adopting either definition of $C_1$), that the scaling behavior of low-order statistics for passive scalars and LWC are similar. From the consistency of our results, we also conclude that either definition of $C_1$ can be used in an intercomparison study, even though they may be hard to reconcile with each other.

6. Marine Sc Structure from 100 to 103 m, Using Various LWC Data Sets

It is of interest to systematically compare results from our three-fold analysis (spectral, structure functions, and singularities) of SOCEX data with corresponding scale-by-scale results for ASTEX data [Davis et al., 1994a] and FIRE data [Davis et al., 1996a; Marshak et al., 1997a]. This exercise can be viewed as an attempt at using scale-invariance to establish a climatology of internal cloud structure, a counterpart of the extensive satellite-based investigation of broken cloud fields by Cahalan and Joseph [1989] and Joseph and Cahalan [1990]. Values of $B$, $H_1$, and $C_1$ are entered for all three experiments in Table 1 along with the associated scaling ranges that are graphically illustrated in Figure 7 (with wavenumber spectra). For $H_1$ and $C_1$, only the large-scale regime is considered in SOCEX. We start with these low-order statistics (§6.1), before moving on to higher-orders (§6.2).

6.1. First- and Second-Order Statistical Properties of LWC

6.1.1. Bi-fractal properties of cloud structure.

We show in appendix B that, by using a pair of judiciously chosen exponents, one can avoid at once the well-documented

![Figure 7. Wavenumber spectra for three field experiments. ANTTEX results from Davis et al. [1994a] and FIRE results from Davis et al. [1996a]. Note the relatively small differences in the scaling exponents and the notable differences in scaling range (compare Table 1). Recent instrumental development has enabled observation of the scale break at 2–5 m in the SOCEX data, but it is likely to be present in other types of cloud.](image-url)
shortcoming of a monofractal approach (its ambiguity), the potential pitfalls in a multifractal approach (spurious high-order scaling due to finite data sets), and the possible redundancy between structure functions and singularity analysis. We work with $H_1$ and $C_1$, defined in equations (6) and (11), respectively, and call this "bifractal" analysis.

Figure 8 is an $(H_1, C_1)$ scatterplot for LWC fluctuations, for scales ranging from a few tens of meters to a few tens of kilometers enhanced with LWP and passive scalars. All presently studied cases are illustrated, as follows: the bold dot represents SOCEX (large scales) while other symbols represent FIRE (squares) and ASTEX (diamonds). Some remarks are in order:

1. The three LWC points for ensemble averages are very close in the bifractal plane:

$$\begin{align*}
0.28 & \leq H_1 \leq 0.31 \\
0.08 & \leq C_1 \leq 0.12
\end{align*}$$

This suggests that the physical processes that shape the internal structure of Sc-type clouds are essentially "universal," meaning that the local climate seems to determine primarily the range of scales (compare Table 1) and the prefactors (overall amplitudes), but not these two important scaling exponents. Davis et al. [1996a] note that, for the FIRE and ASTEX case studies of marine Sc (representative of more typical conditions than those that prevailed when our SOCEX data were collected), the scaling range is displaced roughly in proportion to the boundary layer's thickness.

2. Because of their proximity in the bifractal plane, the low-order statistical similarity between LWC and a passive scalar in a turbulent flow established in the previous sections for SOCEX carries over to FIRE and ASTEX.

3. As expected, the vertically integrated quantity LWP (stars in figure) is somewhat smoother than LWC. Using the results of Wiscombe et al. [1995] collected at the ARM Central Great Plains Site in Oklahoma and based on all types of cloud cover, we find $H_1(HWP) \sim 0.37 \geq H_1(LWC)$. Spectral exponents occur in the same order: $\beta_{LWP} \sim 1.7 \geq \beta_{LWC} \geq 1.5$. However, there is no significant difference between LWC and LWP in the sense of intermittency, the two fields having almost identical $C_1$ values.

4. We have indicated in Figure 8 the loci of the better-known scale-invariant models. None of them fall inside the bifractal domain, where the data lives. Indeed, fractional Brownian motion (fBm) is not intermittent enough ($C_1 = 0$), multiplicative cascades are too stationary ($H_1 = 0$), Devil's staircases are too smooth ($H_1 = 1$), and isolated singularities are too intermittent ($C_1 \rightarrow 1$). This underscores the need for a new class of stochastic models to simulate properly the nonstationary and intermittent behavior we observe in LWC signals. An example of such a model is a "fractionally integrated" multiplicative cascade [Schertzer and Lovejoy, 1987]; the list of models that live inside the unit square is still quite short but steadily growing [cf. Davis et al., 1996b; Marshak et al., 1997a].

Finally, we note that the counterparts for the two other field programs of equation (15) for SOCEX are $C_1 \approx 0.08$ from the ASTEX data of Davis et al. [1994a], and $C_1 \approx 0.11$ from the FIRE data of Marshak et al. [1997a]. In comparison with equation (16), these estimates, based only on structure function analyses, rearrange and slightly shift downward the LWC cluster in Figure 8. However, this alternate choice of definition for $C_1$ does not change the LWC cluster's dispersion in the bifractal plane.

6.1.2. Challenges to cloud probing and modeling efforts. Until now, the emphasis is on the apparent robustness of $(H_1, C_1)$ for LWC in boundary-layer clouds under a broad range of meteorological conditions. This is not to be taken as a prediction that every time we measure $H_1$ and $C_1$ for LWC we will find numbers compatible with equation (16), on the contrary. In their comprehensive study of LWC records from FIRE '87, Marshak et al. [1997a] actually obtain a "cloud" of $(H_1, C_1)$ points quite broadly scattered around the position of the ensemble-average used in Figure 8, especially in the $C_1$ direction. This underscores the need for vast amounts of data to defeat the blatant violation of "ergodicity" by each individual data set (i.e., there is no convergence to the ensemble mean). Even an apparently well-defined dynamical regime such as the recurrent/persistent marine Sc investigated during FIRE '87, as many samples as possible are required to obtain meaningful averages.

