

## Cellular Statistical Models of Broken Cloud Fields. Part I: Theory

MIKHAIL D. ALEXANDROV

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

ALEXANDER MARSHAK

*NASA Goddard Space Flight Center, Greenbelt, Maryland*

ANDREW S. ACKERMAN

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

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

### ABSTRACT

A new analytical statistical model describing the structure of broken cloud fields is presented. It depends on two parameters (cell size and occupancy probability) and provides chord distributions of clouds and gaps between them by length, as well as the cloud fraction distribution. This approach is based on the assumption that the structure of a cloud field is determined by a semiregular grid of cells (an abstraction of the atmospheric convective cells), which are filled with cloud with some probability. First, a simple discrete model is introduced, where clouds and gaps can occupy an integer number of cells, and then a continuous analog is developed, allowing for arbitrary cloud and gap sizes. The influence of a finite sample size on the retrieved statistics is also described.

### 1. Introduction

Clouds play an important role in the earth's climate. Classification of extensive satellite observations of clouds, as well as modeling of cloud dynamics and radiation processes in general circulation models, demands simple statistical parameterizations of broken cloud fields. Observations of cloud fields often are made along lines that transect the field. Such observations include, for example, lidar and ground-based sun-photometric measurements. The latter provide a binary time series indicating either presence or absence of cloud at a given moment in the line connecting the instrument and the sun [see the descriptions of cloud screening algorithms by Alexandrov et al. (2004) for the Multifilter Rotating Shadow-Band Radiometer (MFRSR) (cf. Harrison et al. 1994) and by Smirnov et al. (2000) for the CIMEL sun/sky radiometer operated by the Aerosol Robotic Network (AERONET)

(cf. Holben et al. 1998)]. These time series have been used to provide statistical distributions of cloud cover and cloud overpass times, which can be converted into cloud chord length (CCL) distributions using wind speed estimates. If horizontal shapes of clouds can be assumed to be round (e.g., in the cumulus case), a CCL distribution can be converted into area distribution (Yau and Rogers 1984; Sauvageot et al. 1999). Similarly to the cloud chord length, distributions of lengths of clear-sky gaps between clouds can be also derived. Theoretical statistical studies of geophysical fields distributed along a line were performed by Sánchez et al. (1994) and Astin and Di Girolamo (1999). Cloud statistics from linear transects play an important role in stochastic radiative transfer models (Byrne 2005; Lane-Veron and Somerville 2004). Several studies have been published describing the cloud chord statistics based on both ground-based measurements (Lane et al. 2002; Berg and Kassianov 2008) and aircraft and satellite data (Plank 1969; Cahalan and Joseph 1989; Joseph and Cahalan 1990; Rodts et al. 2003). Some of these studies, especially those restricted to fair weather cumulus fields, report exponential cloud chord length distributions (Astin and Latter 1998; Lane et al. 2002), while others report

---

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

power-law distributions (Cahalan and Joseph 1989; Koren et al. 2008). The observed cloud chord distribution depends on the dominant cloud type(s) during an observation period.

There are two main categories of cloud models widely used in simulating the spatial distribution of cloud condensate: dynamical and stochastic cloud models. Dynamical cloud models are physically based but require a lot of atmospheric parameters to be specified and are often computationally expensive. Stochastic cloud models are based on satellite, aircraft, or ground measurements and generate cloud fields close to the observed ones (Evans and Wiscombe 2004; Hogan and Kew 2005; Venema et al. 2006; Prigarin and Marshak 2009) and are computationally inexpensive. Some stochastic models generate cloud fields with an internal cloud structure (Schertzer and Lovejoy 1987; Cahalan 1994; Marshak et al. 1994; Schmidt et al. 2007), while others generate cloud fields as a binary mixture of cloudy and cloud-free areas (Su and Pomraning 1994; Zuev and Titov 1995; Prigarin et al. 2002). The latter have been used for stochastic radiative transfer modeling, which provides a radiation field averaged over many realizations of a stochastic cloud model (Sánchez et al. 1994; Titov 1990; Malvagi et al. 1993; Kassianov 2003; Lane-Veron and Somerville 2004).

Unlike a completely deterministic model of a cloud field (Aida 1977; Kite 1987; Breon 1992), which consists of identical clouds with centers on a regular grid, in a stochastic model the most straightforward assumption (approximation) of the distribution of cloudy and clear areas in a binary mixture is a homogeneous Markovian distribution. For this distribution, statistics of the medium are everywhere the same; thus, the probability of moving from cloudy areas to clear (or the reverse) along a line is independent of the position on the line (e.g., Byrne 2005). This assumption leads to an exponential distribution of cloud chord (and cloud separations) lengths, though not necessarily with the same characteristic lengths. For example, in the stochastic model of Su and Pomraning (1994) nonoverlapping clouds of fixed length randomly placed in an infinite 1D atmosphere exhibited exponential intercloud spacing distribution. Other models (Randall and Huffman 1980; Ellingson 1982; Ramirez and Bras 1990) allow for a variety of cloud sizes, dimensions, and altitudes.

We introduce a cloud-field model based on an alternative to the above described statistical models. Instead of randomly creating and then randomly placing a cloud into the model “atmosphere” (e.g., Su and Pomraning 1994), we consider the latter as a lattice with cells (of size  $l$ ) that may either be occupied by a cloud with probability  $p$  or be clear with probability  $q = 1 - p$ . Lattice models like this are widely used in percolation theory (cf. Isichenko 1992). Such an approach was used for cloud

modeling by Nagel and Raschke (1992), but in a different form than that presented here. The advantage of a lattice model is in parameterizing the cloud field as a whole, considering clouds and gaps between them together, and allowing the relation of such quantities as cloud cover (or fraction) and distributions of clouds and gaps by size, which appear at first glance to be independent. We further develop our model by allowing clouds to take arbitrary size and shape (instead of being clusters of several square cells) while retaining the basic statistical properties of the discrete-cell prototype (i.e., the probability of a cell of size  $l$  being occupied by cloud still equals  $p$ ). Computations in the framework of the continuous model were performed in 1D case; however, their results are also applicable to the statistics of 2D cloud fields (which need to be reasonably statistically isotropic) sampled along 1D transects.

Physically, the cell model is justified by the observed structure of cumulus [e.g., Wielicki and Welch (1986)] and stratocumulus (Nicholls 1989) cloud fields. In particular, Wielicki and Welch (1986) report that the fair weather cumulus clouds (even as small as 1 km in diameter) observed in their Landsat-based study were often multicelled, with cell sizes ranging from 250 m to 1 km. The aircraft-based observations by Nicholls (1989) show the chord distribution of convective cells in stratocumulus cloud fields to be consistent with the hexagonal cell structure having average cell diameter ranging from 200 to 600 m (i.e., being approximately half of the boundary layer depth). A convective cell can be filled (completely or partially) with cloud or remain cloud-free depending on atmospheric conditions. While we do not pursue their physical nature in this study, the cells in our model can be considered as abstractions of the atmospheric convective cells. We will refer to them as cellular statistical model (CSM) cells.

In section 2 we describe the basic discrete cellular cloud model. The statistics of its continuous extension are presented in section 3. Section 4 describes the construction of numerical examples in 1D and 2D cases. The relationship of our construction to the Ising model is explained in section 5. We conclude with discussion and a summary in sections 6 and 7.

## 2. Discrete cloud field model with fixed occupancy probability

We start with presentation of the basic discrete-cell model, and then construct its continuous extension, which is more capable of describing realistic cloud fields.

### a. Infinite lattice

Let us consider first an infinite 1D lattice, each cell of which can be occupied by cloud with probability

$p$  ( $0 < p < 1$ ) or remain empty with probability  $q = 1 - p$ . We define a cloud of the size  $m$  as a sequence of  $m$  neighboring cloudy cells separated from other clouds by at least one clear cell at each end of the cloud. The statistical distribution of clouds by size is geometric:

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

where  $m = 1, 2, 3, \dots$ . Indeed, the probability of finding a cloud of size  $m$  is  $q^2p^m$  and that of finding any cloud at all is  $qp$ . Thus, the conditional probability of finding one, given that there is cloud there, is  $qp^{m-1}$ . Note, that the smallest cloud size is 1, even if  $p \rightarrow 0$ . This distribution can be also derived from Bernoulli's scheme of independent trials starting at a cloudy cell and repeated until the first failure (clear cell). It satisfies the normalization condition

$$\sum_{m=1}^{\infty} n_c(m) = 1 \tag{2}$$

and has the mean

$$\bar{m}_c = \frac{1}{q}. \tag{3}$$

The statistical distribution of clear intervals (gaps) between clouds in our model can be obtained from Eq. (1) by interchanging  $p$  and  $q$ :

$$n_g(m) = pq^{m-1}. \tag{4}$$

It is normalized similarly to Eq. (2) and leads to the mean gap size

$$\bar{m}_g = \frac{1}{p}. \tag{5}$$

*b. Finite-length samples*

In the real-world observations, cloud field samples have finite size. As we will show below, the observed cloud length statistics are influenced by the sample size, especially for large clouds. Also, we need to consider overcast clouds (i.e., those with the length equal to or exceeding the sample size), which do not exist on the infinite lattice.

In the case of finite sample consisting of  $N$  cells, the cloud size distribution takes the form (derived in appendix A)

$$n_c(m) = \left[ 1 + \frac{1 - qm}{1 + q(N - 1)} \right] qp^{m-1} \tag{6}$$

for  $m < N$ , while

$$n_c(N) = \frac{p^{N-1}}{1 + q(N - 1)} \tag{7}$$

for  $m = N$ . By construction,  $n_c$  is normalized by

$$\sum_{m=1}^N n_c(m) = 1 \tag{8}$$

and has the mean

$$\bar{m}_c = \left[ q + \frac{p}{N} \right]^{-1}. \tag{9}$$

In the limit  $N \rightarrow \infty$  ( $N \gg m$ ), the distribution (6) becomes geometric (1), while  $n_c(N)$  vanishes and the mean cloud size takes the form (3). In the opposite limit case of  $N = 1$ , the mean cloud size (9) is expectedly equal to 1, and  $n_c$  has only one value  $n_c(N) = 1$ . In the limiting case of  $p = 0$  ( $q = 1$ , no chance of a cloud),  $\bar{m}_c = 1$ , which is the minimal possible size in the discrete model. Indeed, in the discrete model as  $p$  tends to 0, it becomes increasingly difficult to find a cloud, but if we succeed in finding one, it cannot be smaller than the cell size. In the overcast case  $p = 1$  ( $q = 0$ , no chance of clearing),  $\bar{m}_c = N$ , which is the full length of the sample.

The size distribution for gaps between clouds can be obtained from Eqs. (6) and (7) by replacing  $p$  with  $q$ :

$$n_g(m) = \left[ 1 + \frac{1 - pm}{1 + p(N - 1)} \right] pq^{m-1} \tag{10}$$

for  $m < N$  and

$$n_g(N) = \frac{q^{N-1}}{1 + p(N - 1)} \tag{11}$$

for  $m = N$ . This distribution is normalized by the same condition as  $n_c$  [Eq. (8)] and has the mean

$$\bar{m}_g = \left[ p + \frac{q}{N} \right]^{-1}. \tag{12}$$

*c. Cloud cover*

Another important characteristic of a cloud field is the cloud cover (also referred to as cloud fraction),

$$c = \frac{k}{N}, \tag{13}$$

where  $k$  is the number of all cloudy cells (not necessarily consecutive) in a sample of size  $N$ . As follows from combinatorics, the probability of having  $k$  occupied cells in a sample obeys the binomial distribution

$$n_{\text{cf}}(k) = \binom{N}{k} p^k q^{N-k}, \quad k = 0, 1, \dots, N, \quad (14) \quad n_c(m_s) = (1 - p_s) p_s^{m_s - 1} = (1 - p^{1/M}) p^{(m_s - 1)/M}. \quad (22)$$

where

$$\binom{N}{k} = \frac{N!}{k!(N-k)!} \quad (15)$$

is the binomial coefficient. The distribution (14) is normalized by

$$\sum_{k=0}^N n_{\text{cf}}(k) = 1 \quad (16)$$

and has the mean

$$\bar{k} = Np, \quad (17)$$

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

### 3. Continuous limit

As mentioned above, in reality atmospheric cells are not necessarily completely full with cloud or completely empty. They can be occupied partially, which means that the observed cloud/gap size distributions are continuous rather than discrete. To reflect this property in our model, we make the transition from the discrete to continuous cloud length distribution. To do this, let us divide each cell of size  $l$  into  $M$  subcells of the size

$$l_s = \frac{l}{M}. \quad (18)$$

Then, to keep the occupation probability of a whole cell equal to  $p$  we need to assume that the occupation probability of a subcell is

$$p_s = p^{1/M}, \quad \text{so that} \quad p_s^M = p. \quad (19)$$

We define the continuous cloud length  $x$  corresponding to  $m_s$  subcells as

$$x = m_s l_s = \frac{m_s}{M} l, \quad (20)$$

and the corresponding increment  $dx$  for the continuous probability density is

$$dx = l_s = \frac{l}{M}. \quad (21)$$

#### a. Infinite lattice

The geometric cloud length distribution (1) expressed in terms of subcell number is

To obtain the continuous probability density we need to take the limit  $M, m_s \rightarrow \infty$ , while preserving the value of  $x$ . Taking into account that  $e^y \approx (1 + y)$  when  $y = 1/M \ll 1$ , we find that

$$1 - p^{1/M} = 1 - e^{(\ln p)/M} \xrightarrow{M \rightarrow \infty} -\frac{\ln p}{M}, \quad (23)$$

and Eq. (22) takes the form

$$n_c(x) = -\frac{\ln p}{M} p^{x/l} = \frac{-\ln p}{l} e^{-[x(\ln p)/l]} \frac{l}{M}, \quad (24)$$

which can be written in terms of exponential distribution density  $f_c(x)$ :

$$n_c(x) = f_c(x) dx, \quad (25)$$

where

$$f_c(x) = \frac{1}{L_c} e^{-x/L_c} \quad (26)$$

with the cloud mean length

$$L_c = -\frac{l}{\ln p}. \quad (27)$$

Similar computations for the gap distribution (4) lead to the gap distribution density

$$f_g(x) = \frac{1}{L_g} e^{-x/L_g} \quad (28)$$

with the gap mean length

$$L_g = -\frac{l}{\ln q}. \quad (29)$$

It follows from Eqs. (27) and (29) that  $p$  and  $l$  are uniquely determined by the observed values of  $L_c$  and  $L_g$ . Indeed,  $p$  can be found as a (numerical) solution of

$$\frac{\ln p}{\ln(1-p)} = \frac{L_g}{L_c}, \quad (30)$$

and then  $l$  can be determined as

$$l = -L_c \ln p. \quad (31)$$

Thus,  $L_c$  and  $L_g$  provide an alternative (but equivalent) parameterization of the model, and in some approaches

(e.g., Astin and Di Girolamo 1999) the use of  $p$  and  $l$ , as well as the whole “cellular” concept, is avoided.

*b. Finite-length samples*

Computations similar to those presented above can be performed in the case of a 1D sample of a finite length:

$$L = Nl = NML_s. \tag{32}$$

In this case Eq. (6) transforms into the probability density

$$f_c(x) = \frac{1}{L_c} \left[ 1 + \frac{1 - (x/L_c)}{1 + a_c} \right] e^{-x/L_c} \tag{33}$$

for  $x < L$ , while the overcast contribution does not vanish with  $dx \rightarrow 0$ , becoming

$$f_c^o = \frac{e^{-a_c}}{1 + a_c}. \tag{34}$$

Here we used the following notations:

$$a_c = \frac{L}{L_c} = -\ln(p^N) \quad \text{and} \quad a_g = \frac{L}{L_g} = -\ln(q^N). \tag{35}$$

The fact that the overcast contribution does not scale with  $dx$  means that the complete length distribution of clouds has a singularity at  $x = L$ :

$$F_c(x) = f_c(x) + f_c^o \delta(x - L), \quad x \in [0, L]. \tag{36}$$

It is easy to show (see appendix B) that  $F_c$ , rather than  $f_c$ , is normalized:

$$\int_0^L F_c(x) dx = 1. \tag{37}$$

The mean cloud length for the distribution (36) is (see appendix B)

$$\bar{x}_c = \frac{LL_c}{L + L_c}, \tag{38}$$

which implies that

$$\frac{1}{\bar{x}_c} = \frac{1}{L_c} + \frac{1}{L}. \tag{39}$$

It follows from this formula that  $\bar{x}_c \rightarrow L_c$  as  $L \rightarrow \infty$ , while for finite  $L$  we always have  $\bar{x}_c < L_c$ ; in particular, for a very small sample ( $L \ll L_c, a_c \ll 1$ ) we expect  $\bar{x}_c \approx L$ .

