

Cellular Statistical Models of Broken Cloud Fields. Part II: Comparison with a Dynamical Model and Statistics of Diverse Ensembles

MIKHAIL D. ALEXANDROV

Department of Applied Physics and Applied Mathematics, Columbia University, and NASA Goddard Institute for Space Studies, New York, New York

ANDREW S. ACKERMAN

NASA Goddard Institute for Space Studies, New York, New York

ALEXANDER MARSHAK

NASA Goddard Space Flight Center, Greenbelt, Maryland

(Manuscript received 29 October 2009, in final form 17 February 2010)

ABSTRACT

Cellular statistical models are designed to provide a simple two-parameter characterization of the structure of broken cloud fields described through distributions of cloud fraction and of chord lengths for clouds and clear gaps. In these analytical models cloud fields are assumed to occur on a semiregular grid of cells (which can be vaguely interpreted as atmospheric convective cells). In a simple, discrete cell model, cell size is fixed and each cell can either be completely filled with cloud with some probability or remain empty. Extending the discrete model to a continuous case provides more realism by allowing arbitrary cloud and gap sizes. Here the continuous cellular model is tested by comparing its statistics with those from large-eddy simulations (LES) of marine boundary layer clouds based on case studies from three trade-cumulus field projects. The statistics largely agree with some differences in small sizes approaching the LES model grid spacing. Exponential chord-length distributions follow from the assumption that the probability of any cell being cloudy is constant, appropriate for a given meteorological state (narrow sampling). Relaxing that assumption, and instead allowing this probability to have its own distribution, leads to a power-law distribution of chord lengths, appropriate to a broader sample of meteorological states (diverse sampling).

1. Introduction

Well-defined cellular structures are often observed in cumulus and stratocumulus cloud fields. The study by Wielicki and Welch (1986) based on four Landsat satellite scenes demonstrated that small (1 km in diameter) fair-weather cumulus clouds can consist of several cells with sizes ranging from approximately 250 m to 1 km. The aircraft-based observations by Nicholls (1989) indicated chord distributions of convective cells in stratocumulus cloud fields consistent with hexagonal cell structures with an average cell diameter around half of the depth of

the planetary boundary layer (i.e., diameters ranging from 200 to 600 m).

In Part I of this paper (Alexandrov et al. 2010, hereafter Part I) two cellular statistical models (CSMs) of cloud fields were documented. In the first, “discrete” model, our approach is similar to that in percolation theory (Isichenko 1992), in which we consider a cloud-embedded layer of the atmosphere as a 1D or 2D horizontal lattice with cells (of size l) that are either cloudy with probability p or clear with probability $q = 1 - p$. One may loosely think of these cells as atmospheric convective cells. In this study, however, the cellular structure of cloud fields is used as an abstract modeling concept, whose physical nature we do not pursue. A similar concept in a different form was used in the statistical cloud model of Nagel and Raschke (1992). We have extended this approach in our model by allowing clouds to take arbitrary size and

Corresponding author address: Mikhail Alexandrov, NASA Goddard Institute for Space Studies, 2880 Broadway, New York, NY 10025.
E-mail: malexandrov@giss.nasa.gov

shape (instead of being clusters of several square cells), while retaining the basic assumption of the discrete-cell prototype that the probability of a cell of size l being fully cloudy is p . This extension leads to our second, “continuous” model, which is formulated to better describe statistics of real cloud fields. The statistics of this model, such as distributions of the cloud fraction and chord length distributions of clouds and gaps between them, can be computed for the 1D case. Such 1D models are also applicable to 1D transects (chords) of 2D cloud fields (e.g., Sánchez et al. 1994; Astin and Di Girolamo 1999; Byrne 2005). Part I also presents an iterative procedure that allows for construction of a continuous 2D cloud field from a realization of the discrete model; here we consider only 1D models.

Computationally inexpensive stochastic cloud models based on satellite, aircraft, or ground measurements have been used to generate cloud fields resembling observations (e.g., Evans and Wiscombe 2004; Hogan and Kew 2005; Venema et al. 2006; Prigarin and Marshak 2009). Some stochastic models generate cloud fields with internal cloud structure (e.g., Schertzer and Lovejoy 1987; Cahalan 1994; Marshak et al. 1994; Schmidt et al. 2007), while others treat cloud fields as a binary mixture of cloudy and clear areas (e.g., Su and Pomraning 1994; Zuev and Titov 1995; Prigarin et al. 2002). In contrast to their statistical counterparts, dynamical cloud models use a more first-principles approach and are more physically based. The objective of this study is to demonstrate that our simple stochastic model can generally reproduce the statistics of shallow, broken cloud fields obtained from large-eddy simulation (LES). In the comparisons here we use LES results from case studies of shallow, maritime convection based on observations from three trade-cumulus field projects: the Atlantic Trade Wind Experiment (ATEX; Augstein et al. 1973), the Barbados Oceanographic and Meteorological Experiment (BOMEX; Nitta and Esbensen 1974), and Rain in Cumulus over the Ocean (RICO; Rauber et al. 2007).

Several studies have been published describing cloud chord statistics based on ground-based measurements (e.g., Lane et al. 2002; Berg and Kassianov 2008) as well as aircraft and satellite data (e.g., Plank 1969; Cahalan and Joseph 1989; Joseph and Cahalan 1990; Rodts et al. 2003). Some of these studies—notably for this study, those restricted to fair-weather cumulus fields—report exponential cloud and gap chord length distributions (Astin and Latter 1998; Lane et al. 2002), while others find power-law distributions (Cahalan and Joseph 1989; Koren et al. 2008). Here we argue that the type of the observed cloud chord distribution depends not only on the dominant cloud type during an observation period, but also on the size and diversity of the sample. Cloud properties in our

statistical model are described by two parameters: l and p . In Part I of this paper we demonstrate that the cloud and gap length distributions are exponential if p and l (or equivalently the mean cloud and gap lengths) are uniform for the whole sample. However, as we show below, in the case of a diverse sample, in which p is not fixed but rather has a statistical distribution of its own, the cloud and gap length distributions instead follow a power law. Such power-law distributions are applicable to datasets with large coverage (spatial, temporal, or both) that include a diversity of cloud types and atmospheric conditions.

In Section 2 we present an overview of our cellular cloud models, in section 3 we compare statistics from our continuous cellular model with those from large-eddy simulations, and in section 4 we discuss power-law cloud fields. We conclude with a summary in section 5.

2. Cellular models of cloud fields

In Part I of this paper we describe two types of cellular models of broken cloud fields: 1) a simple discrete lattice model and 2) its continuous generalization. Below we present a brief overview of those models.

a. Discrete model

In the discrete cellular model we consider a 1D lattice, each cell of which can be occupied by cloud with probability p ($0 < p < 1$) or remain clear with probability $q = 1 - p$. We define a cloud of size m as a sequence of m neighboring cloudy cells separated from other clouds by at least one clear cell at each end. If the lattice is infinite, the statistical distribution of clouds by size in this model is geometric:

$$n_c(m) = qp^{m-1}, \tag{1}$$

$m = 1, 2, 3, \dots$, with the mean

$$\bar{m}_c = \frac{1}{q}. \tag{2}$$

We show in Part I that the observed cloud and gap length statistics are influenced by the sample size, especially for relatively large clouds and gaps. For finite samples the model also needs to take into account overcast clouds (i.e., those with length equal to or exceeding the sample size) as well as entirely clear samples, cases that do not exist on an infinite lattice. In the case of a finite sample consisting of N cells, the cloud length distribution takes the form

$$n_c(m) = \left[1 + \frac{1 - qm}{1 + q(N - 1)} \right] qp^{m-1} \tag{3}$$

for $m < N$, while

$$n_c(N) = \frac{p^{N-1}}{1 + q(N-1)} \tag{4}$$

for $m = N$. The mean cloud size for this distribution is

$$\bar{m}_c = \left[q + \frac{p}{N} \right]^{-1}. \tag{5}$$

The statistical distributions of clear intervals (gaps) between clouds can be obtained from the corresponding expressions for clouds the by interchanging p and q .

The cloud cover (fraction) is defined as

$$c = \frac{k}{N}, \tag{6}$$

where k is the number of all ‘‘cloudy’’ cells (not necessarily consecutive) in the sample of the size N . The probability of having k occupied cells in this sample obeys the binomial distribution

$$n_{cf}(k) = \binom{N}{k} p^k q^{N-k}, \quad k = 0, 1, \dots, N, \tag{7}$$

which has the mean

$$\bar{k} = Np, \tag{8}$$

corresponding to $\bar{c} = p$.

b. Continuous model

The continuous model is derived by subdividing each discrete cell into a number (infinite in the limit) of small subcells and defining their occupation probability such that the average occupation probability of each greater discrete cell remains p . In this model, cloud and gap length distributions for an infinite 1D space are exponential,

$$f_{c,g}(x) = \frac{1}{L_{c,g}} e^{-x/L_{c,g}} \tag{9}$$

with the mean cloud and gap lengths being respectively

$$L_c = -\frac{l}{\ln p} \quad \text{and} \quad L_g = -\frac{l}{\ln q}. \tag{10}$$

In the case of a 1D sample of finite length $L = Nl$ (in which N is not necessarily an integer), the length distributions become singular at $x = L$:

$$F_{c,g}(x) = f_{c,g}(x) + f_{c,g}^o \delta(x - L), \quad x \in [0, L]. \tag{11}$$

Here

$$f_{c,g}(x) = \frac{1}{L_{c,g}} \left[1 + \frac{1 - (x/L_{c,g})}{1 + a_{c,g}} \right] e^{-x/L_{c,g}} \tag{12}$$

is the probability density for $x < L$, and

$$f_{c,g}^o = \frac{e^{-a_{c,g}}}{1 + a_{c,g}} \tag{13}$$

gives the entirely cloudy (overcast) and clear sky contributions at $x = L$. Here we have used the following notation:

$$a_c = \frac{L}{L_c} = -\ln(p^N) \quad \text{and} \quad a_g = \frac{L}{L_g} = -\ln(q^N). \tag{14}$$

It is easily confirmed that $F_{c,g}$, rather than $f_{c,g}$, is normalized

$$\int_0^L F_{c,g}(x) dx = 1 \tag{15}$$

and that the mean \bar{x} obey the relations

$$\frac{1}{\bar{x}_{c,g}} = \frac{1}{L_{c,g}} + \frac{1}{L}. \tag{16}$$

Note that in the infinite sample case discussed above, $F_{c,g}(x) \equiv f_{c,g}(x)$.

