4-1-2000

On the spatial distribution of cloud particles

Alexander Kostinski
Michigan Technological University

A. R. Jameson
RJH Scientific, Inc.

Recommended Citation

Kostinski, A., & Jameson, A. R. (2000). On the spatial distribution of cloud particles. *Journal of the Atmospheric Sciences, 57*(7), 901-915. http://dx.doi.org/10.1175/1520-0469(2000)057<0901:OTSDOC>2.0.CO;2
Retrieved from: https://digitalcommons.mtu.edu/physics-fp/243

Follow this and additional works at: https://digitalcommons.mtu.edu/physics-fp

 Part of the Physics Commons

On the Spatial Distribution of Cloud Particles

A. B. KOSTINSKI

Department of Physics, Michigan Technological University, Houghton, Michigan

A. R. JAMESON

RJH Scientific, Inc., Alexandria, Virginia

(Manuscript received 8 February 1999, in final form 25 May 1999)

ABSTRACT

Recent studies have led to the statistical characterization of the spatial (temporal) distributions of cloud (precipitation) particles as a doubly stochastic Poisson process. This paper arrives at a similar conclusion (larger-than-Poissonian variance) via the more fundamental route of statistical physics and significantly extends previous findings in several ways. The focus is on the stochastic structure in the spatial distribution of cloud particles.

A new approach for exploring the stochastic structure of clouds is proposed using a direct relation between number density variance and the pair correlation function. In addition, novel counting diagrams, particularly useful for analyzing counts at low data rates, demonstrate droplet clustering and striking deviations from Poisson randomness on small (centimeter) scales. These findings are shown to agree with pair correlation functions calculated for droplet counts obtained from an aircraft-mounted cloud probe. Time series of the arrival of each droplet are used to bin the data evenly so as to avoid corruption of the statistics through the operations of multiplication and division. Furthermore, it is shown that statistically homogeneous series of particle counts exhibit super-Poissonian variance.

Since it is not always practical or feasible to obtain such direct measurements, the possibility of studying cloud texture using a revival of the idea of coherent microwave scatter from cloud droplets is discussed, including a more complete interpretation of Bragg scatter that seems to explain some recent observations in clouds. Finally, the appearance of clustering and the subsequent geometric distribution of droplet counts are interpreted using basic considerations of turbulence.

1. Introduction

The statistical laws governing spatial distribution of particles in clouds are, obviously, of great interest and have been studied for decades. Perhaps the simplest and most basic distribution law to adopt for particle counting statistics in clouds is the Poisson distribution. Many texts and monographs have done so (e.g., Rogers and Yau 1989, p. 134) However, there is by now a considerable body of evidence in conflict with the Poissonian statistics (which tends to underestimate true variability) because clouds appear patchy on many length scales down to centimeters (e.g., Paluch and Baumgardner 1989; Baker 1992; Pawlowska et al. 1997; Pinsky and Khain 1997; and Davis et al. 1999). For a concise summary see Pruppacher and Klett (1997, chapter 2, pp. 28–30) and their appendix A-2.1.5, in particular.

What distinguishes our approach from the previous work in this field is the use of the correlation function formalism to characterize the stochastic structure of clouds. In particular, it was shown that ''patchiness'' can be quantitatively characterized by correlations of particle counts and that the mere presence of patchiness is often inconsistent with the Poissonian statistics (Kostinski and Jameson 1997; Jameson et al. 1998). These spatial correlations of droplet concentration may exist on several length scales. Perhaps the simplest multiscale approach is to assume a fractal (or multifractal) model, as has been done in, for example, Hentschel and Procaccia (1984) or Davis et al. (1999). At the other end of the spectrum, we recently advanced the Poisson mixture approach based on the assumption of a single correlation length scale and wide separation of measurement and correlation scales. One of our goals here is to suggest a deeper and more fundamental statistical physics approach that yields super-Poissonian variance (larger than Poissonian variance for the same mean) of particle counts without the restrictive assumption of wide scale separation and single correlation length. We accomplish this by using the correlation-fluctuation theorem (CFT), which is a direct relation between concentration variance and a two-point correlation function.

Corresponding author address: Prof. A. B. Kostinski, Department of Physics, Michigan Technological University, Houghton, MI 49931. E-mail: kostinsk@mtu.edu

Note that the previously advanced Poisson mixture interpretation does not enter this argument at all.

Another major goal of this paper is to demonstrate the possibility of observing the cloud texture (finescale correlation structure) by calculating the two-point correlation function directly and through the CFT, both of which are computed by analyzing aircraft-mounted cloud probe data of time series of particle arrival times. In the process, we introduce new tools for processing and displaying droplet time-of-arrival data. We focus on the following question: is the probability of detecting a droplet within next, say, 10 cm, enhanced by having just detected a droplet?

The defining feature of the Poisson process or of ''pure'' randomness is that such a conditional probability equals the unconditional probability of simply detecting a droplet within the next 10 cm of the probe path. In other words, there is no enhancement in the Poisson case. In this paper, by the direct use of coincidence counting, we demonstrate a striking pattern of droplet clustering and deviations from Poissonian behavior. This confirms previous reports of finescale patchiness from a different point of view and with different data and techniques. Furthermore, we argue that the use of CFT simplifies and widens the range of possible cloud probe experiments by removing the need for consecutive time series of droplet counts.

We also suggest coherent Bragg scatter by droplets themselves as a possible way to measure cloud texture. This was suggested already in the 1960s but was not actively pursued. We argue for the revival of that idea since the CFT provides an important new link to the Fourier transform of the correlation function, which determines the radar Bragg scatter off cloud droplets. We then proceed to elaborate on the previously suggested model of a Poisson mixture through an incorporation of recent results in fluid dynamics of particles in a turbulent flow such as vorticity-induced patchiness caused by turbulent entrainment, and by providing justification for the previously derived geometric distribution of counts via exponential tails of passive scalars in turbulence.

2. Definitions, terminology, and background theory of the Poisson process

This section is of a rather tutorial nature but we feel that it is essential for clarifying our differences from the work of others and for providing an important guide for subsequent sections. Let a number of cloud particles in a given unit volume be our random variable of interest (call it n) and assume that the random process is statistically homogeneous, that is, the moments of the random process $n(x)$ such as the mean and the variance are unaffected by shifts in the choice of origin. Below, in the data analysis section, we provide evidence for the statistical homogeneity of the data analyzed here (see appendix A).

It is important to understand that finescale physical inhomogeneities (patchiness of clouds) do not preclude statistical homogeneity. This further distinguishes our approach from previous work such as Pawlowska et al. (1997) who model clouds as an inhomogeneous Poisson process. To illustrate our approach, we can do no better than to quote from Landau and Lifshitz (1980, p. 351).

> The assertion that the particles of a homogeneous isotropic body (liquid or gas) are equally likely to be at any position in space, applies to each separate particle on condition that all the other particles can have arbitrary positions. This assertion certainly does not contradict the fact that, owing to their interaction, there must exist some correlation between the relative positions of the different particles. This means that if we consider, say, two particles at the same time, then for a given position of one particle, different positions of the other will not be equally probable.

Hence, unlike Pawlowska et al. (1997) or Davis et al. (1999), we treat cloud particles that are randomly distributed in space as a homogeneous but correlated random process. The so-called small-scale inhomogeneities often mentioned in the cloud physics literature are manifestations of correlation (stochastic structure) in this framework. Let us make this precise.

Let two volume elements dV_1 and dV_2 be sufficiently small so that they can contain either zero or one particle only and the probability of containing two or more particles is negligible. Hence, the average number of particles $\overline{n}dV$ is also the probability that a particle is in the volume element dV (Landau and Lifshitz 1980, p. 351).

Then, for a general statistically homogeneous random field, the joint probability $P(1, 2)$ of finding two particles in the two volumes dV_1 and dV_2 (one particle in each) is (e.g., see Green 1969, pp. 62–63)

$$P(1, 2) = \overline{n}^2 dV_1 dV_2[1 + \eta(l)], \tag{1}$$

where $\eta(l)$ is the so-called pair correlation function (in theory liquids) or two-point correlation function (in astronomy) and l is the separation distance between two volumes [statistical isotropy is implied by $\eta = \eta(l)$]. The $\eta(l)$ is defined operationally as

$$\eta(l) \equiv \frac{[\overline{N(l)N(0)} - \overline{N}^2]}{(\overline{N})^2} = \frac{\overline{N(l)N(0)}}{(\overline{N})^2} - 1, \tag{2}$$

where $N = nV$ and V is the test volume. We see from Eq. (1) that the assumption of statistical independence of counts in nonoverlapping volumes implies $\eta(l) = 0$ because only in this case is the joint probability simply a product of the individual ones. However, in the presence of correlations, the conditional probability of finding the second particle, given that the first one is there, is enhanced (or inhibited) by a factor of $(1 + \eta)$. Equivalently, from the definition of $\eta(l)$ in (2), it can be seen that the two-point autocorrelation function is identically zero in the absence of correlations (the Poissonian case).