Technically, small  $l/M = dx$  in Eq. (24) comes from the factor  $(1 - p_s) \ll 1$  in Eq. (22), which is absent in the

expression for overcast probability; however, the singularity of the cloud length distribution can also be explained by the independence of all-clear and overcast statistics on the size of the histogram bin. Note also that in the case of a single cell box ( $N = 1, L = l$ ) the value of  $f_c^o = p/(1 - \ln p)$  is not equal to 1 (unless  $p = 1$ ), as it would be in the discrete case. In fact, it is always less than 1, allowing for a continuous population  $f_c$  of subcell clouds.

The corresponding gap distribution statistics can be obtained from the above formulas by replacing  $L_c$  and  $a_c$  with  $L_g$  and  $a_g$ .

*c. Cloud cover*

A continuous analog of the binomial cloud cover distribution (14) can be obtained in a similar way; however, the computations in this case are substantially more complicated (see appendixes C and D). The resulting singular density

$$F_{cf}(c) = f_{cf} + f_{cf}^{(0)} \delta(c) + f_{cf}^{(1)} \delta(1 - c), \tag{40}$$

obeying

$$\int_0^1 F_{cf}(c) dc = 1, \tag{41}$$

has the following components: the continuous density

$$f_{cf}(c) = \frac{2a_c a_g}{a_c + a_g} e^{-[a_c c + a_g(1-c)]} \times \left\{ I_0(Z) + [a_g c + a_c(1 - c)] \frac{I_1(Z)}{Z} \right\}, \tag{42}$$

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

$$Z = 2[a_c a_g c(1 - c)]^{1/2}; \tag{43}$$

and the all-clear sky and overcast probabilities

$$f_{cf}^{(0)} = \frac{a_c}{a_c + a_g} e^{-a_g} \quad \text{and} \tag{44}$$

$$f_{cf}^{(1)} = \frac{a_g}{a_c + a_g} e^{-a_c}. \tag{45}$$

Note that the latter probabilities are fractions of the corresponding probabilities in the discrete model, respectively

$$q^N = e^{-a_g} \quad \text{and} \quad p^N = e^{-a_c}. \tag{46}$$

This fact indicates that fractional cloud cover is more probable in a continuous model than a discrete one. In particular, when  $N = 1$  the all-clear and overcast probabilities in discrete model add up to unity, while in the continuous case they allow for subcell clouds.

The mean cloud cover for the distribution (40) has the following form (see appendix D), which appears to be independent of the sample size  $L$ :

$$\bar{c} = \frac{a_g}{a_c + a_g} = \frac{L_c}{L_c + L_g}. \quad (47)$$

For an infinite sample this formula is quite intuitive (the mean cloud length divided by the sum of mean lengths of cloud and gap) and has been published in the literature (Sánchez et al. 1994; Astin and Di Girolamo 1999; Byrne 2005); however, it is far from obvious that this relation holds in a finite sample case. Equation (47) can be also be written in the form

$$\bar{c} = \frac{\ln(1-p)}{\ln p + \ln(1-p)}, \quad (48)$$

which can be used to derive  $p$  from the observed cloud fraction. The plot of  $\bar{c}$  versus  $p$  is shown in Fig. 1. We see that  $\bar{c} \neq p$ , unless  $p = 1/2, 0$ , or  $1$ : for  $p < 1/2$  cloud cover is always smaller than  $p$ , and vice versa for  $p > 1/2$ .

The distribution (C77) with the density (C78) and end components (C80) [leading to Eqs. (40), (42), (44), and (45)] has been also derived by Astin and Di Girolamo (1999) using an approach inspired by queue theory (Takacs 1957; Conolly 1971); however, they did not compute the mean in the finite sample case, and thus the initial probabilities  $u$  and  $v = 1 - u$  remained undetermined. It also was not clearly spelled out in that study that the end values (C80) are the coefficients at the corresponding  $\delta$  functions [as in Eq. (40)].

When  $L$  is considerably larger than the mean sizes of clouds and gaps ( $L \gg L_c, L_g$ ), the width of the distribution density (42) becomes very narrow, while both completely clear sky and overcast cases become improbable. Thus, in the limit case of an infinite sample ( $L = \infty$ , assuming that  $p \neq 0$  and  $p \neq 1$ , so that  $a_c, a_g \rightarrow \infty$ ) we have

$$F_{\text{cf}}(c) \equiv f_{\text{cf}}(c) = \delta(c - \bar{c}), \quad (49)$$

and

$$f_{\text{cf}}^{(0)} = f_{\text{cf}}^{(1)} = 0 \quad (50)$$

(see appendix D for details).

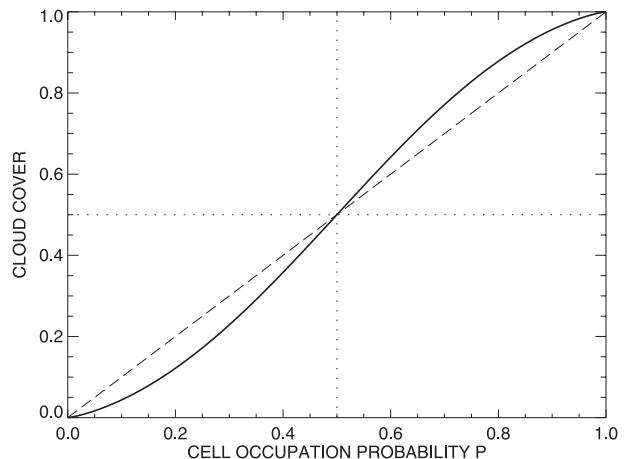


FIG. 1. Plot of  $\bar{c}$  in a continuous model vs  $p$  from Eq. (48) (solid curve). The identity  $\bar{c} = p$  is shown by the dashed line; the dotted lines correspond to  $p = 0.5$  and  $\bar{c} = 0.5$ .

#### 4. Construction of examples

Generation of an example of a 1D or 2D cloud field according to a discrete model is quite straightforward since the cells are filled independently one from each other. Thus, we do not show 1D examples from the discrete model, while a 2D example is presented (Fig. 8, top left). In a continuous model the situation is different since the subcells are no longer independent; that is, the filling probability of a subcell is conditioned on the clear/cloudy statuses of its neighbors.

##### a. 1D case

In the 1D case we define the subcell filling probabilities

$$p_s = p^{1/M} \quad \text{and} \quad q_s = q^{1/M} \quad (51)$$

to preserve the original filling probabilities  $p$  and  $q$  for the whole cells. As  $M \rightarrow \infty$ ,  $p_s, q_s \rightarrow 1$ . Note, that  $p_s + q_s \neq 1$ , since in distinction from  $p$  and  $q$  in the original (discrete) cellular model,  $p_s$  and  $q_s$  are conditional probabilities:  $p_s$  ( $q_s$ ) is the probability of a subcell to be cloudy (clear) given that the previous subcell is cloudy (clear). Thus, to construct a realization of a continuous model (which is not strictly speaking “continuous,” since in such an example  $M$  is finite) we move through the sample lattice from left to right using the following rule: if the  $i$ th subcell is cloudy we use  $p_s$  for cloudy and  $1 - p_s$  for clear to find the status of the  $(i + 1)$ th subcell, and conversely, if the  $i$ th subcell is clear we use probability  $q_s$  for clear and  $1 - q_s$  for cloudy. Obviously, the switch between cloud and clear sky appears when the  $(i + 1)$ th subcell status differs from that of the  $i$ th subcell. The procedure continues until the end of the sample is reached. The probability of the first subcell to be cloudy should be

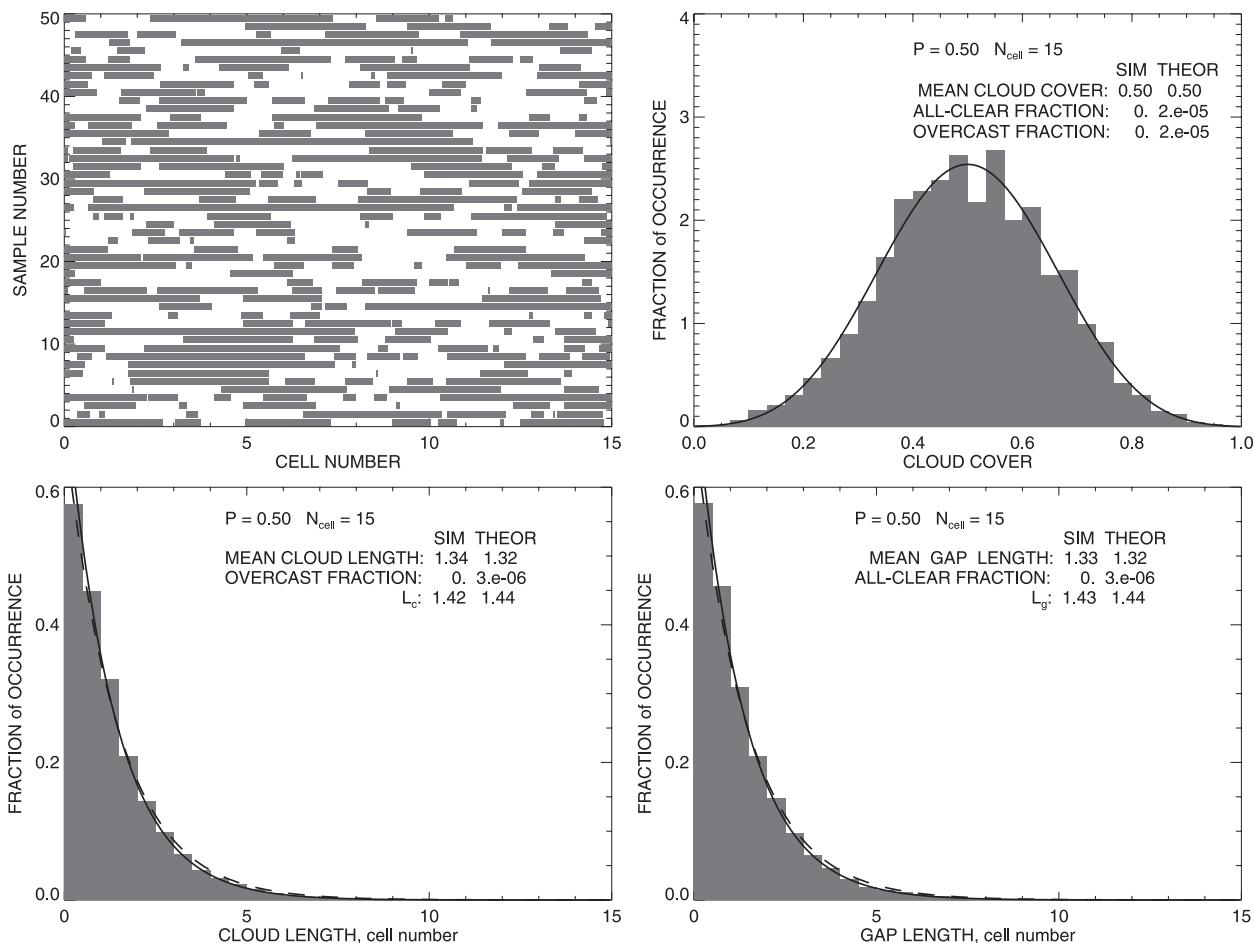
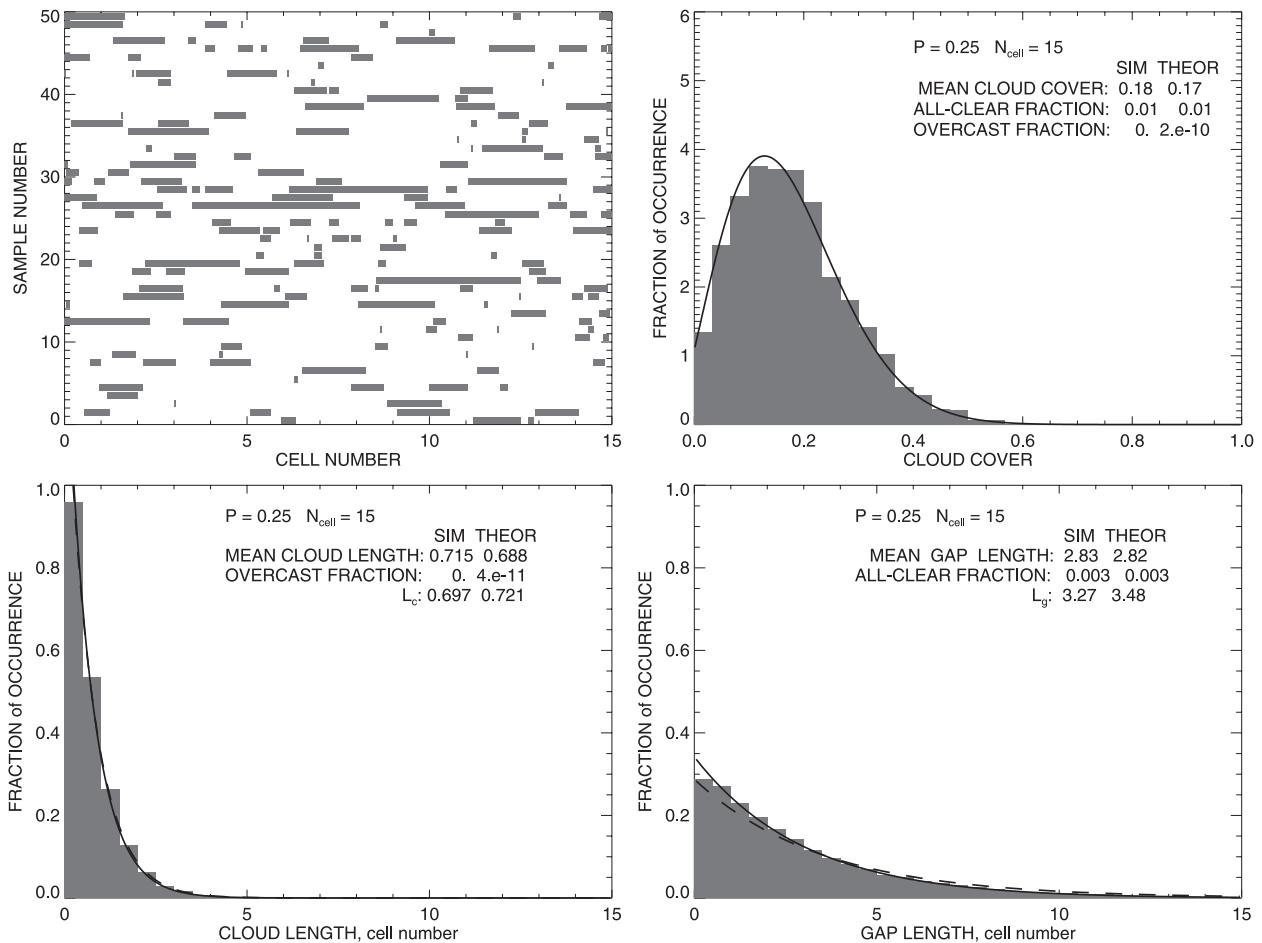


FIG. 2. Statistics of 1D cloud fields modeled according to our continuous cellular model with  $p = 0.5$  and  $L = 15l$  ( $M = 20$ ). (top left) The first 50 of 5000 samples used. (top right) Cloud cover histogram, in which the solid line is the theoretical probability density Eq. (42), the theoretical mean is computed using Eq. (47), and all-clear and overcast fractions are computed using Eqs. (44) and (45). (bottom left) Cloud length histogram, in which the solid line is the theoretical density Eq. (33) and the dashed line depicts the infinite-sample density Eq. (26). The theoretical overcast fraction is derived from Eq. (34). The theoretical mean cloud length is computed using Eq. (38), while its infinite-sample value  $L_c$  is determined from Eq. (27). The estimate of  $L_c$  from simulations is also presented. (bottom right) As in bottom left, but for gap lengths.

the same as of any other cell in the sample and thus equal to the mean cloud cover (47). The first subcell status can be selected directly according to this probability. Another way to do this (which was actually used to create the examples described below) is to create a longer sample (e.g., 3 times longer than required), with an arbitrary selection of the first subcell, and then take the last third of it as our example. This procedure is equivalent to the direct status selection, since Eq. (47) also gives the mean cloud cover of a long sample with an arbitrary first subcell status (the sample statistics “forget” that status after a certain length). We should mention that while the probability of transition from cloudy to clear is different from the probability of transition from clear to cloudy (if  $p \neq 1/2$ ), the chord statistics do not depend on the direction of filling the sample (Sánchez et al. 1994).