A continuous analog of the binomial cloud cover distribution (7) is also a singular density:

$$F_{cf}(c) = f_{cf}(c) + f_{cf}^{(0)} \delta(c) + f_{cf}^{(1)} \delta(1 - c), \tag{17}$$

obeying

$$\int_0^1 F_{cf}(c) dc = 1. \tag{18}$$

It has the following components: first, a continuous density,

$$f_{cf}(c) = \frac{2a_c a_g}{a_c + a_g} e^{-[a_c c + a_g (1-c)]} \times \left\{ I_0(Z) + [a_g c + a_c (1 - c)] \frac{I_1(Z)}{Z} \right\}, \tag{19}$$

where I_0 and I_1 are modified Bessel functions of the argument,

$$Z = 2\sqrt{a_c a_g c(1 - c)}; \tag{20}$$

and second, the entirely clear sky and overcast probabilities:

$$f_{cf}^{(0)} = \frac{a_c}{a_c + a_g} e^{-a_g} \quad \text{and} \quad (21)$$

$$f_{cf}^{(1)} = \frac{a_g}{a_c + a_g} e^{-a_c}. \quad (22)$$

The mean cloud cover for the distribution has the following form:

$$\bar{c} = \frac{a_g}{a_c + a_g} = \frac{L_c}{L_c + L_g}, \quad (23)$$

which is in principle independent of L . For an infinite sample this formula is somewhat intuitive—the mean cloud cover is the mean cloud length divided by the sum of the mean cloud and gap lengths—and has been published elsewhere (Sánchez et al. 1994; Astin and Di Girolamo 1999; Byrne 2005).

c. Derivation of model parameters from observations

If both L_c and L_g are obtained from observed cloud and gap length histograms (or as simulated by a dynamical model), they uniquely determine model parameters l and p . We find p as a (numerical) solution of

$$\frac{\ln p}{\ln(1-p)} = \frac{L_g}{L_c}, \quad (24)$$

and then l can be determined as

$$l = -L_c \ln p. \quad (25)$$

Alternatively, Eq. (23) can be written in the form

$$\bar{c} = \frac{\ln(1-p)}{\ln p + \ln(1-p)}, \quad (26)$$

which can be used to derive p from the “observed” cloud fraction.

The simplest method to derive L_c and L_g even from noisy data is to use the mean cloud and gap lengths \bar{x}_c and \bar{x}_g computed from the sample and Eq. (16):

$$\frac{1}{L_{c,g}} = \frac{1}{\bar{x}_{c,g}} - \frac{1}{L}. \quad (27)$$

In the case of infinite sample size, determination of the cloud and gap effective lengths also can be made by regressing the logarithms of the corresponding histograms against the cloud and gap lengths. However, this procedure produces systematic errors when the sample is finite. To show this, we write Eq. (12) in the form of

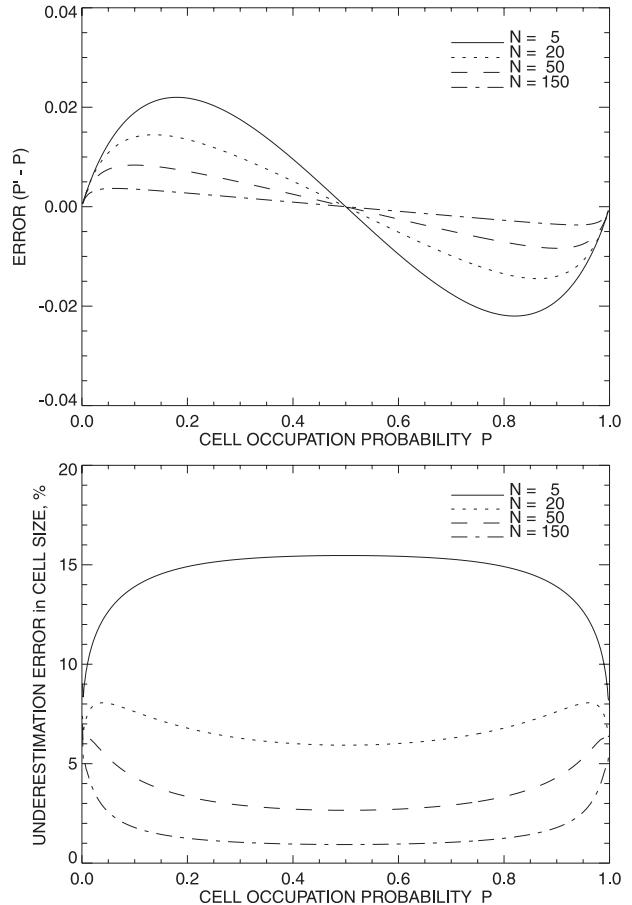


FIG. 1. Errors in (top) derived occupation p and (bottom) l that result from neglecting the influence of finite N .

$$f_{c,g}(x) = \frac{2L_{c,g} + L}{L_{c,g}(L_{c,g} + L)} \left(1 - \frac{x}{2L_{c,g} + L}\right) e^{-x/L_{c,g}} \quad (28)$$

and note that for sufficiently small cloud sizes [when we can use the fact that $1 + y \approx e^y$, that is, for $x \ll 2L_{c,g} + L$] this distribution density is close to exponential, but with effective lengths $L'_{c,g} \neq L_{c,g}$:

$$\frac{1}{L'_{c,g}} = \frac{1}{L_{c,g}} + \frac{1}{2L_{c,g} + L}. \quad (29)$$

The true values of $L_{c,g}$ can be derived from Eq. (29) because L is assumed to be known:

$$L_{c,g} = \frac{1}{4} \left(3L'_{c,g} - L + \sqrt{9L_{c,g}'^2 + 2LL'_{c,g} + L^2} \right). \quad (30)$$

Estimates of errors in p and l that result from neglecting the influence of finite sample size [i.e., by using the $L'_{c,g}$ values in Eq. (25) instead of the corrected $L_{c,g}$ from Eq. (30)] are presented in Fig. 1. It is seen that the errors

are rather small. We find that p is overestimated when it is less than 0.5 (and underestimated otherwise) by less than 0.02 in value for reasonable samples with more than five cells. The cell size is always underestimated: by up to 15% for $N = 5$ and by 7%–8% for $N = 20$. When the sample consists of a large number of cells and $L_{c,g} \ll L/2$, the relation (29) can be simplified

$$\frac{1}{L'_{c,g}} \approx \frac{1}{L_{c,g}} + \frac{1}{L} = \frac{1}{\bar{x}_{c,g}} \quad (31)$$

so that $L'_{c,g} \approx \bar{x}_{c,g}$, and the analysis is similar to Eq. (27) can be applied.

The overcast contribution (13) also can be used for estimation of L_c when $L_c \gg L$ (large p or small N or both), while the corresponding singular part of the gap length distribution (contribution from entirely clear sky) can be used for determination of L_g in the case when $L_g \gg L$ (small p or small N or both).

3. Comparison with a dynamical model

To test our continuous cellular model, we next evaluate its ability to describe the statistics from large-eddy simulations of shallow, maritime convection. The LES model (Ackerman et al. 2004) treats three-dimensional fluid dynamics of the atmosphere and incorporates a bin microphysics model that resolves the size distributions of aerosol and cloud droplets in each grid cell. It also includes a two-stream radiative transfer model that treats the vertical transport of radiation in each model column. Three LES datasets are used, from simulations of idealized environmental conditions based on case studies from trade-cumulus field projects. The ATEX simulations, which are characterized as cumulus rising into a thin stratocumulus layer, are based on an idealization of measurements obtained during the Atlantic Trade Wind Experiment (Stevens et al. 2001). The other two cases are more characteristic of pure trade cumulus, with lesser fractional cloud coverage. The BOMEX and RICO simulations are respectively based on idealizations of measurements obtained during the Barbados Oceanographic and Meteorological Experiment (Siebesma et al. 2003) and the Rain in Cumulus over the Ocean project (M. van Zanten 2008, unpublished manuscript). The ATEX and BOMEX datasets consist of 33 scenes each representing evolution of a cloud field with 15-min temporal resolution on a 96×96 grid, and the RICO dataset consists of 19 scenes with 1-h temporal resolution and a 128×128 grid. Horizontal grid spacing is 100 m in all cases. Here we simply define a column as cloudy if the liquid water path (LWP) exceeds 10 g m^{-2} . (We plan to investigate the dependence of these statistics on the

definition of cloud in future work.) The cloud masks for the first 12 scenes of each dataset are shown in Figs. 2, 3, and 4 and the last 12 scenes of the RICO dataset are shown in Fig. 5.

To compute chord statistics we form an ensemble of all 1D sections of each scene's cloud mask (taken along both north–south and east–west directions). Thus, if we have an $n \times n$ grid and k scenes, our ensemble consists of $2 \times n \times k$ samples of length n . It is essential for agreement with our statistical model that the cloud conditions (and the model parameters p and l) do not vary significantly within a comparison dataset, which is not the case for that ATEX and RICO simulations, during which the cloud fields substantially evolve. Thus we limited the ATEX dataset to the first 12 scenes and split the RICO dataset into two overlapping subsets (Figs. 4 and 5). The resulting cloud/gap length and cloud fraction distribution histograms are presented in Fig. 6 for the ATEX and BOMEX datasets and in Fig. 7 for the two subsets of the RICO dataset. The bin sizes for the cloud and gap length histograms are the same as the spacing of the LES grid mesh (100 m). For the cloud cover histogram the number of bins covering the interval between 0 and 1 is equal to the number of LES grid cells in a 1D transect sample: 96 for ATEX and BOMEX and 128 for RICO; thus, the respective bin sizes are $1/96$ and $1/128$ (both close to 0.01). Such narrow bins result in noisy histograms (especially for cloud cover); however, adopting the discretization of the LES grid mesh reveals narrow features, such as the maxima in some cloud length histograms, which would not be seen if, for example, logarithmically sized bins were adopted. The theoretical distributions shown in the figures are based on the parameters L_c and L_g derived from the mean cloud and gap lengths in the sample datasets according to Eq. (27). Note that the CSM l is denoted in Figs. 6 and 7 by L_{cell} .

To evaluate the level of agreement between statistics from the LES results and from our cellular model, we tried Pearson's χ^2 and Kolmogorov's goodness-of-fit tests (cf. Conover 1980). The χ^2 test is too weak to show differences between the distributions in our case: it indicates no disagreement between cloud cover distributions from the LES and our cellular model at all significance levels, even if our cellular model distribution is replaced with uniform density. Kolmogorov's test is useful, however. The maximum significance level for which this test shows no disagreement between the data and the model is provided in Figs. 6 and 7 as α_k . This value gives the probability that it is wrong to reject the hypothesis that the simulated data statistics agree with our model. Thus, $\alpha_k = 1$ indicates perfect agreement and smaller values indicate poorer agreement. We also define a simple metric

