

Supporting Information for “Cloud regimes as phase transitions”

Samuel N. Stechmann^{1,2} and Scott Hottovy¹

¹Department of Mathematics, University of Wisconsin–Madison, Madison, Wisconsin, USA.

²Department of Atmospheric and Oceanic Sciences, University of Wisconsin–Madison, Madison, Wisconsin, USA.

Contents of this file

1. Text S1 to S8
2. Figures S1 to S3

Introduction

This Supporting Information describes details of the methods for analyzing the model and the observational data, and it describes derivations of mathematical formulas. The text sections are organized as follows:

Text S1. Model Derivation

Text S2. Spatial Discretization

Text S3. Stationary Distribution

Text S4. Derivation of Formula for Mean Cloud Area Fraction

Text S5. Methods for Model Cloud Scenes in Figure 1

Text S6. Asymptotic Formulas for $\text{Var}(q_{i,j})$, $\bar{\sigma}$, and χ in Limit of Small Δx

Text S7. Parameter Sensitivity Studies

Text S8. Observational Data

Text S1. Model Derivation

The model

$$\partial_t q = b\nabla^2 q - \frac{1}{\tau} q + D\dot{W} + F \quad (\text{S1})$$

can be related to atmospheric fluid dynamics in the following way, as described by *Hottovy and Stechmann* [2015a]. The water vapor mass concentration evolves according to

$$\frac{\partial q}{\partial t} + (uq)_x + (vq)_y + (wq)_z = S, \quad (\text{S2})$$

where u, v , and w are velocity components, and S represents any source or sink such as precipitation. Next q is decomposed as $q = \bar{q} + q'$ into a large-scale component \bar{q} and a small-scale component q' . The large-scale component \bar{q} represents a vertical integral of q over the depth of the atmosphere and a horizontal average over a scale similar to the lattice grid spacing Δx . The dynamics of \bar{q} is then found from (S2) to be

$$\frac{\partial \bar{q}}{\partial t} = \bar{S} - [(\bar{u}\bar{q})_x + (\bar{v}\bar{q})_y] - [(\overline{u'q'})_x + (\overline{v'q'})_y]. \quad (\text{S3})$$

The relationship with (S1) can then be seen after two common simplifying assumptions for turbulent flows: (i) the small-scale flux convergence, $-(\overline{u'q'})_x - (\overline{v'q'})_y$, is modeled as eddy diffusion, $b\nabla^2 q$, and (ii) the nonlinear turbulent effects of $-(\bar{u}\bar{q})_x - (\bar{v}\bar{q})_y$ are modeled with additional turbulent damping, $-\tau^{-1}q$, and stochastic forcing, $D\dot{W}$ [*DelSole, 2004; Majda and Grote, 2007*]. With this connection to atmospheric dynamics, the terms of the model can be identified with physical processes of precipitation, evaporation, and turbulent advection-diffusion. The various physical processes involved are illustrated schematically in Fig. S1.

The representation of turbulent advection-diffusion here, as in *Hottovy and Stechmann* [2015a], is in a highly idealized form as eddy diffusion, stochastic forcing, and damping.

Hottovy and Stechmann [2015a] noted that more sophisticated representations could also be used to provide additional realism, as described, e.g., by *Majda and Kramer* [1999] and *Majda and Gershgorin* [2013] and references therein. For instance, these more sophisticated models can reproduce the fat-tailed probability density functions (pdfs) that are commonly seen in situations of turbulent advection-diffusion with a background mean gradient, including the case of water vapor pdfs with exponential tails. With water vapor, an added complication are the source terms due to precipitation and evaporation, and *Stechmann and Neelin* [2011] and *Hottovy and Stechmann* [2015b] suggested that the source terms may play a key role in accounting for the exponential tails analyzed by *Neelin et al.* [2009]. It would be interesting to possibly extend the model used here to explore the contributions of source terms versus turbulent advection-diffusion in accounting for the exponential tails in water vapor pdfs.