We have demonstrated the current lack of good empirical (i.e., stochastic) models for the internal structure of marine boundary layer clouds. In Figures 7–8 there is also a challenge to the current thrust in dynamical cloud modeling. As the resolution of Large Eddy Simulation (LES) and Cloud-Resolving Models (CRMs) increases, statistics sensitive to spatial correlations will soon pass simple "scale-by-scale" approaches or later ("power law scaling" approaches) become necessary to...
6.2. Higher-Order Statistical Properties of LWC

6.2.1. Dynamical meanings of lower- and higher-order statistics. Using first order exponents only, we have uncovered in Figure 8 a remarkable statistical similarity in the structure of various types of marine Sc: this carries over, more approximately, to second-order statistics. Notably, multifractal analysis recognizes that the basic physics of marine boundary layers are the same everywhere and that there is a rather well-defined class of (top-cooled) stratiform clouds that occur there. There are also remarkable differences in the first- and second-order statistics: the scaling range (compare Figure 7 here) and the prefactors [cf. Figure 7 in Davis et al., 1996a].

There is another significant difference between the various sources of LWC data for higher-order statistics. The hold symbols in Figure 9 represent the \( \zeta(q) \) scaling exponents for qth-order structure functions obtained from FIRE, ASTEX, and SOCEX: the higher-order moments diverge significantly (see Figure 4b for typical error bars). We see, however, that the scaling exponents for ASTEX and SOCEX are in better agreement between each other than with those of FIRE. This is not surprising: we have already commented on the visual similarity between the (coarse-grained) SOCEX data in Figure 2 and the ASTEX data of Davis et al. [1994a, Figure 10] (at 8-m resolution); the FIRE data of Marshak et al.'s [1997a, Figure 1] (at 5-m resolution) is strikingly different. It is noteworthy that comparative multifractal analysis corroborates visual examination. Moreover, this is a multifractal reminder that FIRE was designed to study "pure" marine Sc, whereas ASTEX was targeting Sc-to-Cu "transition" regimes and that the SOCEX data analyzed here was captured in a "mixed" cloud system.

In summary, the events, hence the "dynamics," that dominate the means \( (q = 1) \) are similar from one locale to another; the strongest/rarer events that determine moments of order \( q > 1 \) are not found to be similar from one locale/experiment to another.

6.2.2. Do we need more parameters? Pursuing the idea of two parameter characterizations of multifractal data in section 6.1, Pierrehumbert [1996] devised a scheme using structure functions alone that highlights the extreme values of \( q \). In this respect, it is complementary to bifractal analysis where the focus is on \( q - 1 \). Pierrehumbert proposes to fit the empirical \( \zeta(q) \) curve with a hyperbola branch:

Figure 9. Scaling exponents \( \zeta(q) \) of qth-order structure functions for LWC from three field programs with analytical representations. The ASTEX, FIRE and SOCEX \( \zeta(q) \) functions are plotted with symbols (same as in Figure 8) in all three panels: low-order moments agree, and higher-order moments differ. (a) Numerical (\( H_1, C_1 \)) values from Table 1 were used in equation (15b) to infer the two parameters, \( \zeta_0 \) and \( \zeta_\alpha \), of Pierrehumbert's [1996] model in equation (17): because of the \( (H_1, C_1) \) clustering in Figure 8, the high-\( q \) predictions are similar and fall in the middle of the data rather than follow \( \zeta(q) \) for any particular experiment. (b) \( \zeta_0 \) and \( \zeta_\alpha \) are now determined by a nonlinear least squares fit for \( 0 \leq q \leq 4 \); because of the weight given to the high \( q \) values in this procedure, these values lead via equation (18a) to more dispersion of \( (H_1, C_1) \) points than seen in Figure 8. (c) Here the four-parameter model for \( \zeta(q) \) in equation (13), with equation (19) for \( K(q) \), is used in the nonlinear regression: goodness of fit (as measured by the regression coefficient R) is generally improved, as expected (in view of the smoothness of the \( \zeta(q) \) data).
\[
\zeta(q) = \frac{q}{q \zeta_0 + 1} \zeta_0', \quad (17)
\]

where the two parameters have simple geometrical (and statistical) meanings: \( \zeta_0' = \zeta'(0) \in [0, \infty) \) is the slope at \( q = 0 \) (dominated by low \( q \) values); and \( \zeta_0 = \zeta(\infty) \in (0, \infty) \) is the horizontal asymptote (dominated by the highest \( q \)'s available). Figure 9b shows the results of nonlinear least squares fits to the \( \zeta(q) \) data for LWC from the three field programs using equation (17).

Equations (6) and (14), with \( b = 1 \), connect ours and Pierehumbert's parameterizations:

\[
\begin{align*}
H_1 &= \frac{\zeta_0' \zeta_0/\zeta_0^2 + \zeta_0}{(C_1 - 1)^2}, \\
C_1 &= \frac{\zeta_0/(\zeta_0^2 + 1)}{(C_1 - 1)^2}
\end{align*}
\]

and conversely,

\[
\begin{align*}
\zeta_0' &= H_1[1 - (C_1 - 1)^2], \\
\zeta_0 &= H_1^{-1/2} C_1
\end{align*}
\]

Note how the latter quantity controls explicitly the multiscaling since \( \zeta_0 \rightarrow \infty \) \( (C_1 \rightarrow 0) \) leads to \( \zeta(q) = \zeta_0' q^0 \) (i.e., monoscaling).