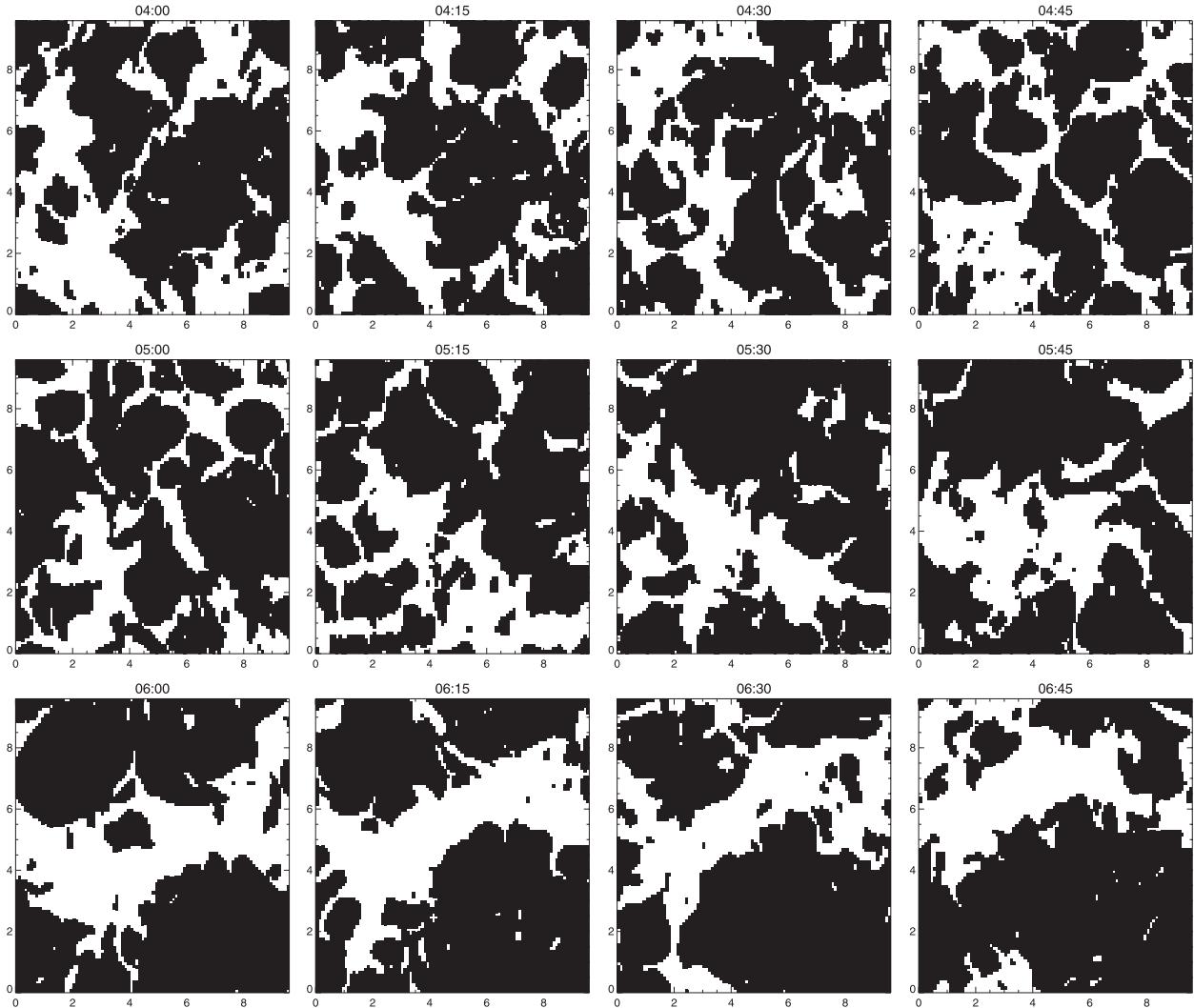


FIG. 2. Cloud mask from LES for the ATEX case of cumulus rising into stratocumulus (simulated times 0400–0645 h). Cloudy columns defined as those with liquid water path $>10 \text{ g m}^{-2}$. Distances on x and y axes given in km. LES grid spacing is 100 m on a 96×96 mesh.

$$\Delta = \frac{1}{2} \int |F_s(x) - F_m(x)| dx, \quad (32)$$

where F_s and F_m are, respectively, the simulated (from the LES) and the modeled (cellular) distribution densities normalized to unity (note that for proper normalization the singular parts associated with entirely clear and overcast samples must be included). This metric indicates the fraction of cases in the simulated ensemble that are “misplaced” relative to the cellular model distribution. Here $\Delta = 0$ indicates that $F_s \equiv F_m$, while $\Delta = 1$ indicates that F_s and F_m do not overlap [i.e., $F_s(x) = 0$ when $F_m(x) \neq 0$ and vice versa]. This metric responds to both systematic and random disagreements between two distributions. We see that the statistical noise is quite notable in the cloud cover histograms,

while the cloud and gap lengths histograms are relatively smooth. To reduce the effect of the noise on the comparisons we also compute the smoothed cloud cover densities (depicted by dashed–dotted curves) using a 15-bin moving average procedure (we also could have chosen broader histogram bins for a comparable effect). The difference between the smoothed and model distributions is characterized in the discussion below by Δ_{sm} . To avoid confusion we will also explicitly call the cells of our statistical model “CSM cells” to distinguish them from LES model cells.

For the ATEX case of cumulus rising into stratocumulus, our analysis indicates an average CSM cell size of 854 m, while the probability of a cell being cloudy is 0.6. For the purer cumulus cloud cases of BOMEX and RICO, the average CSM cell sizes are a bit larger (914 m

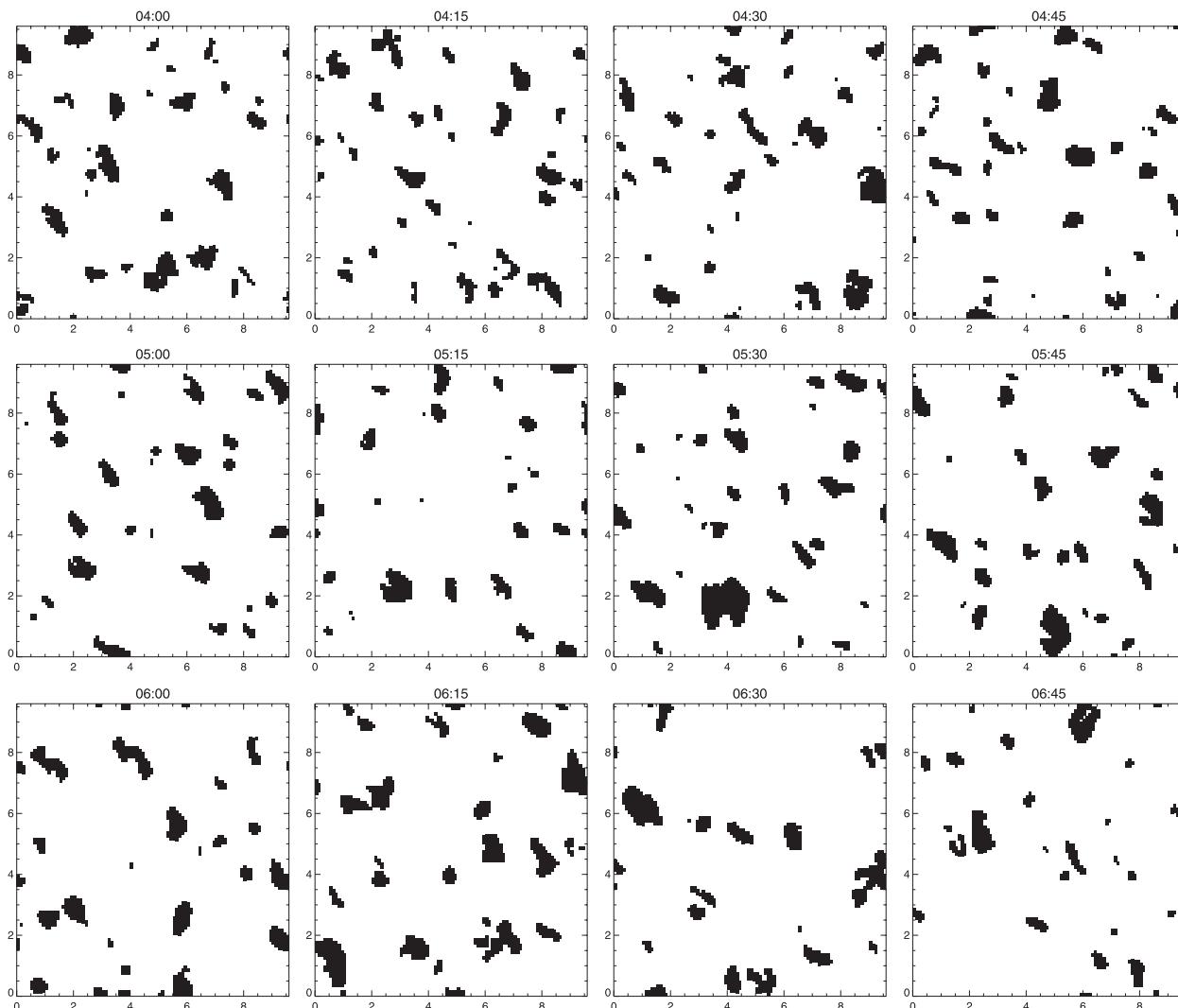


FIG. 3. As in Fig. 2, but for BOMEX trade cumulus at times 0400–1030 h.

for BOMEX and 907–1018 m for RICO), while the cell occupation probability is much smaller than for ATEX, as expected: 0.13 for BOMEX and 0.19–0.24 for RICO. The ATEX data also indicate larger clouds (1443-m mean length; $L_c = 1699$ m) and shorter gaps between them (840-m mean length; $L_g = 920$ m) than for the purer cumulus cases. [Recall that the differences between mean cloud/gap lengths and the corresponding $L_{c,g}$ are due to the finite sample size; see Eq. (16).] For BOMEX the mean cloud length is 435 m ($L_c = 456$ m), and the mean gap length is 3807 m ($L_g = 6310$ m), while for RICO the mean cloud length are 522 and 669 m ($L_c = 544$ and 706 m) and the mean gap lengths are 3234 and 2918 m ($L_g = 4328$ and 3779 m).

Figures 6 and 7 show reasonably good agreement between statistics from the LES dataset and our cellular

model. The mean cloud cover, as well as the entirely clear and overcast fractions, shows agreement with the model within one percentage point or better for all four datasets considered. The cloud cover distributions are closer to the model for ATEX and RICO ($\alpha_k = 1$, $\Delta < 10\%$, $\Delta_{sm} < 5\%$) than for BOMEX ($\alpha_k = 0.77$, $\Delta = 11\%$, $\Delta_{sm} = 9.2\%$). The latter disagreement is likely attributable to the larger deviation of the simulated cloud length distributions from an exponential form in the BOMEX dataset. This deviation, however, is also clearly seen in both parts of the RICO dataset (although much less pronounced in ATEX data): the cloud length histograms for BOMEX and RICO both have a maximum in the third bin (corresponding to 300 m), occupied with about twice the probability of the first bin (100 m). This deviation results in notable disagreements with our model

