

## Nature versus Nurture in Shallow Convection

DAVID M. ROMPS AND ZHIMING KUANG

*Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts*

(Manuscript received 1 September 2009, in final form 3 December 2009)

### ABSTRACT

Tracers are used in a large-eddy simulation of shallow convection to show that stochastic entrainment (and not cloud-base properties) determines the fate of convecting parcels. The tracers are used to diagnose the correlations between a parcel's state above the cloud base and both the parcel's state at the cloud base and its entrainment history. The correlation with the cloud-base state goes to zero a few hundred meters above the cloud base. On the other hand, correlations between a parcel's state and its net entrainment are large. Evidence is found that the entrainment events may be described as a stochastic Poisson process. A parcel model is constructed with stochastic entrainment that is able to replicate the mean and standard deviation of cloud properties. Turning off cloud-base variability has little effect on the results, which suggests that stochastic mass-flux models may be initialized with a single set of properties. The success of the stochastic parcel model suggests that it holds promise as the framework for a convective parameterization.

### 1. Introduction

The motivation for this paper is to answer a very simple question about atmospheric convection: Which more strongly affects a cloudy updraft's state in the free troposphere: its initial conditions at the cloud base or the sequence of entrainment events above the cloud base? For example, were the most buoyant parcels in a cloud destined to be so at birth (e.g., because they were born with a high equivalent potential temperature  $\theta_e$  at the cloud base) or did they become the most buoyant, because they were treated well during the course of their life (e.g., because they experienced a minimum of entrainment)? This is a question of nature versus nurture.

This question can also be phrased in terms of the variability that is observed within convecting clouds. As has been observed, clouds are highly heterogeneous entities, with temperature, humidity, buoyancy, and vertical velocity varying significantly among the parcels that make up a single cloud (e.g., Warner 1977; Jonas 1990). This variability among the cloudy parcels can be generated in two ways: from the original heterogeneity among the

cloud-base parcels or by stochastic entrainment above the cloud base, which introduces new variability.

If the existing mass-flux models of clouds are any indication, there is little agreement regarding which of these two processes is more important. In mass-flux models that use a deterministic entrainment rate—such as the constant rate used by Tiedtke (1989) or the velocity-dependent rate of Neggers et al. (2002)—the variability of cloudy updrafts in the free troposphere is owed entirely to the variability at the cloud base. In mass-flux models that use a stochastic entrainment mechanism—such as the model of Raymond and Blyth (1986)—variability in the free troposphere can be generated without any cloud-base variability.

To tease out the relative importance of nature and nurture, we present results from a large-eddy simulation (LES) of shallow convection during the Barbados Oceanographic and Meteorological Experiment (BOMEX; Holland and Rasmusson 1973). Using tracers, we are able to calculate the squared correlation coefficients—roughly, the explained variance—between the free-tropospheric parcel properties and both the parcel's cloud-base properties (nature) and the parcel's entrainment history (nurture). To further test the relative importance of nature versus nurture, we build a stochastic parcel model in which the two sources of variability can be turned off independently.

---

*Corresponding author address:* David M. Romps, Dept. of Earth and Planetary Sciences, Harvard University, 416 Geological Museum, 24 Oxford St., Cambridge, MA 02138.  
E-mail: davidromps@gmail.com

A brief outline of the subsequent sections is as follows: Section 2 describes the cloud-resolving model (CRM) used for the LES, the setup of the simulation, and the tracers method. Section 3 presents the results from the CRM. Section 4 describes the stochastic parcel model, and section 5 describes the results of three runs: with both sources of variability, with only cloud-base variability, and with only entrainment variability. Section 6 summarizes the conclusions.

## 2. The LES model

The model used here to replicate the shallow convection during the BOMEX field campaign is Das Atmosphärische Modell (DAM; Romps 2008). DAM is a three-dimensional, finite-volume, fully compressible, nonhydrostatic, cloud-resolving model. In its standard implementation, DAM uses the six-class Lin–Lord–Krueger microphysics scheme (Lin et al. 1983; Lord et al. 1984; Krueger et al. 1995). To better simulate the nonprecipitating convection observed during BOMEX, autoconversion of cloud water to rain is turned off. Instead of using an interactive radiation scheme, the radiative cooling is specified as the fixed profile used in the intercomparison study of Siebesma et al. (2003).