In Figure 9a, we have overlaid the predictions of equations (17) and (18b) using \( (H_1, C_1) \) values from Table 1 or Figure 8. The disagreement between models and data is large at high \( q \) values, especially for the FIRE data. This is not a failing of the model in equation (17) because better parameter values were obtained (compare Figure 9b) by going back to the \( \zeta(q) \) data rather than using equation (18b). Rather, we consider this as a shortcoming of equation (13) viewed as a means of predicting high-\( q \) behavior from a low-\( q \) anchor, at least with the chosen \( (a, b) \) values. Conversely, if we insert into equation (18a) the parameters \( (\zeta_0', \zeta_0) \) from the nonlinear fits in Figure 9b, we find a larger dispersion of the \( (H_1, C_1) \) points than seen in Figure 8, especially along the \( C_1 \) axis. This is traceable to the uniform weight given to all the values of \( q \) in the nonlinear regression and therefore does not change our conclusions about the remarkable clustering of the bone fide first-order exponents in equation (16).

There are several closed-form one- and two-parameter representations of \( K(q) \) in the literature that we can use in equation (13). For illustrative purposes, we will use Schertzer and Lovejoy's [1987] model:

\[
K(q) = C_1 X \left( \frac{q^\alpha - q}{\alpha - 1} \right), \quad \alpha \in (0, 1) \cup (1, 2), \quad q > 0.
\]

Letting \( a \) and \( b \) in equation (13) vary along with \( C_1 \) and \( \alpha \) here, we have a four-parameter model for \( \zeta(q) \). Given the smoothness of the \( \zeta(q) \) data sets, we expect and obtain good agreement (compare Figure 9c) when this model is used in the nonlinear regression; this proves little about the inherent validity of equations (13) and/or equation (19). Notwithstanding, we notice that (1) \( 1/a \) is numerically close to the new estimates of \( H_1 - 1/a = K(1/b) \) because \( K(1/b) \approx C_1 \times (1/b - 1) \) is small; indeed, (2) \( 1/b \) is quite close to unity, justifying a posteriori the assumption of \( b = 1 \) used in deriving equations (15) and (18a, b); (3) \( C_1 \) values are similar to those in Figure 9b and more dispersed than in Figure 8, again due to the influence of the large \( q \) values; and (4) \( \alpha \) is found well within the range \((0, 2)\) prescribed in equation (19) which is associated with “log-Lévy” or “log-stable” statistics for the small-scale increments.

In summary, we draw the following conclusions about cloud LWC data processing: (1) model-assisted extrapolation from low- to high-order exponents using equation (13) is risky; (2) being outside of pure turbulence (where \( a \) and \( b \) are known), it is not recommended to use equation (13) to do away with either structure functions or singularity analysis; (3) a wide variety of \( \zeta(q) \) functions can be well approximated analytically with a relatively small number of parameters. However, this last point has more to do with the inherent smoothness and the concavity of empirical \( \zeta(q) \) functions than with the validity of the analytical model. Moreover, parameters of the analytical model can often be mapped to a specific value of \( q \), but they are determined numerically by all values of \( q \) (through nonlinear regression), and the uncertainty in the data for large \( q \) is not easy to estimate. The above conclusions are likely to carry over to many other atmospheric and geophysical signals.

All things considered, it is our opinion that a simple low-order two-exponent characterization such as (in bifractal analysis) is a safe strategy when the limitations of a one-exponent method (such as in spectral analysis) become problematic.

7. Discussion: Radiative Consequences of LWC Variability

7.1. Large Scales and the “Independent Pixel Approximation”

Cahalan et al. [1994] developed a random fractal cloud model to emulate fluctuations of LWP in marine Sc, which is proportional to optical depth \( \tau \). The targeted scales range from tens of meters to tens of kilometers and Cahalan and Snider's [1989] LWP data were used to assign numerical values to the model’s three parameters, essentially \( \mu_{\text{WP}}, \sigma_{\text{WP}}, \) and \( B_{\text{WP}} \). The authors found dramatically reduced domain-average albedoes at nonabsorbing solar wavelengths for their internally variable cloud models with respect to the standard (homogeneous plane-parallel) model that uses only the mean optical depth \( \bar{\tau} \), dependent only on \( \mu_{\text{WP}} \). The corresponding enhancement of transmittance, \( \langle T(\tau) \rangle > \langle T(\bar{\tau}) \rangle \), is explained by the averaging over many pixel values of \( T(\tau) \), as obtained from plane-parallel theory; indeed, \( T(\tau) \) is a nonlinear function, with \( \langle d^2/dx^2 \rangle > 0 \), at least in the two-flux approximation. This independent pixel approximation (or “IPA”) to 3-D radiative transfer is now being used to capture cloud variability effects in large-scale atmospheric environments. For instance, Tiedke [1996] does a GCM impact study that accounts for the systematic effects of subgrid scale variability in Sc using Cahalan et al.'s [1994] parameterization of “effective” optical depth \( \tau_{\text{eff}} \), that is, the unique solution of \( T(\tau_{\text{eff}}) = \langle T(\tau) \rangle \), see Barker [1996] for another semianalytic approach with two variability parameters. IPAs are also applied deterministically (i.e., on a single-pixel basis) in cloud remote sensing applications as well as radiation budget computations.

As a general approach, IPA is currently under close scrutiny [e.g., Marshak et al., 1995a; Chambers et al., 1997; Davis et al., 1997a; Marshak et al., 1998; Zuidema and Evans, 1998]. In the special case of single-layer stratus however, the theory of IPA break down is well-developed:

1. For remote sensing, where only reflected photons are used, Marshak et al. [1995b] and Davis et al. [1997a] advance theoretical and empirical evidence that, for stratiform clouds, the
IPA is only valid for pixel sizes several times the radiative “smoothing scale,” equal to the harmonic mean of physical cloud thickness and the photon transport mean-free-path, typically 0.2–0.3 km.

2. Empirical estimation of solar absorbance in the atmospheric column call for both reflected and transmitted fluxes. Recent efforts in this area [Cess et al., 1996, Imre et al., 1996, and references therein] have encountered considerable spatial averaging problems due to IPA breakdown, as documented by Marshak et al. [1997b], Titov [1998], Davis et al. [1997b], and Barker and Li [1997].