Text S2. Spatial Discretization

A discrete version of (S1) is used here by dividing the two-dimensional (x, y) domain of size $L \times L$ into a two-dimensional (i, j) lattice of size $N \times N$, using a grid spacing of $\Delta x = \Delta y = L/N$. Define $q_{i,j}(t)$ to be the discrete version of $q(x, y, t)$ that evolves according to the following stochastic differential equation (SDE)

$$\begin{aligned} \frac{dq_{i,j}(t)}{dt} = & \tilde{b} [q_{i+1,j}(t) + q_{i-1,j}(t) + q_{i,j+1}(t) + q_{i,j-1}(t) - 4q_{i,j}(t)] - \frac{1}{\tau} q_{i,j}(t) \\ & + \tilde{D}\dot{W}_{i,j}(t) + F, \end{aligned} \quad (\text{S4})$$

for $i, j = 1, \dots, N$, and $\dot{W}_{i,j}(t)$ are independent white noises. This equation arises from (S1) upon using a standard centered discretization of the Laplacian operator, ∇^2 , with $\tilde{b} = b/(\Delta x)^2$. Also note that the natural discretization of $D\dot{W}(x, y, t)$ is $\tilde{D}\dot{W}_{i,j}(t)$, with

$\tilde{D} = D/(\sqrt{\Delta x}\sqrt{\Delta y}) = D/\Delta x$. For simplicity, periodic boundary conditions are imposed for the above equation. That is, $q_{N+1,j}(t) = q_{1,j}(t)$ for all $j = 1, 2, \dots, N$ and $q_{i,N+1}(t) = q_{i,1}(t)$ for all $i = 1, 2, \dots, N$.

Note that the continuum limit of (S4) involves some mathematical subtleties. In fact, while we have drawn a natural connection between (S1) and (S4), the connection is only formal, as it is well-known that the continuum model (S1) does not have a finite variance in two spatial dimensions, and, correspondingly, the limit of (S4) for $\Delta x \rightarrow 0$ leads to a diverging variance (see Text S6 below). Nevertheless, the discrete model (S4) is well-defined for finite Δx , and its dependence on Δx is very weak, so it is reasonable to make the connection between (S1) and (S4) despite these subtleties. Furthermore, the continuum model (S1) would, in fact, be well-defined with a finite variance if the uncorrelated noise \dot{W} were replaced by a reasonable spatially correlated noise term; such a choice leads to a slightly more complicated model, since the noise is then characterized not by a single parameter, D , but by the entire correlation function (or Fourier spectral density) of the noise.

Also note that slightly different notation is used here compared with *Hottovy and Stechmann* [2015a]. Here, the notation for b was chosen to be simplest for the continuum model (S1), whereas in *Hottovy and Stechmann* [2015a] the notation was chosen to be simplest for the discrete version of the model. In particular, the parameters b and \tilde{b} here correspond with b_0 and b , respectively, of *Hottovy and Stechmann* [2015a]. Furthermore, the parameter \tilde{D} here corresponds with D_* of *Hottovy and Stechmann* [2015a]. Also note that Eq. 1 of *Hottovy and Stechmann* [2015a] includes a typo, since the D_* in Eq. 1 there is

not the same as the D_* that appears in Eq. 2 there (in the same way that, here, D in (S1) is different from \tilde{D} in (S4)).

Text S3. Stationary Distribution

To analyze the model solutions, the stationary distribution of the model is used. As described by *Hottovy and Stechmann* [2015a], the stationary state of each Fourier mode of $q(x, y)$ is a Gaussian distribution with mean 0 and variance $\tilde{D}^2/(2c_{k,l})$, where

$$c_{k,\ell} = \tilde{b} \left[4 + (\tilde{b}\tau)^{-1} - 2 \cos\left(\frac{2\pi k \Delta x}{L}\right) - 2 \cos\left(\frac{2\pi \ell \Delta y}{L}\right) \right]. \quad (\text{S5})$$

For $k = \ell = 0$, however, the mean is not 0 but is instead $F\tau N^2$. The wavenumbers k and l correspond to the x and y spatial coordinates, respectively. Therefore, to sample the stationary distribution, one draws an independent Gaussian random variable for each Fourier mode, with variance $\tilde{D}^2/(2c_{k,l})$ for Fourier wavenumber (k, l) . Taking an inverse Fourier transform then leads to the solution $q(x, y)$ in physical space.

In short, the stationary distribution of the model is known analytically and can therefore be sampled by drawing independent Gaussian random variables. For example, the sample solutions in Fig. 1 are created numerically by drawing pseudorandom numbers for each Fourier mode; in this way, no time-stepping is needed to create Fig. 1. Furthermore, the stationary distribution can also be studied analytically, as in the formulas for $\bar{\sigma}$ and χ in Eqs. 3 and 4.