The simulation was performed on a doubly periodic square domain with a 12.8-km width and a model top at 3 km. The grid spacing in both the horizontal and vertical dimensions was set to 50 m. The run was initialized using the initial profiles of water vapor, temperature, and horizontal velocities as described in appendix B of Siebesma et al. (2003), with a small amount of random noise in temperature added to break translational symmetry. In particular, the temperature profile consists of a dry lapse rate from the surface to 520 m, a conditionally unstable lapse rate from 520 to 1480 m, and a stable inversion from 1480 to 2000 m. During the simulation, a large-scale vertical velocity, radiative cooling, advective moisture sink, and horizontal pressure gradient were applied as described by Siebesma et al. (2003). Unlike that intercomparison study, a bulk parameterization was used to calculate the surface fluxes off the 300.4-K ocean surface. The CRM was run for 3 h to reach a steady state followed by 5 h during which statistics were collected.

DAM is outfitted with a variety of tracers to diagnose convection, the most important of which is the purity tracer. The purity tracer  $\phi$  is a mixing ratio that records the fraction of dry air in a grid cell that was recently advected from the cloud base. Here, the “cloud base” is defined as the minimum height at which the area-averaged and time-averaged cloud cover has a local maximum; in the simulation, this gives a height of 625 m.

At every time step, the purity tracer is set to one below the cloud base and to zero everywhere above the cloud base except in the vicinity of a cloudy updraft. In this way, clouds begin their life at the cloud base with a purity  $\phi$  of 1 and they advect the purity tracer passively, all the while entraining environmental air with a purity of 0. Therefore, the purity of a cloudy updraft is equal to the fraction of air that came directly from the sub-cloud layer; 1 minus the purity is the fraction of air that was entrained.

A grid cell is defined here as a cloudy updraft when its liquid-water mass fraction is greater than  $10^{-5} \text{ kg kg}^{-1}$  and its vertical velocity is greater than  $0.5 \text{ m s}^{-1}$ . This choice is but one of many possible definitions. For example, in a study of shallow convection, Siebesma and Cuijpers (1995) define a grid cell to be a cloudy updraft if it has a positive liquid-water mass fraction and a positive vertical velocity. The definition we have chosen differs by the use of nonzero thresholds for both quantities, which avoids mistakenly identifying as updrafts any detrained air with very slight amounts of liquid. It also avoids counting cloudy Brunt–Väisälä oscillations as updrafts, which could lead to an overestimation of the cloudy-updraft mass flux. Another definition used by Siebesma and Cuijpers (1995) is that of a “cloud core,” which has a positive liquid content, vertical velocity, and buoyancy. Our standard definition of cloudy updraft avoids any requirement of buoyancy so as to allow parcels to qualify as cloudy updrafts even as they are punching through an inhibition layer or overshooting their level of neutral buoyancy.

As already mentioned, the purity tracer is zeroed out at every time step above the cloud base except in the vicinity of a cloudy updraft.<sup>1</sup> We define a grid cell to be in the vicinity of a cloudy updraft if a 350-m-wide cube, centered on the grid cell, contains any cloudy updraft. By not zeroing out the purity tracer in these protected volumes, the purity tracer is preserved in eddies of size  $\lesssim 300 \text{ m}$  even if the grid cells in the downward branch do not qualify as a cloudy updraft. A cartoon of this approach is depicted in Fig. 1, which shows an  $x$ – $z$  slice of convection at a snapshot in time.

As discussed by Romps and Kuang (2010), it is possible to encode information into parcels using additional tracers. In particular, we can encode  $N$  real numbers into cloudy updrafts by using the purity tracer plus  $N$  additional tracers. Above the cloud base, these additional tracers are treated in exactly the same way as the purity

<sup>1</sup> Note that this method of zeroing out the purity tracer differs from the one used by Romps and Kuang (2010), which is not applicable in the presence of a mean advection.