7.2. Small Scales and Sub-Mean-Free-Path Variability Effects

Small-scale variations in cloud LWC also affect the flow of solar radiation through the atmosphere, but in a rather different way than their large-scale counterparts. If LWC fluctuates at scales smaller than a typical photon mean-free-path, the distribution of photon free paths between scatterings can no longer be exponential, even locally. As far as we know, this break down of a basic tenet of standard radiative transfer theory was first documented by Romanova [1975] in the framework of atmospheric optics as a general consequence of 3-D cloud geometry. Nonexponential photon free-path distributions are actually a recurring theme in the radiation transport literature at large; see Koyacik et al. [1998] for a statement on the same issue in plant canopies. The effect of sub mean free path LWC variability on large-scale cloud radiative properties is unknown at present. Because the perturbation is systematic in the sense of promoting longer paths between scatterings [Davis, 1992], it may have a significant role in resolving the current issue of enhancement by clouds of shortwave column absorption [Stephens and Tsay, 1990; Cess et al., 1996, and references therein].

Marshak et al.’s [1998] numerical case studies of sub-mean-free-path variability in stratus layers are based on the same SOCEX PVM data as used here; they show small bulk effects in albedo, transmittance, and absorbance, somewhat larger ones in the satellite radiances field. In contrast, an analytic study of sub-mean-free-path variability by Davis and Marshak [1997], more representative of broken cloudiness than stratiform cloud cover—suggests that rather strong effects should occur in transmission and the pathlength distribution (at conservative wavelengths), hence absorption (at nonconservative ones). This model, which uses power-law free path distributions instead of the standard exponential ones, has received support from recent oxygen A band observations [Pfeilsticker, 1999]. We note however that, in presence of extensive brokenness, the intercloud gaps dominate the mean in the overall (not just “in-cloud”) photon free-path distribution so the mean-free-path is commensurate with cloud size, possibly even larger.

8. Conclusions

Liquid water content data captured at 4-cm resolution with a PVM–100A probe during SOCEX in 1993 exhibit two distinct regimes for different scales (denoted by r). One regime goes from r ~ 8 cm to r ~ 2–5 m, and the other goes from r ~ 5 m to at least 2 km. See section 2.

Conclusions concerning scales r > 2–5 m are as follows:
1. Nonstationarity and intermittency of cloud LWC data are quantified and qualified in this large-scale regime which is clearly dominated by turbulent dynamics. See sections 3–5.
2. The “bifractal” properties (H1,C1) of large-scale SOCEX fields are remarkably similar to those of other nonstationary LWC signals measured during FIRE and ASTEX, as well as those of passive scalar density fluctuations in fully developed turbulence. This reminds us that passive advection is a reasonable first-order model for cloud LWC and suggests that there is a degree of universality in the turbulent dynamics that determines the horizontal structure of Se in three remote locales (California, Azores, Tasmania). See appendix B and sections 4–6.
3. Interesting differences between the three field programs arise in the scaling ranges, prefactors, and higher-order multifractal properties. This is also how the similarity of LWC and passive scalars in turbulence breaks down. At least one more parameter than used in bifractal analysis is needed to describe the multifractal behavior, but sampling of the strong events that dominate higher-order statistical moments is generally poor; so caution is required. See sections 4–6.
4. Neither of the standard classes of scale-invariant stochastic model, intermittent multiplicative cascades and Gaussian additive processes, are adequate to simulate the internal structure of boundary layer clouds; a new kind of hybrid (multiplicative/additive) models is needed. See section 6.

Conclusions concerning scales r < 2–5 m are as follows:
1. The small-scale LWC fields from SOCEX, and probably elsewhere, are stationary and show enhanced variance with respect to the level expected by (fractally interpolating the large-scale field, meaning that we assume the spectral scaling law obeyed by the large scale fluctuations can be prolonged down to mm-scales. This conclusion holds even after accounting for Poissonian noise resulting from the finite instantaneous sampling volume of the PVM–100A. See appendix A.
2. The existence of this component in the LWC fluctuations that exceeds the natural Poissonian variability at scales of a few cm’s confirms similar findings by Baker [1992] and Brenner [1993] based on FSSP data. See section 5.

This last point implies that the (large) scales deemed “homogeneous” by Baker’s [1992] standard are still highly variable from the standpoint of multifractal analysis, for example, higher-order structure functions. The effects of this large-scale variability on radiation transport in stratocumulus have been studied extensively (this research is briefly surveyed in section 7). In contrast, radiative consequences of the small-scale, essentially sub photon mean free path, variability are an open question that is partially addressed elsewhere for stratus layers [Marshak et al., 1998] and for broken cloud fields [Davis and Marshak, 1997].

Appendix A: Simulation of PVM Operation in a Non-Poissonian Environment

The goal of this simulation is to answer the following question. Suppose the structure of SOCEX (and other) clouds is in reality scale-invariant down to the 4–5 mm scales, roughly the Kolmogorov dissipation scale for turbulent velocity. If so, can the scale break at 2–5 m in the 4-cm (2 kHz) PVM–100A data be traced to the finite size of the sampling volume and/or instrumental filtering? (An aircraft speed of 80 m/s is used throughout this appendix.) If the answer is “yes,” the scale break is an instrumental artifact (and we have in essence developed a way of statistically removing the uninteresting Poisson noise superimposed on the actual liquid water fluctuations). If it is “no,” Poissonian fluctuations of droplet number and size
in the sampling volume do not explain the observed small-scale variability; consequently, the scale break is real, and another physical explanation must be sought.