A set of examples generated according to this procedure is presented in Figs. 2–7. The plots are generated for  $p = 0.5, 0.25$ , and  $0.1$  and  $L = 15l$  (Figs. 2–4, subdivision number  $M = 20$ ) and  $L = l$  (Figs. 5–7, subdivision number  $M = 200$ ). Although, strictly speaking, one needs an infinite number of samples in order to use statistics for such a model, in practice we have to use a finite number of simulations. This number should be sufficiently large to closely approximate the “true” statistical distributions (represented by the theoretical curves in our case). We used 5000 samples in each of our simulations. The first 50 of them are shown in the top left panel of the corresponding figure, while the histograms in the other three panels are based on the whole 5000-sample datasets. The cloud cover histograms are in the top right panels with the solid lines depicting the theoretical probability

FIG. 3. As in Fig. 2, but for  $p = 0.25$ .

density (42). The theoretical mean cloud covers are computed using Eq. (47), while all-clear and overcast fractions are computed using Eqs. (44) and (45). The bottom left panels show cloud length histograms with the theoretical densities (33) depicted by solid lines, while the dashed lines correspond to the infinite-sample density (26). The difference between the finite and infinite-sample densities is better seen in the case of short sample (Figs. 5–7), where the former shows better agreement with the simulations. The theoretical all-clear and overcast fractions are derived from Eq. (34). Expectedly, for short samples these fractions are relatively large. For example, for  $p = 0.5$  they both are 25% in cloud cover for  $N = 1$  and virtually zero for  $N = 15$ . Thus, in the former case 50% of samples have fractional cloud cover (note that in the discrete model such samples are not allowed). The theoretical mean cloud lengths are computed using Eq. (38), while the values of  $L_c$  are determined from Eq. (27). The bottom right panels provide analogous descriptions of the gap length distributions. We see from these plots that the simulation histograms constructed using the “long sample” method

described above agree with the theoretical curves. This agreement indicates that the theoretically derived first subcell status probabilities are correct, and also that our simulations are not biased. Examples for  $p > 1/2$  (not shown) are similar to those shown: one just needs to interchange clouds and gaps (and, therefore,  $p$  and  $q$ ) and the cloud cover distribution histogram flips around the vertical axis at  $c = 0.5$ .

#### b. 2D case

In the 2D case, the discrete example of a cloud field (Fig. 8, top left) is a lattice with square cells of edge size  $l$  randomly filled with cloud (with probability  $p$ ). The cloud (chord) length statistics are the same as in 1D case, since each row or column can be considered as a 1D sample with the same  $l$  and  $p$ . Construction of subgrid examples in 2D is more complicated than in 1D because there is no clear direction for filling the sample. In the 2D case an unfilled subcell may have already filled neighbors with different statuses, making the decision of its own status ambiguous. While the correct way to fill

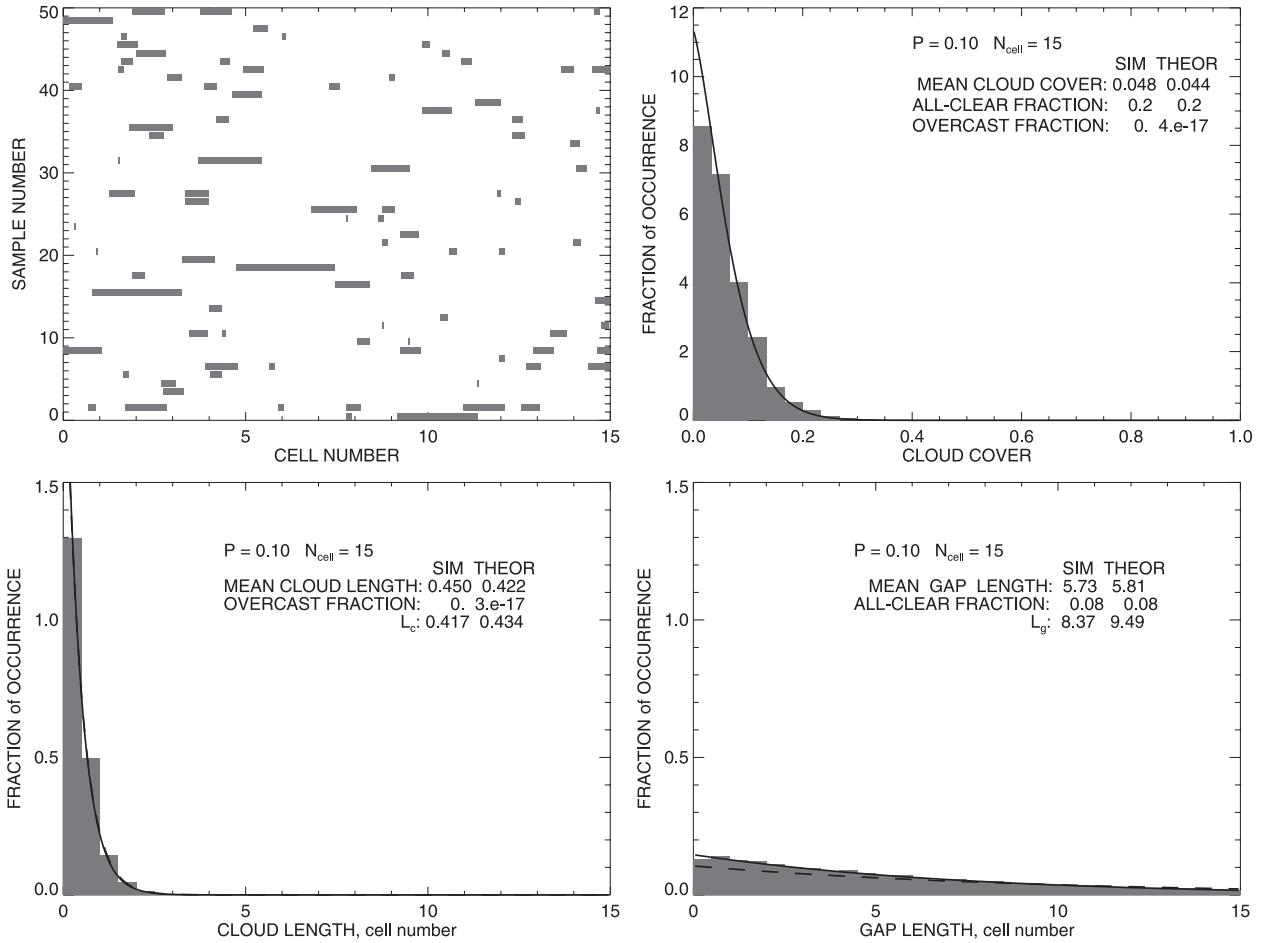


FIG. 4. As in Fig. 2, but for  $p = 0.1$ .

a 2D sample may well exist, we have not found it yet. Thus, instead of a filling procedure, we use an ad hoc iterative “mixing” algorithm to construct 2D broken cloud fields. This algorithm follows the general ideology of a continuous cell model; however, agreement of the statistics of its output with those prescribed by the model still remains to be checked (this requires extensive simulations). We start by creating a discrete model realization described above on the  $N \times N$  grid with spacing  $l$ . Then, we divide each cell into  $M^2$  subcells, creating the initial  $NM \times NM$  state for our iteration process. As in the 1D case, we define the subgrid filling probabilities

$$p_s = p^{1/M^2} \quad \text{and} \quad q_s = q^{1/M^2}. \quad (52)$$

To simplify the simulation we leave the subcells on the sample borders intact (more rigorously, after completing the simulation we should take a smaller sample, which does not contain border cells). Each subcell inside the sample has eight neighbors. If on the  $i$ th iteration step  $n$  of these neighbors are cloudy and  $(8 - n)$  are

clear, then the probability of our cell being cloudy on the next  $(i + 1)$ th step is

$$p_{i+1} = \frac{np_s + (8 - n)(1 - p_s)}{8}. \quad (53)$$

To suppress creation of an unphysical “fog” of small clouds (most of them consisting of a single subcell) and similar small holes inside clouds, we impose the following condition: if on the  $i$ th iteration step a subcell had seven or eight cloudy (clear) neighbors, it will be cloudy (clear) on the next step (with probability 1). If  $p < 1/2$  ( $p > 1/2$ ) this iteration procedure leads to monotonic decrease (increase) in cloud cover. Thus, to obtain an example consistent with our continuous model, the iterations should be stopped when the cloud cover comes close to the value prescribed by Eq. (48). To be precise, the ensemble average of cloud cover over a number of samples constructed in the same way (with the same number of iterations) should agree with the value from Eq. (48). However, if only one example needs to be

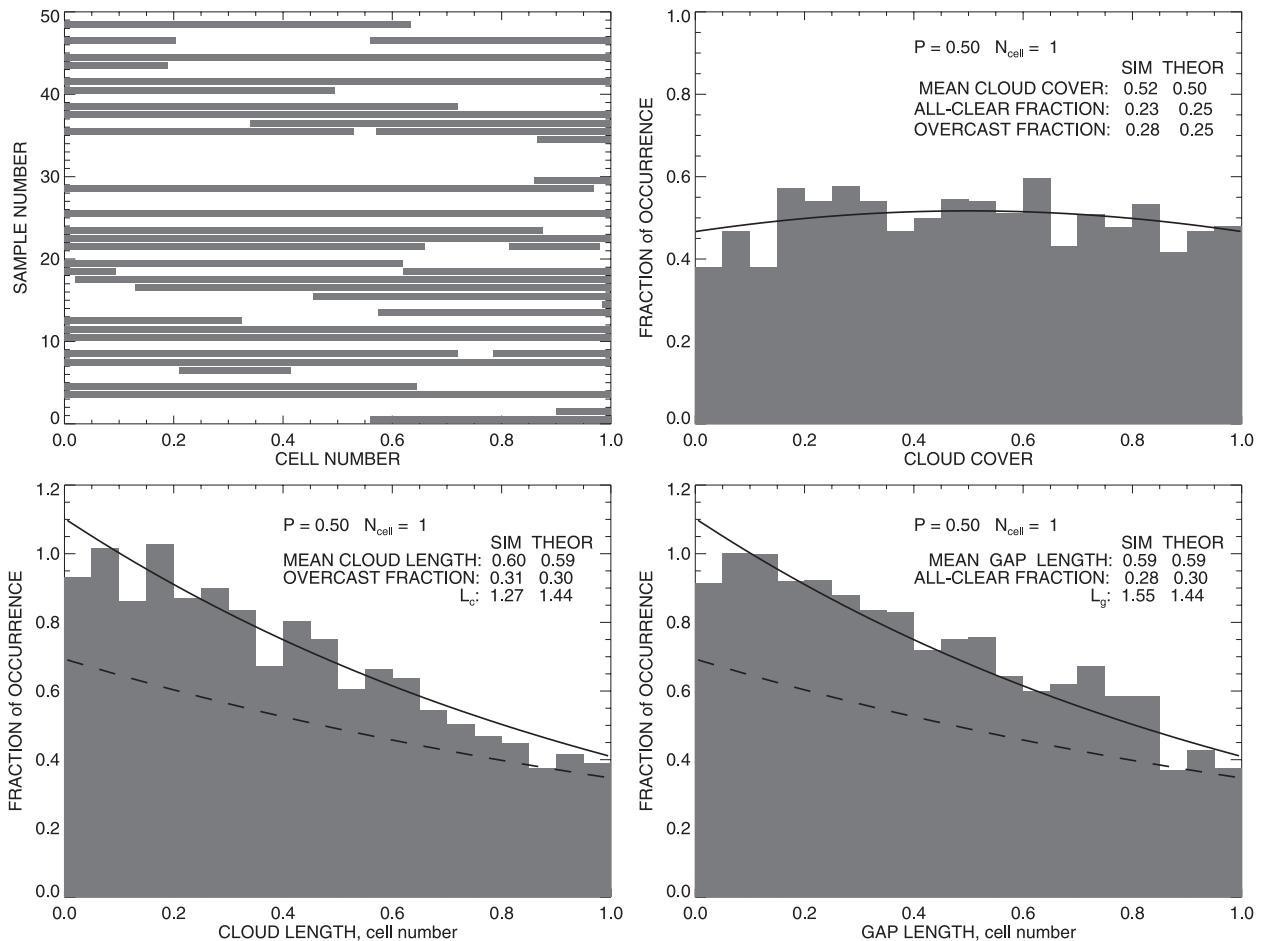


FIG. 5. As in Fig. 2, but for  $L = 1$  and  $M = 200$ .

constructed, reaching this value can be used as a signal to stop the iterative process.

An example of the above described construction of a 2D sample cloud field is shown in Fig. 8. For this sample we took  $p = 0.24$  and a  $15 \times 15$ -cell grid. Each of the cells was divided into  $M \times M$  subcells with  $M = 20$ . The cloud cover 0.276 of the initial discrete-model simulation (top left) is close to  $p$ , while during the mixing process it drops to 0.16 after 180 iterations (bottom right). We stop the iterations at this point, since this cloud cover value corresponds to our  $p$  according to Eq. (48). We see that the resemblance to the initial state gradually decreases during the mixing iterations. The final cloud field looks more realistic than the initial sample, since the clouds now have more detailed structure rather than being clusters of a few square cells.

## 5. Relationship to the Ising model

The 2D Ising model is widely studied both analytically and numerically with applications in phase transitions

and critical phenomena. The classical approach to Monte Carlo simulations of this and other statistical lattice models is the Metropolis “importance sampling” algorithm (Metropolis et al. 1953). In this classical approach the states of the sample are generated recursively one from the other with some transition probability. The difference between the neighboring states is small: usually a single spin flip at a random location. Recently more computationally efficient methods have been developed—for example, cluster flip algorithms (Swendsen and Wang 1987; Wolff 1989), in which large regions of spins are flipped instead of single spins.