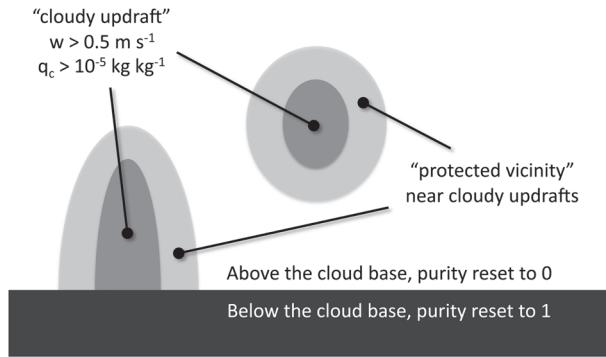


FIG. 1. At every time step, the purity tracer is set to 1 below the cloud base and to 0 outside protected volumes above the cloud base. Protected volumes are defined as grid cells for whom the surrounding 350-m-wide cube contains a cloudy updraft. A grid cell is defined as hosting a cloudy updraft if it has more than  $10^{-5} \text{ kg kg}^{-1}$  of liquid water and a vertical velocity greater than  $0.5 \text{ m s}^{-1}$ .

tracer: they are advected conservatively within protected volumes and they are zeroed out at every time step outside those protected volumes. This guarantees that cloudy updrafts entrain environmental air that has 0 values of all  $N + 1$  tracers. The only differences between the  $N + 1$  tracers are the different values they are assigned below the cloud base. As discussed above, the purity tracer is reset to 1 everywhere below the cloud base at every time step. The other  $N$  tracers are reset—below the cloud base at every time step—to whatever value we want to encode in the cloudy updrafts. Since cloudy updrafts entrain air that has 0 values of all  $N + 1$  tracers, entrainment reduces the values of all tracers by the same factor. Therefore, the original cloud-base values of the  $N$  additional tracers can be decoded by dividing by the value of the purity tracer.

Since this is a new technique, it is worth giving an example. As a very simple example, consider two tracers: a purity tracer and a “time tracer.” At every time step and everywhere below the cloud base, we set the purity tracer to 1 and the time tracer to the current model time. Now, consider a parcel in a cloudy updraft that passes through the cloud base at 1140 s. As it leaves the subcloud layer, the parcel carries with it a purity tracer equal to 1 and a time tracer equal to 1140. Let us imagine that the parcel reaches a height of 1500 m after having entrained 4 times its (dry air) mass in (dry) environmental air. Since the environmental air has 0 values of both tracers, the new mixing ratios are  $(1 \times 1 + 4 \times 0)/(1 + 4) = 0.2$  and  $(1 \times 1140 + 4 \times 0)/(1 + 4) = 228$ . In practice, we do not track parcels in the LES as they move from the cloud base to, say, 1500 m. Instead, we simply observe parcels at 1500 m. If we were to observe this parcel at 1590 s, we could divide the time tracer by the purity tracer to get  $228/0.2 = 1140$ , which tells us that it took

the parcel  $1590 - 1140 = 450 \text{ s}$  to travel there from the cloud base.

Of course, there is no reason why the additional  $N$  tracers must be set to a *uniform* value below the cloud base. Instead, we can set tracers equal to the *local* properties of the air below the cloud base. For example, we could use a “ $\theta_e$  tracer” that is set in each grid cell below the cloud base to the value of  $\theta_e$  in that grid cell. Thereby, a parcel with  $\theta_e = 350 \text{ K}$  as it passes through the cloud base will carry a purity tracer equal to 1 and a  $\theta_e$  tracer equal to 350. If the parcel is later observed above the cloud base with a purity tracer equal to 0.16 and a  $\theta_e$  tracer equal to 56, then its original cloud-base value of  $\theta_e$  can be decoded by dividing:  $56/0.16 = 350 \text{ K}$ . Similarly, a “velocity tracer” can be set to the local value of vertical velocity below the cloud base; when a parcel is observed in the free troposphere with a purity of 0.64 and a velocity tracer equal to 2.08, we can infer that it had a velocity of  $2.08/0.64 = 3.25 \text{ m s}^{-1}$  at the cloud base. In the simulations presented here, we have used tracers to record the cloud-base vertical velocity, water vapor, liquid water, temperature, potential temperature, and equivalent potential temperature.