If we interpret a ''blob'' (''void'') as a region of ap-

preciable positive (negative) correlation of particle counts then *the Poisson process is ideally random in the sense that only in the case of the Poisson process are there no blobs or voids at any length scale. Hence, the Poisson field is as even as randomness allows.*

The assumptions behind the Poisson distribution (e.g., Ochi 1990) are that 1) the process is statistically homogeneous; 2) the probability of detecting more than one particle in a given volume δV is vanishingly small for sufficiently small δV; and 3) particle counts in nonoverlapping volumes are statistically independent random variables (at any length scale). It is the last assumption of statistical independence at any length scale that is the true origin of the Poisson distribution because it prevents blobs or voids from appearing. Given the assumptions, the probability distribution of counts is given by

$$p(N) = \frac{\overline{N}^N \exp(-\overline{N})}{N!}, \tag{3}$$

where N is the random variable of interest (number of particles in a test volume), $p(N)$ is a probability of finding N particles in a test volume, and $\overline{N}$ is the mean rate. It is important to note that the variance of counts equals the mean ($\sigma_N^2 = \overline{N}$) for the Poisson distribution. Given our interpretation of a blob, we expect deviations from the Poisson statistics on these scales, for particles immersed in or entrained by vigorous turbulence because of spatial correlations of concentration introduced by uneven stirring. In this sense, Poissonian clouds are devoid of turbulence and resemble an ideal gas of molecules more than they resemble real clouds.

The Poisson distribution is a function of one parameter, the mean value (equal to the variance). There are two natural generalizations of the Poisson process: 1) inhomogeneous Poisson process and 2) doubly stochastic Poisson process (where the local mean is itself a random variable on a longer spatial scale). In the former case [adopted, e.g., in Pawlowska et al. (1997)], the mean is viewed as a spatially dependent but predictable deterministic variable so it does not contribute to variance. For example, in Fig. 1, the time series on the rhs has variations about the global mean. If these variations are treated as deterministic (predictable at least in principle), then they do not contribute to the variance (only random variables do).[1] Therefore, the distribution for the number of droplets observed during a spatial interval $(0, L)$ is given by (e.g., see Ochi 1990, 437–439)

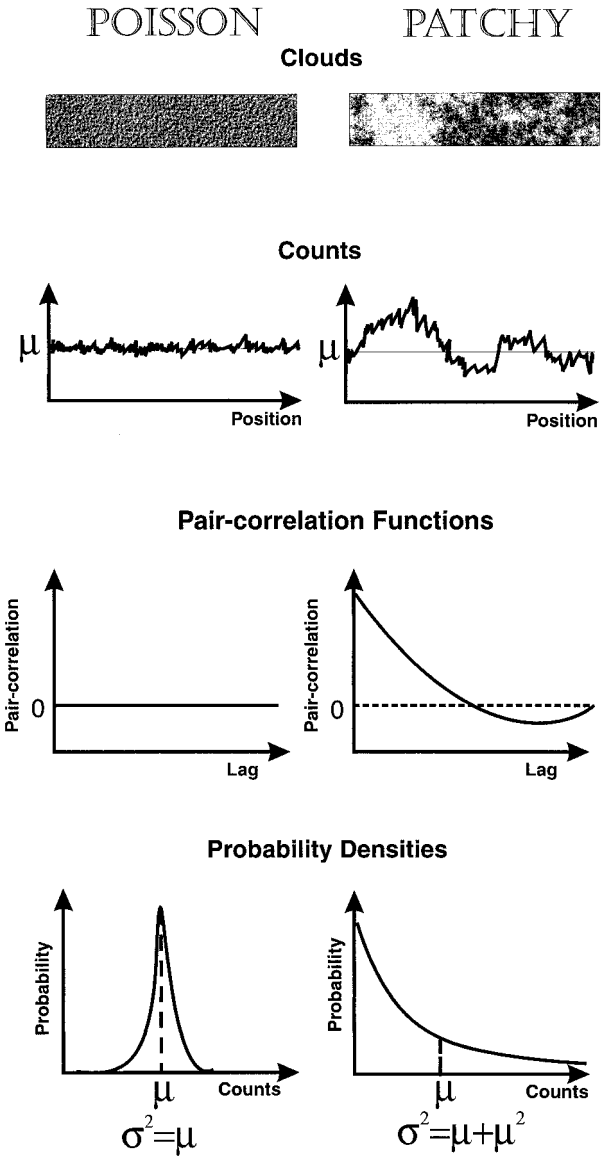

FIG. 1. The rhs of this figure is a schematic depiction of the chain of reasoning that leads to super-Poissonian fluctuations via the doubly stochastic random process model. It is essential to understand that the time sequence on the right is viewed as a stationary (homogeneous) but correlated process as opposed to an inhomogeneous Poisson process. Hence, ''stochastic structure'' is characterized by correlation functions. The Poissonian ''droplet-to-droplet'' fluctuations are riding on top of the ''patch-to-void'' variations. The independent variances add. The patch-to-void fluctuations are unpredictable and therefore are treated as random (i.e., any deterministic variations, of course, do not contribute to the variance).

$$p(n) = \frac{\left[\int_0^L \overline{n}(x)\, dx\right]^n}{n!} \exp\left[-\int_0^L \overline{n}(x)\, dx\right], \tag{4}$$

and this distribution still has the property that the mean and variance are equal.

To avoid this restriction (not satisfied by the data we

[1] Dispersion about a mean of a random process (function) is understood here as a measure of randomness and should not be confused with variation about a mean of deterministic functions in space or time.

have analyzed so far; see below), we adopt the doubly stochastic Poisson process model (e.g., see Ochi 1990). Instead of being inhomogeneous and Poissonian, it is statistically homogeneous (stationary) but not Poissonian as is confirmed by the data analysis in appendix A. The most important distinction is that the element of ''structure'' of clouds is viewed stochastically rather than deterministically; that is, rather than assuming predictability of large scale structures, we characterize it in terms of correlation functions. The situation is somewhat similar to that in modern statistical communication theory, where deterministic signals carry no information and all real messages are treated as random but with preassigned probabilities and intersymbol correlations.

Perhaps it is worth pointing out that models of homogeneous random processes are used to describe non-Poissonian statistics in other branches of science, for example, bunching of photons (Loudon 1983) or clustering of galaxies (Peebles 1980). These phenomena are in conflict with the assumption of statistical independence of spatially separated volumes. While in these examples spatial correlations appear because of interaction among objects (e.g., exchange interaction of photons or gravitational attraction of galaxies), in our case the cause of correlations is the intermittency of turbulent mixing. In our opinion, this makes clouds a particularly interesting and unique example of non-Poissonian statistics because it makes a point of fundamental importance: *correlation between positions of spatially separated objects does not necessarily imply physical interaction between these objects.*

For example, there is no direct physical interaction between drops separated by 100 m but their spatial locations are correlated to the extent that both belong to the same cloud patch. As another example in time, an arrival of a raindrop makes it more likely that the next drop will arrive shortly. The conditional part of the joint probability in (1) is enhanced by $(1 + \eta)$ because the first drop has been detected. In other words, raindrops arrive in bunches (patches of clouds) as was elaborated on in Kostinski and Jameson (1997) and Jameson and Kostinski (2000). From this point of view, the well-known expression ''it never rains but it pours'' is a manifestation of raindrop bunching. As we show below, this applies to clouds as well.

3. The correlation-fluctuation theorem and cloud texture measurements

Above, we argued that turbulent patches understood as regions of enhanced (blobs) or decreased (voids) concentration, represent long-range positive (negative) correlation of particle count fluctuations and therefore violate the statistical independence assumption behind the Poisson process. Presumably, the longer and stronger the correlation, the more substantial the violation of statistical independence assumption required in Poissonian statistics and, therefore, the larger the deviations from the Poissonian variance. In other words, a natural question arises: is there a relationship between the variance of counts in a fixed sampling volume and the spatial correlation of particle positions?

It turns out that the answer is yes. A powerful formula is available in statistical physics that relates the variance of counts in a given volume to the pair correlation function, integrated over the same volume. It was originally developed for the case of density fluctuations in gases and liquids but the derivation is completely general as can be found in Landau and Lifshitz [1980, p. 352, Eq. (116.5)] or Green (1969, pp. 62–63). It reads

$$\frac{\overline{(\delta N)^2}}{\overline{N}} - 1 = \frac{\overline{N}}{V}\int_V \eta(l)\, dV, \tag{5}$$

where, as in Eq. (1), $\eta(l)$ is the pair correlation function between particle counts in volumes V_1 and V_2, separated by distance l; δN is the deviation from the mean count in a given volume V; and $\overline{N} = \overline{n}V$, where $\overline{n}$ is the local mean concentration. Note that our η differs from the ν in Landau and Lifshitz (1980) by the factor $\overline{N}/V$. Also note that in the limiting case of no correlation we recover the Poisson relation $\sigma_N^2 = \overline{N}$. Thus, according to (5), the mean squared fluctuation $\overline{(\delta N)^2}$ of particle counts (variance of N) in a given volume is related to the pair correlation function integrated over the same volume.

Then the physical meaning of Eq. (5) is as follows: given a fixed mean concentration, the variance of particle counts is enhanced by the presence of correlations throughout the sampling volume V. This is illustrated in Fig. 1. One of the goals of section 6 is to carefully elaborate on this interpretation.

Note, however, that (5) is completely general and needs no additional assumptions or blob and void interpretation, such as introduced in Kostinski and Jameson (1997) and Jameson et al. (1998). After multiplying both sides of (5) through by $\overline{N}$ and rearranging, we obtain

$$\overline{(\delta N)^2} = \overline{N} + \overline{\eta}\overline{N}^2, \tag{6}$$

where the volume-averaged pair correlation function $\overline{\eta}$ is defined as

$$\overline{\eta} = \frac{1}{V}\int_V \eta(l)\, dV. \tag{7}$$

[From the ''random walk'' point of view, the quadratic term on the rhs of (6) can be viewed as the coherent displacement of the walker if N is interpreted as a number of steps.]