Our above described method for construction of a “realistic” 2D broken cloud field, while showing some resemblance to Ising model simulations, is different from the Metropolis algorithm and its relatives. While our method is also an iterative procedure, only its final state is considered as a representation of the continuous cloud model. Another representation should be generated independently. Our method also cannot be used for simulation of phase transitions, since we directly prescribe on

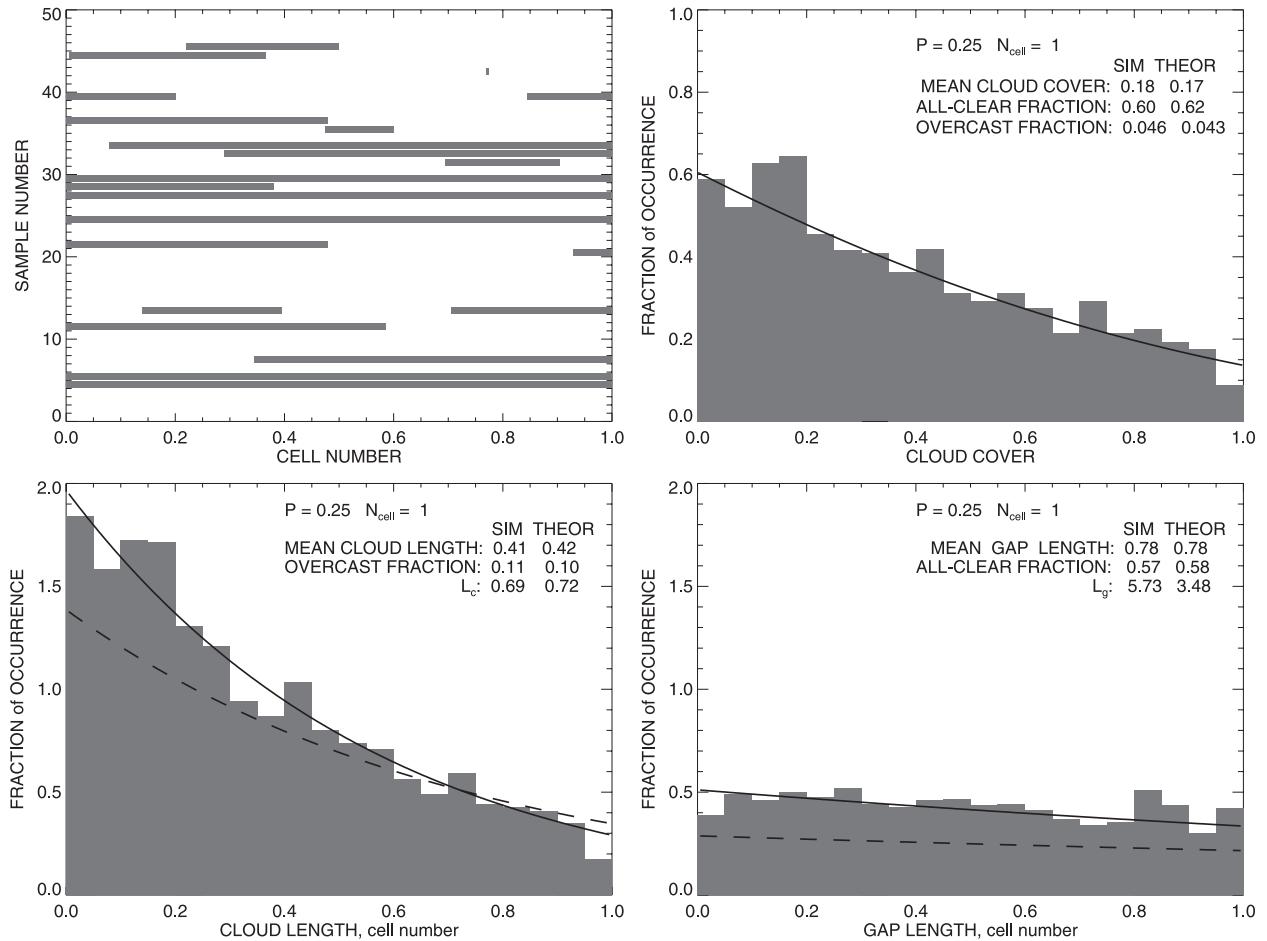


FIG. 6. As in Fig. 5, but for  $p = 0.25$ .

the first step the positions of what would later become “domains” rather than letting them form naturally during a simulation.

However, there is certainly similarity between our continuous model and the Ising model in the 1D case, based on interpretation of both these models in terms of a Markov process. Indeed, a realization of our continuous model constructed according the procedure described in section 4a is a binary Markov chain with the transition matrix

$$\mathbf{P} = \begin{pmatrix} P_{++} & P_{+-} \\ P_{-+} & P_{--} \end{pmatrix} = \begin{pmatrix} p_s & 1 - p_s \\ 1 - q_s & q_s \end{pmatrix}, \quad (54)$$

where  $P_{ij}$  is the probability of transition from state  $i$  into state  $j$ . Both  $i$  and  $j$  can take two values: a plus sign = +1 (cloudy) or a negative sign = -1 (clear). Each row of  $\mathbf{P}$  sums to unity.

A state of the 1D Ising model (cf. Baxter 1982) is described by values of a sequence  $\sigma = \{\sigma_1, \sigma_2, \dots, \sigma_n\}$  of

$n$  random variables (spins)  $\sigma_i$ , each of which can take values of  $\pm 1$  (indicative of cloudy or clear subcell in our model). The probability of the system being in a particular state  $\sigma$  is

$$P(\sigma) = Z^{-1} e^{-E(\sigma)/kT}, \quad (55)$$

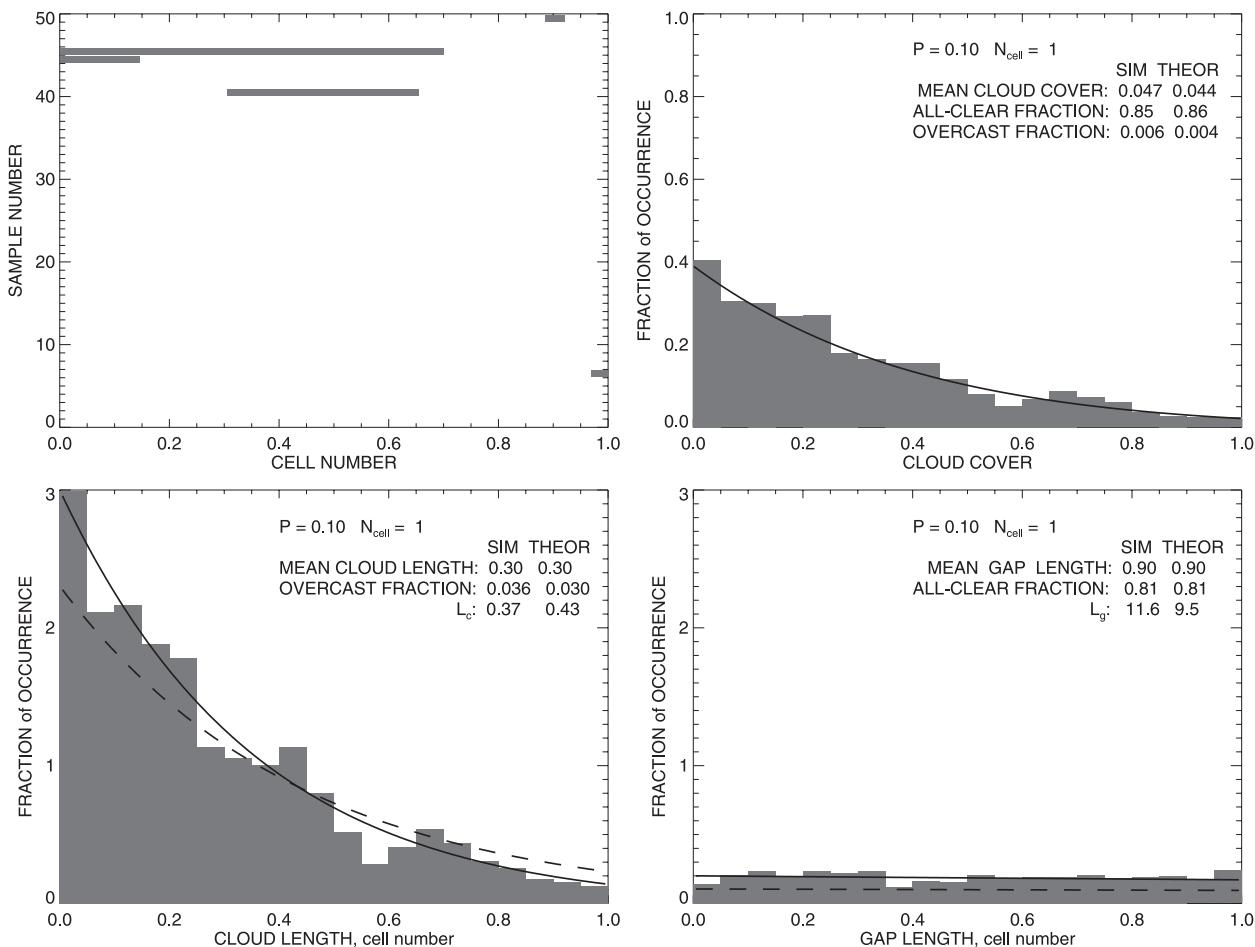
where  $E(\sigma)$  is the state’s energy,

$$E(\sigma) = -J \sum_{j=1}^n \sigma_j \sigma_{j+1} - H \sum_{j=1}^n \sigma_j, \quad (56)$$

and  $Z$  is the partition function,

$$Z = \sum_{\sigma} e^{-E(\sigma)/kT}, \quad (57)$$

where summation is taken over all possible states, and  $T$  is the system’s temperature,  $k$  is Boltzmann’s constant,  $H$  is the magnetic field, and  $J$  is the coupling constant.

FIG. 7. As in Fig. 5, but for  $p = 0.1$ .

Here we do not discuss boundary conditions, which may alter the summation in Eq. (56). It can be shown (cf. Baxter 1982) that  $P(\sigma)$  can be represented as a product of probabilities involving only neighboring spins, which have the form

$$P_j = \exp\left[K\sigma_j\sigma_{j+1} + \frac{h}{2}(\sigma_j + \sigma_{j+1}) + \frac{a}{2}(\sigma_j - \sigma_{j+1}) + b\right].$$

Here

$$K = \frac{J}{kT} \quad \text{and} \quad h = \frac{H}{kT}, \quad (58)$$

and  $a$  and  $b$  are constants necessary for consistency with the definition of transition probability. Indeed,  $P_j$  can be considered as the probability of transition from the state of the  $j$ th spin to a state of its right neighbor. This means that the values of  $P_j$  form the transition matrix of a binary Markov process, which has the form Eq. (54):

$$\mathbf{P}' = \begin{pmatrix} e^{K+h+b} & e^{-K+a+b} \\ e^{-K-a+b} & e^{K-h+b} \end{pmatrix}. \quad (59)$$

The relationship between Ising model parameters and  $p_s$  and  $q_s$  (or equivalently,  $p$  and  $M$ ) can be determined from the equality of  $\mathbf{P}$  and  $\mathbf{P}'$ :

$$K = \frac{1}{4} \ln\left[\frac{p_s q_s}{(1-p_s)(1-q_s)}\right] \quad \text{and} \quad (60)$$

$$h = \frac{1}{2} \ln\left(\frac{p_s}{q_s}\right), \quad (61)$$

while

$$a = \frac{1}{2} \ln\left(\frac{1-p_s}{1-q_s}\right) \quad \text{and} \quad (62)$$

$$b = \frac{1}{4} \ln[p_s q_s (1-p_s)(1-q_s)]. \quad (63)$$

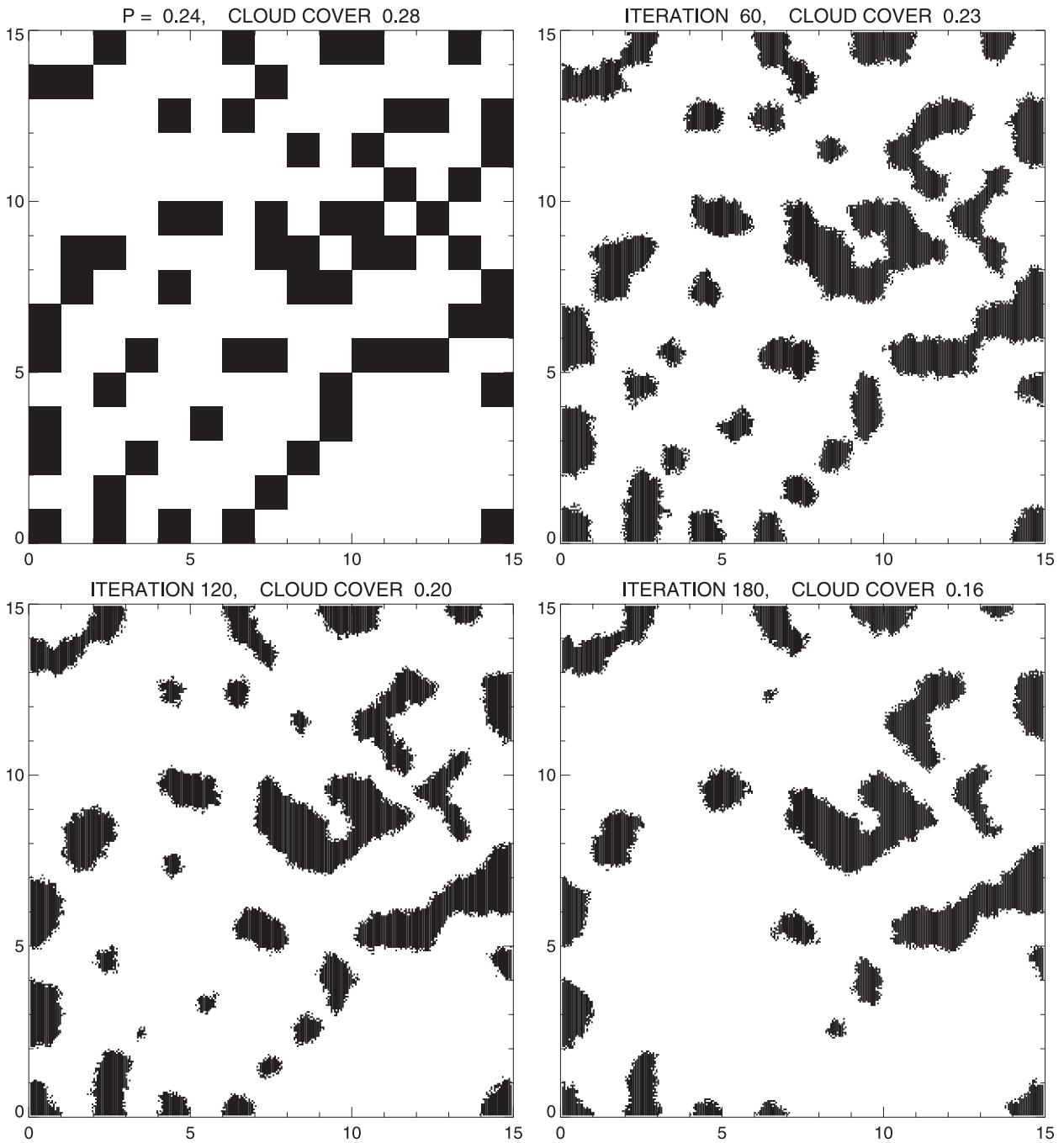


FIG. 8. Construction of 2D cloud field example using the “mixing” process described in section 4b. For this sample,  $p = 0.24$  and a  $15 \times 15$  cell grid is taken with a  $20 \times 20$  cell subdivision. (top left) The initial discrete-model simulation (the cloud cover 0.28 is close to  $p$ ). Cloud field at the iteration steps (top right) 60, (bottom left) 120, and (bottom right) 180. The cloud cover value in the last step (0.16) is equal to that derived from Eq. (48) for  $p$ .

Note that in the symmetric case when  $p = q = 1/2$  both  $h$  and  $a$  vanish. Recalling the relationship in Eq. (51) of  $p_s$  and  $q_s$  to  $p$  and  $M$ , we can see that in the continuous limit as  $M \rightarrow \infty$  the coupling constant  $K$  grows logarithmically with  $M$ , while  $h$  decreases as  $M^{-1}$ .

A relationship, similar to that described above, between the parameters of our cellular cloud model and those of the Ising model is likely to hold in the 2D case. This opens the possibility of considering an extensive arsenal of simulation methods and analytical techniques

developed for the 2D Ising model for construction of broken cloud field samples and to study their statistical properties, which we leave for our future studies.

## 6. Discussion

We next discuss the differences between our approach and that described in the series of papers by Astin et al. (2001), van de Poll et al. (2006), and Settle and van de Poll (2007), which are also devoted to the subject of cloud transect statistics. In these studies the authors solve an inverse statistical problem: they use a fixed observed sample (or a relatively small set of samples) and use Bayesian statistics to estimate the underlying cloud model parameters, such as the cloud and gap exponents (equivalent to  $L_c$  and  $L_g$  in our terminology) and the true (infinite space) cloud fraction  $\bar{c}_\infty = L_c/(L_c + L_g)$ . This estimation results in statistical distributions of the model parameters. These estimates, therefore, depend on the particular structure of the observed sample, such as the numbers of clouds and gaps, and when these numbers become large, the probability density of, say,  $\bar{c}_\infty$  becomes narrowly peaked. Entirely clear and overcast samples require special consideration in this approach. Conversely, our study is focused on a forward problem, which means that we start with fixed model parameters (such as  $L_c$  and  $L_g$ ) and study statistical properties of the ensemble of observations, which consists of all possible realizations of the model (of course, we have to use a limited but large subset of this ensemble in numerical simulations). Thus, in our approach we do not look for, for example, a distribution of  $\bar{c}_\infty$  conditioned by the observed sample (since  $\bar{c}_\infty$  is predetermined by the fixed model parameters); instead, we derive the distribution of the cloud cover values observed in the samples.