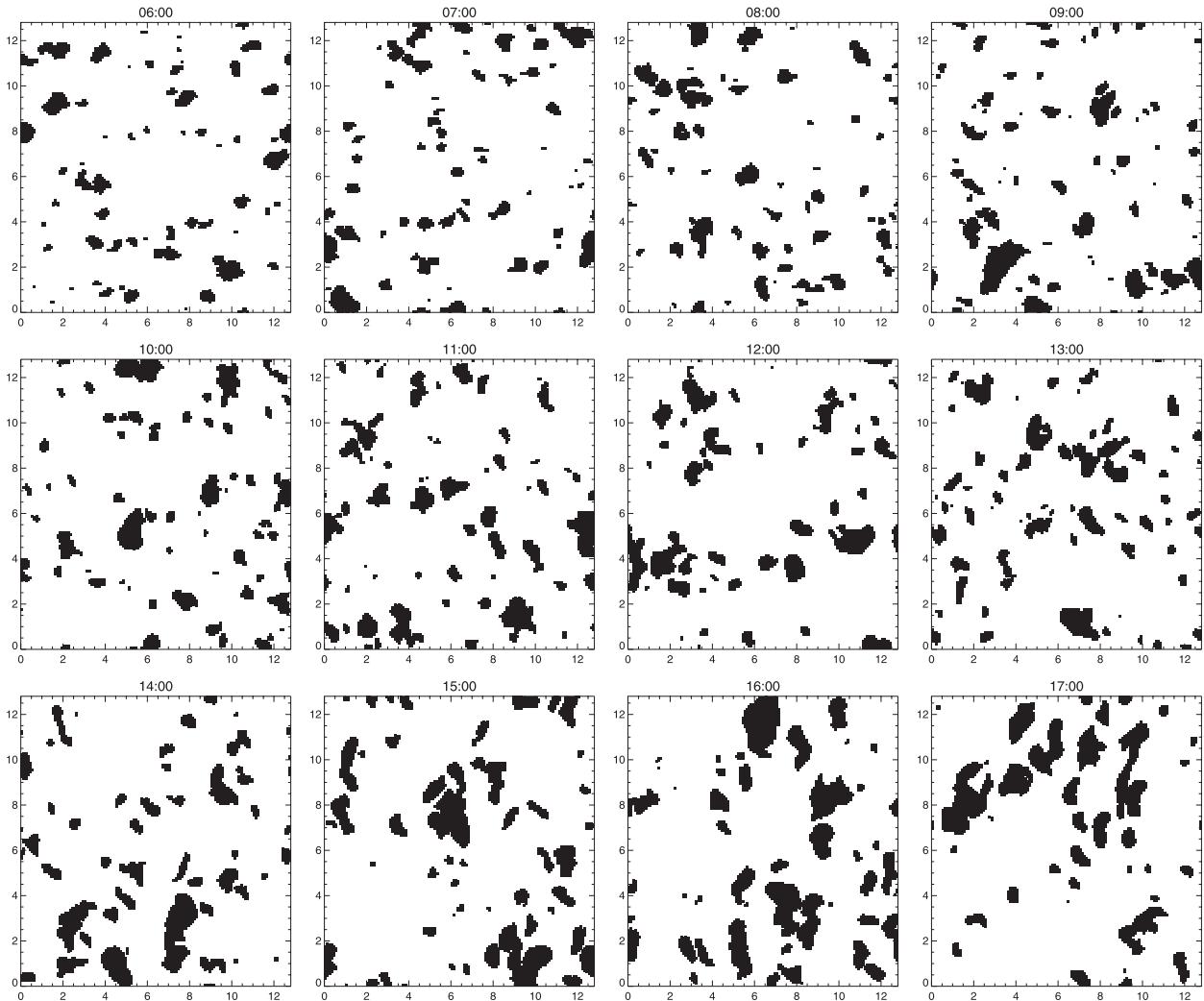


FIG. 4. As in Fig. 2, but for RICO trade cumulus at times 0600–1700 h. Note that the LES grid mesh is larger for RICO, at 128×128 , while the LES model pixel size remains equal to 100 m.

cloud length distributions and also a decrease in the fraction of small cloud cover values (top panels of Figs. 6 and 7). While the mean cloud and gap lengths in our cellular model and the LES dataset are guaranteed to be identical by CSM construction, differences in the cloud length distributions are large: $\Delta = 20\%$, $\alpha_k = 0.01$ for BOMEX, and $\Delta = 16\%$, $\alpha_k = 0.014$ for the first part of the RICO dataset (it is slightly better for the second part: $\Delta = 14\%$, $\alpha_k = 0.11$). At the same time the gap length distributions are in good agreement with the model: $\alpha_k = 1$ for both datasets, while $\Delta = 4.5\%$ for BOMEX and 6% for RICO. We also see a better agreement with LES of the CSM distributions that account for the finite sample length (12), shown with solid lines, compared with their infinite-sample analogs (9), shown with dashed lines.

It is likely that the LES statistics become unreliable for scales approaching the grid spacing size; this behavior can be expected to result from smoothing by high-order advection, numerical diffusion, and explicit subgrid-scale diffusion. Further understanding of the lack of exponentiality in the LES statistics might be gained from comparisons with real observations. The satellite datasets from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) onboard the NASA *Terra* spacecraft (Abrams 2000) may be particularly useful for this purpose because of exceptionally high (15 m) resolution of these measurements. Unfortunately for our purposes, currently published statistical analyses of ASTER data (e.g., Dey et al. 2008; Zhao and Girolamo 2007) deal with cloud area-equivalent diameters rather than with chord statistics, so the issue is presently unresolved.

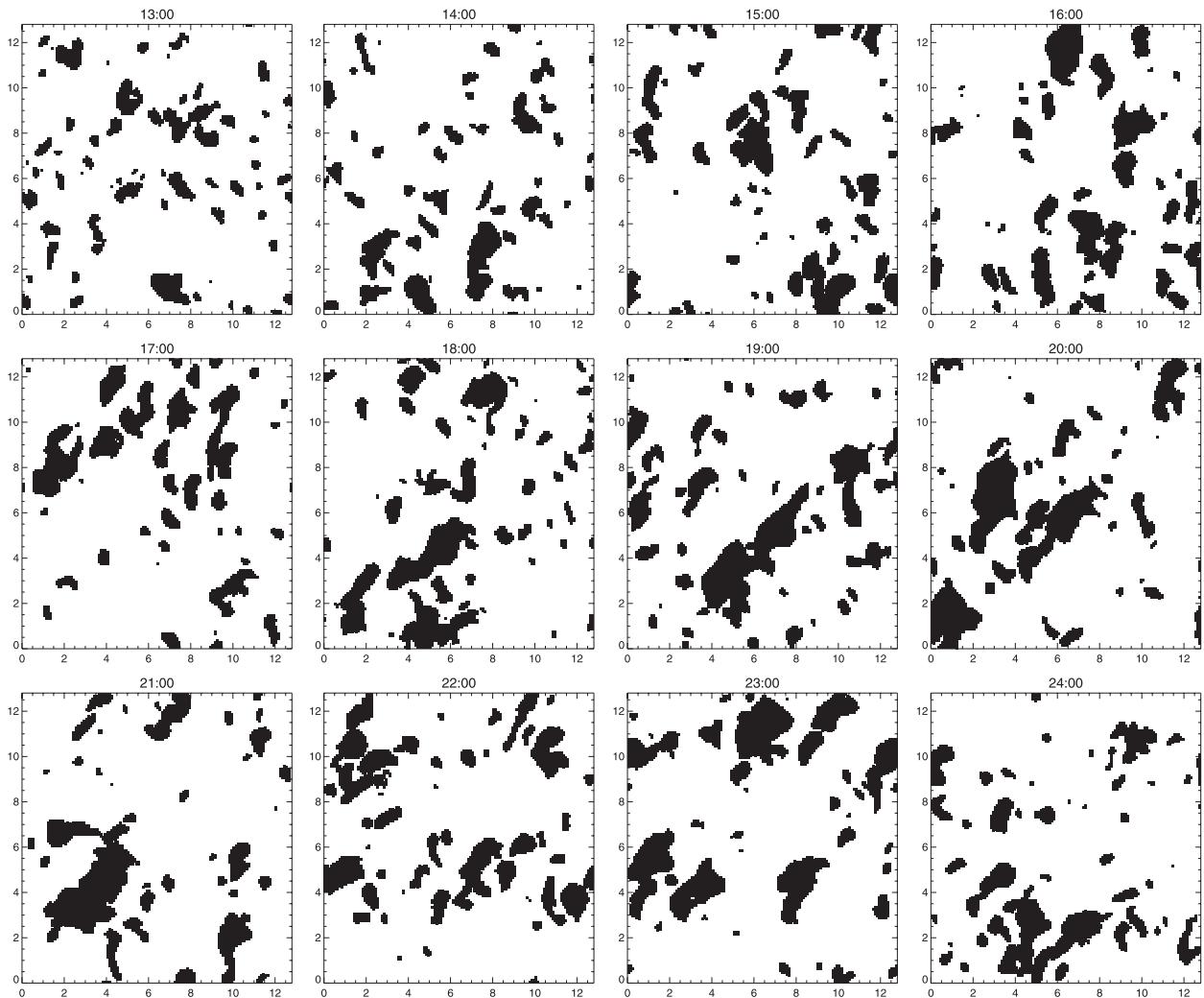


FIG. 5. As in Fig. 4, but for times 1300–2400 h.

The ATEX case of cumulus rising into stratocumulus does not indicate pronounced underestimation (relative to our CSM) of small clouds, and thus the agreement of the LES cloud length statistics with our cellular model is good in this case: $\Delta = 6.9\%$, $\alpha_k = 0.81$. The gap length histogram, however, indicates some overestimation of the likelihood of small gaps at the expense of gaps at moderate sizes (approximately 1–2 km), thereby diminishing the agreement ($\Delta = 10\%$, $\alpha_k = 0.17$ for the gaps). A possible source of deviation between the LES and cellular model statistics is the gradual change in the cloud field during the simulated period, which results in a non-homogeneous statistical ensemble. While we reduced this inhomogeneity somewhat by considering only the first part of the ATEX dataset (including the second part results in a bimodal cloud cover distribution), it is likely that residual problems associated with the gradual evolution remain.

4. Diverse ensembles and power-law chord length distributions

In many observations, especially those with extensive coverage (temporal, geographic, or both), cloud length distributions are predominantly observed to follow a power-law form (Cahalan and Joseph 1989; Koren et al. 2008), rather than the exponential form that emerges from our cellular model. To obtain power-law distributions with our cellular model, we generalize the statistical ensemble by allowing the cell cloudiness probability to vary. This is a reasonable extension, since the probability of cloudiness varies with meteorological conditions (e.g., consider the differences among the small set of cases presented in the previous section). Thus we assume that p is itself a random variable with its own statistical density function $g(p)$ ($0 \leq p \leq 1$), which is normalized on the unit interval:

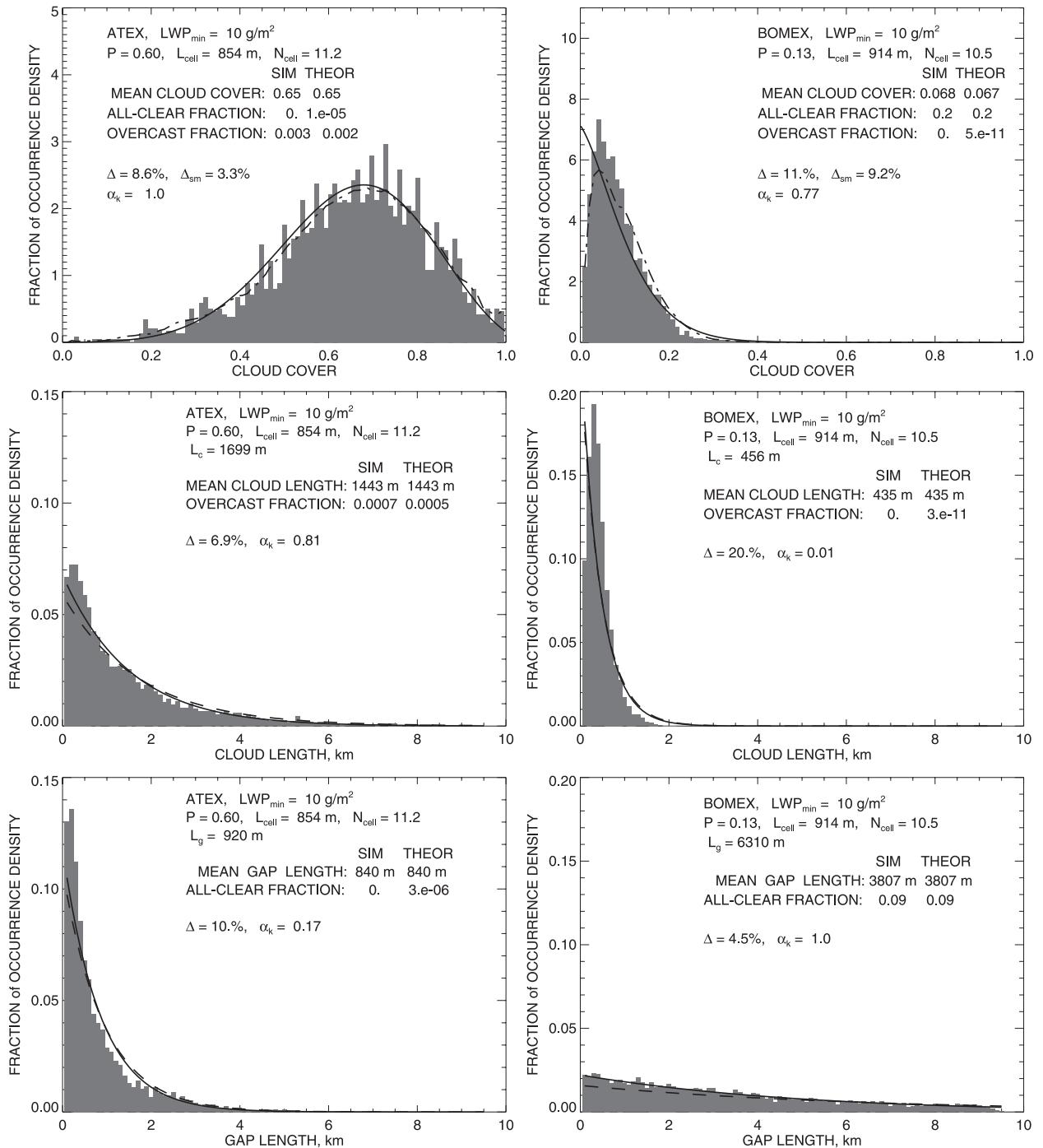


FIG. 6. Cloud and gap chord statistics from LES results for the (left) ATEX (simulation times 0400–0645 h) and (right) BOMEX (0400–1030 h) cases. Cloudy columns are defined as those with LWP > 10 g m⁻². (top) Cloud cover distributions shown as histograms, 15-bin moving window smoothing over histograms shown as dashed–dotted lines, and theoretical density [Eq. (19)] shown as solid lines. (middle), (bottom) Cloud and gap length distributions, in which solid lines show corresponding theoretical densities [Eq. (12)] for finite samples and dashed lines show theoretical densities for infinite samples [Eq. (9)]. Theoretical distributions are based on parameters derived from mean cloud and gap lengths. Histograms bin widths are described in text.

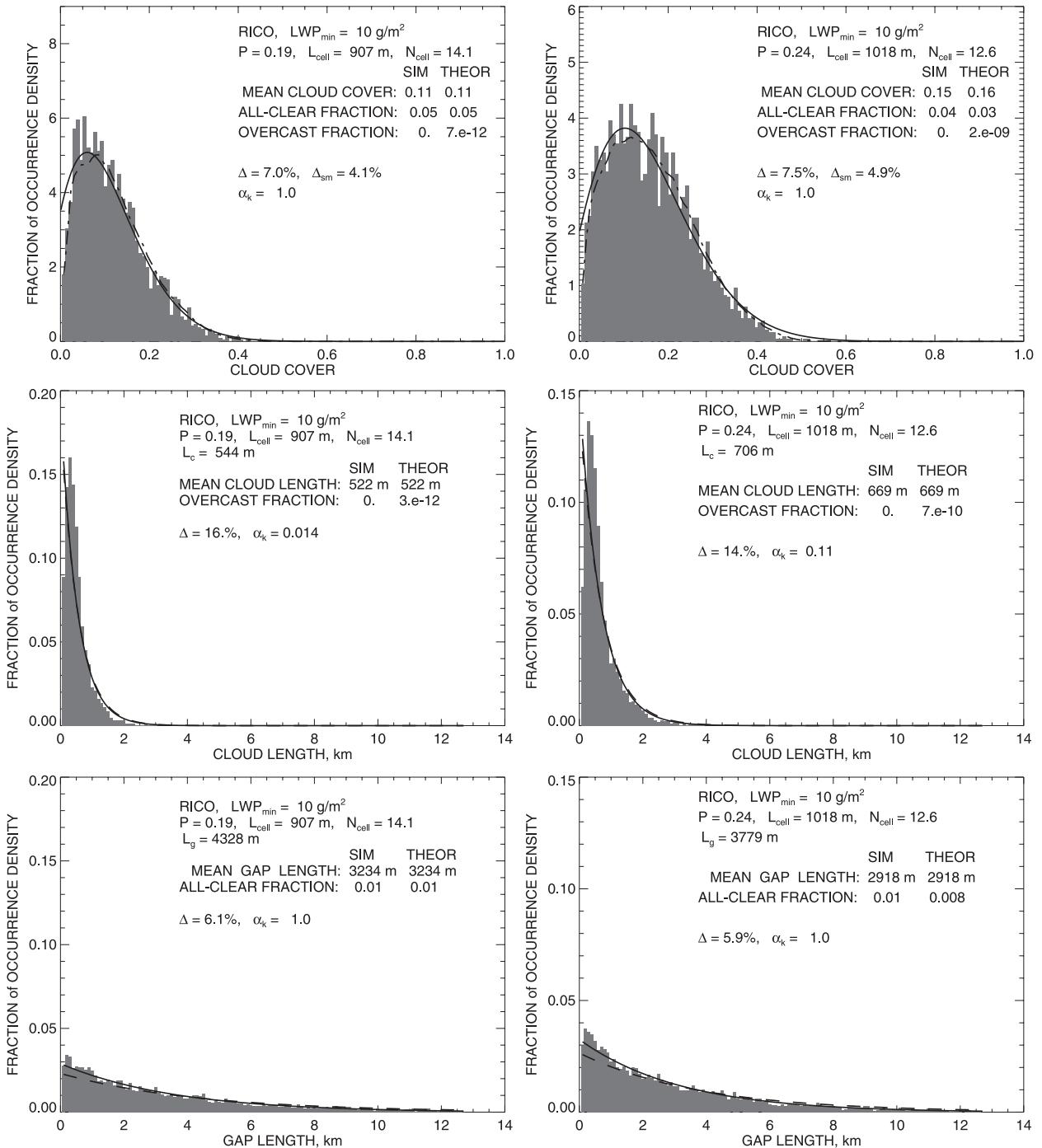


FIG. 7. As in Fig. 6, but for RICO field campaign simulations: (left) 0600–1700 h, (right) 1300–2400 h.

$$\int_0^1 g(p) dp = 1. \quad (33)$$

This means that each sample in a diverse ensemble is generated assuming a fixed p (which may differ, however, from sample to sample). We do not consider a more complex situation when p is allowed to vary within a

sample. We also do not consider the possibility of variation of the CSM cell size l because of the difficulty of corresponding analytical computations. Note that this generalization is much more natural in our (p, l) formulation of the problem than in the equivalent (L_c, L_g) formulation adopted, for example, by Astin and Di Girolamo (1999). Below we will present examples for several simple forms

of $g(p)$, such as uniform and beta distributions, in both discrete and continuous models. We will also describe (in the framework of the continuous model) the conditions on $g(p)$ in general case, which are sufficient for the corresponding cloud and/or gap distributions to have a power-law form.

a. Discrete model

In the discrete model the generalized cloud length distribution is computed by integration of the distribution (1) with $g(p)$:

$$n_c(m) = \int_0^1 g(p)(1-p)p^{m-1} dp. \tag{34}$$

In the simplest case of $g(p) \equiv 1$ this distribution takes the form

$$n_c(m) = \frac{1}{m} - \frac{1}{m+1} = \frac{1}{m^2 + m}, \tag{35}$$

which is normalized by the condition

$$\sum_{m=1}^{\infty} n_c(m) = 1 \tag{36}$$

and has a power-law form at large m :

$$n_c(m \gg 1) \approx m^{-2}. \tag{37}$$

To chose a more general form of $g(p)$ we refer to the literature on cloud cover, which is a natural proxy for cloud probability. In a number of studies the beta distribution

$$g_{ab}(p) = \frac{1}{B(a, b)} p^{a-1} (1-p)^{b-1}, \tag{38}$$

with two parameters $a, b > 0$, has been used to model the distribution of cloud cover (e.g., Falls 1974; Jones 1992). Here

$$B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \tag{39}$$

is the beta function, and $\Gamma(x)$ is the gamma function. Obviously, $g(p) \equiv 1$ is a specific instance of Eq. (38), with $a = b = 1$. For other examples of beta distributions see Fig. 9 (left panels). Substituting Eq. (38) into Eq. (34), we obtain

$$\begin{aligned} n_c(m) &= \frac{1}{B(a, b)} \int_0^1 (1-p)^b p^{m+a-2} dp \\ &= \frac{B(b+1, m+a-1)}{B(a, b)}. \end{aligned} \tag{40}$$

To estimate the behavior of the distribution (40) at large m , we use the Stirling's asymptotic formula for the beta function:

$$B(x \gg 1, y) \approx \Gamma(y)x^{-y}, \tag{41}$$

where y is fixed. This yields

$$\begin{aligned} n_c(m \gg 1) &\approx \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \Gamma(b+1)(m+a-1)^{-(b+1)} \\ &\approx b \frac{\Gamma(a+b)}{\Gamma(a)} m^{-(b+1)}, \end{aligned} \tag{42}$$

showing that at large m our cloud length distribution has a power-law form. The chord length distribution of gaps is the same as Eq. (42) only with a and b interchanged:

$$n_g(m) = \frac{B(a+1, m+b-1)}{B(a, b)}, \tag{43}$$

which at large m takes the form

$$n_g(m \gg 1) \approx a \frac{\Gamma(a+b)}{\Gamma(b)} m^{-(a+1)}. \tag{44}$$

Note that the exponent in the cloud length distribution relates to b , governing the behavior of $g(p)$ when p is close to unity (i.e., when clouds are expected to be large). Conversely, the gap length distribution exponent relates to a , governing $g(p)$ at small p , which corresponds to small clouds and large gaps. If $g(p)$ is symmetric relative to $p = 1/2$ (i.e., $a = b$), cloudy and clear intervals have the same asymptotic distribution,

$$n_c(m) \equiv n_g(m) \approx a \frac{\Gamma(2a)}{\Gamma(a)} m^{-(a+1)},$$

while for the uniform distribution ($a = b = 1$) we obtain Eq. (37).

Similar computations for cloud cover (see the appendix) show that for $N \gg 1, k \gg 1$, and $N - k \gg 1$, cloud fraction has the same beta distribution as the cell cloudiness probability:

$$n_{ct}(c) = \frac{1}{B(a, b)} c^{a-1} (1-c)^{b-1}. \tag{45}$$

b. Continuous model

In the continuous model, unfortunately we are able to perform few computations analytically (see the appendix for details). The cloud and gap length distributions of the diverse ensemble can be computed by integrating Eq. (9) or (11) with $g(p)$:

$$F_{c,g}^{(d)}(x) = \int_0^1 g(p) F_{c,g}(x, p) dp. \tag{46}$$

Here $F_{c,g}(x, p)$ is the generally singular length distribution for fixed occupation p . Below we will drop the superscript (d) from the diverse sample distributions when doing so does not cause confusion.