Text S4. Derivation of Formula for Mean Cloud Area Fraction

The mean cloud area fraction, $\bar{\sigma}$, can be computed analytically in the following way. First, for a single lattice site (x, y) or (i, j) , the value of $E[\sigma(x, y, t)]$, which we will denote

as $E[\sigma_{i,j}]$, can be computed as

$$E[\sigma_{i,j}] = 1 \cdot P(\sigma_{i,j} = 1) + 0 \cdot P(\sigma_{i,j} = 0) \quad (\text{S6})$$

$$= P(q_{i,j} > 0) \quad (\text{S7})$$

$$= \int_0^\infty \frac{1}{\sqrt{2\pi\text{Var}(q_{i,j})}} \exp\left(-\frac{(\xi - F\tau)^2}{2\text{Var}(q_{i,j})}\right) d\xi \quad (\text{S8})$$

$$= \frac{1}{2} \left[1 + \text{erf}\left(\frac{\tau F}{\sqrt{2\text{Var}(q_{i,j})}}\right) \right]. \quad (\text{S9})$$

In this calculation, we have used the definition of $\sigma_{i,j}$ from Eq. 2 as a Heaviside function in terms of $q_{i,j}$, and we have used the fact that $q_{i,j}$ is a Gaussian random variable with mean $F\tau$ and variance [Hottovy and Stechmann, 2015a]

$$\text{Var}(q_{i,j}) = E[(q_{i,j} - \tau F)^2] = \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{\ell=0}^{N-1} \frac{\tilde{D}^2}{2c_{k,\ell}}, \quad (\text{S10})$$

where the form of $c_{k,\ell}$ was given in (S5). Finally, the mean cloud fraction $\bar{\sigma}$ is defined as the average cloudiness over all lattice sites in the domain:

$$\bar{\sigma} = E \left[\frac{1}{N^2} \sum_{i,j} \sigma_{i,j} \right] = \frac{1}{N^2} \sum_{i,j} E[\sigma_{i,j}] = E[\sigma_{i,j}] \quad (\text{S11})$$

which reduces to $E[\sigma_{i,j}]$, which was given above in (S9) and reported in the main text as Eq. 3. Note that this formula corresponds to a particular finite grid spacing, with value Δx , not to a continuum limit. The dependence of $\bar{\sigma}$ on Δx is through $\text{Var}(q_{i,j})$, which, in turn, depends on Δx through $c_{k,\ell}$, as shown in (S5) and (S10).

Text S5. Methods for model cloud scenes in Figure 1

In Fig. 1e-h, model cloudiness is illustrated, and it is computed from the model variable $q(x, y)$ in the following way. Rather than plotting the binary variable $\sigma(x, y)$, which was defined in Eqn. 2 as $\sigma(x, y) = H(q(x, y))$ in terms of the Heaviside function H , a smoothed

version of $\sigma(x, y)$ is plotted to allow a more realistic transition between cloudy ($\sigma = 1$) and cloud-free ($\sigma = 0$). For the smoothed version, we define a point (x, y) to be completely cloudy if $q(x, y) > +0.33$ mm, cloud-free if $q(x, y) < -0.33$ mm, and partially cloudy otherwise. Cloudy points are displayed as a bluish-white color, cloud-free points are displayed as a bluish-black color, and partially cloudy points are displayed using a linear transition between these two extremes in red-green-blue (RGB) space. The bluish-white and bluish-black colors are used instead of white and black in order to be more comparable to the colors in the satellite images in Fig. 1a-d.

The parameter values used in this figure are (e) $D = 1.55$ mm km h^{-1/2}, $F = 0.12$ mm d⁻¹, (f) $D = 1.94$ mm km h^{-1/2}, $F = 0.048$ mm d⁻¹, (g) $D = 1.55$ mm km h^{-1/2}, $F = -0.12$ mm d⁻¹, and (h) $D = 11.62$ mm km h^{-1/2}, $F = -0.72$ mm d⁻¹. These values are also illustrated by the four circles in Fig. 2a of the main text.