The physical meaning of the Poissonian limit is as follows. As the dimensions of the sampling volume V become much larger than the correlation distance, most of the integration region in $\overline{\eta}$ does not contribute and $\overline{\eta}$ approaches zero. Then the Poissonian relation $\sigma_N^2 = \overline{N}$ is recovered. This can also happen if negative and positive correlation regions nearly cancel for a given

sampling volume (wave phenomena such as Kelvin–Helmholtz instability could result in sinusioidally decaying η). At the other end, if $V^{1/3}$ is much smaller than the correlation distance, the quadratic term is always there and $\overline{\eta}$ can be approximated by $\eta(0)$. Hence, if the average number of particles $\overline{N}$ is much larger than unity, the quadratic term in (6) dominates. This result is quite general and is not model dependent. Note that, depending on the resolution, some experiments will pick up the super-Poissonian variance and some will not.

Equation (5) suggests a possibly simple means of measuring cloud texture (small-scale variability) as expressed by the pair correlation function $\eta(l)$ via a direct measurement of the lhs of (5). For example, the following procedure might be used in aircraft-mounted cloud probe studies.

1) Pick a small (compared with a small turbulent eddy) sampling volume dV. Here small might mean comparable to the viscous dissipation (Kolmogorov) scale of turbulence; for example, dV can be set to several cubic centimeters. Note that much larger sampling volumes may ''wash out'' finer scale variability.
2) Count particles in dV for many consecutive ''snapshots'' (realizations of a random process) along the flight path to accumulate a time series of measurements of N. These measurements then form an ensemble of realizations of a random variable N.
3) Compute the lhs of (5), which is a simple calculation of the variance, and the mean of the time series of counts (''relative'' to the Poissonian relation $\sigma_N^2 = \overline{N}$). Such processing has been done in the past (e.g., Baker 1992; Jameson et al. 1998).
4) Using the same measurements, increase the size of sampling volume and repeat the procedure to get $\int_v \eta(l)\, dV$ on larger spatial scales. This way η as a function of l can be deduced by approximating the integrals by their mean values.

One can also monitor directly the average product of volume counts separated by a fixed distance (e.g., coincidence counters such as the ones used in nuclear physics), but this involves mounting two probes and may be more difficult. The above steps can also be applied to the analysis of recorded particle interarrival times, as we will now demonstrate.

4. Observations of small-scale clustering and pair correlation functions

During the Tropical Ocean Global Atmosphere Coupled Ocean–Atmosphere Response Experiment (TOGA COARE) in 1992–93, the National Center for Atmospheric Research Electra aircraft flew at a speed of 130 m s^{-1} in a saturated environment about 1 km (3 km above the ground level) above the cloud base of raining, warm tropical clouds. Attached to the wing, was a Particle Measuring System (PMS) probe capable of imaging the shadows cast by cloud droplets passing through a wing-oriented linear array of downward-directed laser beams as the aircraft moved horizontally. The probe cross-sectional area was 9.6 mm^2. For this study, we chose data on 50-μm diameter droplets imaged with the PMS 2D-C instrument.[2] Furthermore, a traverse without cloud breaks was chosen.

The data consist of time intervals between successive droplet arrivals over a 20-s total duration that traversed about 2.75 km and included slightly over 1400 50-μm diameter droplets. While these observations were used to study interdrop distances, as described in greater detail in Jameson et al. (1998), here we use these observations differently.

Specifically, beginning with the first we measure the distances to all the subsequent droplets. We then divide the total 2.718-km path into 10-cm increments and assign each droplet to one of these bins. This, in effect, transforms the data from a time series of the number of 50-μm droplets into a spatial series of the number per 10-cm path increment (corresponding to the sampling volume element V of about 10 $mm^2 \times 10$ cm $= 1$ cm^3). Later on, in order to implement the CFT, we process the data for a series of increasing ''sampling volumes'' traversing distances from 10 cm to 1 m in 10-cm increments.

For the 10-cm resolution, $\overline{N}$ then is a number of 50-μm particles per cubic centimeter, averaged over the entire pass through the cloud. The $\overline{N}$ is about 0.05 per each V; that is, only one 10-cm interval out of 20 has a particle in it and the total path contains 27 180 intervals. To test for clustering at such small scales and low counts, we first ask the following question: is the probability of detecting more than one droplet within a 10-cm interval larger than the one given by the law of independent events?

Figure 2 clearly shows that the answer is yes. For example, the expected value of 10-cm intervals containing three and four particles for independently distributed particles is shown in Fig. 2 as about 1 and 0, respectively. This is obtained by simply drawing 1400 uniformly distributed random numbers on the (0, 1) interval and histograming into 27 000 bins. This is clearly much lower than the actual data show. The same conclusion holds (although slightly less dramatic) when the procedure is repeated for even a finer resolution of 1 cm (270 000 intervals and 1 in 200 intervals containing more than zero particles), as is shown Fig. 3. This type of exploratory data analysis and display is designed to achieve the finest possible resolution and it is suitable only for very low counts when the vast majority of occurrences are 0 or 1 events.

To gain further insight into clustering among con-

[2] Data on particle diameters from 25 to 800 μm in 25-μm intervals were available and we chose the second smallest size bin as the most reliable one (see Jameson et al. 1998, for details).

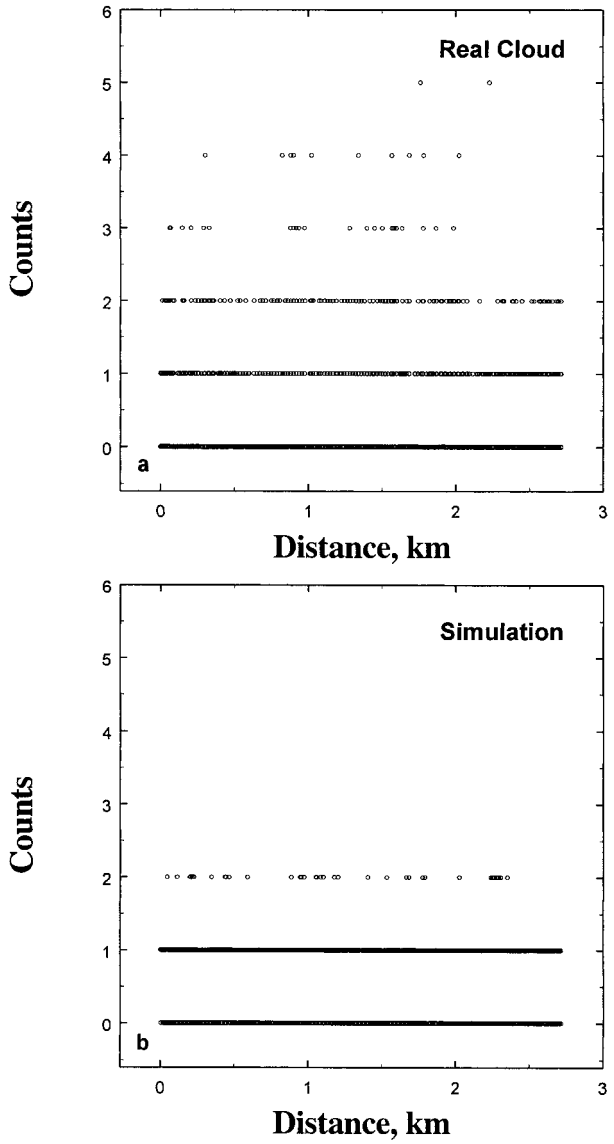

FIG. 2. (a) Beginning with the first droplet, the distances to all the subsequent 1404 droplets were measured. The total 2.718-km path was then divided into 10-cm increments and each droplet is assigned to one of these bins, which transforms the data into a time series of the number of 50-μm droplets per 10-cm path increment (corresponding to the sampling volume element V of about 10 mm^2 $\times$ 10 cm = 1 cm^3). These data are compared with the Poissonian simulation (b) based on independent trials (Bernoulli- or binomial-type experments for very small expected value) simply by drawing 1400 uniformly distributed random numbers on the (0, 1) interval and histograming into 27 200 bins.

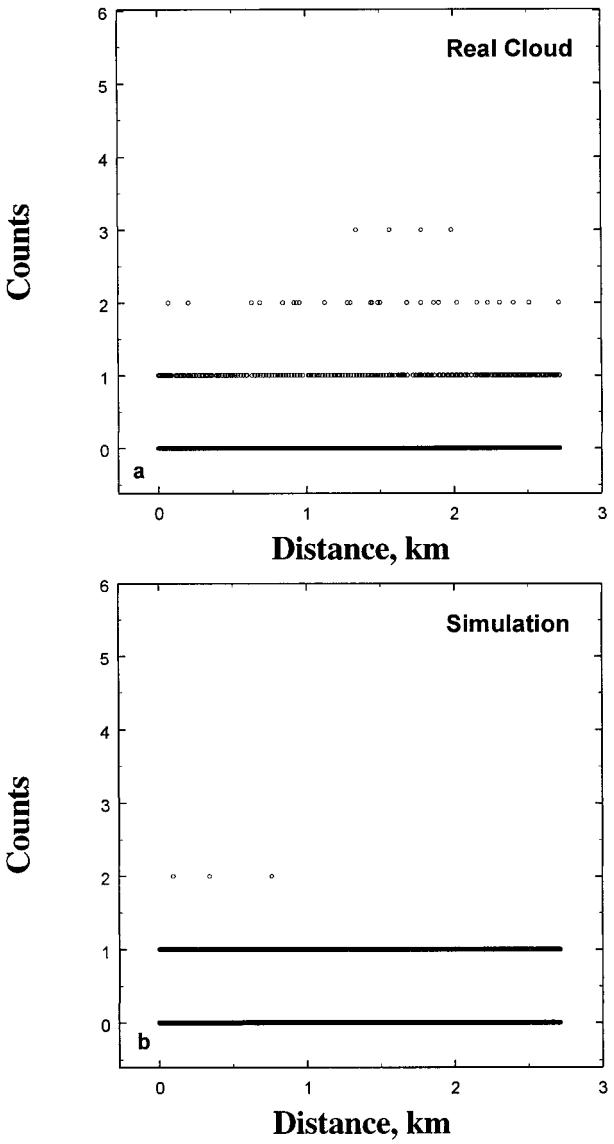

FIG. 3. The same comparison as for Fig. 2 but for intervals of only 1 cm. The Poisson simulation (b) is now performed on 272 000 bins.