1) INFINITE SAMPLES

In the infinite sample case Eq. (46) takes the following form:

$$f_c(x) = \int_0^1 \frac{g(p)}{L_c(p)} e^{-x/L_c(p)} dp$$

$$= -\frac{1}{l} \int_0^1 g(p) p^{x/l} \ln p dp, \tag{47}$$

and

$$f_g(x) = \int_0^1 \frac{g(p)}{L_g(p)} e^{-x/L_g(p)} dp$$

$$= -\frac{1}{l} \int_0^1 g(p) (1-p)^{x/l} \ln(1-p) dp$$

$$= -\frac{1}{l} \int_0^1 g(1-p) p^{x/l} \ln p dp. \tag{48}$$

In the simplest case $g(p) \equiv 1$ these integrals can be easily computed and $f_c \equiv f_g$ takes a power-law form:

$$f_{c,g}(x) = \frac{1}{l} \left(\frac{x}{l} + 1 \right)^{-2}. \tag{49}$$

It is easily verified that this distribution is properly normalized.

We can compute the cloud length distribution in the more complicated case of rectangular $g(p)$:

$$g(p) = \begin{cases} 1/2w & \text{for } p \in [p_0 - w, p_0 + w] \\ 0 & \text{otherwise} \end{cases}, \tag{50}$$

where p_0 is the mean probability and w is the distribution's half-width (obviously $w \leq p_0 \leq 1 - w$). If $p_0 = w = 1/2$ we obtain the case of $g \equiv 1$ discussed above. In the contrasting case of a very narrow distribution ($w \ll 1$), $g(p) \approx \delta(p - p_0)$, and f_c is the same as for the case of a single-probability model with $p = p_0$. In the general case f_c from Eq. (47) takes the form

$$f_c(x) = -\frac{1}{2wl} \int_{p_0-w}^{p_0+w} p^{x/l} \ln p dp,$$

$$= \frac{1}{2w} e^{-(x+l)/L_+} \left[\frac{l}{L_+(x+l)} + \frac{l}{(x+l)^2} \right]$$

$$- \frac{1}{2w} e^{-(x+l)/L_-} \left[\frac{l}{L_-(x+l)} + \frac{l}{(x+l)^2} \right], \tag{51}$$

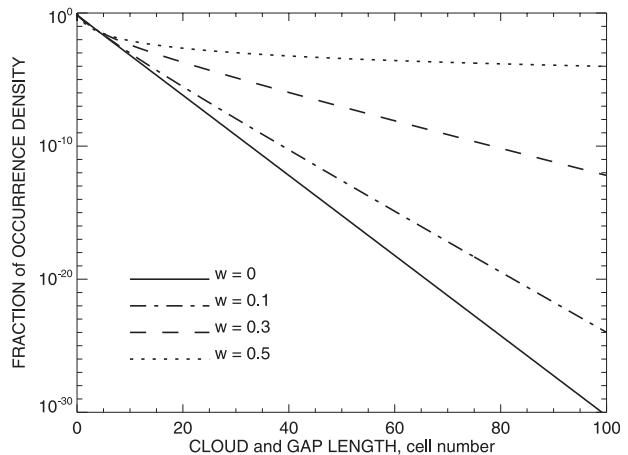


FIG. 8. Cloud length distribution densities [Eq. (51)] for $p_0 = 1/2$ and different values of the width w . The distribution of gap lengths is identical to that for clouds in this case.

where

$$L_{\pm} = -\frac{l}{\ln(p_0 \pm w)}. \tag{52}$$

It is not difficult to show that the distribution (51) has exponential asymptotic unless $p_0 + w = 1$, in which case $f_c(x) \propto x^{-2}$, as in Eq. (49). However, the transformation of the distribution shape with increase of w is continuous and takes the form of gradual fattening of the distribution tail, as shown in Fig. 8 for $p_0 = 1/2$.

In the case when $g(p)$ is described by a beta distribution [Eq. (38)], the cloud length distribution takes the following form (see the appendix):

$$f_c(x) = \frac{B\left(\frac{x}{l} + a, b\right)}{lB(a, b)} \left[\psi\left(\frac{x}{l} + a + b\right) - \psi\left(\frac{x}{l} + a\right) \right], \tag{53}$$

where ψ is the digamma function (logarithmic derivative of the gamma function; cf. Gradshteyn and Ryzhik 1965). The corresponding gap length distribution can be obtained from Eq. (53) by interchanging the indices a and b [since $g_{ab}(1-p) = g_{ba}(p)$]:

$$f_g(x) = \frac{B\left(\frac{x}{l} + b, a\right)}{lB(a, b)} \left[\psi\left(\frac{x}{l} + a + b\right) - \psi\left(\frac{x}{l} + b\right) \right]. \tag{54}$$

Examples of cloud and gap length distributions for various beta distribution parameters are shown in Fig. 9. As expected, the cloud and gap length densities coincide for symmetric distributions ($a = b$, Fig. 9 top). We also see that variation of parameters of strongly asymmetric $g(p)$ (Fig. 9 bottom) leads to larger changes in the corresponding cloud and gap distributions than in the symmetric

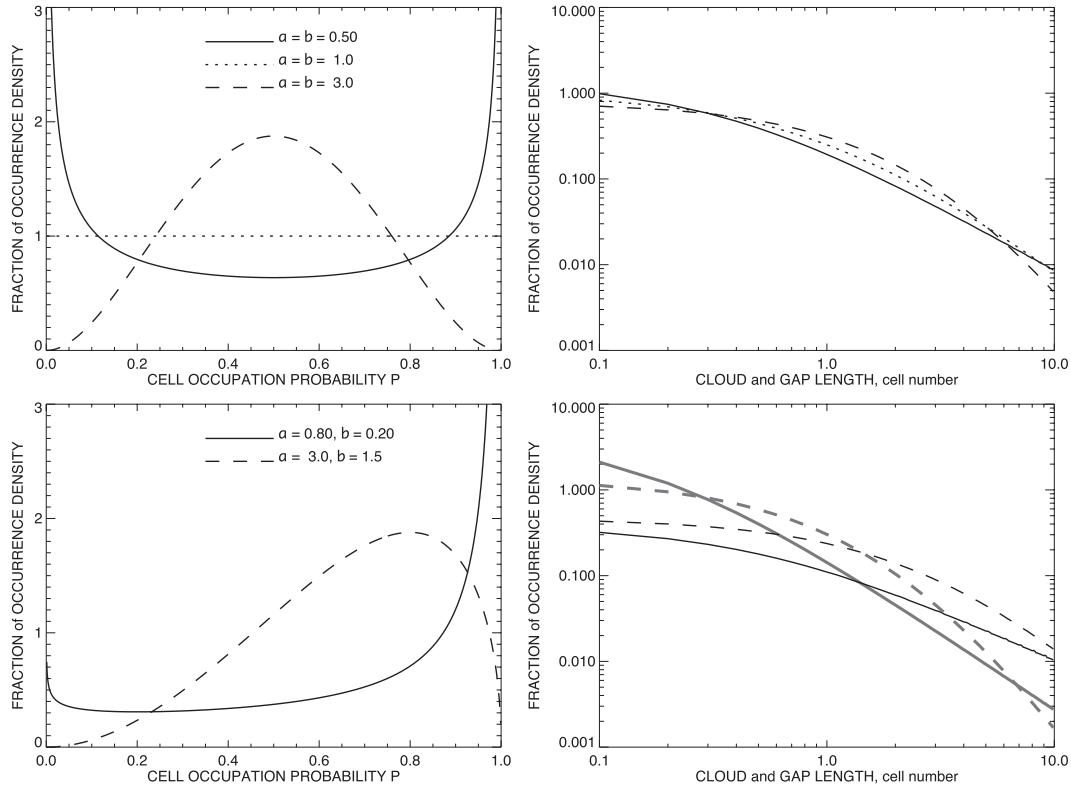


FIG. 9. (left) Examples of beta distributions of occupation probability for different values of parameters a and b . (right) Corresponding cloud and gap distribution densities computed according to Eqs. (53) and (54). Since the distributions in the top left plot are symmetric with respect to $p \leftrightarrow (1 - p)$, the corresponding cloud and gap length distributions coincide. In the bottom right, cloud and gap length distributions are shown respectively by thin black and thick gray curves.

case. The power-law asymptotic behavior in the displayed cloud/gap size range is better pronounced for $a, b < 1$.

The asymptotic behavior of f_c for large $x \gg l$ has a power-law form

$$f_c(x \gg l) \propto \frac{1}{x^{b+1}} \quad (55)$$

that is consistent with the discrete model result (42). In the degenerate case $b = 1$ when

$$g_{a1}(p) = ap^{a-1}, \quad (56)$$

Eq. (53) then takes the simple form

$$f_c(x) = \frac{a}{l} \left(\frac{x}{l} + a \right)^{-2}, \quad (57)$$

which coincides with Eq. (49) when $a = 1$.

In general, the asymptotic behavior of the cloud and gap length distributions as $x \rightarrow \infty$ is determined by the integration range where p is close to 1 (otherwise the factor $p^{x/l}$ is very small). Therefore, asymptotes of f_c and

f_g depend on the way in which $g(p)$ and $g(1 - p)$ respectively approach $p = 1$ [in the latter case this is the same as the expansion of $g(q)$ at $q = 0$]. Indeed, the largest clouds expectedly come from samples with large p , while the largest gaps are from those with large q . For example, when $g(p) \propto (1 - p)^{b-1} = q^{b-1}$ for $q \ll 1$ (i.e., has a power-law asymptotic form, as in the case of beta distribution), we can write Eq. (47) in the following form (noticing that in this limit case $\ln p \approx q$):

$$\begin{aligned} f_c(x \gg l) &\approx \frac{1}{l} \int_0^1 q^b (1 - q)^{x/l} dp \propto B\left(b + 1, \frac{x}{l} + 1\right) \\ &\approx \Gamma(b + 1) \left(\frac{x}{l} + 1\right)^{-(b+1)} \\ &\sim x^{-(b+1)}. \end{aligned} \quad (58)$$

Here we replaced $g(p)$ by its expansion on the whole unit interval, since only small values of q contribute to the integration, and the values of g elsewhere do not matter. We also used the asymptotic formula for the beta function (41). Computations for gap size distribution are similar.

We certainly cannot consider all possible shapes of $g(p)$ here; however, some observations are not difficult to make based on the examples that we presented. In particular, we saw in the case of rectangular $g(p)$ that if the distribution density of p is zero in a neighborhood of $p = 1$, the corresponding cloud length distribution (51) has exponential asymptote. On the other hand, in the case when $g(p)$ approaches $p = 1$ as a power-law function of q , the cloud length distribution also has a power-law form. Other types of behavior of g require individual investigations.

2) FINITE SAMPLES

In the finite sample case, when the cloud and gap length distributions are described by Eqs. (11)–(13), and p obeys the beta distribution, it can be demonstrated (see the appendix) that the expression for the cloud length distribution consists of the infinite-sample term (53) and a term proportional to the inverse sample length (and therefore vanishing as $L \rightarrow \infty$). However, a complete analytical formula for this distribution can be obtained only in the degenerate case of $b = 1$ mentioned above:

$$f_c(x) = \frac{a}{l} \left(1 - \frac{x}{L}\right) \left(\frac{x}{l} + a\right)^{-2} + \frac{a}{L} \left(1 + \frac{x}{L}\right) \times \left[\frac{l}{L} e^{(x+al)/L} \text{Ei}\left(-\frac{x+al}{L}\right) + \left(\frac{x}{l} + a\right)^{-1} \right], \quad (59)$$

where Ei is the exponential integral. As $L \rightarrow \infty$ this expression converges to Eq. (57).