We also emphasize that in our computations of statistics for finite samples we do not imply that the sky outside the sample is clear (or cloudy). On the contrary, we assume that our finite sample is taken at random from a large, possibly infinite, statistically homogeneous sample. This means that the cloudy (clear) intervals at the ends of a sample may be (and most likely are) parts of clouds (gaps), rather than whole clouds (gaps) in that large sample. An overcast (entirely clear) condition appears when a sample is taken from within a cloud (gap) that is larger than its size. The partial clouds and gaps are treated in our approach on an equal footing with the whole ones. Note that the computations (presented in appendix A) of cloud and gap length distributions for a finite sample [Eqs. (6) and (7)] are made in the framework of the discrete model, where cells are independent; thus, the states of the cells outside the sample do not influence this distribution or its continuous limit [Eqs. (33)–(36)]. In the computation of the cloud cover distribution

[Eqs. (40)–(45)], the choice of initial probabilities is made based on the requirement of homogeneity of the sample (the probability that the first subcell is cloudy is the same as for any other subcell in the sample, i.e., equal to the mean cloud cover; see appendixes C and D for details). If we had instead postulated clear conditions outside the sample (i.e., if the subcell to the left of the first subcell in the sample was assumed to be always clear), the initial probability would be  $u = q_s$ . This probability is close to unity for a large enough  $M$ ; thus, the “clear outside” condition would create a bias toward clear between the beginning of the sample and the rest of it, thereby making it statistically inhomogeneous. The homogeneity condition appeared to be equivalent to the assumption that samples are picked from an infinite field, since it has been demonstrated that the mean cloud cover is the same in finite and infinite cases. Our analytical results are consistent with the statistics of our simulated examples, for the construction of which the samples were taken from the ends of sufficiently long records. Finally, all expressions derived in this study are invariant under simultaneous interchange of  $p \leftrightarrow q$  and cloud  $\leftrightarrow$  gap. This invariance contradicts any assumption of a fixed (e.g., clear) condition outside of the sample, since such a transformation would change this condition to its opposite.

## 7. Concluding remarks

We have introduced two types of cellular models providing statistical descriptions of broken cloud-field properties. Computations were performed in 1D to derive statistical distributions of clouds and gaps between them by length, as well as of the cloud cover. The resulting formulas also describe cloud field properties measured along 1D transects in a 2D case. The first (discrete) model describes statistics of cloud fields on a discrete grid, each cell of which can be filled with cloud with a certain prescribed probability  $p$  (which does not depend on the state of the other cells in the sample). This basic model is too simplistic for adequate description of realistic cloud fields, since cloud-field parameters can take only discrete values within its framework. The second (continuous) model is an extension of the discrete model, and allows for continuous distributions of cloud/gap lengths and cloud fraction. We show in Part II of this paper (Alexandrov et al. 2010, hereafter Part II) that this model is in agreement with statistics of large-eddy cloud-field simulations. Note that 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, such as seen in Figs. 2–8. In both discrete and continuous models we assume that the observed cloud/gap statistics are determined from samples of finite size.

This means, in particular, that the samples within a cloud (gap), which is larger than the sample size, are interpreted as overcast (entirely clear) and are assigned cloud (gap) lengths equal to the sample size and the cloud fraction of 1 (0). In the continuous model such samples contribute to the singularities ( $\delta$  functions) in the length distributions (at the length equal to the sample size) and in the cloud cover distribution (at the fractions of 0 and 1).

If  $p$  does not change across the statistical ensemble, the cloud/gap lengths distributions in our continuous model are essentially exponential. The opposite “diverse ensemble” situation with power-law clouds and gaps is described in Part II of this paper. We should note that according to Eqs. (27) and (29) the set of cellular model parameters  $p$  and  $l$  is in a one-to-one correspondence with the set of the exponents ( $L_c^{-1}, L_g^{-1}$ ) of the cloud and gap length distributions. This means that in the 1D case some of our results, such as the cloud cover distribution [Eqs. (40)–(45)], can also be obtained using the approach of Astin and Di Girolamo (1999). In this approach the exponential cloud/gap lengths distributions are postulated, and the samples are constructed by sequential alternating drawing of clouds and gaps from their respective distributions (without invocation of a cellular structure). The situation, however, becomes quite different when we attempt to construct a cloud field in the 2D case, where we lack the concept of a filling direction necessary for a sequential method described above. We currently do not have a rigorous continuous cellular model in 2D case; however, our preliminary analysis (section 4b; Fig. 8) shows that realistic examples of cloud fields in this case can be constructed through an iterative process starting at a realization of the discrete model. The cellular structure is essential for such a procedure.

In Part II of this paper we present quantitative comparisons demonstrating that our model is generally in good agreement with the statistics of the cloud fields obtained using large-eddy simulations. We will also describe statistical properties of diverse cloud datasets obtained from observations with wide temporal and/or geographical coverage. Our future plans include development of a 2D analog of the continuous model described in this paper, as well as extension of the presented techniques from binary (cloud/clear) datasets to those with continuous values (such as LWP or cloud optical thickness). The latter will be achieved through consideration of multilayer binary models, where the statistics in each layer depend on the states of the layers below.

*Acknowledgments.* This research was supported in part by the NASA Glory Project. The authors thank E. Kassianov, A. Korolev, A. Del Genio, B. Cairns, and A. Lacis for useful discussions, and also A. S. Schwarz for

pointing our attention at the similarity between our continuous model and the 1D Ising model. We are also indebted to the reviewers of this paper for their detailed comments, which helped us to improve our work.

## APPENDIX A

### Discrete Cloud Length Distribution for Finite Samples

To derive the cloud size distribution for a 1D lattice with  $N$  cells, we compute statistical sums, to which cloud field realizations contribute with the weights equal to their probabilities. For each cloud length  $m < N$  we consider two cases: when the cloud under consideration is at the end of the observation interval

$$\bullet - \bullet - \dots - \bullet - \circ - \dots - \circ \quad (A1)$$

$\underbrace{\hspace{10em}}_{m+1 \text{ cells}} \quad \underbrace{\hspace{10em}}_{N-m-1 \text{ cells}}$

and the generic case when this cloud is inside the interval

$$\dots - \circ - \bullet - \bullet - \dots - \bullet - \circ - \dots \quad (A2)$$

$\underbrace{\hspace{10em}}_{i-2 \text{ cells}} \quad \underbrace{\hspace{10em}}_{i-1, i, \dots, i+m-1, i+m} \quad \underbrace{\hspace{10em}}_{N-i-m \text{ cells}}$

Here  $i = 2, \dots, N - m$  and the cells occupied by clouds are depicted by bullets, while the clear cells are shown by open circles. In the first case each realization contributes to the statistical sum with the weight  $p^m q \times$  weight of the last  $(N - m - 1)$  cells. The sum of the weights of the realizations of the last  $(N - m - 1)$  cells is 1, since we sum up the probabilities of all possible cases. Thus, the contribution of the end-interval case (A1) is just  $p^m q$ , and there are two such cases (at each end of the interval). Similarly, the contribution of a midinterval case (A2) is  $p^m q^2$ , and we have  $(N - m - 1)$  cases like this. Thus, the statistical sum for the cloud of the size  $m < N$  is

$$S_m = 2p^m q + p^m q^2(N - m - 1), \quad (A3)$$

while for the overcast cloud ( $m = N$ )

$$S_N = p^N. \quad (A4)$$

To find the normalization constant  $S$  in the cloud size distribution we sum up all  $S_m, m = 1, \dots, N$ :

$$\begin{aligned} S &= p^N + \sum_{m=1}^{N-1} [2p^m q + p^m q^2(N - m - 1)] \\ &= p^N + q[2 + q(N - 1)] \sum_{m=1}^{N-1} p^m - q^2 \sum_{m=1}^{N-1} m p^m \\ &= p[1 + q(N - 1)]. \end{aligned} \quad (A5)$$

Here we use the standard formula (cf. Gradshteyn and Ryzhik 1965) for finite geometric series and its derivative with respect to  $p$ . Thus, the cloud size distribution takes the form

$$n_c(m) = \left[ 1 + \frac{1 - qm}{1 + q(N-1)} \right] \frac{q}{p} p^m \quad (\text{A6})$$

for  $m < N$  and

$$n_c(N) = \frac{p^{N-1}}{1 + q(N-1)} \quad (\text{A7})$$

for  $m = N$ . By construction,  $n_c$  is normalized by the condition

$$\sum_{m=1}^N n_c(m) = 1. \quad (\text{A8})$$

To derive the mean cloud length (number of cells in the cloud)

$$\begin{aligned} \bar{m}_c &= \sum_{m=1}^N m n_c(m) = \sum_{m=0}^N m n_c(m) \\ &= q \left[ 1 + \frac{1}{1 + q(N-1)} \right] \sum_{m=0}^{N-1} m p^{m-1} \\ &\quad - \frac{q^2}{1 + q(N-1)} \sum_{m=0}^{N-1} m^2 p^{m-1} + \frac{Np^{N-1}}{1 + q(N-1)} \end{aligned} \quad (\text{A9})$$

we use the first and second derivatives (with respect to  $p$ ) of the finite geometric series formula, and, after some algebraic transformations, obtain the following simple expression for the cloud mean length:

$$\mathcal{N}_F = \frac{2 + a_c}{(1 + a_c)L_c} \int_0^L e^{-x/L_c} dx - \frac{1}{(1 + a_c)L_c} \int_0^L \frac{x}{L_c} e^{-x/L_c} dx + \frac{e^{-a_c}}{1 + a_c}, \quad (\text{B5})$$

which after the substitution  $t = x/L_c$  transforms into

$$\begin{aligned} \mathcal{N}_F &= \frac{2 + a_c}{1 + a_c} \int_0^{a_c} e^{-t} dt - \frac{1}{1 + a_c} \int_0^{a_c} t e^{-t} dt + \frac{e^{-a_c}}{1 + a_c} \\ &= \frac{2 + a_c}{1 + a_c} (1 - e^{-a_c}) - \frac{1}{1 + a_c} [1 - e^{-a_c}(1 + a_c)] + \frac{e^{-a_c}}{1 + a_c} \\ &= \frac{2 + a_c - 1}{1 + a_c} = 1. \end{aligned} \quad (\text{B6})$$

$$\bar{x} = \frac{2 + a_c}{(1 + a_c)L_c} \int_0^L x e^{-x/L_c} dx - \frac{1}{(1 + a_c)L_c} \int_0^L \frac{x^2}{L_c} e^{-x/L_c} dx + L \frac{e^{-a_c}}{1 + a_c}, \quad (\text{B7})$$

$$\bar{m}_c = \left[ q + \frac{p}{N} \right]^{-1}. \quad (\text{A10})$$

## APPENDIX B

### Cloud Length Statistics in the Continuous Model with Finite Samples

#### a. Normalization condition

To verify that the cloud length distribution (36)

$$F_c(x) = f_c(x) + f_c^o \delta(x - L), \quad x \in [0, L], \quad (\text{B1})$$

is properly normalized [according to Eq. (37)], we need to compute the following integral

$$\mathcal{N}_F = \int_0^L F_c(x) dx = \int_0^L f_c(x) dx + f_c^o, \quad (\text{B2})$$

where

$$\begin{aligned} f_c(x) &= \frac{1}{L_c} \left[ 1 + \frac{1 - (x/L_c)}{1 + a_c} \right] e^{-x/L_c} \\ &= \frac{1}{L_c} \left[ \frac{2 + a_c}{1 + a_c} - \frac{(x/L_c)}{1 + a_c} \right] e^{-x/L_c} \end{aligned} \quad (\text{B3})$$

is the probability density (33) for  $x < L$ , and

$$f_c^o = \frac{e^{-a_c}}{1 + a_c} \quad (\text{B4})$$

is the overcast ( $x = L$ ) contribution. Thus, the integral (B2) takes the form

Thus, our singular density (B1) is properly normalized.

#### b. Computation of the mean

Computation of the mean cloud length for the density (B1) is similar to that of the norm presented above. We need to take the following integral

which after the substitution  $t = x/L_c$  (and using  $L = a_c L_c$  in the last term) becomes

$$\begin{aligned} \bar{x} = & L_c \frac{2 + a_c}{1 + a_c} \int_0^{a_c} t e^{-t} dt - \frac{L_c}{1 + a_c} \int_0^{a_c} t^2 e^{-t} dt + L_c \frac{a_c}{1 + a_c} e^{-a_c} = L_c \frac{2 + a_c}{1 + a_c} [1 - e^{-a_c} (1 + a_c)] \\ & - \frac{L_c}{1 + a_c} \{2 - e^{-a_c} [(1 + a_c)^2 + 1]\} + L_c \frac{a_c}{1 + a_c} e^{-a_c} = \frac{L_c a_c}{1 + a_c} = \frac{L}{1 + (L/L_c)}. \end{aligned} \tag{B8}$$

The integrals involved in this calculation can be found in a standard table (e.g., Prudnikov et al. 1986). The above expression is equivalent to Eqs. (38) and (39).

APPENDIX C

**Cloud Cover Distribution in the Continuous Model**

To derive the statistical distribution of cloud fraction in continuous limit we use the approach outlined in section 3 [although the approach of Astin and Di Girolamo (1999) may be computationally simpler, we prefer to use the cellular model formalism in our derivations]. In the  $N$ -cell box we subdivide each cell into  $M$  subcells (we will take the limit  $M \rightarrow \infty$  to obtain the continuous case). Each cell has size  $l$ , so the subcell size is  $l/M$ . The total number of subcells and the total length of the sample are, respectively,

$$n = NM \quad \text{and} \quad L = Nl. \tag{C1}$$

We define the subcell filling probabilities  $p_s$  and  $q_s$  according to Eq. (51) (note that  $p_s, q_s \rightarrow 1$  as  $M \rightarrow \infty$ ). We remind the reader that to construct a realization of the model, we move through our sample from left to right using the following rule: if the  $i$ th subcell is cloudy we use a draw with probability  $p_s$  for cloudy and  $1 - p_s$  for clear to find the status of the  $(i + 1)$ th subcell, and conversely, if the  $i$ th subcell is clear we use a draw with probability  $q_s$  for clear and  $1 - q_s$  for cloudy. The procedure continues until the end of the box is reached. In general, the status of the first cell is selected by a draw with an initial probability  $u$  to be cloudy (we also define  $v = 1 - u$ ), which is a free parameter of the model influencing the structure of the sample. To determine this initial probability we impose the following homogeneity condition: the probability of the first cell to be cloudy is the same as of any other cell in the sample (thus, equal to the mean cloud cover).

*a. Computation of the statistical sum*

For each  $k$  cloudy subcells out of  $n$  subcells in the box ( $c = k/n$ ), we compute the statistical sum as expansion

over the number  $r = 1, 2, 3, \dots$  of ‘‘quasi-cells’’ defined as interchanging clear and cloudy intervals of arbitrary length. For example, for  $r = 1$  we have two situations: completely overcast

$$\bullet \bullet \bullet \bullet \dots \bullet \bullet \bullet = \bullet - \bullet, \tag{C2}$$

with the weight

$$w_1 = u p_s^{n-1} \xrightarrow{M \rightarrow \infty} u p^N = u e^{-L/L_c}, \tag{C3}$$

and a completely clear sample

$$\circ \circ \circ \circ \dots \circ \circ \circ = \circ - \circ, \tag{C4}$$

with the weight

$$w_0 = v q_s^{n-1} \xrightarrow{M \rightarrow \infty} v q^N = v e^{-L/L_c}. \tag{C5}$$

Note that these weights do not scale with  $M$  and therefore will contribute to the singular part of the probability density. For  $r = 2$  we also have two cases:

$$\bullet - \bullet \circ - \circ \quad \text{and} \quad \circ - \circ \bullet - \bullet, \tag{C6}$$

as well as for  $r = 3$ :

$$\bullet - \bullet \circ - \circ \bullet - \bullet \quad \text{and} \quad \circ - \circ \bullet - \bullet \circ - \circ. \tag{C7}$$

In general, for even  $r = 2i$  the two cases are

$$\bullet - \bullet \circ - \circ \dots \bullet - \bullet \circ - \circ \quad \text{and} \tag{C8}$$

$$\circ - \circ \bullet - \bullet \dots \circ - \circ \bullet - \bullet, \tag{C9}$$

while for odd  $r = 2i + 1$  they are

$$\bullet - \bullet \circ - \circ \dots \circ - \circ \bullet - \bullet \quad \text{and} \tag{C10}$$

$$\circ - \circ \bullet - \bullet \dots \bullet - \bullet \circ - \circ. \tag{C11}$$

To construct the statistical sum we assign weights  $p_s$  to all cloudy subcells, except for those where the switch from clear to cloudy happened (first subcells of cloudy intervals), which are assigned with the weight  $(1 - q_s)$ .