### 3. The LES results

During the last 5 h of the LES, snapshots were recorded at 5-min intervals of the layer at 1275 m, which was chosen for its proximity to the 1250-m height studied by Neggers et al. (2002). The grid cells at this height are then plotted on axes of total water  $q_t$  and liquid-water potential temperature  $\theta_l$ , both of which are exactly conserved for adiabatic and reversible transformations. For small water-vapor mixing ratios (less than about  $0.02 \text{ kg kg}^{-1}$ ),  $\theta_l$  is roughly equal (within about 1 K) to the potential temperature of the parcel before any water vapor has been condensed. (See the appendix for a definition and derivation of  $\theta_l$ .) Plotting the 1275-m grid cells on axes of  $q_t$  and  $\theta_l$ , we obtain the Paluch diagram displayed in Fig. 2. The solid line in Fig. 2 corresponds to the mean environmental profile with the mean surface-air value in the upper left at high  $q_t$  and low  $\theta_l$ . The dashed line separates the region above, where parcels with the corresponding  $q_t$  and  $\theta_l$  would be saturated at the mean pressure at 1275 m, from the region below, where parcels would be unsaturated. The dotted line denotes the parcels whose  $q_t$  and  $\theta_l$  would make them neutrally buoyant at 1275 m; the region above the dotted line corresponds to positive buoyancy and the region below corresponds to negative buoyancy. The mean values of  $q_t$  and  $\theta_l$  at 1275 m occur where the dotted line intersects the solid line.

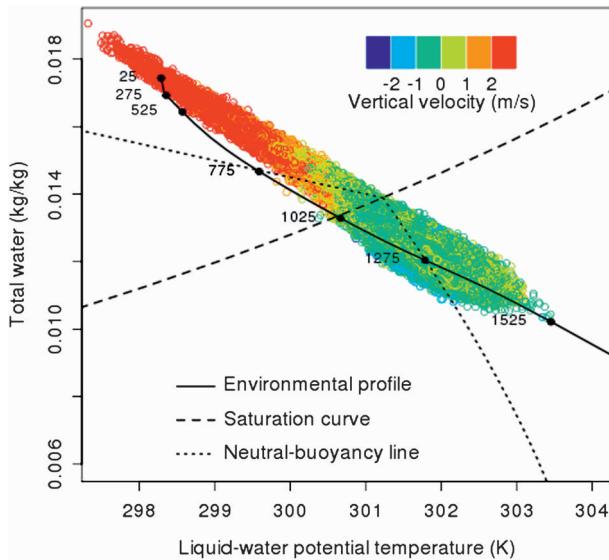


FIG. 2. The Paluch diagram at 1275 m from the cloud-resolving simulation. The circles on the mean environmental profile are labeled with the corresponding heights in meters.

The points in Fig. 2 may be thought of as comprising two groups. First, there is the nearly motionless grouping of unsaturated points roughly centered around the mean value of  $q_i$  and  $\theta_i$  at 1275 m, which corresponds to the environmental air. Second, there is the long “tail” of saturated points extending from the saturation curve all the way up to the properties of pure surface air. This long tail is indicative of a large amount of variability among cloudy parcels.