secutive intervals (rather than inside each one), it is natural to turn to the pair correlation function $\eta(l)$ as defined by Eq. (2), which is specifically designed to monitor correlation among counts whose intervals are separated by a specified distance (lag). However, we operate in the domain of such small counts (19 out of 20 values are 0) that one must be cautious in interpreting the results; for example, $\eta(l)$ is expected to be erratic. The directly computed and the average pair correlation functions (denoted as $\overline{\eta}$) are shown in Fig. 4, for the 1- and 10-cm resolutions. The $\overline{\eta}$'s were computed from the CFT as

$$\overline{\eta} = \frac{1}{V}\int_V \eta(l)\, dV = \frac{\overline{(\delta N)^2}}{\overline{N}^2} - \frac{1}{\overline{N}}. \qquad (8)$$

The zero-lag values are excluded in the calculations because unlike the conventional autocovariance function, η is not bounded by unity at zero separation. In fact, as one decreases the interval and as $\overline{N}$ approaches zero, $\eta(0)$ diverges as $1/\overline{N}$ in the nearly Poissonian regime. This is why $\overline{\eta}$ is expected to be a bit higher than η; the large values at small lags make a large contri-

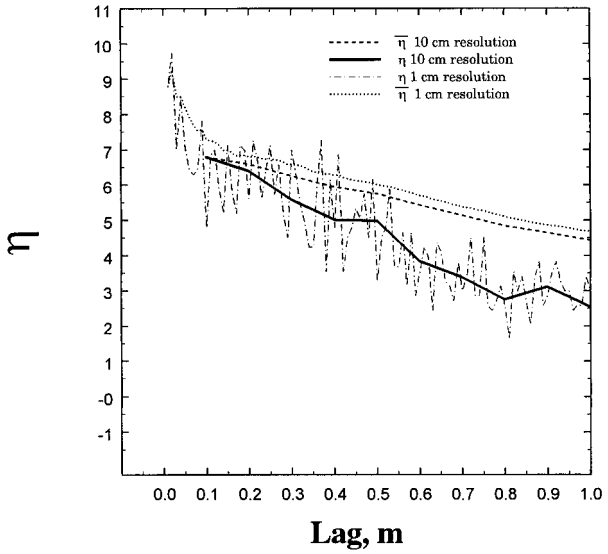

FIG. 4. Here we plot the pair correlation functions based on the 1- and 10-cm-interval data and for lags up to 1 m (excluding zero lag; see text). The correlation functions are computed on the basis of the direct definition [Eq. (2)] and the average correlation functions are derived using the CFT relation [Eq. (8)].

bution to the integral. In spite of the difficulties of interpreting correlation functions when counts are so low, Fig. 4 confirms our previous conclusions: for both 1- and 10-cm lag increments we see correlations at centimeter separations that decay at longer distances. Also, it is encouraging that the results of Fig. 4 agree on 1- and 10-cm resolutions.

The $\overline{\eta}$'s are, of course, smoother than the directly computed η's because of the huge number of realizations. Indeed, every 10-cm interval of the entire traverse is regarded as a new realization of the volume count. This is an important advantage of the CFT: the number of realizations is so large that it is possible to obtain statistically meaningful results from one path through a cloud. Furthermore, and perhaps more importantly, one need not record spatially nor temporally arranged series, in order to compute $\overline{\eta}$'s. For example, one can leave several probes of different volumes on a mountain top (high enough for frequent cloudiness, e.g., Mauna Loa) and make automated particle measurements, say, every hour for many days. This is sufficient to estimate $\overline{\eta}$'s at different spatial scales (different volumes of probes) through the CFT, especially when $\overline{N} \gg 1$. (In passing, we also note that the ''box counting method'' applied to the 50-μm particle data on resolution from centimeters to 100 m reveals no signs of fractal scaling.)

In addition to the simplicity of the proposed approach, we stress the fact that it provides subresolution information; that is, the pair correlation function *inside the sampling volume* is measured. This is illustrated by the 10-m sampling data of appendix A. Figures A1 and A2 establish statistical homogeneity of the data as reflected in the constancy of the local mean and the variance. We see from Fig. A3 that the data pair correlation function is barely significant for the 10-m and longer separations because it is so similar to the equivalent Poisson case. Yet Fig. A4 shows that the variance of the data is 4.25 times the variance of the equivalent Poisson distribution! This is in complete agreement with CFT because in these data it is the sub-10-m correlations that make a significant contribution to the 10-m volume integral.

Note that the relevant physical mechanisms enter the picture through the functional form of the pair correlation function versus the separation distance. Thus one would expect η to depend on particle size, the state of turbulence (e.g., the Reynolds number, energy dissipation rate), etc. Hence, the underlying physics can be studied by means of particle counting. Aside from direct observations, however, it should be possible in principle to observe stochastic structure of clouds using remote sensing based on the idea of Bragg scatter in liquids.

5. Remote sensing of cloud texture: Coherent scattering from cloud particles

While the previous section shows that spatial correlation can be observed directly, such observations may not always be convenient or plausible. Is there, then, a way to detect spatial correlations of droplets remotely? An approach to measure $\eta(l)$ in physical science has been known since the discovery of coherent scattering of X rays by liquids or gases near phase transition. There, the scatterers form a random but spatially correlated array and it is the Fourier transform of the pair correlation function (sometimes called the structure factor) that is measured via Bragg scattering (e.g., see Green 1969, pp. 59–63). The key physical idea is to employ a spectral representation of a random process because the random but correlated collection of scatterers can be viewed as a superposition of diffraction gratings, one of which is in ''resonance'' with the incident radiation. A mathematical manifestation of this statement is the Wiener–Khintchin theorem about the autocorrelation function and the power spectrum being a Fourier pair. Hence, the critical role is played by the Fourier transform of the pair (or auto-) correlation function. A particularly striking example is that of critical opalescence where a transparent gas becomes turbid and opaque as a phase transition is approached because molecules develop long-range correlations (blobs) comparable to the wavelength of visible light.

However, it appears that coherent scattering has been reserved in radar meteorology for scattering off the continuous fluctuations of index of refraction caused by turbulence (temperature and humidity) but not so much for particulate matter. Specifically, in the most detailed analysis we know, Gossard and Strauch (1983, p. 60) state the following.

> In calculating the scatter from cloud and precipitation particles it is usually assumed that the scatterer concen-

> tration is random in the sense that the particle concentration in neighboring parcels is completely uncorrelated where the meaning of ''parcel'' is an atmospheric volume of size approximately $(\lambda/2)^3$. If the number of scatterers $\delta N(r_1)dr_1$ is independent of $\delta N(r_2)dr_2 \ldots$ then
>
> $$I = \int_V \overline{(\delta N)^2(r)}\, dr = \int_V \overline{N}(r)\, dr = N_T. \qquad [9]$$

Note that the Poissonian-like process $\sigma_N^2 = \overline{N}$ is implied in the above equation. A little later on in Gossard and Strauch (1983, p. 61) we find the following.

> The possibility that number density in neighboring parcels is not completely uncorrelated in precipitation was considered briefly . . . and rejected as an important consideration. However, it is not so evident that a coherent (Bragg) scatter is negligible in clouds, where number densities can be as high as 1000 cm^{-3} and this possibility was considered from a theoretical standpoint by Smith (1964), by Naito and Atlas (1964) and by Chernikov (1968). The subject was apparently never pursued further because of the difficulty of measuring cloud number density spatial spectra. . . .

We agree with the authors that ''it is not so evident'' that Bragg scatter of cloud droplets is negligible, and perhaps this is the right time to reexamine the issue.[3]

Indeed, we argued above that Eq. (5) provides relatively simple means of measuring the correlation function and the related Fourier transform (cloud number density spatial spectrum). Furthermore, as discussed in the introduction, recent experimental evidence and results of this study suggest that spatial correlations between cloud droplet counts exist on the centimeter scale. This is also suggested by the theoretical arguments described in the following section. On the other hand, common ground-based radar wavelengths are in the range of 3–10 cm.

Thus, it is plausible that coherent radar scattering off cloud droplets is detectable because the random but correlated collection of cloud scatterers is equivalent to a superposition of diffraction gratings, one of which is likely to be in resonance with the incident radar wavelength. Note that even if N is decomposed as $N = \overline{N} + \delta N$, where $\overline{N}$ is assumed deterministic and therefore does not contribute to the scattering (as is done, e.g., by Gossard and Strauch), the coherent contribution is proportional to $(\delta N)^2$. On the other hand, whenever $\overline{N} \gg 1$ and correlations are present, $(\delta N)^2 \propto \overline{N}^2$ according to Eq. (6), so that coherent intensity ends up quadratic in $\overline{N}$. This is in contrast to the incoherent intensity, which is linear in $\overline{N}$.[4]

More precisely, the coherent scattered intensity I is given by

$$I = \overline{(\delta N)^2} \int_0^\infty lC(l)\, \sin(2\kappa l)\, dl, \qquad (10)$$

(e.g., see Gossard and Strauch 1983, p. 58), where $\kappa = 2k \sin\theta$, θ is the scattering angle, k is the wavenumber, and $C(l)$ is the spatial correlation function of particle counts similar to our $\eta(l)$ [see Eq. (2–12) of Gossard and Strauch 1983].