Similarly, the first subcells of clear intervals are assigned the weight  $(1 - p_s)$ , while all other clear subcells will enter the sum with weight  $q_s$ . The first subcell of the whole sample is assigned the initial probability weight  $u$  if it is cloudy and  $v = 1 - u$  if it is clear (these weights are free parameters to be determined from additional constraints described further below).

For each of the realization types (“diagrams”) (C8)–(C11), the number of switch subcells of both kinds (cloud to clear and clear to cloud) is fixed, as are the total numbers of cloudy and clear subcells [ $k$  and  $(n - k)$ , respectively], while the lengths of particular clear and cloudy intervals may vary. Thus, the weight of each diagram is multiplied by the number of ways the cloudy and clear subcells (only those that are not switch points) can be placed into the respective sets of intervals. To compute this number we solve a simple combinatorial problem about the number of ways to put  $m$  identical balls into  $n$  numbered boxes. This problem can be solved by treating the balls and the walls separating the boxes [ $(n - 1)$  of them] on equal footing (e.g., considering the walls as balls of a different color), and the solution is

$$K_m^n = \frac{(n + m - 1)!}{(n - 1)!m!} = \binom{n + m - 1}{n - 1}. \tag{C12}$$

Let us start with the diagram (C8) for  $r = 2i$ . For this diagram we have  $(i - 1)$  cloudy switch subcells (switching from clear to cloud) and  $i$  clear switch subcells (switching from cloud to clear). We also have  $k - i + 1$  free (nonswitch) cloudy subcells that should be distributed between the  $i$  intervals and  $(n - k - i)$  free clear subcells also to be put into  $i$  intervals. Thus the weight of the first diagram will be

$$uK_{k-i}^i K_{n-k-i}^i (1 - q_s)^{i-1} (1 - p_s)^i p_s^{k-i} q_s^{n-k-i}. \tag{C13}$$

Similarly, for diagram (C9) we have  $i$  cloudy and  $i$  clear intervals,  $i$  cloudy and  $(i - 1)$  clear switch subcells, and  $(k - i)$  cloudy and  $(n - k - i + 1)$  clear free subcells. Thus, the weight of this diagram is

$$vK_{k-i}^i K_{n-k-i}^i (1 - q_s)^i (1 - p_s)^{i-1} p_s^{k-i} q_s^{n-k-i}. \tag{C14}$$

For the first “odd” diagram (C10), we have  $(i + 1)$  cloudy and  $i$  clear intervals,  $i$  cloudy and  $i$  clear switch subcells, and  $(k - i)$  cloudy and  $(n - k - i)$  clear free subcells. This leads to the weight

$$uK_{k-i-1}^{i+1} K_{n-k-i}^i (1 - q_s)^i (1 - p_s)^i p_s^{k-i-1} q_s^{n-k-i}. \tag{C15}$$

The second odd diagram (C11) has  $i$  cloudy and  $i + 1$  clear intervals, and, as well as the first odd diagram,  $i$  cloudy and  $i$  clear switch subcells, and  $(k - i)$  cloudy and  $(n - k - i)$  clear free subcells. Thus, its weight is

$$vK_{k-i}^i K_{n-k-i-1}^{i+1} (1 - q_s)^i (1 - p_s)^i p_s^{k-i} q_s^{n-k-i-1}. \tag{C16}$$

Next we sum up these weights for  $i$  running from 1 (we consider the all-clear/overcast case  $r = 1$  separately) to  $\min(k, n - k)$ . The upper limit goes to infinity with the cell subdivision number  $M$ ; thus, we can replace the finite sums by the infinite series. Let us first compute the coefficients  $K$  in the limit of  $k, (n - k) \rightarrow \infty$ . To avoid repetition of computations similar for the coefficients  $K$  in Eqs. (C13)–(C16) we denote by  $m$  the number going to infinity in the continuous limit [ $k$  or  $(n - k)$ ], and by  $j$  the number that remains finite (combinations of  $\pm i$  and  $\pm 1$ ). Then, the majority of the coefficients  $K$  above have the form

$$\begin{aligned} K_{m+j}^i &= \binom{m + i + j - 1}{i - 1} \\ &= [(m + i + j - 1)B(i, m + j + 1)]^{-1} \\ &\approx [mB(i, m)]^{-1} \approx [m\Gamma(i)m^{-i}]^{-1} = \frac{m^{i-1}}{(i - 1)!}. \end{aligned} \tag{C17}$$

Here we represent the binomial coefficient (15) in terms of beta function and use the asymptotic formula for the latter:

$$B(x, y \gg 1) \approx \Gamma(x)y^{-x}. \tag{C18}$$

Similarly, we obtain that

$$K_{m+j}^{i+1} = \frac{m^i}{i!}. \tag{C19}$$

Thus, the coefficients in Eqs. (C13)–(C16) can be expressed as follows:

$$K_{k-i}^i = \frac{k^{i-1}}{(i - 1)!}, \quad K_{n-k-i}^i = \frac{(n - k)^{i-1}}{(i - 1)!}, \tag{C20}$$

$$K_{k-i-1}^{i+1} = \frac{k^i}{i!}, \quad K_{n-k-i-1}^{i+1} = \frac{(n - k)^i}{i!}. \tag{C21}$$

Because of the similarity of the structure of the terms (C13) and (C14) of the sums corresponding to the two even diagrams (C8) and (C9), it is convenient to write these sums together as

$$\begin{aligned} \Sigma_0 &= [u(1 - p_s) + v(1 - q_s)] \sum_{i=1}^{\infty} K_{k-i}^i K_{n-k-i}^i (1 - q_s)^{i-1} (1 - p_s)^{i-1} p_s^{k-i} q_s^{n-k-i} \\ &= [u(1 - p_s) + v(1 - q_s)] p_s^{k-1} q_s^{n-k-1} \sum_{i=1}^{\infty} \frac{k^{i-1}}{(i-1)!} \frac{(n-k)^{i-1}}{(i-1)!} \frac{(1 - q_s)^{i-1} (1 - p_s)^{i-1}}{p_s^{i-1} q_s^{i-1}} \\ &= [u(1 - p_s) + v(1 - q_s)] p_s^{k-1} q_s^{n-k-1} \sum_{i=0}^{\infty} \frac{[Pk(n-k)]^i}{(i!)^2}. \end{aligned} \tag{C22}$$

Here we have changed the summation index  $i \rightarrow (i - 1)$  and introduced the notation

$$P = \frac{(1 - q_s)(1 - p_s)}{p_s q_s}. \tag{C23}$$

In a similar way we compute the sum of the terms (C15) corresponding to the first odd diagram (C10)

$$\begin{aligned} \Sigma_{1a} &= u \sum_{i=1}^{\infty} K_{k-i-1}^{i+1} K_{n-k-i}^i (1 - q_s)^i (1 - p_s)^i p_s^{k-i-1} q_s^{n-k-i} \\ &= u \sum_{i=1}^{\infty} \frac{k^i (n-k)^{i-1}}{i!(i-1)!} (1 - q_s)^i (1 - p_s)^i p_s^{k-i-1} q_s^{n-k-i} \\ &= uk p_s^{k-1} q_s^{n-k} P \sum_{i=0}^{\infty} \frac{[Pk(n-k)]^i}{i!(i+1)!}, \end{aligned} \tag{C24}$$

and the sum of the terms (C16) corresponding to the second odd diagram (C11)

$$\begin{aligned} \Sigma_{1b} &= v \sum_{i=1}^{\infty} K_{k-i}^i K_{n-k-i-1}^{i+1} (1 - q_s)^i (1 - p_s)^i p_s^{k-i} q_s^{n-k-i-1} \\ &= v \sum_{i=1}^{\infty} \frac{k^{i-1} (n-k)^i}{(i-1)! i!} (1 - q_s)^i (1 - p_s)^i p_s^{k-i} q_s^{n-k-i-1} \\ &= v(n-k) p_s^k q_s^{n-k-1} P \sum_{i=0}^{\infty} \frac{[Pk(n-k)]^i}{i!(i+1)!}. \end{aligned} \tag{C25}$$

Thus, the sum of the two odd diagrams becomes

$$\begin{aligned} \Sigma_1 &= \Sigma_{1b} + \Sigma_{1a} = [uk p_s^{-1} + v(n-k) q_s^{-1}] \\ &\quad \times p_s^k q_s^{n-k} P \sum_{i=0}^{\infty} \frac{[Pk(n-k)]^i}{i!(i+1)!}. \end{aligned} \tag{C26}$$

The series entering Eqs. (C22), (C24), and (C25) are related to representations of the modified Bessel functions on the orders 0 and 1 (cf. Gradshteyn and Ryzhik 1965):

$$I_0(z) = \sum_{i=0}^{\infty} \frac{(z/2)^{2i}}{(i!)^2} \quad \text{and} \tag{C27}$$

$$I_1(z) = \sum_{i=0}^{\infty} \frac{(z/2)^{2i+1}}{i!(i+1)!}. \tag{C28}$$

In particular,

$$\sum_{i=0}^{\infty} \frac{z^i}{(i!)^2} = I_0(2\sqrt{z}) \quad \text{and} \tag{C29}$$

$$\sum_{i=0}^{\infty} \frac{z^i}{i!(i+1)!} = \frac{I_1(2\sqrt{z})}{\sqrt{z}}. \tag{C30}$$

In Eqs. (C22) and (C26) the argument  $z$  of Eqs. (C29) and (C30) is

$$z = Pk(n-k) = \frac{(1 - q_s)(1 - p_s)}{p_s q_s} k(NM - k). \tag{C31}$$

In the limit  $M \rightarrow \infty$

$$p_s = p^{1/M} \rightarrow 1, \quad \text{and} \quad q_s = q^{1/M} \rightarrow 1, \tag{C32}$$

while, as we have shown in Eq. (23),

$$1 - p_s \rightarrow -\frac{\ln p}{M} = -\frac{\ln p}{l} \frac{l}{M} = \frac{dx}{L_c} = \frac{L}{L_c} dc = a_c dc \quad \text{and} \tag{C33}$$

$$1 - q_s \rightarrow -\frac{\ln q}{M} = \frac{dx}{L_g} = \frac{L}{L_g} dc = a_g dc, \tag{C34}$$

where

$$dx = \frac{l}{M} \quad \text{and} \quad dc = \frac{dx}{L} \tag{C35}$$

are the respective differentials of total cloud (or gap) length and cloud cover, and we have introduced the notation

$$a_c = \frac{L}{L_c} = -\ln(p^N) \quad \text{and} \quad a_g = \frac{L}{L_g} = -\ln(q^N). \tag{C36}$$

Using these results we find the limit value of  $z$ :

$$\begin{aligned} z &\sim \frac{\ln p \ln q}{M} k(NM - k) = \frac{\ln p \ln q}{l} \frac{kl}{M} \left( Nl - \frac{kl}{M} \right) \\ &= \frac{1}{L_c L_g} x(L - x) = \frac{L^2}{L_c L_g} c(1 - c) \\ &= a_c a_g c(1 - c). \end{aligned} \tag{C37}$$

Here  $l$  is the cell size,

$$x = kl_s = \frac{kl}{M} \tag{C38}$$

is the total cloud length in the sample, and  $c = k/n = x/L$  is the cloud cover. Using Eqs. (C33)–(C37), and also noticing that

$$\Sigma_1 = \sqrt{\frac{(1 - p_s)(1 - q_s) ukp_s^{-1} + v(n - k)q_s^{-1}}{p_s q_s}} \frac{1}{\sqrt{k(n - k)}} p_s^k q_s^{n-k} \sqrt{Pk(n - k)} \sum_{i=0}^{\infty} \frac{[Pk(n - k)]^i}{i!(i + 1)!}. \tag{C40}$$

Now, using Eqs. (C28), (C33)–(C37), (C38), and (C40) and observing that

$$\begin{aligned} \sqrt{(1 - p_s)(1 - q_s)} &\rightarrow \frac{\sqrt{\ln p \ln q}}{M} = \sqrt{\frac{\ln p \ln q}{l} \frac{l}{M}} \\ &= \frac{dx}{\sqrt{L_c L_g}} = \frac{Ldc}{\sqrt{L_c L_g}} = \sqrt{a_c a_g} dc \end{aligned} \tag{C41}$$

and

$$\begin{aligned} \frac{uk + v(n - k)}{\sqrt{k(n - k)}} &= \frac{uk/M + v(N - k/M)}{\sqrt{(k/M)(N - k/M)}} \\ &= \frac{ukl/M + v(Nl - kl/M)}{\sqrt{(kl/M)(Nl - kl/M)}} \\ &= \frac{ux + v(L - x)}{\sqrt{x(L - x)}} = \frac{uc + v(1 - c)}{\sqrt{c(1 - c)}}, \end{aligned} \tag{C42}$$

we can write the continuous limit of the odd diagrams as

$$\begin{aligned} dS_1 &= [uc + v(1 - c)] e^{-[a_c c + a_g(1-c)]} \frac{\sqrt{a_c a_g}}{\sqrt{c(1 - c)}} \frac{I_1(2\sqrt{z})}{\sqrt{c(1 - c)}} dc \\ &= [uc + v(1 - c)] e^{-[a_c c + a_g(1-c)]} a_c a_g \frac{I_1(2\sqrt{z})}{\sqrt{z}} dc, \end{aligned} \tag{C43}$$

where  $z$  is defined by Eq. (C37). Thus, the expression for the total density can be written in the form

$$p_s^k = p^{k/M} = e^{(\ln p/l)(kl/M)} = e^{-x/L_c} = e^{-cL/L_c} = e^{-a_c c}$$

and

$$\begin{aligned} q_s^{n-k} &= q^{N-(k/M)} = e^{(\ln q/l)[Nl-(kl/M)]} = e^{-(L-x)/L_g} = e^{-(1-c)L/L_g} \\ &= e^{-a_g(1-c)}, \end{aligned}$$

we can derive the continuous forms of Eqs. (C22) and (C26). The sum corresponding to the even diagrams takes the form of the probability density

$$dS_0 = [a_c u + a_g v] e^{-[a_c c + a_g(1-c)]} I_0(2\sqrt{z}) dc. \tag{C39}$$

The sum (C26) corresponding to the odd diagrams can be written as

$$dS = dS_0 + dS_1 = S(c) dc, \tag{C44}$$

where

$$\begin{aligned} S(c) &= e^{-[a_c c + a_g(1-c)]} \left\{ (a_c u + a_g v) I_0(2\sqrt{z}) \right. \\ &\quad \left. + a_c a_g [uc + v(1 - c)] \frac{I_1(2\sqrt{z})}{\sqrt{z}} \right\}. \end{aligned} \tag{C45}$$