The expression for the overcast fraction in this case takes the form

$$f_c^o = -\frac{al}{L} e^{(L+al)/L} \text{Ei}\left(-\frac{L+al}{L}\right). \quad (60)$$

Note that $\text{Ei}(x) < 0$ for negative x , thus the overcast fraction is always positive. Clearly $f_c^o \rightarrow 0$ as $L \rightarrow \infty$. Plots of the cloud and gap length distributions (59) for $L = l, 5l$, and ∞ are presented in Fig. 10 for the uniform $g(p) \equiv 1$ ($a = b = 1$), together with the corresponding values of f_c^o . These values show significant increase from 0 to 0.36 as L/l decreases from ∞ to 1, while the differences in density functions themselves (likely to be driven by the corresponding change in normalization conditions) are most pronounced at the shortest cloud and gap lengths. As expected, the clearest asymptotic power-law behavior is seen in the infinite sample case [when Eq. (59) coincides with Eq. (57)].

5. Summary

Part I documents two cellular models intended to describe cloud field statistics such as the distributions of

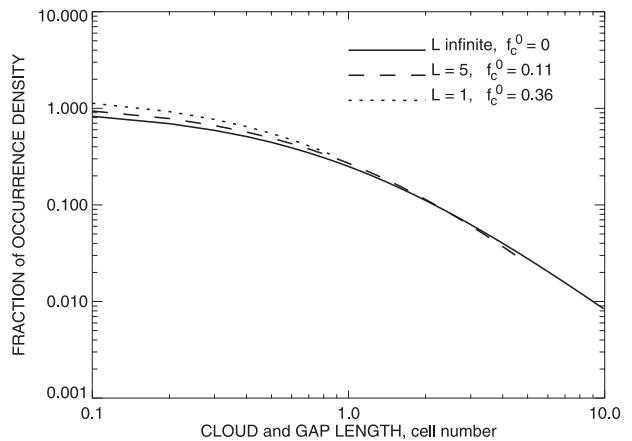


FIG. 10. Cloud and gap length distribution densities [Eq. (59)] for finite sample lengths $L = l, 5l$, and ∞ and for uniform $g(p) \equiv 1$. The corresponding values of f_c^o are computed according to Eq. (60).

cloud and gap lengths and of cloud fraction. Our discrete model can be used to construct examples of cloud fields in both 1D and 2D, since the cell occupation probabilities are independent of each other; however, since our continuous model is well defined only in 1D (see Part I for elaboration), we have restricted our analysis to cloud statistics measured along 1D transects. Thus we have considered distributions of cloud and gap chord lengths rather than cloud and clear areas.

While our 1D cellular cloud models are mathematically well defined, their fidelity can be determined only through comparison with statistics of either observed cloud fields or those simulated with a dynamical model, as done here using the output of large-eddy simulations (LES) based on three case studies of maritime trade-cumulus cloud fields. We have considered only the continuous model, since distributions from its discrete analog are not realistic by virtue of all its cloud and gap lengths being integer multiples of the CSM cell length l . The continuous model is not equivalent to a discrete model just with a larger number of smaller cells: the latter would produce rather homogeneous samples with no clumpy clouds. Defining cloudy columns as those with more than 10 g m^{-2} liquid water path, 1D transect ensembles were produced from the 2D LES fields by taking all horizontal samples in the north–south and east–west directions. Kolmogorov’s goodness-of-fit test and another simple difference metric (32) were used to compare the LES statistics with those from our cellular model. The observed average CSM cell sizes range from 850 to 1000 m, while the cell occupation (cloudiness) probability is in the range of 13%–24% for the purer cumulus cases (BOMEX, RICO) and much greater, at 60%, for the case of cumulus rising into stratocumulus

(ATEX). The chord length and cloud fraction distributions from our cellular model generally agree with those from the LES datasets, with the main differences found in the population of the cloud lengths approaching the LES grid spacing (100 m) in the BOMEX and RICO datasets. The cloud length distributions in these cases have a peak at 200–300 m, which does not fit the exponential distribution of the cellular model. This discrepancy merits further study focusing on the structure of cloud boundaries, from which the shortest cloud transects come.

We also have demonstrated how at certain conditions the cloud and gap statistics in our cellular model can change from an exponential distribution for a narrow sample to a power-law distribution for a diverse sample that includes a variety of meteorological states. We incorporate meteorological diversity into our cellular model by allowing the cell cloudiness probability p to have its own distribution rather than assuming a uniform value. This approach can play an important role in statistical analysis of global satellite cloud datasets including a large diversity of cloud types and atmospheric conditions.

Acknowledgments. This research was supported by NASA’s Glory Project and Radiation Science Program. We thank E. Kassianov, A. Korolev, A. Del Genio, B. Cairns, A. Lacis, and D. Stevens for useful discussions. We are also indebted to the reviewers of this paper for their detailed comments, which helped us to improve our work.

APPENDIX

Statistics of Diverse Ensembles

a. Cloud cover in discrete model

The computations in this section show that cloud cover distribution in the discrete model statistically resembles that of the cell-filling probability p . To demonstrate this we perform an integration over p of the binomial distribution (7) with $g(p)$, as done above for the cloud length distribution:

$$n(k) = \binom{N}{k} \int_0^1 g(p) p^k (1-p)^{N-k} dp. \quad (A1)$$

In the general case when $g(p)$ is a beta distribution (38), we obtain

$$\begin{aligned} n(k) &= \binom{N}{k} \frac{1}{B(a, b)} \int_0^1 p^{k+a-1} (1-p)^{N-k+b-1} dp \\ &= \frac{N}{k(N-k)} \frac{B(N, a+b)}{B(k, a)B(N-k, b)}. \end{aligned} \quad (A2)$$

Now, assuming that $N \gg 1, k \gg 1$, and $N - k \gg 1$, we can use the asymptotic formula for the beta function [Eq. (41)] to derive

$$\begin{aligned} n(k) &\approx \frac{N}{k(N-k)} \frac{\Gamma(a+b)N^{-(a+b)}}{\Gamma(a)k^{-a}\Gamma(b)(n-k)^{-b}} \\ &= \frac{1}{N} \frac{1}{B(a, b)} \left(\frac{k}{N}\right)^{a-1} \left(1 - \frac{k}{N}\right)^{b-1}. \end{aligned} \quad (A3)$$

We notice that

$$n(c) = Nn(k) \quad (A4)$$

since

$$\int_0^N n(k) dk = \int_0^1 n(c) dc = 1, \quad (A5)$$

while $dk = Ndc$. Thus, Eq. (A3) translates into the beta distribution (45) for the cloud cover c , which is identical to the distribution of occupation probability.

b. Cloud length statistics in continuous model

To compute the cloud length distribution according to Eq. (47) in the general case, it is convenient to use the substitutions

$$u = -\ln p, \quad p = e^{-u}, \quad \text{and} \quad dp = -e^{-u} du. \quad (A6)$$

Note that u runs from ∞ to 0 as p runs from 0 to 1. We also introduce the following notation:

$$z = \frac{x}{l} \quad \text{and} \quad h(u) = g(e^{-u}). \quad (A7)$$

In this notation Eq. (47) in the infinite sample case can be written as

$$lf_c(lz) = - \int_0^1 g(p) p^z \ln p dp = \int_0^\infty uh(u) e^{-(z+1)u} du, \quad (A8)$$

which is essentially the Laplace transform (cf. Prudnikov et al. 1992)

$$\Phi(s) = \mathcal{L}[\phi(u)] = \int_0^\infty \phi(u) e^{-su} du$$

of $\phi(u) = uh(u)$.

In the case when $g(p)$ is a beta distribution

$$h_{ab}(u) = \frac{1}{B(a, b)} e^{-(a-1)u} (1 - e^{-u})^{b-1}, \quad (A9)$$

and we take a slightly different approach, writing

$$lf_c(lz) = \frac{1}{B(a, b)} \int_0^\infty ue^{-(z+a)u}(1 - e^{-u})^{b-1} du$$

$$= \frac{\mathcal{F}(z + a)}{B(a, b)}, \tag{A10}$$

where

$$\mathcal{F}(s) = \int_0^\infty u(1 - e^{-u})^{b-1} e^{-su} du \tag{A11}$$

is the Laplace transform of

$$\phi(u) = u(1 - e^{-u})^{b-1}. \tag{A12}$$

This transform appears in the tables of Prudnikov et al. (1992), yielding

$$\mathcal{F}(s) = B(s, b)[\psi(s + b) - \psi(s)], \tag{A13}$$

where B is the beta function and ψ is the digamma function (logarithmic derivative of the gamma function), which leads to Eq. (53).

In the degenerate case $b = 1$ we use the relations

$$\psi(x + 1) = \psi(x) + \frac{1}{x} \quad \text{and} \quad B(x, 1) = \frac{1}{x}$$

to derive

$$\mathcal{F}(s) = \frac{1}{s^2}, \tag{A14}$$

and $f_c(x)$ takes the form (57).

To see the asymptotic behavior of the distribution (53) for large $x \gg l$ ($s \gg 1$), we use the asymptotic approximation for the digamma function (cf. Gradshteyn and Ryzhik 1965):

$$\psi(x \gg 1) \approx \ln x - \frac{1}{2x} - \dots$$

Thus,

$$\psi(s + b) - \psi(s) \approx \ln\left(1 + \frac{b}{s}\right) \approx \frac{b}{s}$$

[the contribution of the second term in the expansion is of the order of s^{-2}]. Combining this result with Eq. (41) for $B(s, b)$ we find that

$$\mathcal{F}(s \gg 1) \propto \frac{1}{s^{b+1}};$$

thus, f_c has a power-law form (55).