In the backscatter case this reduces to

$$I = \overline{(\delta N)^2} \int_0^\infty lC(l)\, \sin(-2kl)\, dl. \qquad (11)$$

Hence, the radar echo is proportional to the variance of the particle number density. This fact renders super-Poissonian counting statistics [presence of the $(\overline{N})^2$ term] essential for the existence of Bragg scattering.

In our opinion, recent dual-wavelength (3 and 10 cm) radar observations of small cumulus clouds by Knight and Miller (1993) and Knight and Miller (1998) appear to have several ''symptoms'' of particulate Bragg scatter. In particular, they report that Bragg and Rayleigh returns are correlated. This is to be expected if both types of return come from the same source: cloud particles.

Also, Knight and Miller (1998) report that the Bragg component is larger for the longer wavelength. This is plausible from the particulate Bragg scatter point of view as the spatial correlation of concentration may have a stronger Fourier component (''diffraction grating'') at 10 cm than at 3 cm. For example, vorticity-induced patchiness (Shaw et al. 1998), which is discussed further in the next section, occurs on the scale of about 10 Kolmogorov scales or about a few centimeters. To take a concrete example, consider a correlation function model of a sinusoidally decaying exponential (x is in the direction of propagation of the radar waves),

$$C(x) = \exp\left(\frac{x}{L_1}\right) \cos\left(\frac{x}{L_2}\right), \qquad (12)$$

whose Fourier transform is

$$\phi(k) = \frac{L_1}{1 + (L_1)^2(k - 1/L_2)^2} + \frac{L_1}{1 + (L_1)^2(k + 1/L_2)^2} \qquad (13)$$

[3] Our framework differs from that of Naito and Atlas (1966) or Smith (1964) in several essential aspects. Previous studies did not appreciate the central role of the Poisson distribution in quantifying deviations from ideal randomness and equated lack of correlation with complete and latticelike, rather than statistical, uniformity (e.g., see Smith 1964, p. 204). Furthermore, these studies treated local mean counts and gradients as deterministic functions. Negative correlations were not allowed (our voids), etc.

[4] Recall that the number of cloud droplets in the radar resolution volume ($\overline{N}$) might be on the order of 10^{15} for 100 particles per cubic centimeter and radar resolution volume of 10^7 m^3. Under the right circumstances, that is, proper spatial correlation, squaring this number may more than offset the D^6 size dependence of the radar Rayleigh backscatter cross section of these cloud droplets.

and has a peak at L_2 (spatial period of a wavelike disturbance, e.g., or preferred spacing of vorticity-induced filaments) that yields stronger radar return at that wavelength.[5] If L_2 is closer to 10 cm, S-band radar echo is stronger. This is a speculation, of course, in the absence of reliable (higher counts) data on correlation functions at the centimeter scale but it does yield a consistent picture and suggests intriguing measurements.[6] We are still left with the question of the physical origin of the super-Poissonian statistics. One possible interpretation based on the doubly stochastic Poisson process model is considered in the next section.

6. On the physical interpretation of the Poisson mixture

The purpose of this section is to elaborate on the previously developed Poisson mixture interpretation (Kostinski and Jameson 1997; Jameson et al. 1998) in light of recent results in fluid dynamics and turbulent mixing. Consider a simple distribution of blobs, all roughly of a size L and relative voids of about the same size, for example, such as depicted at the top right of Fig. 1, and imagine traversing any one of such patches and measuring cloud particle counts per unit volume whose dimension is much smaller than L. Then, particle counts will obey (3) as the local mean concentration ($\overline{n}$) remains constant because the experiments are confined to a given patch. However, on longer spatial scales (larger than L), the concentration itself will fluctuate as measurements move from patch to patch. Hence, the time series depicted on the right of Fig. 1 results. Kostinski and Jameson (1997) interpreted the distribution of $\overline{n}$, denoted $p(\overline{n})$, of the patches as a random variable in such circumstances. Thus, (3) holds *only when conditioned upon a constant mean concentration,* and to obtain the total distribution, one must integrate over $p(\overline{n})$,

$$p(n) = \int_0^\infty p(n|\overline{n})p(\overline{n})\, d\overline{n} = \int_0^\infty \frac{\overline{n}^n \exp(-\overline{n})}{n!} p(\overline{n})\, d\overline{n}, \qquad (14)$$

where the vertical bar denotes conditional probability. Hence, the process is doubly stochastic; that is, Poissonian fluctuations ''ride on top'' of the longer scale blob-to-void fluctuations as depicted in Fig. 1. Again note that despite the presence of physical inhomogeneities (patches), the random process is regarded as statistically homogeneous because $p(\overline{n})$ does not depend on the choice of origin. Realizations of such process on scales, comparable to correlation length, may appear to have ''trends'' but these are random as well and are to be viewed as correlations rather than predictable deterministic structures. This is an essential difference between the doubly stochastic and inhomogeneous Poisson process interpretations.

For conditional random variables (e.g., Ochi 1990, pp. 65–66),

$$\sigma_n^2 = E_{\overline{n}}[\sigma^2(n|\overline{n})] + \sigma_{\overline{n}}^2(E[n|\overline{n}]), \qquad (15)$$

where E denotes the expectation value. Thus, the variance of a Poisson mixture is enhanced beyond that of a pure Poisson probability distribution function (pdf) by the variance of $\overline{n}$ (the blob-to-void contribution), that is, $\sigma_n^2 = \sigma_P^2 + \sigma_{\overline{k}}^2$ where the first term is the pure Poisson contribution, that is, $\sigma_{\overline{k}}^2 = \mu \equiv \int_0^\infty \overline{k} f(\overline{k})\, d\overline{k}$, and $\mu = E(\overline{n})$ is the global (averaged over many correlation lengths) mean concentration of cloud particles.

To make further progress with the doubly stochastic Poisson process approach, we need to know more about the physics of the formation of the enhanced concentration filaments and their probability distribution $p(\overline{n})$. A physical discussion of this distribution is the main purpose of this section. In the language of fluid dynamics, this is a question of probability density function of the particle concentration fluctuations in a turbulent flow.

Let us begin by pointing out that even on a small scale (say, 10 cm), patchiness of cloud particles is likely to be ubiquitous to the extent that turbulent air, in which these particles are immersed, is ubiquitous. The physical picture that emerges from recent fluid dynamics studies is that the formation of patches and filaments is due to the interplay of intense and spotty random vorticity fields of atmospheric turbulence and a small amount of particle inertia as they slowly fall through the eddies while being only partially entrained by them.

Indeed, many recent studies have indicated that the spottiness of a concentration field remains there even when sedimentation or gravitational settling take place. For example, Maxey and Corrsin (1986) pointed out in their abstract that ''particles with weak inertia show a strong tendency to collect along isolated paths.'' Squires and Eaton (1991) echoed (we quote from their abstract) that ''turbulence may actually inhibit rather than enhance mixing of particles.'' See also Fig. 5, of Wang and Maxey (1993, p. 42) for the visual evidence of clumping, and Shaw at al. (1998). Basically, turbulent eddies push such particles away from the center of the vortex, and they tend to collect in high-strain and low-vorticity regions. Some implications of this phenomenon in cloud physics have already been examined by Pinsky and Khain (1998) and Shaw et al. (1998). The

[5] Note that the pair correlation function differs from the autocorrelation function simply by normalization ($\overline{n}^2$ vs $\overline{n^2}$) and subtracting off the unity.

[6] In this connection, it is interesting to note that Bragg scattering from raindrops may also be present in radar scattering off precipitation because smallest precipitation patches are likely to exist on spatial scales of several centimeters (Jameson and Kostinski 2000). Hence, one might be able to observe the coherent scatter with an S-band or longer radar waves (e.g., wind profilers).

important point here is that *even an initially homogeneous purely Poissonian distribution of cloud particles with small inertia quickly evolves into collection of filaments simply by virtue of falling through a random vorticity field created by atmospheric turbulence.*

Having established the presence of patchiness (and, therefore, correlation function η), we conclude via Eqs. (5) and (6) that the mere presence of the pair correlation function η is sufficient to indicate super-Poissonian variance. Thus, non-Poissonian statistics follow from presence of turbulence, at least on some spatial scales. This chain of reasoning is depicted schematically in Fig. 1.

Vorticity and partial entrainment need not be the dominant mechanism. For example, imagine an initially homogeneous blob of light cloud droplets is twisted and distorted by a succession of turbulent eddies as time progresses. The resultant patches must be viewed stochastically (e.g., see Brodkey 1995, p. 329, Fig. 14-48 in particular). Let us regard cloud droplets as passive tracers in a turbulent velocity field and apply recent results in the field of turbulent mixing (e.g., see Hozler and Siggia 1994, particularly images of patchy mixing on pp. 1826–1830). These clearly show that vigorous turbulence is a quick, but not a thorough, mixer in the sense that patches persist for a long time until molecular diffusion finishes the job.