*b. Normalization condition*

Being a probability density  $S(c)$  by construction should satisfy the normalization condition

$$\int_0^1 S(c) dc + w_0 + w_1 = 1, \tag{C46}$$

where  $w_0$  and  $w_1$  are the respective weights (C5) and (C3) of all-clear and overcast samples, which in the notation of Eq. (C36) have the form

$$w_0 = v e^{-a_g} \quad \text{and} \quad w_1 = u e^{-a_c}. \tag{C47}$$

An essential step of verifying this condition is taking the integral of  $S(c)$  between 0 and 1. To do this it is convenient to split the integral into three parts:

$$\mathcal{N} = \int_0^1 S(c) dc = \mathcal{N}_0 + \mathcal{N}_1 + \mathcal{N}_2, \tag{C48}$$

where

$$\mathcal{N}_0 = (a_c u + a_g v) \int_0^1 e^{-[a_c c + a_g(1-c)]} I_0(2\sqrt{z}) dz, \tag{C49}$$

$$\mathcal{N}_1 = v a_c a_g \int_0^1 e^{-[a_c c + a_g(1-c)]} \frac{I_1(2\sqrt{z})}{\sqrt{z}} dz, \tag{C50}$$

and

$$\mathcal{N}_2 = (u - v) a_c a_g \int_0^1 c e^{-[a_c c + a_g(1-c)]} \frac{I_1(2\sqrt{z})}{\sqrt{z}} dz, \tag{C51}$$

and to use the following substitution:

$$t = 2\sqrt{c(1-c)}, \tag{C52}$$

which runs from 0 to 1 when  $c \in [0, 1/2]$  and back from 1 to 0 when  $c \in [1/2, 1]$ . Thus,  $c$  can be expressed through  $t$  as

$$c = \frac{1}{2} \mp \frac{1}{2} \sqrt{1-t^2}, \tag{C53}$$

where the minus sign is used when  $c < 1/2$ , and the plus sign is used when  $c > 1/2$ , and the integral over  $t$  includes both terms. We also have

$$1 - c = \frac{1}{2} \pm \frac{1}{2} \sqrt{1-t^2}, \tag{C54}$$

$$a_c c + a_g(1-c) = \frac{a_c + a_g}{2} \mp \frac{a_c - a_g}{2} \sqrt{1-t^2}, \tag{C55}$$

$$uc + v(1-c) = \frac{u+v}{2} \mp \frac{u-v}{2} \sqrt{1-t^2}, \tag{C56}$$

$$2\sqrt{z} = t\sqrt{a_c a_g}, \text{ and} \tag{C57}$$

$$dz = \pm \frac{t dt}{2\sqrt{1-t^2}}. \tag{C58}$$

Substituting the above expressions into Eq. (C45) leads to the following:

$$\begin{aligned} \mathcal{N}_0 &= (a_c u + a_g v) e^{-(a_c + a_g)/2} \\ &\times \int_0^1 \frac{t}{\sqrt{1-t^2}} \cosh\left(\frac{a_c - a_g}{2} \sqrt{1-t^2}\right) I_0(t\sqrt{a_c a_g}) dt, \end{aligned} \tag{C59}$$

$$\begin{aligned} \mathcal{N}_1 &= 2v\sqrt{a_c a_g} e^{-(a_c + a_g)/2} \\ &\times \int_0^1 \frac{1}{\sqrt{1-t^2}} \cosh\left(\frac{a_c - a_g}{2} \sqrt{1-t^2}\right) \\ &\times I_1(t\sqrt{a_c a_g}) dt, \text{ and} \end{aligned} \tag{C60}$$

$$\mathcal{N}_2 = \mathcal{N}_{2a} + \mathcal{N}_{2b}, \tag{C61}$$

where

$$\begin{aligned} \mathcal{N}_{2a} &= (u - v) \sqrt{a_c a_g} e^{-(a_c + a_g)/2} \\ &\times \int_0^1 \frac{1}{\sqrt{1-t^2}} \cosh\left(\frac{a_c - a_g}{2} \sqrt{1-t^2}\right) I_1(t\sqrt{a_c a_g}) dt \end{aligned}$$

is similar to  $\mathcal{N}_1$ , and

$$\begin{aligned} \mathcal{N}_{2b} &= (v - u) \sqrt{a_c a_g} e^{-(a_c + a_g)/2} \\ &\times \int_0^1 \sinh\left(\frac{a_c - a_g}{2} \sqrt{1-t^2}\right) I_1(t\sqrt{a_c a_g}) dt. \end{aligned} \tag{C62}$$

We can combine two of the above integrals into one:

$$\begin{aligned} \mathcal{N}_3 = \mathcal{N}_1 + \mathcal{N}_{2a} &= \sqrt{a_c a_g} e^{-(a_c + a_g)/2} \\ &\times \int_0^1 \frac{1}{\sqrt{1-t^2}} \cosh\left(\frac{a_c - a_g}{2} \sqrt{1-t^2}\right) I_1(t\sqrt{a_c a_g}) dt. \end{aligned} \tag{C63}$$

To compute this expression we use a table integral (Prudnikov et al. 2002)

$$\begin{aligned} &\int_0^\alpha \frac{\cosh(\beta\sqrt{\alpha^2 - x^2})}{\sqrt{\alpha^2 - x^2}} I_\nu(\gamma x) dx \\ &= \frac{\pi}{2} I_{\nu/2}\left(\frac{\alpha\sqrt{\beta^2 + \gamma^2} + \alpha\beta}{2}\right) I_{\nu/2}\left(\frac{\alpha\sqrt{\beta^2 + \gamma^2} - \alpha\beta}{2}\right) \end{aligned} \tag{C64}$$

[note that there is a misprint in Prudnikov et al. (2002): the factor  $1/2$  in the arguments of both Bessel functions are missing]. In our case ( $\nu = 1, \alpha = 1$ ) this expression can be simplified using the relation [cf. Gradshteyn and Ryzhik (1965)]

$$I_{1/2}(z) = \sqrt{\frac{2}{\pi z}} \sinh z, \tag{C65}$$

resulting in

$$\begin{aligned} \mathcal{J}_1 &= \int_0^1 \frac{\cosh(\beta\sqrt{1-x^2})}{\sqrt{1-x^2}} I_1(\gamma x) dx \\ &= \frac{2}{\gamma} \sinh\left(\frac{\sqrt{\beta^2 + \gamma^2} + \beta}{2}\right) \sinh\left(\frac{\sqrt{\beta^2 + \gamma^2} - \beta}{2}\right). \end{aligned} \tag{C66}$$

Using this formula with

$$\beta = \frac{a_c - a_g}{2}, \quad \gamma = \sqrt{a_c a_g}, \quad \text{and} \quad \sqrt{\beta^2 + \gamma^2} = \frac{a_c + a_g}{2}, \tag{C67}$$

we find that Eq. (C63) for  $\mathcal{N}_3$  takes the following form:

$$\begin{aligned} \mathcal{N}_3 &= 2e^{-(a_c+a_g)/2} \sinh \frac{a_c}{2} \sinh \frac{a_g}{2} \\ &= e^{-(a_c+a_g)/2} \left[ \cosh \left( \frac{a_c+a_g}{2} \right) - \cosh \left( \frac{a_c-a_g}{2} \right) \right]. \end{aligned} \tag{C68}$$

In the important case of  $p = 1/2$  when  $a_c = a_g \equiv a$ , it looks even simpler:

$$\mathcal{N}_3 = 2e^{-a} \sinh^2 \frac{a}{2} = e^{-a} (\cosh a - 1). \tag{C69}$$

To take the integral (C62) for  $\mathcal{N}_{2b}$  we notice that the integral (C66) can be written in the form

$$\mathcal{J}_1 = \frac{1}{\gamma} (\cosh \sqrt{\beta^2 + \gamma^2} - \cosh \beta), \tag{C70}$$

which after differentiation with respect to  $\beta$  yields the following formula:

$$\begin{aligned} \frac{\partial \mathcal{J}_1}{\partial \beta} &= \int_0^1 \sinh(\beta \sqrt{1-x^2}) I_1(\gamma x) dx \\ &= \frac{1}{\gamma} \left( \frac{\beta}{\sqrt{\beta^2 + \gamma^2}} \sinh \sqrt{\beta^2 + \gamma^2} - \sinh \beta \right), \end{aligned} \tag{C71}$$

allowing us to compute  $\mathcal{N}_{2b}$  using the substitution (C67):

$$\begin{aligned} \mathcal{N}_{2b} &= (v-u) e^{-(a_c+a_g)/2} \left[ \frac{a_c-a_g}{a_c+a_g} \sinh \left( \frac{a_c+a_g}{2} \right) \right. \\ &\quad \left. - \sinh \left( \frac{a_c-a_g}{2} \right) \right]. \end{aligned} \tag{C72}$$

Unfortunately we cannot find a table integral corresponding to  $\mathcal{N}_0$  of a general case. Thus, we guess the form of the solution and then verify the result numerically. As a basis for our guess we take the solution in the specific case of  $p = 1/2$  ( $a_c = a_g \equiv a$ ), which can be found using another table integral (Prudnikov et al. 2002):

$$\begin{aligned} \mathcal{N}_0 &= ae^{-a} \int_0^1 \frac{t}{\sqrt{1-t^2}} I_0(at) dt \\ &= e^{-a} \sinh a = 2e^{-a} \sinh \frac{a}{2} \cosh \frac{a}{2}. \end{aligned} \tag{C73}$$

A comparison of Eqs. (C69) and (C68) suggests that the generalization of Eq. (C73) to the case of an arbitrary  $p$  may have the following form:

$$\begin{aligned} \mathcal{N}_0 &= 2 \frac{a_c u + a_g v}{a_c + a_g} e^{-(a_c+a_g)/2} \sinh \left( \frac{a_c+a_g}{2} \right) \\ &= 2 \frac{a_c u + a_g v}{a_c + a_g} e^{-(a_c+a_g)/2} \left[ \sinh \frac{a_c}{2} \cosh \frac{a_g}{2} \right. \\ &\quad \left. + \cosh \frac{a_c}{2} \sinh \frac{a_g}{2} \right]. \end{aligned} \tag{C74}$$

Extensive numerical tests verify that this formula is indeed correct, which allows us to compute the total integral (C48):

$$\begin{aligned} \mathcal{N} &= \mathcal{N}_0 + \mathcal{N}_{2b} + \mathcal{N}_3 = e^{-(a_c+a_g)/2} \left[ 2 \frac{a_c u + a_g v}{a_c + a_g} \sinh \left( \frac{a_c+a_g}{2} \right) + (v-u) \frac{a_c-a_g}{a_c+a_g} \sinh \left( \frac{a_c+a_g}{2} \right) \right. \\ &\quad \left. + (u-v) \sinh \left( \frac{a_c-a_g}{2} \right) + \cosh \left( \frac{a_c+a_g}{2} \right) - \cosh \left( \frac{a_c-a_g}{2} \right) \right] = 1 - ue^{-a_c} - ve^{-a_g}. \end{aligned} \tag{C75}$$

Thus, the normalization condition (C46) holds for any initial probabilities  $u$  and  $v$ .

*c. General form of the singular density*

The complete cloud cover singular density that is normalized by the condition

$$\int_0^1 F_{cf}(c) dc = 1 \tag{C76}$$

takes the form

$$F_{cf}(c) = f_{cf}(c) + f_{cf}^{(0)} \delta(c) + f_{cf}^{(1)} \delta(1-c) \tag{C77}$$

with

$$\begin{aligned} f_{cf}(c) &= e^{-[a_c c + a_g (1-c)]} \left[ (a_c u + a_g v) I_0(Z) \right. \\ &\quad \left. + 2a_c a_g [uc + v(1-c)] \frac{I_1(Z)}{Z} \right], \end{aligned} \tag{C78}$$

where  $Z$  is defined as

$$Z = 2\sqrt{a_c a_g c(1-c)} \tag{C79}$$

and

$$f_{cf}^{(0)} = ve^{-a_g} \quad \text{and} \quad f_{cf}^{(1)} = ue^{-a_c}. \quad (C80)$$

$$\mathcal{M}_2 = (u - v)a_c a_g \int_0^1 c^2 e^{-[a_c c + a_g(1-c)]} \frac{I_1(2\sqrt{z})}{\sqrt{z}} dc. \quad (D5)$$

Under the substitution (C53)

$$c^2 = c - \frac{t^2}{4}; \quad (D6)$$

thus, we can write  $\mathcal{M}_2$  as

$$\mathcal{M}_2 = \mathcal{M}_{2a} + \mathcal{M}_{2b}, \quad (D7)$$

where

$$\mathcal{M}_{2a} = (u - v)a_c a_g \int_0^1 ce^{-[a_c c + a_g(1-c)]} \frac{I_1(2\sqrt{z})}{\sqrt{z}} dc \quad \text{and} \quad (D8)$$

$$\mathcal{M}_{2b} = \frac{v - u}{2} a_c a_g \int_0^1 t^2 e^{-[a_c c + a_g(1-c)]} \frac{I_1(2\sqrt{z})}{2\sqrt{z}} dc. \quad (D9)$$

Also,  $\mathcal{M}_{2a}$  is similar to  $\mathcal{M}_1$ , so the two can be combined into

$$\mathcal{M}_3 = \mathcal{M}_1 + \mathcal{M}_{2a} = ua_c a_g \int_0^1 ce^{-[a_c c + a_g(1-c)]} \frac{I_1(2\sqrt{z})}{2\sqrt{z}} dc. \quad (D10)$$

After the substitution (C53), we have

## APPENDIX D

### Mean Cloud Cover in the Continuous Model

To compute the mean cloud cover

$$\bar{c} = \int_0^1 cF_{cf}(c) dc = \int_0^1 cf_{cf}(c) dc + f_{cf}^{(1)} \quad (D1)$$

for the singular density (C77), we use the same substitution (C52) that was used for computing the norm. Under this substitution  $c$  is expressed according to Eq. (C53); as for the normalization, we express the integral in Eq. (D1) as a sum of three terms:

$$\mathcal{M} = \int_0^1 cf_{cf}(c) dc = \mathcal{M}_0 + \mathcal{M}_1 + \mathcal{M}_2, \quad (D2)$$

where

$$\mathcal{M}_0 = (a_c u + a_g v) \int_0^1 ce^{-[a_c c + a_g(1-c)]} I_0(2\sqrt{z}) dc, \quad (D3)$$

$$\mathcal{M}_1 = va_c a_g \int_0^1 ce^{-[a_c c + a_g(1-c)]} \frac{I_1(2\sqrt{z})}{\sqrt{z}} dc, \quad \text{and} \quad (D4)$$

---


$$\begin{aligned} \mathcal{M}_0 &= (a_c u + a_g v) e^{-(a_c + a_g)2} \int_0^1 \frac{t dt}{2\sqrt{1-t^2}} \left[ \cosh\left(\frac{a_c - a_g}{2} \sqrt{1-t^2}\right) - \sqrt{1-t^2} \sinh\left(\frac{a_c - a_g}{2} \sqrt{1-t^2}\right) \right] I_0(t\sqrt{a_c a_g}) \\ &= \frac{1}{2} N_0 - \mathcal{R}_0, \end{aligned} \quad (D11)$$

where  $\mathcal{N}_0$  is defined by Eqs. (C59) and (C74), while

$$\begin{aligned} \mathcal{R}_0 &= \frac{a_c u + a_g v}{2} e^{-(a_c + a_g)2} \\ &\times \int_0^1 t \sinh\left(\frac{a_c - a_g}{2} \sqrt{1-t^2}\right) I_0(t\sqrt{a_c a_g}) dt. \end{aligned} \quad (D12)$$

Similarly,

---


$$\begin{aligned} \mathcal{M}_3 &= u\sqrt{a_c a_g} e^{-(a_c + a_g)2} \\ &\times \int_0^1 \frac{dt}{\sqrt{1-t^2}} \left[ \cosh\left(\frac{a_c - a_g}{2} \sqrt{1-t^2}\right) \right. \\ &\quad \left. - \sqrt{1-t^2} \sinh\left(\frac{a_c - a_g}{2} \sqrt{1-t^2}\right) \right] I_1(t\sqrt{a_c a_g}) \\ &= u\mathcal{N}_3 - \frac{u}{v-u} \mathcal{N}_{2b}, \end{aligned}$$

where  $\mathcal{N}_3$  is defined by Eqs. (C63) and (C68), while  $\mathcal{N}_{2b}$  is given by Eqs. (C62) and (C72). Here  $\mathcal{M}_{2b}$  takes the following form:

$$\mathcal{M}_{2b} = \frac{v-u}{2} \sqrt{a_c a_g} e^{-(a_c+a_g)/2} \times \int_0^1 \frac{t^2 dt}{\sqrt{1-t^2}} \cosh\left(\frac{a_c - a_g}{2} \sqrt{1-t^2}\right) I_1(t\sqrt{a_c a_g}).$$

To compute the integral for  $\mathcal{R}_0$  we need to extend [reversing the substitution (C67)] our guess (C74) to the following relation:

$$\mathcal{J}_0 = \int_0^1 x \frac{\cosh(\beta\sqrt{1-x^2})}{\sqrt{1-x^2}} I_0(\gamma x) dx = \frac{\sinh(\sqrt{\beta^2 + \gamma^2})}{\sqrt{\beta^2 + \gamma^2}} \quad (\text{D13})$$

(again conducting a numerical verification). Taking a partial derivative of it with respect to  $\beta$ , we come to the following:

$$\begin{aligned} \frac{\partial \mathcal{J}_0}{\partial \beta} &= \int_0^1 x \sinh(\beta\sqrt{1-x^2}) I_0(\gamma x) dx \\ &= \frac{\beta}{\beta^2 + \gamma^2} \left[ \cosh(\sqrt{\beta^2 + \gamma^2}) - \frac{\sinh(\sqrt{\beta^2 + \gamma^2})}{\sqrt{\beta^2 + \gamma^2}} \right]. \end{aligned} \quad (\text{D14})$$