A1. Instrument Description

The optical PVM-100A probe [Gerber et al., 1994] has an instantaneous sampling volume \( V_s \) defined by a diode laser beam of circular section \((0.37 \text{ cm in diameter, 11.43 cm long})\) oriented at a right angle to the direction of flight. In the following simulations, the PVM is viewed as a device operating in two stages:

1. It samples the second and third moments of the droplet-size distribution analogically in a long cylindrical volume defined by a laser beam aligned perpendicular to the flight direction; the volume is \( V_s = (\pi/4) \times (0.37)^2 \times 11.43 = 125 \text{ cm}^3 \). The sampling rate 20 kHz, that is, every 0.4 cm (so there is no volumetric intersection from one observation to the next).

2. It then filters out the highest frequencies in that dataset and with an analog low-pass filter and samples the output at an A/D conversion rate of 2kHz. Digitization uses 13 bits over an output in the range 0 5 V: after calibration, the LWC sensitivity is \( 1.6 \times 10^{-4} \text{ g/cm}^3 \).

3. For typical densities of CCN (cloud concentration nuclei), the number and sizes of droplets present in \( V_s \), about one cubic centimeter of cloudy air, will fluctuate according to Poissonian statistics.

There will be no correlations in this noise because there is no volume overlap (i.e., the cloud droplets irradiated by the laser are always new). However, the sampling volume is weakly dependent on droplet size, due to “vignetting.” Droplets have size-dependent diffraction patterns (measured at a random distance from the scattering center), so slightly different “effective” sampling volumes need to be assigned to droplets of different sizes. Figure A1a shows measured values of \( f_s(d) = V_s \delta n(d) / V_s \) versus droplet diameter \( d \). For future reference, the empirical frequency attenuation function \( a(f) \) of the A/D converter is plotted in Figure A1b.

A2. Droplet-Size Distribution

We need to assume some droplet-size distribution \( n(x;d) \), not necessarily uniform in space \((x \text{ denotes position})\). In the following, we use a lognormal distribution for simplicity; however, experimentation showed this assumption to be unimportant for the outcome of the simulation. We therefore write the number of droplets with diameter between \( d \) and \( d+\delta d \) in a unit volume as

\[
\delta n(x;d) = \frac{n(x)}{\sigma \sqrt{2\pi}} \exp \left[ -\frac{(\ln(d/\mu)-\mu)^2}{2\sigma^2} \right] \delta d, \quad d > 0.
\]  
(A1a)

where, apart from finite sampling effects, the spatial variability is assumed to be carried entirely by (total) droplet density or concentration

\[
n(x) = \int_0^\infty n(x;d) \delta d.
\]  
(A1b)

The normalized droplet distribution

\[
\frac{\delta n}{\delta d}(d) = \frac{n(x;d)}{n(x)}
\]  
(A1c)

has parameters

\[
\mu = \ln(\delta/d) = \int_0^\infty \ln(\delta/d) n(d) \delta d,
\]  
(A2a)

\[
\sigma^2 = [\ln(\delta/d)-\mu]^2 = \int_0^\infty [\ln(\delta/d) - \mu]^2 n(d) \delta d.
\]  
(A2b)

\[
\frac{\mu}{2} = \int_0^\infty \frac{d^2 n(d)}{d^2} \delta d,
\]  
(A3)

\[
\rho_w \text{ being the density of water. Instead of } \mu \text{ and } \sigma^2, \text{ it is conventional to parameterize the droplet distribution with effective droplet radius,}
\]

\[
r_e = \left( \frac{d^3}{d^2} \right)^{1/2} = \frac{3}{2\mu_w} \left( \frac{\text{LWC}}{\text{PSA}} \right),
\]  
(A5a)

and (dimensionless) effective variance [Hansen and Travis, 1974],

\[
v_e = \left( \frac{d^4}{d^2} \right)^{1/2} = 1.
\]  
(A5b)

For the lognormal distribution, we have

\[
\frac{d^n}{d^n} = \exp \left[ \mu_n + (\sigma^2/2)n^2 \right]
\]  
(A6)

with \( \mu_n = \ln(1+v_e) \) and \( \mu = \ln(\epsilon_n - 2.5 \ln(1+v_e)) \); therefore

\[
\frac{d^n}{d^n} = (2\mu_n)^n (1+v_e)^{n(n-5)/2}
\]  
(A7)

We use \( r_e = 14.8 \mu m \) and \( v_e = 0.015 \), but vary these parameters within the observed range does not change our results.

A3. PVM Operation in a Variable Environment

On the basis of the above instrumental and microphysical characteristics, we describe the five steps of our simulation and illustrate them with figures in physical and Fourier spaces. As a starting point, we use the LWC data collected and archived at 2 kHz. A 41-m long sample is presented in Figure A2a, extracted from the long flight segment 5 in Table 2 with typically low LWC values but high variability, including the intriguing up-spikes. We denote this field

\[
\varphi_i = \varphi(x_i), x_i = i\ell \quad (i = 1, ..., N)
\]  
(A8)

where: \( \ell = 4 \text{ cm} \) is the grid constant (pixel size); \( N = L/\ell \) is the total number of points \((N = 6.3 \times 10^4 \text{ here})\), \( L \) being the overall length of the record \((L = 26 \text{ km in this case})\).

A3.1. First step: Coarse-graining to a scale \( L_s \), roughly where the scale break occurs. We know that \( \varphi(x) \) is scale-invariant at least down to the scale break at \( L_s \approx 2-5 \text{ m} \) (see Figures 1 and 4a in the main text). To describe the LWC at this relatively coarse scale, we need to degrade the
high-resolution LWC data in equation (A8) by averaging over $m_p = \ell_p / \ell = 2^b = 64$ points if we take $\ell_p = 2.56$ m:

$$\bar{\varphi}_j = \frac{1}{m_p} \sum_{i=1}^{m_p} \varphi_i \quad (j = 1, \ldots, N_p),$$  \hspace{1cm} (A9)

where $N_p = [N/m_p]$, $[\cdot]$ denoting integer part. Using $m_p = 128$, hence $\ell_p = 5.12$ m, yields very similar end results. The resulting field $\bar{\varphi}$, displayed in Figure A2b, obeys equation (2) in the main text:

$$\langle |\Delta \bar{\varphi}(r)|^q \rangle \sim r^{\zeta(q)}, \quad q \geq 0, \quad \ell_p \leq r \leq L,$$  \hspace{1cm} (A10)

with, in particular, a spectral exponent $\beta = 2(2)+1$ in theory (Weiner-Khinchin theorem); in practice, we find $\beta = 1.6\pm0.1$ and $\zeta(2) = 1.53\pm0.02$.
A3.2 Second step: Fractal interpolation, down to a very small scale $\ell^*$. We now assume that the scaling prevails down to a scale $\ell^* = 0.005 \text{ m}$ at least for $q = 2$. (This means we are now dealing with $N^* = 512N_o = 8N$ points in all, corresponding to the 16 kHz collection rate used further on, close to the actual 20 kHz but in a convenient power of 2 ratio with 2 kHz.) How do we restore the nonarchived data? We used fractal interpolation [Ramsley, 1988c] with, as an interpolator, a nine-step bounded cascade model [Marshak et al., 1994]. This multifractal model was actually developed and used elsewhere [Cahalan et al., 1994c] to reproduce cloud liquid water path fluctuations at much larger scales (10 m to 10 km). This choice of model is not critical for the outcome: fractionally integrated singular cascades [Schertzer and Lovejoy, 1987], a linear spline, or even a constant, work just as well. At any rate, we now have

$$\tilde{\varphi}_i = \varphi_{j(i)} \prod_{k=1}^{i} W_k[j(k)] (i = 1, \ldots, N^*), \quad (A11)$$

where $j(k) = \lfloor \beta k \rfloor + 1, i = 1, \ldots, N^*$, and $n_c = \log_2(N^*/N_o) = 9$ is required to the number of (discrete, two-fold) cascade steps. We use mass-conserving bounded cascades, so that averaging $\tilde{\varphi}$ over scale $\ell_s$ gives exactly $\varphi_{00}^*$. The two parameters that control the random (unit mean) multiplicative weights $W$ were chosen so that the scaling in equation (A10) remains valid, at least for $q = 2$; in other words, the field $\tilde{\varphi}$ is guaranteed to be scale-invariant down to $\ell = \ell^*$. ($|\Delta \tilde{\varphi}(r)|^2 \propto r^{-1}, \ell^* \leq r \leq L.$) (A12)

A sample of the field $\tilde{\varphi}$ at $\approx$ mm resolution is plotted in Figure A2c.

A3.3 Third step: Generating the Poisson noise present in the PVM’s internal 16 kHz dataset. We now simulate the PVM–100A’s actual estimate $\tilde{\varphi}^*(x)$ of the physical LWC value $\varphi(x)$ for $x = i\ell^* (i = 1, \ldots, N^*)$, subject to its finite sampling volume. We first need to evaluate the ensemble-average droplet density from equation (A4b), with LWC(x) = $\varphi(x)$, and equation (A7):

$$n(x) = \bar{\varphi}(x) \left( \frac{\pi}{6} \rho_w \bar{d}^3 \right) = \left( \frac{3(1+\nu^2)}{4\pi \rho_w e^3} \right) \tilde{\varphi}(x) \quad (A13)$$

Next, we obtain the expected number of droplets $N_d(x)$ in volume $V_S$ at position $x$ from

$$N_d = V_S n, \quad (A14)$$

dropping the explicit dependence on position.

The actual number of droplets $v$ is then generated from the discrete Poisson distribution of mean $N_d$:

$$\text{Prob}(v = X) = \frac{N_d^X}{X!} \exp(-N_d), X = 0, 1, 2, \ldots \quad (A15)$$

Operationally, we generate the Poissonian deviate $v$ exactly if $N_v < 16$ by answering the following question: given an arbitrarily long sequence of pseudorandom numbers uniformly distributed on the unit interval, how many do we need to sequentially multiply before crossing the threshold $\exp(-N_d)$? For $N_d \geq 16$ (i.e., 4 standard deviations $\sigma_{\bar{d}} = \sqrt{N_d}$), we use a Gaussian approximation.

For each of the $v$ droplets, we assign a diameter drawn randomly from a population that follows the (normalized) lognormal distribution in equation (A1c): $d_i (i = 1, \ldots, v)$; explicitly, we take $d_i = 2 \times \exp[\mu + \sigma N(0,1)] (i = 1, \ldots, v)$ where $N(0,1)$ is a zero-mean, unit-variance normal deviate. We can then compute $\tilde{\varphi}$, the actual 16 kHz LWC measurement, from first principles:

$$\tilde{\varphi} = \left( \frac{4 \pi}{6} \rho_w \sum_{i=1}^{v} d_i^3 \right), \quad (A16)$$

with the understanding that $v = 0$ means no sum, hence $\tilde{\varphi} = 0$. Finally, we use the rejection method to simulate the effect of vignetting, as described in Figure A1a. A sample of this noisy $\tilde{\varphi}$ data is plotted in Figure A2d.

A3.4 Fourth step: Low-pass filtering of the simulated 16 kHz dataset. The frequency response $a(k) = \exp(-j2\pi k/80) (k = 0, \ldots, 80$ m/s) in Figure A1b is now used, to filter $\tilde{\varphi}$ in Fourier space. As a result, we have

$$\tilde{\varphi} \rightarrow \tilde{\varphi} = \left[ \frac{\pi}{80} \right] \tilde{\varphi} \quad (A17)$$

where $\Gamma[]$ is the Fourier transform. A fragment of filtered $\tilde{\varphi}$ ($i = 1, \ldots, N^*$) is plotted in Figure A2e.