As Neggers et al. (2002) illustrated quite nicely, this variability *cannot* be reproduced using an ensemble of Lagrangian parcels initialized with the observed cloud-base values and made to entrain with a fixed, constant rate. In other words, they found that the variability in cloud-base properties was not sufficiently large to explain the variability further up in the atmosphere. This fact would seem to suggest that nurture, not nature, is responsible for the long Paluch tail. Instead, Neggers et al. (2002) further pursued the nature explanation by postulating a mechanism that would deterministically grow the cloud-base variability. In particular, they argued that the fractional entrainment rate for parcels goes like one over the vertical velocity,  $\epsilon \sim 1/w$ . This argument would imply that the more buoyant updrafts, which will move more quickly, would entrain less and so would become even more buoyant than the other updrafts. Acting between the cloud base and 1275 m, this positive feedback on buoyancy would spread the distribution of cloudy updrafts into this long tail, from low buoyancy near the dotted line to high buoyancy near the surface values. Based on the ability of the  $w$ -dependent  $\epsilon$  to reproduce

the Paluch tail, this model of entrainment has been implemented in the Integrated Forecasting System of the European Centre for Medium-Range Weather Forecasts (ECMWF) (Neggers et al. 2009).

If the explanation of Neggers et al. (2002) were correct, then there would be excellent correlations between the buoyancy of a parcel at 1275 m and the values of  $w$  and  $\theta_e$  that the parcel had at the cloud base. To test this concept, we have used the tracers that record  $w$  and  $\theta_e$  at the cloud base. By binning cloudy-updraft mass flux at 1275 m on a two-dimensional histogram with axes corresponding to 1275-m buoyancy and cloud-base  $w$ , we produce the contour plot in the left panel of Fig. 3. The same plot using cloud-base  $\theta_e$  is shown in the middle panel of Fig. 3. Instead of excellent correlations, there is almost no correlation in either of these figures: 1275-m buoyancy is correlated with cloud-base  $w$  with an  $R^2$  of 0.01 and with cloud-base  $\theta_e$  with an  $R^2$  of 0.00. The contour plots of 1275-m  $q_i$  and 1275-m  $\theta_i$  versus cloud-base  $w$  and cloud-base  $\theta_e$  (not shown) are qualitatively identical; there is no correlation.

Since this is an important point, it is worth checking that this result does not depend on the definition of cloudy updrafts. To this end, we can repeat the simulation using a more restrictive definition of cloudy updraft that includes grid cells whose liquid-water mass fraction is greater than  $10^{-5} \text{ kg kg}^{-1}$ , whose vertical velocity is greater than  $0.5 \text{ m s}^{-1}$ , and whose buoyancy is positive; this is similar to the cloud-core category defined by Siebesma and Cuijpers (1995). The results, displayed in the left and middle panels of Fig. 4, show that the lack of correlation persists even for this more restrictive class of updrafts. Again, the contour plots with 1275-m buoyancy replaced by 1275-m  $\theta_i$  or  $q_i$  exhibit the same lack of correlation, which shows, rather convincingly, that the variability in cloud-base properties is not responsible for the variability at 1275 m. In other words, nature is not responsible for the long Paluch tail.

Since cloud-base properties are uncorrelated with 1275-m properties, the variability in cloudy updrafts at 1275 m must be generated from one of two remaining sources of variability: either the random properties of the clear-air environment through which the updraft ascends or randomness in when and how much clear air an updraft entrains. We can easily rule out the first possibility by calculating the magnitude of the clear-air moisture variability. For this purpose, we define the effective water-vapor mass fraction between the cloud base and 1275 m as

$$q_{v,\text{effective}} \equiv \frac{\left( \int_{\text{cloud base}}^{1275 \text{ m}} dz q_v \rho \right)}{\left( \int_{\text{cloud base}}^{1275 \text{ m}} dz \rho \right)}.$$