Thus, using the substitution (C67), we can write  $\mathcal{R}_0$  as

$$\begin{aligned} \mathcal{R}_0 &= \frac{(a_c u + a_g v)(a_c - a_g)}{(a_c + a_g)^2} e^{-(a_c+a_g)/2} \left[ \cosh\left(\frac{a_c + a_g}{2}\right) \right. \\ &\quad \left. - \frac{2}{a_c + a_g} \sinh\left(\frac{a_c + a_g}{2}\right) \right]. \end{aligned} \quad (\text{D15})$$

To take the integral for  $\mathcal{M}_{2b}$  we again use Eq. (D13). Differentiating Eq. (D13) with respect to  $\gamma$  and noticing that (cf. Gradshteyn and Ryzhik 1965)

$$I_1(x) = \frac{dI_0(x)}{dx}, \quad (\text{D16})$$

we obtain the following relation:

$$\begin{aligned} \frac{\partial \mathcal{J}_0}{\partial \gamma} &= \int_0^1 x^2 \frac{\cosh(\beta\sqrt{1-x^2})}{\sqrt{1-x^2}} I_1(\gamma x) dx \\ &= \frac{\gamma}{\beta^2 + \gamma^2} \left[ \cosh(\sqrt{\beta^2 + \gamma^2}) - \frac{\sinh\sqrt{\beta^2 + \gamma^2}}{\sqrt{\beta^2 + \gamma^2}} \right]. \end{aligned} \quad (\text{D17})$$

Note that the rhs of Eq. (D13) is symmetric with respect to the  $\beta \leftrightarrow \gamma$  interchange, so the result of Eq. (D17) can be also obtained from Eq. (D14) simply by interchanging these variables. The expression for  $\mathcal{M}_{2b}$  now takes the form

$$\begin{aligned} \mathcal{M}_{2b} &= 2 \frac{(v-u)a_c a_g}{(a_c + a_g)^2} e^{-(a_c+a_g)/2} \left[ \cosh\left(\frac{a_c + a_g}{2}\right) \right. \\ &\quad \left. - \frac{2}{a_c + a_g} \sinh\left(\frac{a_c + a_g}{2}\right) \right], \end{aligned} \quad (\text{D18})$$

and we can collect the terms in Eq. (D2):

$$\mathcal{M} = \frac{1}{2} \mathcal{N}_0 - \mathcal{R}_0 + u \mathcal{N}_3 - \frac{u}{v-u} \mathcal{N}_{2b} + \mathcal{M}_{2b}. \quad (\text{D19})$$

It is convenient to introduce the notation

$$a_+ = \frac{a_c + a_g}{2} \quad \text{and} \quad a_- = \frac{a_c - a_g}{2}, \quad (\text{D20})$$

in which

$$\begin{aligned} \mathcal{M}_{2b} - \mathcal{R}_0 &= -\frac{a_c u - a_g v}{2a_+} e^{-a_+} \left[ \cosh(a_+) - \frac{\sinh(a_+)}{a_+} \right], \\ \frac{1}{2} \mathcal{N}_0 &= \frac{a_c u + a_g v}{2a_+} e^{-a_+} \sinh(a_+), \\ u \mathcal{N}_3 &= u e^{-a_+} [\cosh(a_+) - \cosh(a_-)], \quad \text{and} \\ \frac{u}{v-u} \mathcal{N}_{2b} &= u e^{-a_+} \left[ \frac{a_-}{a_+} \sinh(a_+) - \sinh(a_-) \right], \end{aligned}$$

and after some algebraic transformations we obtain

$$\mathcal{M} = \frac{a_g}{2a_+} - u e^{-a_c} + \frac{a_c u - a_g v}{4a_+^2} (1 - e^{-2a_+}). \quad (\text{D21})$$

Now we can write the expression for the mean cloud cover:

$$\bar{c} = \mathcal{M} + u e^{-a_c} = \frac{a_g}{2a_+} + \frac{a_c u - a_g v}{4a_+^2} (1 - e^{-2a_+}) \quad (\text{D22})$$

or, explicitly,

$$\bar{c} = \frac{a_g}{a_c + a_g} + \frac{a_c u - a_g v}{(a_c + a_g)^2} [1 - e^{-(a_c+a_g)}]. \quad (\text{D23})$$

Note that in the limiting case  $p = 0$  (no cloud)  $L_c = 0$  and  $a_c = \infty$ , while  $L_g = \infty$  and  $a_g = 0$ , and it is easy to see that  $\bar{c} = 0$ . In the opposite limit  $p = 1$  (completely overcast)  $L_c = \infty$  and  $a_c = 0$ , while  $L_g = 0$  and  $a_g = \infty$ , so  $\bar{c} = 1/2 + 1/2 = 1$ .

#### a. Infinite sample

In the limit  $L \rightarrow \infty$  when the sample length is considerably larger than the mean sizes of clouds and gaps (here we assume that  $0 < p < 1$ , so that  $L_c$  and  $L_g$  are finite), the

distribution (C77) becomes very narrow, making both completely clear and overcast cases impossible:

$$f_{\text{cf}}^{(0)} = f_{\text{cf}}^{(1)} = 0 \tag{D24}$$

(this follows from  $a_c, a_g \rightarrow \infty$ ). Equation (D23) for the mean cloud fraction takes the form

$$\bar{c}_\infty = \frac{a_g}{a_c + a_g} = \frac{L_c}{L_c + L_g} = \frac{\ln q}{\ln p + \ln q}, \tag{D25}$$

regardless of the values of the initial probabilities  $u$  and  $v$ .

The density (C78) in this limiting case can be simplified by using the asymptotic representation of Bessel functions (Gradshteyn and Ryzhik 1965):

$$I_\nu(x \rightarrow \infty) \approx \frac{e^x}{\sqrt{2\pi x}}, \tag{D26}$$

which is independent of the index  $\nu$ :

---


$$f_{\text{cf}}(c) = (a_c a_g)^{1/4} e^{-[\sqrt{a_c c} - \sqrt{a_g(1-c)}]^2} \frac{A\sqrt{c(1-c)} + [uc + v(1-c)]}{2\sqrt{\pi}[c(1-c)]^{3/4}}, \tag{D27}$$


---

where

$$A = \frac{a_c u + a_g v}{\sqrt{a_c a_g}} = u\sqrt{\frac{L_c}{L_g}} + v\sqrt{\frac{L_g}{L_c}}. \tag{D28}$$

The last factor of Eq. (D27) does not contain  $L$ , while  $a_c$  and  $a_g \propto L$ , and therefore the argument of the exponent is also proportional to  $L$ . Thus, if for some  $c$  this argument is not equal to 0, then  $f_{\text{cf}}(c) = 0$ , since the exponent declines with  $L \rightarrow \infty$  faster than the power-law factor  $(a_c a_g)^{1/4} \propto L^{1/2}$  grows. And when this argument equals 0 we have  $f_{\text{cf}}(c) = \infty$ . This happens when

$$a_c c = a_g (1 - c), \tag{D29}$$

that is, when  $c = \bar{c}_\infty$ . These observations, together with the normalization condition (C76) and Eq. (D24), imply that in the limit of an infinite sample the singular density for cloud cover does not depend on  $u$  or  $v$  and has the form

$$F_{\text{cf}}(c) \equiv f_{\text{cf}}(c) = \delta(c - \bar{c}_\infty), \tag{D30}$$

where  $\bar{c}_\infty$  is defined by Eq. (D25).

*b. Choice of initial probabilities*

As discussed above, the initial probabilities  $u$  and  $v$  are free parameters and are specified by the way in which the samples are generated. However, in a realistic situation when the samples are taken at random from a long record, we may assume a certain homogeneity of the cloud pattern. This means that the probability  $u$  of the first subcell to be cloudy should be the same as the probability of any subcell to be cloudy, which is equal to the mean cloud fraction  $\bar{c}$ . Thus, we can use Eq. (D23) and the condition

$$u = \bar{c} \tag{D31}$$

to determine the value of  $u$  from the equation

$$u = \frac{a_g}{2a_+} + \frac{a_c u - a_g v}{4a_+^2} (1 - e^{-2a_+}), \tag{D32}$$

which, using  $v = 1 - u$ , can be written as

$$u \left( 1 - \frac{1 - e^{-2a_+}}{2a_+} \right) = \frac{a_g}{2a_+} \left( 1 - \frac{1 - e^{-2a_+}}{2a_+} \right).$$

We notice that  $e^{-2a_+} > 1 - 2a_+$  for all  $a_+ \neq 0$  (since the line  $y = 1 - x$  is tangent to the graph of  $y = e^{-x}$  at  $x = 0$ ), while  $a_+ = 0$  is unphysical; thus, the only solution of Eq. (D31) is

$$u = \frac{a_g}{2a_+} = \frac{a_g}{a_c + a_g} = \bar{c}_\infty; \tag{D33}$$

that is, the probability  $u$  is equal to mean cloud cover for an infinite sample (D25). Correspondingly,

$$v = 1 - u = \frac{a_c}{a_c + a_g}. \tag{D34}$$

Substituting Eqs. (D33) and (D34) into Eq. (D23), we see that, indeed, the second term in this expression vanishes, so the mean cloud cover becomes equal to its limiting value

$$\bar{c} = \bar{c}_\infty = \frac{a_g}{a_c + a_g} = \frac{L_c}{L_c + L_g} \tag{D35}$$

and does not depend on the sample size. Unlike its mean, however, the shape of the cloud cover distribution does depend on  $L$ . Substituting Eqs. (D33) and (D34)

into Eqs. (C78) and (C80) gives Eqs. (40)–(45) for the components of the singular density (C77).

## REFERENCES

- Aida, M., 1977: Scatter of solar radiation as a function of cloud dimensions and orientations. *J. Quant. Spectrosc. Radiat. Transfer*, **17**, 303–310.
- Alexandrov, M. D., A. Marshak, B. Cairns, A. A. Lacis, and B. E. Carlson, 2004: Automated cloud screening algorithm for MFRSR data. *Geophys. Res. Lett.*, **31**, L04118, doi:10.1029/2003GL019105.
- , A. S. Ackerman, and A. Marshak, 2010: Cellular statistical models of broken cloud fields. Part II: Comparison with a dynamical model and statistics of diverse ensembles. *J. Atmos. Sci.*, **67**, 2152–2170.
- 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.
- , —, and H. M. van de Poll, 2001: Bayesian confidence intervals for true fractional coverage from finite transect measurements: Implications for cloud studies from space. *J. Geophys. Res.*, **106**, 17 303–17 310.
- Baxter, R. J., 1982: *Exactly Solved Models in Statistical Mechanics*. Academic Press, 498 pp.
- 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.
- Breon, F.-M., 1992: Reflectance of broken cloud fields: Simulation and parameterization. *J. Atmos. Sci.*, **49**, 1221–1232.
- 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.
- Conolly, B. W., 1971: On randomized random walks. *SIAM Rev.*, **13**, 81–99.
- Ellingson, R. G., 1982: On the effects of cumulus dimensions on longwave irradiance and heating rate calculations. *J. Atmos. Sci.*, **39**, 886–896.
- Evans, K. F., and W. J. Wiscombe, 2004: An algorithm for generating stochastic cloud fields from radar profile statistics. *Atmos. Res.*, **72**, 263–289.
- Gradshteyn, I. S., and I. M. Ryzhik, 1965: *Table of Integrals, Series, and Products*. Academic Press, 1130 pp.
- Harrison, L., J. Michalsky, and J. Berndt, 1994: Automated multifilter rotating shadow-band radiometer: An instrument for optical depth and radiation measurements. *Appl. Opt.*, **33**, 5118–5125.
- 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.
- Holben, B. N., and Coauthors, 1998: Aeronet—A federated instrument network and data archive for aerosol characterization. *Remote Sens. Environ.*, **66**, 1–16.
- Isichenko, M. B., 1992: Percolation, statistical topography, and transport in random media. *Rev. Mod. Phys.*, **64**, 961–1043.
- Joseph, J. H., and R. F. Cahalan, 1990: Nearest neighbor spacing of fair weather cumulus clouds. *J. Appl. Meteor.*, **29**, 793–805.
- Kassianov, E., 2003: Stochastic radiative transfer in multilayer broken clouds. Part I: Markovian approach. *J. Quant. Spectrosc. Radiat. Transfer*, **77**, 373–394.
- Kite, A., 1987: The albedo of broken cloud fields. *Quart. J. Roy. Meteor. Soc.*, **113**, 517–531.
- 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.
- Lane-Veron, D. E., and R. C. J. Somerville, 2004: Stochastic theory of radiative transfer through generalized cloud field. *J. Geophys. Res.*, **109**, D18113, doi:10.1029/2004JD004524.
- Malvagi, F., R. N. Byrne, G. Pomraning, and R. C. J. Somerville, 1993: Stochastic radiative transfer in a partially cloudy atmosphere. *J. Atmos. Sci.*, **50**, 2146–2158.
- Marshak, A., A. Davis, R. F. Cahalan, and W. J. Wiscombe, 1994: Bounded cascade models as nonstationary multifractals. *Phys. Rev. E*, **49**, 55–69.
- Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, 1953: Equation of state calculations by fast computing machines. *J. Chem. Phys.*, **21**, 1087–1092.
- 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.
- 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, 1986: *Elementary Functions*. Vol. 1, *Integrals and Series*, Gordon and Breach, 806 pp.
- , —, and —, 2002: *Special Functions*. Vol. 2, *Integrals and Series*, Taylor and Francis, 758 pp.
- Ramirez, J. A., and R. L. Bras, 1990: Clustered or regular cumulus cloud fields: The statistical character of observed and simulated cloud fields. *J. Geophys. Res.*, **95**, 2035–2045.
- Randall, D. A., and G. J. Huffman, 1980: A stochastic model of cumulus clumping. *J. Atmos. Sci.*, **37**, 2068–2078.
- 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.
- Sauvageot, H., F. Mesnard, and R. S. Tenorio, 1999: The relation between the area-average rain rate and the rain cell size distribution parameters. *J. Atmos. Sci.*, **56**, 57–70.
- 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. Pilewskie, 2007: Reproducing cloud

- microphysics and irradiance measurements using three 3D cloud generators. *Quart. J. Roy. Meteor. Soc.*, **133**, 765–780.
- Settle, J. J., and H. M. van de Poll, 2007: On the Bayesian estimation of cloud fraction from lidar transects. *J. Geophys. Res.*, **112**, D09211, doi:10.1029/2006JD007251.
- Smirnov, A., B. N. Holben, T. F. Eck, O. Dubovik, and I. Slutsker, 2000: Cloud screening and quality control algorithms for the AERONET database. *Remote Sens. Environ.*, **73**, 337–349.
- Su, B. J., and G. C. Pomraning, 1994: A stochastic description of a broken cloud field. *J. Atmos. Sci.*, **51**, 1969–1977.
- Swendsen, R. H., and J.-S. Wang, 1987: Nonuniversal critical dynamics in Monte Carlo simulations. *Phys. Rev. Lett.*, **58**, 86–88.
- Takacs, L., 1957: On certain sojourn time problems in the theory of stochastic processes. *Acta Math. Acad. Sci. Hung.*, **8**, 169–191.
- Titov, G. A., 1990: Statistical description of radiation transfer in clouds. *J. Atmos. Sci.*, **47**, 24–38.
- Van de Poll, H. M., H. Grubb, and I. Astin, 2006: Uncertainty properties of cloud fraction estimates from random transect observations. *J. Geophys. Res.*, **111**, D22218, doi:10.1029/2006JD007189.
- 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.
- Wolff, U., 1989: Collective Monte Carlo updating for spin systems. *Phys. Rev. Lett.*, **62**, 361–364.
- Yau, M. K., and R. R. Rogers, 1984: An inversion problem on inferring the size distribution of precipitation areas from rain-gage measurements. *J. Atmos. Sci.*, **41**, 439–447.
- Zuev, V. E., and G. A. Titov, 1995: Radiative transfer in cloud fields with random geometry. *J. Atmos. Sci.*, **52**, 176–190.