A3.5 Fifth and last step: Resampling of the filtered 16 kHz dataset at 2 kHz. Figure A2f shows $\tilde{\varphi}$ ($i = 1, 9, 17, \ldots, N^*$), one of eight possibilities, with the original 4-cm spatial resolution. The background noise is comparable to that in the original data in Figure A2a; however, all of the up-steps have been obliterated. The whole simulation process was repeated for the short flight segment $6$ ($L = 2.4$ km, $N = 6 \times 10^4$), with high IWC-values and relatively small variability, intermittent occurrence of large jumps in both directions notwithstanding. Input and output products are plotted in Figures A2a and A2f, respectively. Here again the background noise level is adequately reproduced, not the strong/short events that come out of this noise.

We note a =10% discrepancy between the original (0.87 g/m$^3$) and final (0.80 g/m$^3$) spatial averages in Figures A2a and A2f that is traceable to the LWC that remains unobserved due to the empirical vignetting model described in Figure A1a. This bias in our “virtual” PVM-100A has no counterpart in the real instrument because its calibration is blind to vignetting effects. We incorporated vignetting into our simulations only to ensure that all known sources of Poissonian variability are properly accounted for.

A3.6 A summary, in spectral representation. Figure A3 illustrates the successive steps of our simulation with the wavenumber spectrum $E(k)$. We see a clear discrepancy (a factor ranging from 2 to 2.2 – 5.6) between the spectra of the simulated and measured data for scales smaller than $2$–5 m ($k > 2^{-1.10} \approx 0.21 \text{ km}^{-1}$ in Figure A3). This echoes the visually observable differences seen in the spatial samples (Figures A2a and A2a’ versus A2f and A2f’).

A4. Conclusions

We conclude that the Poissonian noise related to the PVM’s finite sampling volume is insufficient to explain the scale break at 2–5 m. Furthermore, the excess variance is contained in intermittently occurring intense events. This amounts to saying that there is more than Poissonian variability in small-scale cloud structure; in this respect, we concur with Raker [1992] and Brenguier [1993].
2.5 m spatial averaging

interpolating with bounded cascade model

low-pass filtering

through Poissonian noise generator

Figure A2.
Appendix B: Monofractal, Multifractal and Bifractal Data Analyses

A convenient device for comparing the basic scaling properties of geophysical signals and/or model calculations is the "bifractal" plane, with coordinates \((H_1, C_1)\) \cite{Davis et al., 1994a, 1996b; Marshak et al., 1997a}. The axes are assigned to two first-order, hence highly robust, scaling exponents that range from 0 to 1 (for 1-D data sets): horizontally, \(H_1\) from equation (6), an index of nonstationarity obtained directly from the data of interest \(\phi(x)\), vertically, \(C_1\) from equation (11), an index of intermittency obtained from the related measure \(\varepsilon_q(x)\) that describes the magnitude of small-scale gradients in \(\phi(x)\) in equation (8).

In the terms of Mandelbrot's [1977] fractal geometry, both exponents are co-dimensions of remarkable sets.

1. Abcissa: \(H_1 = (d+1) - D_6\)

   where \(d \leq D_6 \leq d+1\) is the fractal dimension of the rough graph of \(\phi(x)\) plotted versus \(x \in \mathbb{R}^d\). We are most familiar with the case \(d = 1\) (often just 1-D transects through 3-D fields), but this characterization carries over directly to 2D images (which may be cuts through 3-D fields), \(2 \leq D_6 \leq 3\) and \(H_1 = 3 - D_6\), as well as 3D fields, \(3 \leq D_6 \leq 4\) and \(H_1 = 4 - D_6\).

2. Ordinate: \(C_1 = d - D_1\)

   where \(0 \leq D_1 \leq d\) is the fractal dimension of the sparse set supporting those values of the measure \(\varepsilon_q(x), x \in \mathbb{R}^d\), that dominate its mean \(\langle \varepsilon_q \rangle\), a.k.a. the "information" dimension.

We call "bifractal analysis" the restricted but general scale-invariant characterization of correlations in geophysical data based solely on \((H_1, C_1)\). In other words, this is not a single-exponent or "monofractal" characterization (such as standard spectral analysis) yet not a fully multifractal approach either. In this paper, we actually follow a "bimultifractal" approach that uses both higher-order structure functions, yielding the \(\zeta(q)\) values, and a standard form of singularity analysis, yielding the \(K(q)\) values.

Davis et al. [1993, 1994b, 1996b] and Marshak et al. [1997a] discuss the potential for redundancy between \(\zeta(q)\) and \(K(q)\) exponents and the usual formula that relates them. The outcome is (1) that at least one new exponent is needed to obtain \(\zeta(q)\) from \(K(q)\), a role that can be assigned somewhat arbitrarily to \(H_1\), or the "cancellation" exponent \(1-H_1\) \cite{Ott et al., 1992; Vainshtein et al., 1994}, and (2) that \(K(q)\) is a "detranged" version of \(\zeta(q)\), that is, a linear term in \(q\) is removed.

Problems one can encounter in multifractal analysis are as follows: too much information in the (generally overlooked) proportionality constants and/or the occurrence of spurious multiscale due to limited data resources (accuracy, range of scales, and/or number of realizations) \cite{Davis et al., 1996b}. Even if the above problems did not occur, finding an interpretation for every exponent in a multifractal approach would lead...
to rather laborious modeling of data, hence the appeal of multifractal characterizations with a small number of parameters (compare section 6.2.2).

The most serious weakness of a monofractal approach is its fundamental ambiguity, easily demonstrated by the following example. Suppose we analyze some meteorological data \( q(x) \) and find a spectral exponent \( \beta = 2 \); equivalently (Wiener-Khinchin theorem), second-order structure function analysis using equations (4)-(5) with \( q = 2 \) yields \( \zeta(2) = 1 \). Such findings are representative of standard Brownian motion (Bm), the running sum of independent random variables, alternatively, \( q(x) \) behaves like any coordinate of a particle in a random walk (Figure B1). However, we cannot exclude a Heaviside step (Hs), a finite jump in \( q(x) \) over an infinitesimal distance in \( x \), with constant behavior before and after (Figure B2). We can think of the fingers model as “quiescent” atmospheric turbulence, the latter to an intermittent but strong event such as the passage of a “front,” and we can conceptualize boundary layer dynamics as a alternation and/or combination of Bm and Hs. We naturally wish to statistically distinguish the two situations in a robust manner, so we clearly need to go beyond spectral analysis.