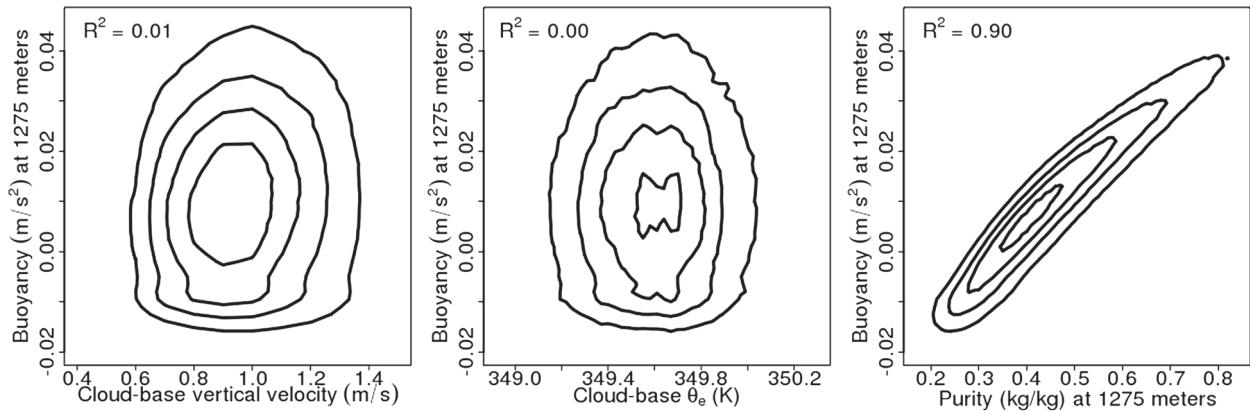


FIG. 3. Contours of 1275-m cloudy-updraft flux on axes of 1275-m buoyancy vs (left) cloud-base velocity, (middle) cloud-base equivalent potential temperature, and (right) purity.

The standard deviation of this quantity among columns that contain no cloud is  $1.5 \times 10^{-4} \text{ kg kg}^{-1}$ , which is an order of magnitude smaller than the standard deviation of total water among cloudy updrafts at 1275 m. Therefore, the variance at 1275 m cannot be explained by the variance in the environment through which the updraft ascends.

We can now use the purity tracer to confirm that the only remaining possibility—stochastic entrainment—is responsible for the variance at 1275 m. Binning the cloudy-updraft mass flux on a two-dimensional histogram with axes of buoyancy at 1275 m and purity at 1275 m, we find the excellent correlation shown in the right panel of Fig. 3. At 1275 m, buoyancy and purity are correlated with an  $R^2$  of 0.90. Therefore, the free-tropospheric buoyancy is highly correlated with net entrainment, as measured by purity. The same contour plot for the more restrictive cloud-core definition of cloudy updraft is shown in the right panel of Fig. 4, which has an  $R^2$  of 0.88. Correlations of 1275-m  $q_i$  and 1275-m  $\theta_i$  with

1275-m purity have similarly high  $R^2$  values for both definitions of cloudy updrafts. We conclude, therefore, that nurture, not nature, is responsible for the variability in cloudy updrafts at 1275 m. In other words, the fate of a parcel depends not on the conditions of its birth but rather on the stochastic entrainment events experienced during its lifetime.

Nevertheless, there must be a height sufficiently close to the cloud base where nature, not nurture, explains more of the variance in cloudy updraft properties. For example, right above the cloud base, the buoyancy of a parcel must correlate strongly with its properties at the cloud base. It is of interest to ask at what height the dependence switches from nature to nurture. For this purpose, we can produce histograms like those in Fig. 3 for every height and record the  $R^2$ . Roughly speaking, this tells us how much of the cloudy-updraft variance can be explained by nature and nurture. The left panel of Fig. 5 shows profiles of the  $R^2$  between buoyancy at each height and three other quantities: purity at that height,