Text S6. Asymptotic Formulas for $\text{Var}(q_{i,j})$, $\bar{\sigma}$, and χ in Limit of Small Δx

In this section, a derivation is presented for the asymptotic formula for $\text{Var}(q_{i,j})$, in terms of the model parameters, in the limit of small grid spacing Δx . The result was presented in the main text below Eq. 3. From this, one can then obtain simplified formulas for $\bar{\sigma}$, and χ , which were presented in the main text in Eqs. 3 and 4.

The starting point is the exact formula for $\text{Var}(q_{i,j})$:

$$\text{Var}(q_{i,j}) = \frac{1}{L^2} \sum_{k=-N/2}^{N/2-1} \sum_{\ell=-N/2}^{N/2-1} \frac{D^2}{2c_{k,\ell}}, \quad (\text{S12})$$

which was presented earlier in (S10), and where the definition of $c_{k,\ell}$ is

$$c_{k,\ell} = \frac{4b}{\Delta x^2} - \frac{2b}{\Delta x^2} (\cos(2\pi k \Delta x / L) + \cos(2\pi \ell \Delta x / L)) + \frac{1}{\tau}, \quad (\text{S13})$$

which was presented earlier in (S5). These formulas have been rewritten in several ways, compared to their presentation in (S10) and (S5), in order to facilitate the approximations described below. For instance, the sum is now written in terms of k and l which run from $-N/2$ to $N/2 - 1$ instead of 0 to $N - 1$, in order to center the sum about the origin for the asymptotics. Also, the factor \tilde{D}^2/N^2 from (S10) has been rewritten here as D^2/L^2 , using the relations $\tilde{D} = D/\Delta x$ and $L = N\Delta x$, to explicitly show the dependence on Δx . Similarly, we have used parameter b instead of $\tilde{b} = b/(\Delta x)^2$ to explicitly show the dependence on Δx .

The asymptotic derivation involves two steps: (i) a Taylor expansion of (S13) and (ii) an approximate formula for the infinite sum in (S12).

First, a Taylor expansion of (S13), assuming $\Delta x/L \ll 1$, leads to

$$c_{k,\ell} \approx b \left(\frac{2\pi k}{L} \right)^2 + b \left(\frac{2\pi \ell}{L} \right)^2 + \frac{1}{\tau},$$

from which the variance can be approximated as

$$\text{Var}(q_{i,j}) \approx \sum_{k=-N/2}^{N/2-1} \sum_{\ell=-N/2}^{N/2-1} \frac{a}{c + k^2 + \ell^2}, \tag{S14}$$

where

$$a = \frac{D^2}{8\pi^2 b}, \quad c = \frac{L^2}{4\pi^2 b \tau}. \tag{S15}$$

Second, an asymptotic formula can be found for the infinite sum on the right-hand side of (S14), as follows. Recall that two sequences $\{x_N\}, \{y_N\}$ are called *asymptotic* to each other as $N \rightarrow \infty$ if

$$\lim_{N \rightarrow \infty} \frac{x_N}{y_N} = 1.$$

It is denoted as $x_N \sim y_N$. Also recall the following well-known calculus result: Let $f(x)$ be a continuous function which is either increasing for all large x or decreasing for all

large x ; provided $f(x) \sim f(x+1)$ as $x \rightarrow \infty$ and $\int_1^N f(x) dx \rightarrow \infty$ as $N \rightarrow \infty$, then

$$\sum_{k=1}^N f(k) \sim \int_1^N f(x) dx.$$

Note that one can also prove two-dimensional versions of this theorem for use on two-dimensional integrals as will be analyzed here. For the sum in (S14), an integral for comparison is

$$\int_1^{N/2} \int_1^{N/2} \frac{a}{c+k^2+\ell^2} dx dy,$$

which diverges logarithmically. Specifically, one can show

$$\int_1^{N/2} \int_1^{N/2} \frac{a}{c+k^2+\ell^2} dx dy \sim \frac{\pi a}{2} \log(N/2),$$

as $N \rightarrow \infty$, where a and c are shown in (S15).

Combining these two steps and summarizing, the approximate form of the variance is

$$\text{Var}(q_{i,j}) = \frac{1}{L^2} \sum_{k=-N/2}^{N/2-1} \sum_{\ell=-N/2}^{N/2-1} \frac{D^2}{2c_{k,\ell}} \quad (\text{S16})$$

$$\approx \sum_{k=-N/2}^{N/2-1} \sum_{\ell=-N/2}^{N/2-1} \frac{a}{c+k^2+\ell^2} \quad (\text{S17})$$

$$\sim \frac{D^2}{4\pi b} \log(N/2), \quad (\text{S18})$$

as $N \rightarrow \infty$. This expression was reported in the main text, below Eq. 3, in simplified form with $\log(N/2)$ replaced by $\log(N)$, since $\log(N/2) \sim \log(N)$.