One solution is to consider one-point Probability Density Functions (PDFs): Bm will tend to look Gaussian, whereas Hs should be distinctly bimodal. However, in practice, this will only work in very “pure” cases. Furthermore, PDFs tell us nothing about spatial correlations; for this, we are still relying entirely on the wavenumber spectrum which, alone, fails to detect intermittency (or lack thereof) in a signal.

---

**Figure A3.** Evolution of the energy spectrum through different stages of the noise simulation. Half of the data set 5 from Table 2 is used, \( \sim 3 \times 10^5 \) points. Three other curves show the energy spectra of the simulated data for two intermediate steps (fractal interpolation and Poissonian noise generation), and for the last step (after low pass filtering and resampling). Fragments of the data are displayed, respectively, in Figures A2a, A2c, A2d, A2c, and A2f.

\[
L'_{m} = 2^{18} \times 4 \text{ cm} \times 10^{-4} \text{ m}
\]

**Figure B1.** Simplified model for “quiescent” atmospheric dynamics. The upper curve is a sample of a Lévy Wiener process, otherwise known as (standard) Brownian motion or “Bm.” It is defined as the running sum of independent normally distributed random variables. Bm is scale invariant between the pixel and outer scales, hence a power law wavenumber spectrum \( E(k) \sim k^{-\beta} \) with \( \beta = 2 \). The lower curve is the absolute pixel scale gradient field, that is, the absolute values of the independent Gaussian deviates. This is essentially a field of white noise, that is, \( \beta = 0 \). Being nonstationary (\( 1 < \beta < 3 \)), we can only study increments of the former. Being stationary (\( \beta > 1 \)), we can legitimately investigate local averages of the latter.

Multifractal analysis easily resolves the ambiguity of spectral analysis:

\[
\begin{align*}
\text{structure functions:} & \quad \xi_{Bm}(q) = q/2 \quad q > 0, \\
\text{singularity analysis:} & \quad K_{Bm}(q) = 0. (B1a)
\end{align*}
\]

and

\[
\begin{align*}
\text{structure functions:} & \quad \xi_{Hs}(q) = 1 \quad q > 0, \\
\text{singularity analysis:} & \quad K_{Hs}(q) = q-1. (B1b)
\end{align*}
\]

**Figure B2.** Simplified model for the passage of a discontinuous “front.” The upper curve is a randomly positioned and randomly strong Heaviside step. The lower curve is the absolute pixel-scale gradient field, that is, a randomly positioned and randomly strong Dirac \( \delta \) function. The wavenumber spectra of these processes are indistinguishable from their counterparts in Figure B1. Indeed, phase information is not used in \( E(k) \), the Fourier space equivalent of spatial averaging in physical space, in this limited statistical approach where only Fourier amplitudes are of interest. However, the only Fourier difference between here and Figure B1 is synchronized rather than random phases. Note the ten-fold increase in vertical scale with respect to Figure B1 required to span the (finite) \( \delta \) function.
noting the identical spectral scaling exponents.

$$\beta_{Bm} = \beta_{Cm}(2\alpha^{2}+1) = \beta_{Cm}(2\alpha^{2}+1) = 2; \quad (B2)$$

Davis et al. [1996b] and Marshak et al. [1997a] give short derivations of equations (B1a)-(B1b). This comprehensive approach certainly sheds new light on the differences between Bm which is monofractal (and Gaussian), and Hs which is multifractal at least for structure functions (and highly non-Gaussian).

Bifractal analysis of this problem is considerably simpler and equally unambiguous:

$$\beta_{Bm} = \beta_{Cm}(2\alpha^{2}+1) = \beta_{Cm}(2\alpha^{2}+1) = 2; \quad (B2)$$

whereas

$$\beta_{Hs} = \beta_{Cm}(1, 1). \quad (B3b)$$

telling us, on the one hand, that the Bm and Hs components are on either end of C1 values intermittency scale and, on the other hand, they have somewhat different degrees of nonstationarity. Considering Figures B1-B2, both of these statistically correct statements also make intuitive sense.

When it comes to modeling a signal, it is generally recommended to focus first on a small number of well-defined (hence low $q$)-parameters rather than try to reproduce a large number of exponents, at least some of which (at high enough $q$) are ill-defined for lack of data (as discussed by Davis et al. [1996b] and others). We propose bifractal analysis as a compromise between too many and too few empirical constants, with simple statistical and geometrical meanings assigned to each axis in the $(H_1, C_1)$ plane.

Acknowledgments. This work was supported by the Environmental Sciences Division of U.S. Department of Energy (under grant DE-AC05-96ER61066 to NASA’s Goddard Space Flight Center) as part of the Atmospheric Radiation Measurement (ARM) program, and by the Physical Meteorology Program of the National Sciences Foundation (under grant ATM-9521073). We thank B. Baker, D. Baumgardner, I.-I. Brenguier, R. Cahalan, C. Durovere, Y. Liu, S. Malinowski, R. Pin cuc, and R. Shaw for many fruitful discussions, as well as two anonymous reviewers for useful comments. H.G. acknowledges the Department of Atmospheric Science of the University of Washington for flying the PVM-100A probe on their aircraft during ASTEX, and to the Division of Atmospheric Research of CSIRO for flying the PVM on their aircraft during SOUCEX. A.D. is grateful for the material and moral support he received for 5 years at NASA’s Goddard Space Flight Center: much of the work was leading to this paper was performed at the Climate and Radiation Branch (a.k.a. “Code 913”).

References


Gollmer, S. M., Harshvardan, R. F. Cahalan, and J. S. Snider, WIndowed