But what can we state quantitatively about the probability distribution $p(\overline{n})$ of the ''patchy'' concentration field? The exponential distribution of concentration is a good starting point because of the Markov ''lack of memory'' property $e^{-(x+y)} = e^{-x}e^{-y}$ so that the joint probability is simply a product of individual probabilities (see Kostinski and Jameson 1999 for more details). Also, several recent and extensive studies indicate that the pdf of passive contaminants $p(\overline{n})$ tends to an exponential form under a rather wide set of conditions (Pumir et al. 1991). Sinai and Yakhot (1989) provide a theoretical framework (based on the underlying fundamental turbulence equations of random advection) for the evidence of exponential pdf of local mean concentration.

We therefore set

$$f(\overline{n}) = \frac{1}{\mu}\exp\left(-\frac{\overline{n}}{\mu}\right). \tag{16}$$

where μ is the global mean concentration; that is, it is averaged over many concentration bursts and voids. We expect fluctuations in particle count to be in excess of that implied by Poisson statistics because of variations in the local mean concentrations (bursts and voids), as shown in Fig. 1. The result of integration in (14) with (16) inside the integral verifies this anticipation. It is the geometric distribution

$$p(n) = \frac{1}{\mu + 1}\left(\frac{\mu}{\mu + 1}\right)^n. \tag{17}$$

Note only a single parameter (as for the Poisson distribution): the global mean particle count μ but the shape of this distribution is completely different from the Poissonian one (much longer tail) and can be easily compared with experimental histograms. Also note that it is the exponential tails of the concentration distribution that are reflected in the geometric tail of the distribution of counts. In other words, the fluctuations are enhanced over those of (3):

$$\sigma_n^2 = \mu + \mu^2. \tag{18}$$

There are two independent sources of randomness whose contributions therefore add: the regular Poisson fluctuations at a fixed concentration ($\sigma_n^2 = \mu$) and the average concentration fluctuations themselves ($\sigma_n^2 = \mu^2$). This is in agreement with the Eq. (5).[7]

We conclude that even in the absence of any physical interaction, such as drop breakup, correlations between droplet counts associated with spatially separated volumes persist. There are at least two independent routes to patchiness from the fluid dynamics point of view: (i) partial entrainment of particles in eddies of the random vorticity field and the tendency of turbulence to collect particles along isolated paths even when the initial distribution is purely random (Poissonian); and (ii) general tendency of passive scalar pdfs produced by turbulent mixing to develop exponential tails. Thus, in our opinion, turbulent mixing renders non-Poissonian statistics ubiquitous because patchiness by its very definition implies spatial correlation of particle counts which, in turn, implies super-Poissonian variance via the relation (5) as depicted in Fig. 1. Furthermore, these arguments are rather general and do not depend on the nature of particles (e.g., can be applied to aerosols) and on the detailed characterization of turbulence such as the exact value of the Reynolds number or degree of homogeneity; for example, they can be applied to convective cores of clouds as well as to the entrainment regions. The form of the pair correlation function η does, of course, depend on such details.

7. Summary and concluding remarks

In this work we attempted to probe deeper into stochastic nature of clouds. Novel counting diagrams, particularly applicable to low count rates, were used to demonstrate droplet clustering on finescale. The results show striking pattern of deviations from the Poissonian statistics.

[7] Note that the result $\sigma_n^2 \propto \mu^2$ applies more generally than would seem from this derivation. Indeed, Eq. (6) is completely model independent. On the other hand, the entire family of two-parameter Γ distributions for $\overline{n}$ results in negative-binomial distributions of counts (see Kostinski and Jameson 1997 for details), which, in turn, yield $\sigma_n^2 = \mu + C\mu^2$, similar to (6). Hence, such models can be used with little loss of generality. Note that Eq. (10) in Kostinski and Jameson (1997, p. 2178) contains a misprint: η should be replaced by μ in one of the sets of parentheses.

We have also introduced a new connection between super-Poissonian statistics and the pair correlation function $\eta(l)$ via the relations (5) and (6). This relation allows a direct measurement of the cloud texture (represented by the pair correlation function) from the data obtained by the airborne cloud probes. This approach is based on rather general methods of statistical physics. For example, the correlation-fluctuation relations (5) and (6) are model independent in the sense that the super-Poissonian variance results from a mere presence of correlations on some scale but is independent of the functional form of $\eta(l)$. The shape of $\eta(l)$ is likely to change from one cloud type to another, depend on particle size, etc., yet the result [Eq. (6)] still holds and only the coefficient in the quadratic term varies. Furthermore, the dependence of the excess variance on the sampling volume is made explicit so that the effects of changes in resolution can now be understood.

We have also suggested a remote sensing approach to study cloud texture by resurrecting the idea of coherent scatter by droplets themselves and offered a more complete interpretation of Bragg scatter that can explain some recent observations in clouds. The basic idea is that the random but nevertheless spatially correlated collection of scatterers can be viewed as a superposition of diffraction gratings, one of which is in resonance with the incident radiation. Hence, the radar ends up measuring the component of the Fourier transform of the spatial correlation function of the cloud droplets at its own wavelength. Measurements at many wavelengths from millimeters to meters and many angles (bistatic radar) can then provide the needed information about the texture remotely. An essential ingredient of this idea is the relation (6), stating that particle concentration variance scales quadratically rather than linearly with the mean concentration.

The quadratic (rather than linear) dependence of the number density variance on the mean density is likely to have other important applications. For example, Poissonian clouds yield Poissonian distribution of the number of collisions experienced by a droplet per unit time (just like in the ideal gas of molecules). Super-Poissonian variance, on the other hand, broadens the collision distribution. This puts ''lucky'' drops moving through patches of locally high droplet concentration and experiencing more than the average number of collisions early in their development, in a position to grow relatively rapidly. The super-Poissonian fluctuations in the number of collisions can greatly accelerate growth of such lucky droplets and help explain the vexing problem of rapid broadening of drop size spectra and appearance of the first raindrops. Similarly, distribution in the number of collisions experienced by a photon going through a non-Poissonian cloud also undergoes the super-Poissonian broadening. This may yield greater transmission as the ''unlucky'' photons undergoing no collisions are more likely to do so in correlated than in Poissonian clouds.

In addition, in this paper we explain previous observations of the apparent applicability of the exponential clustering to droplet statistics using kinematic considerations of turbulence. We have also argued the ubiquity of the super-Poissonian statistics by the following reasoning (depicted in Fig. 1): patchiness (texture) by its very definition implies spatial correlation of particle counts in spatially separated volumes; this, in turn, implies super-Poissonian variance via the relation (6).

But cloud patchiness is likely to be ubiquitous because at least two entirely different mechanisms can lead to regions of preferential concentration in clouds. One applies to particles with some inertia that fall through random vorticity fields and tend to collect along eddy boundaries and in regions of low vorticity and high strain. The other route applies to passive tracers as well, where turbulence begins to mix an initial blob of cloud water and later produces exponential tails of local concentration probability density and yields Eqs. (15) and (18).

In passing, we note that since exponential-type clustering yields variance of concentration which is quadratic in mean liquid water content (LWC), one would expect standard deviation of LWC to correlate strongly with the mean LWC, similar to the distribution of rain rates (Jameson and Kostinski 1999). Finally, in the two appendices we discuss stationarity of data and the effects of clustering on sampling strategies. In particular, it is shown in appendix B that even such a harmless procedure as division of raw particle counts by a number of liters in a sampling volume (to infer concentration) may be misleading with respect to testing the Poissonian character of fluctuations.

Acknowledgments. This work was supported by NSF Grants ATM95-12685, ATM-9419523, and ATM-9708657. ABK would also like to acknowledge sabbatical leave support through the USRA–NASA/Goddard Space Flight Center Fellowship of the Laboratory for Atmospheres.

APPENDIX A

An Example of Statistically Homogeneous Data with Super-Poissonian Variance

It is often difficult to differentiate between data associated with a correlated stationary process from that of an inhomogeneous random process when record duration is limited. For example, a symptom of this ambiguity is the characteristic correlation distance being comparable to the length of the entire data record length. Wunsch (1999, p. 245) recently stated the following.

> But purely random processes, particularly those that have even mildly 'red' spectra, have a behavior that comes as a surprise to many, and there is a great risk of misinterpretation. That is, the purely random behavior of a rigorously stationary process often appears visually inter-

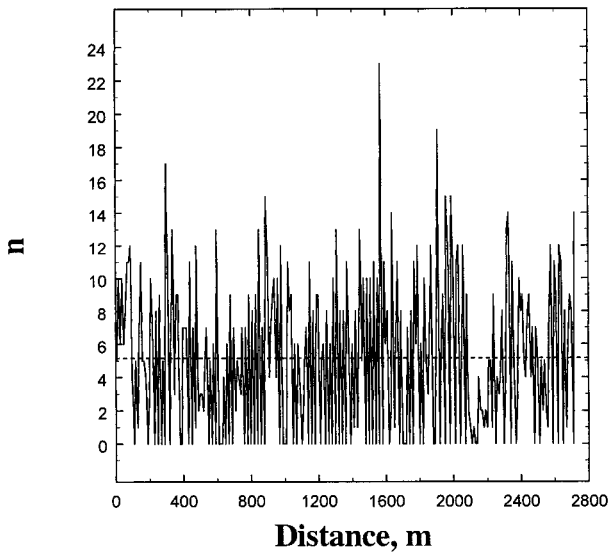

FIG. A1. The number of 50-μm diameter droplets per 10-m path increment (corresponding to the sampling volume of about 10 mm^2 $\times$ 10 m = 100 cm^3) observed along a 2.75-km path in a cumulus cloud. The dashed line denotes the mean value over the entire interval.