To illustrate the accuracy of the approximation, Fig. S2 shows plots of $\text{Var}(q_{i,j})$ using the exact formula from (S16) and using the approximate formula from (S18). The plots are essentially the same, with only some slight deviations for larger values of D . For $\bar{\sigma}$ and χ , based on Eqs. 3 and 4 of the main text, the plots of exact and approximate

formulas are also essentially the same (not shown). Hence the approximate formula in (S18) provides accurate guidance for the parameter dependence of $\text{Var}(q_{i,j})$, $\bar{\sigma}$, and χ .

Text S7. Parameter Sensitivity Studies

In this section, we discuss the sensitivity of the results to changes in model parameters. In particular, changes in b , τ , and Δx will be discussed, since changes in D and F are already demonstrated in detail in the main text. In short, it will be shown that the results here are robust over a wide range of Δx values and over essentially any b and τ values.

Changes in Δx have very little effect on the results here, as can be seen analytically through the formulas in Eqs. 3 and 4 of the main text. In those equations, the dependence of $\bar{\sigma}$ and χ on Δx is shown to be of the form $\log(L/\Delta x)$, which varies very slowly with Δx . To illustrate this, Fig. S3 shows model snapshots with the grid spacing reduced greatly from $\Delta x = 5$ km to $\Delta x = 1$ km. In this figure, the same parameters and random numbers have been used as in Fig. 1 of the main text for all Fourier modes that are shared in common, and the contribution of the newly resolved features have been incorporated using higher Fourier modes to give a physical grid resolution of $\Delta x = 1$ km. The result shows some minor new small-scale features on scales of 1–5 km, but it looks essentially the same as Fig. 1 of the main text because the variance of the high-frequency Fourier modes is small.

Regarding the sensitivity of the results to changes in b and τ , note that the phase transition behavior of the present investigation is robust to changes in the model parameters. For example, this can be seen from the analytical formulas for $\bar{\sigma}$ and χ in Eqs. 3 and 4, which show how $\bar{\sigma}$ and χ depend on all model parameters. For the phase transition

behavior, the key property of Eqs. 3 and 4 is the appearance of the error function and the exponential function, both of which are sharply varying functions, as illustrated by the sharp variations in $\bar{\sigma}$ and χ in certain parts of Fig. 2. Since the functional forms of the error function and the exponential function are present in $\bar{\sigma}$ and χ for any choice of model parameters, the phase transition is also present for essentially any choice of model parameters. If, for example, the values of b and τ were changed, a phase transition would still be seen, although the locations of the phase boundaries in terms of D and F would be different. (For example, from Eq. 3 of the main text, it is seen that the parameter dependence is $\tau F \sqrt{b}/D$; therefore, if \sqrt{b} is increased by a factor of 2, then the phase boundary would occur at a value of D that is twice as large.) This behavior is analogous to the case of the Ising model [Yeomans, 1992; Christensen and Moloney, 2005], where a phase transition occurs for essentially any choice of spatial interaction coefficient J , but the value T_c of the critical temperature of the phase transition changes if J is changed.

Text S8. Observational Data

In the satellite images in Fig. 1a-d, each panel shows a region of 5° latitude x 5° longitude, but the aspect ratio of each panel is different because each panel is taken from a different latitude. Specifically, panel (a) is from 5° West, 25° South, on May 8, 2015 10:15 GMT, panel (b) is from 80° West, 45° South on May 24, 2015 15:15 GMT, panel (c) is from 95° West, 45° South on May 16, 2015 16:05 GMT, and panel (d) is from 35° West, 10° South on May 18, 2015 12:25 GMT. These latitude and longitude values describe the lower-left corner of each panel.