In the case of a finite sample, when the cloud length distribution has the form (11)–(13), the density f_c can be written as

$$lf_c(lz) = -\ln p \left(1 + \frac{1 + z \ln p}{1 - \nu \ln p}\right) p^z$$

$$= u \left(1 + \frac{1 - zu}{1 + \nu u}\right) e^{-u},$$

where we have used the substitution (A6). Here $\nu \equiv N = Ll$, and z is defined by Eq. (A7). In these notations the cloud length density (46) for a diverse ensemble with p obeying beta distribution will have the form

$$lf_c^{(d)}(lz) = -\frac{1}{B(a, b)} \int_0^1 p^{a-1} (1 - p)^{b-1} \ln p \left(1 + \frac{1 + z \ln p}{1 - \nu \ln p}\right) p^z dp = \mathcal{I}_1 + \mathcal{I}_2,$$

where

$$\mathcal{I}_1 = \frac{1}{B(a, b)} \int_0^\infty ue^{-(z+a)u}(1 - e^{-u})^{b-1} du$$

and

$$\mathcal{I}_2 = \frac{1}{B(a, b)} \int_0^\infty ue^{-(z+a)u}(1 - e^{-u})^{b-1} \frac{1 - zu}{1 + \nu u} du. \tag{A15}$$

The first integral is identical to Eq. (A10), which we have already computed in the infinite sample case. Similarly, we can write the second integral as

$$\mathcal{I}_2 = \frac{\mathcal{G}(z + a)}{B(a, b)}, \tag{A16}$$

where

$$\mathcal{G}(s) = \int_0^\infty u(1 - e^{-u})^{b-1} \frac{1 - zu}{1 + \nu u} e^{-su} du \tag{A17}$$

is the Laplace transform of

$$\phi(u) = u(1 - e^{-u})^{b-1} \frac{1 - zu}{1 + \nu u}. \tag{A18}$$

Unfortunately, we have not found or derived an expression for this transform in general case. However, in the special case when $b = 1$, ϕ takes a simple form

$$\phi(u) = \frac{\mu u}{u + \mu} - \frac{\mu z u^2}{u + \mu}, \tag{A19}$$

where $\mu = 1/\nu$, and we can use the formula from the table of Prudnikov et al. (1992):

$$\begin{aligned} \mathcal{L}\left(\frac{x^n}{x + \mu}\right) &= (-1)^{n+1} \mu^n e^{s\mu} \text{Ei}(-s\mu) \\ &+ \sum_{k=1}^n (k-1)! \frac{(-\mu)^{n-k}}{s^k}, \end{aligned}$$

where Ei is the exponential integral. This expression takes the following form for $n = 1$:

$$\mathcal{L}\left(\frac{x}{x + \mu}\right) = \mu e^{s\mu} \text{Ei}(-s\mu) + \frac{1}{s}.$$

For $n = 2$,

$$\mathcal{L}\left(\frac{x^2}{x + \mu}\right) = -\mu^2 e^{s\mu} \text{Ei}(-s\mu) - \frac{\mu}{s} + \frac{1}{s^2}.$$

Thus, the Laplace transform of ϕ takes the form

$$\mathcal{G}(s) = (1 + z\mu) \mu \left[\mu e^{s\mu} \text{Ei}(-s\mu) + \frac{1}{s} \right] - \frac{z\mu}{s^2} \tag{A20}$$

and

$$\begin{aligned} \mathcal{I}_2 &= a(1 + z\mu) \mu \left\{ \mu e^{(z+a)\mu} \text{Ei}[-(z+a)\mu] + \frac{1}{z+a} \right\} \\ &- \frac{az\mu}{(z+a)^2}. \end{aligned} \tag{A21}$$

One can see that this expression is proportional to $\mu = l/L$; thus, it vanishes in the infinite sample limit ($L \rightarrow \infty$). In this case the only contribution to f_c is from \mathcal{I}_1 , which leads to Eq. (53).

The fraction of overcast samples f_c^o in the diverse sample scenario is defined as

$$\begin{aligned} f_c^o &= \frac{1}{B(a, b)} \int_0^1 p^{a-1} (1-p)^{b-1} \frac{p^\nu}{1-\nu \ln p} dp \\ &= \frac{1}{B(a, b)} \int_0^\infty \frac{e^{-(\nu+a)u} (1-e^{-u})^{b-1}}{1+\nu u} du \\ &= \frac{\mathcal{G}_0(\nu+a)}{B(a, b)}, \end{aligned} \tag{A22}$$

where

$$\mathcal{G}_0(s) = \int_0^\infty \frac{(1-e^{-u})^{b-1}}{1+\nu u} e^{-su} du \tag{A23}$$

is the Laplace transform of

$$\phi(u) = \frac{(1-e^{-u})^{b-1}}{1+\nu u}. \tag{A24}$$

Unfortunately, we have not found or derived the solution to this transform in the general case. However, for $b = 1$

$$\phi(u) = \frac{\mu}{u + \mu}, \tag{A25}$$

and we can use the formula (Prudnikov et al. 1992)

$$\mathcal{L}\left(\frac{1}{x + \mu}\right) = -e^{s\mu} \text{Ei}(-s\mu),$$

which yields

$$f_c^o = -\frac{a}{\nu} e^{(\nu+a)/\nu} \text{Ei}\left(-\frac{\nu+a}{\nu}\right). \tag{A26}$$

Note that $\text{Ei}(x) < 0$ for $x < 0$, and that $f_c^o \rightarrow 0$ as $L \rightarrow \infty$, since $\nu \propto L$.

REFERENCES

Abrams, M., 2000: The Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER): Data products for the high spatial resolution imager on NASA's *Terra* platform. *Int. J. Remote Sens.*, **21**, 847–859.

Ackerman, A. S., M. P. Kirkpatrick, D. E. Stevens, and O. B. Toon, 2004: The impact of humidity above stratiform clouds on indirect aerosol climate forcing. *Nature*, **432**, 1014–1017.

Alexandrov, M. D., A. Marshak, and A. S. Ackerman, 2010: Cellular statistical models of broken cloud fields. Part I: Theory. *J. Atmos. Sci.*, **67**, 2125–2151.

Astin, I., and B. G. Latter, 1998: A case for exponential cloud fields? *J. Appl. Meteor.*, **37**, 1375–1382.

—, and L. Di Girolamo, 1999: A general formalism for the distribution of the total length of a geophysical parameter along a finite transect. *IEEE Trans. Geosci. Remote Sens.*, **37**, 508–512.

Augstein, E., H. Riehl, F. Ostapoff, and V. Wagner, 1973: Mass and energy transports in an undisturbed Atlantic trade wind flow. *Mon. Wea. Rev.*, **101**, 101–111.

Berg, L. K., and E. I. Kassianov, 2008: Temporal variability of fair-weather cumulus statistics at the ACRF SGP site. *J. Climate*, **21**, 3344–3358.

Byrne, N., 2005: 3D radiative transfer in stochastic media. *Three-Dimensional Radiative Transfer in Cloudy Atmospheres*, A. Marshak and A. B. Davis, Eds., Springer, 385–424.

Cahalan, R. F., 1994: Bounded cascade clouds: Albedo and effective thickness. *Nonlinear Processes Geophys.*, **1**, 156–167.

—, and J. H. Joseph, 1989: Fractal statistics of cloud fields. *Mon. Wea. Rev.*, **117**, 261–272.

- Conover, W. J., 1980: *Practical Nonparametric Statistics*. 2nd ed. J. Wiley and Sons, 507 pp.
- Dey, S., L. Di Girolamo, and G. Zhao, 2008: Scale effect on statistics of the macrophysical properties of trade wind cumuli over the tropical western Atlantic during RICO. *J. Geophys. Res.*, **113**, D24214, doi:10.1029/2008JD010295.
- Evans, K. F., and W. J. Wiscombe, 2004: An algorithm for generating stochastic cloud fields from radar profile statistics. *Atmos. Res.*, **72**, 263–289.
- Falls, L. W., 1974: The beta distribution: A statistical model for world cloud cover. *J. Geophys. Res.*, **79**, 1261–1264.
- Gradshteyn, I. S., and I. M. Ryzhik, 1965: *Table of Integrals, Series, and Products*. Academic Press, 1130 pp.
- Hogan, R. J., and S. F. Kew, 2005: A 3D stochastic cloud model for investigating the radiative properties of inhomogeneous cirrus clouds. *Quart. J. Roy. Meteor. Soc.*, **131**, 2585–2608.
- Isichenko, M. B., 1992: Percolation, statistical topography, and transport in random media. *Rev. Mod. Phys.*, **64**, 961–1043.
- Jones, P. A., 1992: Cloud-cover distributions and correlations. *J. Appl. Meteor.*, **31**, 732–741.
- Joseph, J. H., and R. F. Cahalan, 1990: Nearest neighbor spacing of fair weather cumulus clouds. *J. Appl. Meteor.*, **29**, 793–805.
- Koren, I., L. Oreopoulos, G. Feingold, L. A. Remer, and O. Altaratz, 2008: How small is a small cloud? *Atmos. Chem. Phys.*, **8**, 3855–3864.
- Lane, D. E., K. Goris, and R. C. J. Somerville, 2002: Radiative transfer through broken clouds: Observations and model validation. *J. Climate*, **15**, 2921–2933.
- Marshak, A., A. Davis, R. F. Cahalan, and W. J. Wiscombe, 1994: Bounded cascade models as nonstationary multifractals. *Phys. Rev. E*, **49**, 55–69.
- Nagel, K., and E. Raschke, 1992: Self-organizing criticality in cloud formation? *Physica A*, **182**, 519–531.
- Nicholls, S., 1989: The structure of radiatively driven convection in stratocumulus. *Quart. J. Roy. Meteor. Soc.*, **115**, 487–511.
- Nitta, T., and S. Esbensen, 1974: Heat and moisture budget analyses using BOMEX data. *Mon. Wea. Rev.*, **102**, 17–28.
- Plank, V. G., 1969: The size distribution of cumulus clouds in representative Florida populations. *J. Appl. Meteor.*, **8**, 46–67.
- Prigarin, S. M., and A. Marshak, 2009: A simple stochastic model for generating broken cloud optical depth and cloud-top height fields. *J. Atmos. Sci.*, **66**, 92–104.
- , T. B. Zhuravleva, and P. B. Volikova, 2002: Poisson model of broken multilayer cloudiness. *Atmos. Ocean. Opt.*, **15**, 947–954.
- Prudnikov, A. P., Y. A. Brychkov, and O. I. Marichev, 1992: *Direct Laplace Transforms*. Vol. 4, *Integrals and Series*, Gordon and Breach, 637 pp.
- Rauber, R. M., and Coauthors, 2007: Rain in Shallow Cumulus over the Ocean: The RICO campaign. *Bull. Amer. Meteor. Soc.*, **88**, 1912–1928.
- Rodts, S. M. A., P. G. Duynkerke, and H. J. J. Jonker, 2003: Size distributions and dynamical properties of shallow cumulus clouds from aircraft observations and satellite data. *J. Atmos. Sci.*, **60**, 1895–1912.
- Sánchez, A., T. F. Smith, and W. F. Krajewski, 1994: A three-dimensional atmospheric radiative transfer model based on the discrete-ordinates method. *Atmos. Res.*, **33**, 283–308.
- Schertzer, D., and S. Lovejoy, 1987: Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes. *J. Geophys. Res.*, **92**, 9693–9714.
- Schmidt, K. S., V. Venema, F. D. Giuseppe, R. Scheirer, M. Wendisch, and P. Pilewski, 2007: Reproducing cloud microphysics and irradiance measurements using three 3D cloud generators. *Quart. J. Roy. Meteor. Soc.*, **133**, 765–780.
- Siebesma, A. P., and Coauthors, 2003: A large-eddy simulation intercomparison study of shallow cumulus convection. *J. Atmos. Sci.*, **60**, 1201–1219.
- Stevens, B., and Coauthors, 2001: Simulations of trade wind cumuli under a strong inversion. *J. Atmos. Sci.*, **58**, 1870–1891.
- Su, B. J., and G. C. Pomraning, 1994: A stochastic description of a broken cloud field. *J. Atmos. Sci.*, **51**, 1969–1977.
- Venema, V., S. Bachner, H. W. Rust, and C. Simmer, 2006: Statistical characteristics of surrogate data based on geophysical measurements. *Nonlinear Processes Geophys.*, **13**, 449–466.
- Wielicki, B. A., and R. M. Welch, 1986: Cumulus cloud properties derived using Landsat satellite data. *J. Climate Appl. Meteor.*, **25**, 261–276.
- Zhao, G., and L. Di Girolamo, 2007: Statistics on the macrophysical properties of trade wind cumuli over the tropical western Atlantic. *J. Geophys. Res.*, **112**, D10204, doi:10.1029/2006JD007371.
- Zuev, V. E., and G. A. Titov, 1995: Radiative transfer in cloud fields with random geometry. *J. Atmos. Sci.*, **52**, 176–190.