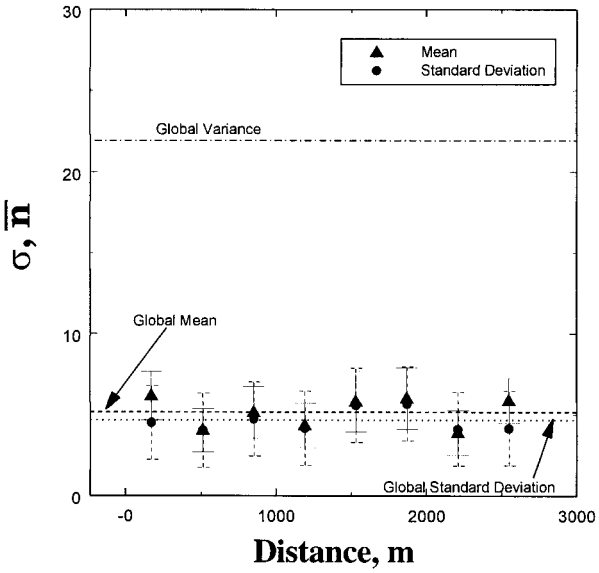

FIG. A2. The ensemble average (over eight groups) of the mean number (computed for each group of 34 10-m path increments) as well as the ensemble standard deviation (σ) along the entire interval in Fig. A1. The constancy of σ and $\overline{n}$ indicates wide-sense homogeneity over the entire interval. Also note that the observed average variance is over four times the mean, i.e., over four times that expected for a Poisson distribution.

esting, particularly over brief time intervals, and creates the temptation to interpret it as arising from specific and exciting deterministic causes.

Fortunately, in the case of the data examined in this paper such differentiation is possible because the correlation length is much smaller than the duration of the entire data record. In fact, the purpose of this appendix is to present examples of statistically stationary data segments that are nevertheless super-Poissonian.

For example, Fig. A1 shows a 2.7-km segment of the 50-μm particle count series corresponding to 10-m intervals. We then divide the segment into eight statistically independent groups, each containing 34 data points. The mean and standard deviations are plotted for each subinterval. Since the observed coherence length is less than 10 m, all 34 samples in each group are independent. The error bars for the mean correspond to one standard deviation computed using the variance measured for each separate group, while the error bars for the standard deviation are calculated using the variance of the variance of the eight sets of 34 points. In Fig. A2, the solid line error bars correspond to the average while the dashed lines correspond to the standard deviation.

Figure A2 shows the results and compares it with the mean and variance of the entire set (''global'' mean and variance.) The series is well approximated as a wide-sense stationary random process because the differences among the means and the variances are within the error bars. Hence, no statistically meaningful trends in either the variance or the mean appear in these data.

Furthermore, Fig. A3 shows that there is little if any correlation among particle counts for volumes separated by 10 m or more. We conclude this because of the similarity of the actual pair correlation function with the one of the Poisson (uncorrelated) simulation with the same parameters as those of the data. Yet, the variance of the data is 4.25 times the equivalent Poisson variance.

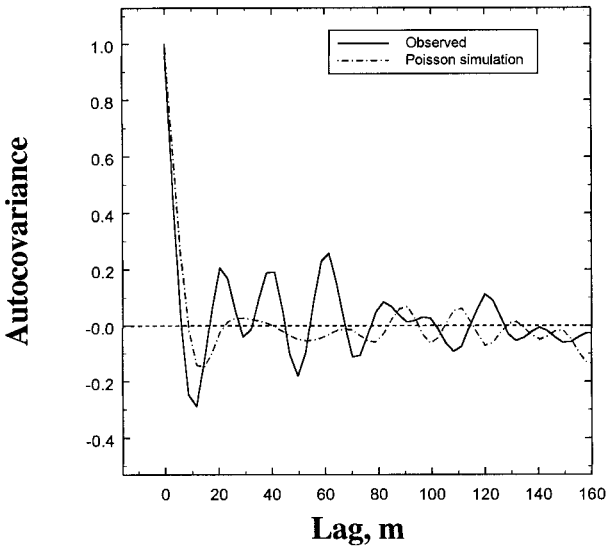

FIG. A3. The autocovariance function for the 10-m counts of 50-μm diameter droplets along the path in Fig. A1. Also shown is the autocovariance function for the ''equivalent'' (same mean and number of points) Poisson sequence of uncorrelated counts. Note the similarity to the Poisson sequence and, hence, the lack of correlation in the data on scales of 10 m and longer.

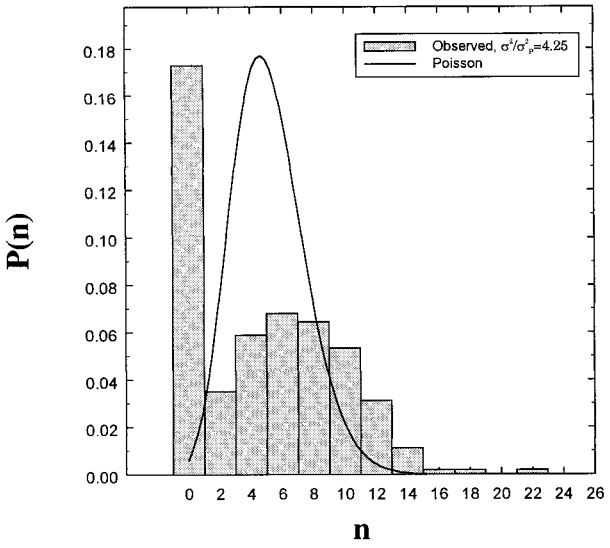

FIG. A4. The histogram of 10-m counts of 50-μm-diameter drops is compared to the equivalent Poisson distribution having the same mean as observed. Note the increased frequencies at both lower and higher counts compared to a Poisson distribution, indicating droplet clustering and significant deviations from Poisson statistics. Also note that the observed variance σ^2 is over four times larger than the Poisson variance (σ_P^2).

This is consistent with the histograms presented in Fig. A4, which again show larger spread of the data as compared with the equivalent Poisson simulation. All of this is in agreement with the physics of the fluctuation-correlation theorem (section 3). Indeed, the 10-m sampling interval contains strong correlations occurring on, say, 10-cm scale. Hence, these contribute to the super-Poissonian variance. Yet by the time we get to 10 m or longer separations, the correlations decay to insignificant levels.

APPENDIX B

On the Importance of Fixed Sampling Volume in Detecting Super-Poissonian Fluctuations with Cloud Probes

Cloud probes often record droplet arrival times (e.g., Baker 1992). The distribution of the number of droplets per unit volume, $p(n)$, can be obtained from such data but the sampling volume varies unless care is taken to divide data into equal intervals. The concentration, for example, number of particles per liter is often obtained by adding up all the particles in a larger volume and dividing the total by the number of liters (denoted as α) in a variable sampling volume V. While the variability of sampling volumes has been discussed by many authors, the specific context of super-Poissonian fluctuations introduces additional concerns. The purpose of this appendix is simply to alert the reader that even the seemingly harmless operation of dividing the raw particle counts by α to infer the concentration is misleading when the sampling volume varies.

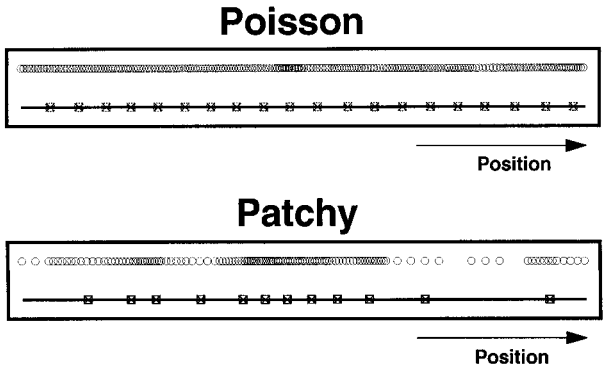

FIG. B1. Illustration of a simple Monte Carlo simulation of two clouds: a homogeneous Poissonian cloud, and a patchy one (consisting of 600 patches whose concentrations are exponentially distributed as described by the geometric model of section 6). Each circle in this figure represents 10 drops while each line with the boxes represents the distance traversed for each sampling volume with each box corresponding to a count reaching 100 droplets.

Let the true distribution be $p(n)$. Multiplying each count by α then causes the mean of the new distribution to increase by α but its variance to increase by α^2 (simply by the definition of variance). This means that the variance-to-mean ratio increases by a factor of α. This ratio is unity for the Poisson process and the relation $\sigma_n^2 = \overline{n}$ is not invariant under the operation of multiplication or division. Hence, this can lead to spurious non-Poissonian statistics and vice versa, underestimate deviations from Poissonian statistics when dividing raw counts by $\alpha > 1$. Furthermore, because of the relation (5), the division by α^2 causes an erroneous appearance of concentration correlations.

To illustrate a possible bias, consider a hypothetical case of the sampling volume being governed by the buffer dumps driven in large part by when the images ''fill up'' the buffer. The data collection then ceases, the buffer is dumped, and the process repeats itself. For example, sampling volumes of actual data reported here varied between 0.5 and 3 l but concentration was recorded as number of droplets per liter.