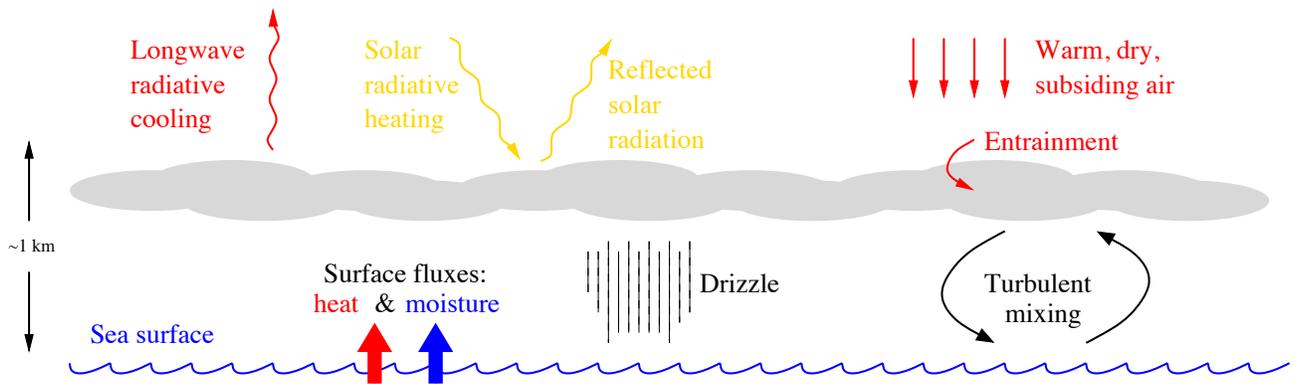


Figure S1. Schematic diagram of physical processes of the stratocumulus-topped boundary layer.

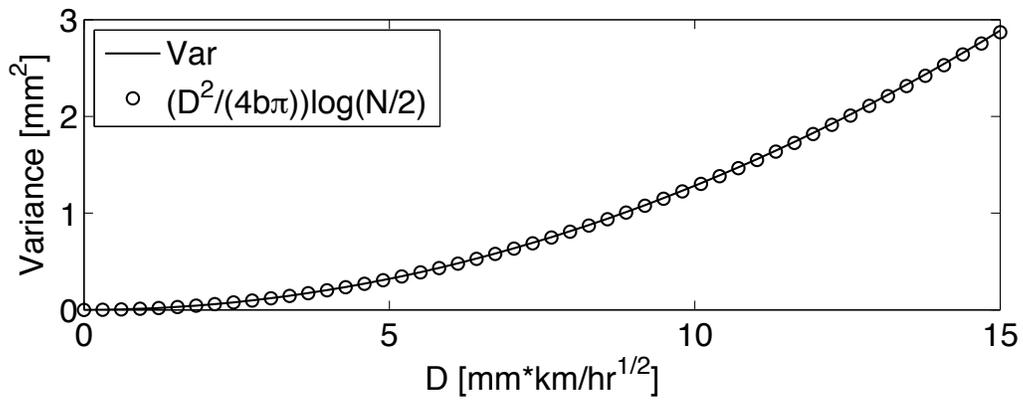


Figure S2. Comparison of exact (solid line) and approximate (circles) formulas for $\text{Var}(q_{i,j})$ as functions of parameter D , using (S16) and (S18), respectively.

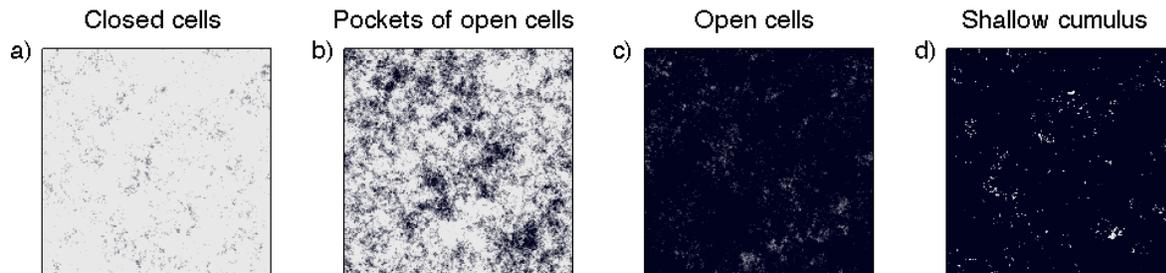


Figure S3. Same as Fig. 1e-h of the main text, except using a grid spacing of $\Delta x = \Delta y = 1$ km instead of $\Delta x = \Delta y = 5$ km.