In order to gain basic understanding of the effects of such operation on the distribution of sampling volumes in patchy clouds, we used the simplest possible Monte Carlo simulation which captures the ''buffer'' problem. We confined ourselves to a one-dimensional case and prepared two clouds: a homogeneous (Poissonian) cloud, and a patchy one (consisting of 600 patches whose concentrations are exponentially distributed as described in section 6). The clouds are illustrated in Fig. B1. We then imposed a 100 droplet ''buffer constraint'' on our probe and let it traverse both clouds. Each circle in the figure represents 10 drops while each line with the boxes represents the distance traversed for each sampling volume, with each box corresponding to a count reaching 100 droplets.

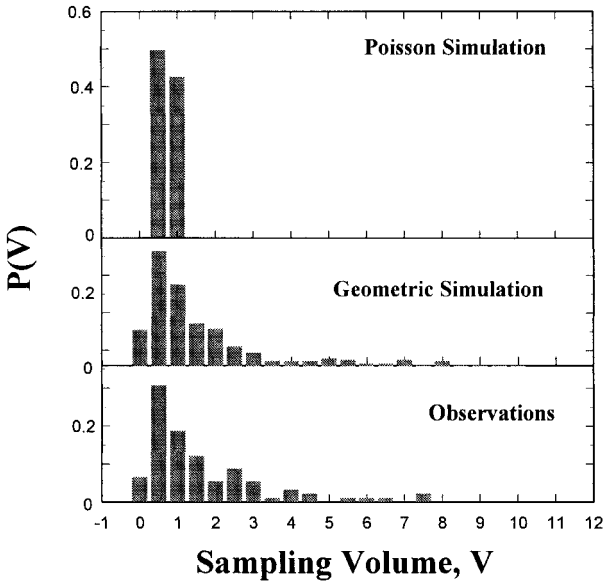

FIG. B2. Sampling volume distribution results of the Monte Carlo simulation of a ''buffer-limited'' cloud probe pass through clouds of Fig. B1. A 100-droplet buffer constraint is imposed on the buffer. The agreement between the patchy (geometric) case and the data is rather striking. Broad distributions of sampling volumes are indicative of small-scale variability (patchiness) of clouds.

The results are illustrated and compared to actual data in Fig. B2. The agreement between the patchy (geometric) case and the data is striking (especially given the simplicity of the simulation). Note that the very fact that sampling volumes have a broad distribution is indicative of small-scale variability (patchiness) of clouds. Also note that high-concentration patches are sampled more often than the relatively sparse regions or voids. Therefore, the usual arithmetic (sample) mean estimator of concentration acquires a bias that approaches a factor of 2 in the ''binary'' case of alternating patches and empty space.

Finally, one should also reconsider the commonly used $1/\sqrt{K}$ convergence rule, where K is the number of realizations or the number of individual experiments. Indeed, consider convergence rates toward the average concentration field for the cases of Poissonian and patchy (geometric) distributions. Let μ represent the global mean number of droplets and let k_i be an ith realization of a randomly varying droplet count. Then, once K small volume measurements have been accumulated, we form the usual arithmetic mean estimate of μ as

$$\hat{\mu} = \frac{1}{K}\sum_{i=1}^{K} k_i. \tag{B1}$$

By the central limit theorem, provided our measurements are statistically independent, the estimate (19) of $\mu_i(D_i)$ converges to the true value as

$$\sigma^2_{\text{est}} \propto \frac{\sigma^2_{\text{pdf}}}{K}, \tag{B2}$$

where σ^2_{est} and σ^2_{pdf} are the estimator and the pdf variances, respectively.

The above formula yields two rather different results for the cases of Poissonian and patchy clouds. In the former case, the $\sigma^2_{\text{pdf}} \propto \mu$, while in the latter case, the pdf is approximately geometric and $\sigma^2_{\text{pdf}} \propto \mu^2$. Therefore, the $1/\sqrt{\mu}$ rule for $\sigma^2_{\text{pdf}}/\mu$ should be viewed with caution in the patchy case; that is, increasing the mean number of droplets per sampling volume is not particularly helpful. More importantly, the patchy case imposes a more stringent constraint on the collection of statistically independent samples because one has to traverse large distances to ensure that many patches have been sampled. Hence, the $1/\sqrt{K}$ convergence of the estimator (19) is to be viewed with caution because K must be interpreted as an ''effective'' number of independent samples (spatially decorrelated patches).

REFERENCES

Baker, B., 1992: Turbulent entrainment and mixing in clouds: A new observational approach. *J. Atmos. Sci.,* **49,** 387–404.

Brodkey, R. S., 1995: *The Phenomena of Fluid Motions.* Dover, 737 pp.

Davis, A. B., A. Marshak, H. Gerber, and W. J. Wiscombe, 1999: Horizontal structure of marine boundary-layer clouds from cm to km scales, *J. Geophys. Res.,* **104** (D6), 6123–6144.

Gossard, E. E., and R. G. Strauch, 1983: *Radar Observations of Clear Air and Clouds.* Elsevier, 280 pp.

Green, H. S., 1969: *The Molecular Theory of Fluids.* Dover, 264 pp.

Hentschel, H. G. E., and I. Procaccia, 1984: Relative diffusion in turbulent media: The fractal dimension of clouds. *Phys. Rev.,* **29A,** 1461–1470.

Hozler, M., and E. D. Siggia, 1994: Turbulent mixing of a passive scalar. *Phys. Fluids,* **6A,** 1820–1828.

Jameson, A. R., and A. B. Kostinski, 1999: Fluctuation properties of precipitation. Part V: Distribution of rain rates—Theory, and observations in clustered rain. *J. Atmos. Sci.,* **56,** 3920–3932.

——, and ——, 2000: Fluctuation properties of precipitation. Part VI: Observations of hyperfine clustering and drop size distribution structures in three-dimensional rain. *J. Atmos. Sci.,* **57,** 373–388.

——, ——, and R. A. Black, 1998: The texture of clouds. *J. Geophys. Res.,* **103,** 6211–6219.

Knight, C. A., and L. J. Miller, 1993: First radar echoes from cumulus clouds. *Bull. Amer. Meteor. Soc.,* **74,** 179–188.

——, and ——, 1998: Early radar echoes from small, warm cumulus: Bragg and hydrometeor scattering. *J. Atmos. Sci.,* **55,** 2974–2992.

Kostinski, A. B., and A. R. Jameson, 1997: Fluctuation properties of precipitation. Part I: Deviations of single size drop counts from the Poisson distribution. *J. Atmos. Sci.,* **54,** 2174–2186.

——, and ——, 1999: Fluctuation properties of precipitation. Part III: On the ubiquity and emergence of the exponential size distribution. *J. Atmos. Sci.,* **56,** 101–111.

Landau, L. D., and E. M. Lifshitz, 1980: *Statistical Physics.* Pergamon Press, 687 pp.

Loudon, R., 1983: *The Quantum Theory of Light.* Clarendon, 393 pp.

Maxey, M. R., and S. Corrsin, 1986: Gravitational settling of aerosol particles in randomly oriented cellular flow fields. *J. Atmos. Sci.,* **43,** 1112–1119.

Naito, K., and D. Atlas, 1966: On microwave scattering by partially

coherent clouds. Preprints, *12th Conf. on Radar Meteorology*, Norman, OK, Amer. Meteor. Soc., 7–12.
Ochi, M., 1990: *Applied Probability and Stochastic Processes.* John Wiley and Sons, 499 pp.
Paluch, I. R., and D. G. Baumgardner, 1989: Entrainment and fine-scale mixing in continental convective cloud. *J. Atmos. Sci.,* **46,** 261–278.
Pawlowska, H., J. L. Brenguier, and G. Salut, 1997: Optimal nonlinear estimation for cloud particle measurements. *J. Atmos. Oceanic Technol.,* **14,** 88–104.
Peebles, P. J. E., 1980: *The Large-Scale Structure of the Universe.* Princeton University Press, 422 pp.
Pinsky, M., and A. Khain, 1997: Formation of inhomogeneity in drop concentration induced by the inertia of drops falling in a turbulent flow, and the influence of the inhomogeneity on the drop spectrum broadening. *Quart. J. Roy. Meteor. Soc.,* **123,** 165–186.
Pruppacher, H. R., and J. D. Klett, 1997: *Microphysics of Clouds and Precipitation.* 2d ed. D. Reidel, 714 pp.
Pumir, A., B. Shraiman, and E. D. Siggia, 1991: Exponential tails and random advection. *Phys. Rev. Lett.,* **66,** 2984–2987.
Rogers, R. R., and M. K. Yau, 1989: *A Short Course in Cloud Physics.* Pergamon Press, 229 pp.
Shaw, R. A., W. C. Reade, L. R. Collins, and J. Verlinde, 1998: Preferential concentration of cloud droplets by turbulence: Effects on the early evolution of cumulus cloud droplet spectra. *J. Atmos. Sci.,* **55,** 1965–1976.
Sinai, Ya. G., and V. Yakhot, 1989: Limiting probability distributions of a passive scalar in a random velocity field. *Phys. Rev. Lett.,* **63,** 1962–1966.
Smith, P. L., 1964: Scattering of microwaves by cloud droplets. Preprints, *11th Conf. on Weather Radar Meteorology,* Boulder, CO, Amer. Meteor. Soc., 202–207.
Squires, K. D., and J. K. Eaton, 1991: Preferential concentration of particles by turbulence. *Phys. Fluids,* **3A,** 1169–1178.
Wang, L.-P., and M. R. Maxey, 1993: Settling velocity and concentration distribution of heavy particles in homogeneous isotropic turbulence. *J. Fluid Mech.,* **256,** 27–68.
Wunsch, C., 1999: The interpretation of short climate records, with comments on the North Atlantic and Southern Oscillations. *Bull. Amer. Meteor. Soc.,* **80,** 245–255.