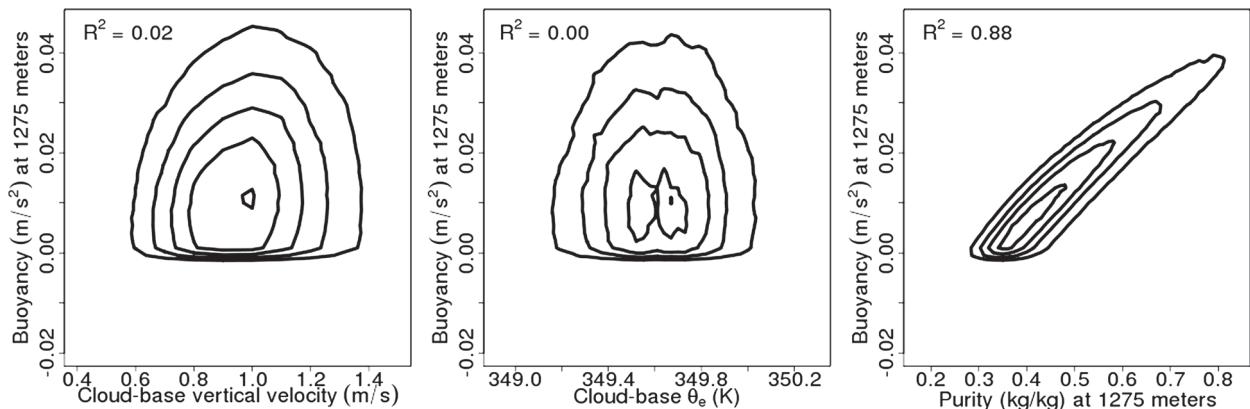


FIG. 4. As in Fig. 3, but with the more restrictive cloud-core definition of cloudy updrafts.

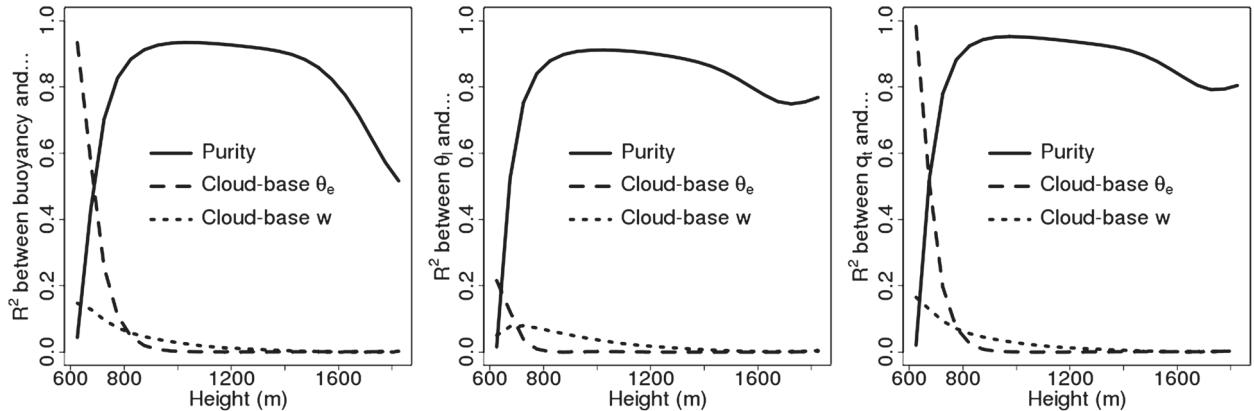


FIG. 5. (left) Value of  $R^2$  at each height for buoyancy vs cloud-base vertical velocity (dotted), buoyancy vs cloud-base  $\theta_e$  (dashed), and buoyancy vs purity (solid). (middle), (right) As at left, but for  $\theta_t$  and  $q_t$ , respectively.

velocity at the cloud base, and  $\theta_e$  at the cloud base. The middle and right panels of Fig. 5 for  $\theta_t$  and  $q_t$  are similar. We see that cloud-base  $\theta_e$  is correlated well with buoyancy and total water only in the first 100 m above the cloud base. There is no height where cloud-base  $w$  correlates well with buoyancy and total water. And neither cloud-base  $\theta_e$  nor cloud-base  $w$  correlates well with  $\theta_t$  at any height. In contrast, the purity correlates very well with buoyancy,  $\theta_t$ , and  $q_t$  at all heights greater than 100 m above the cloud base. Since the purity is a measure of a parcel's entrainment history, this tells us that the entrainment history is the most important predictor of a parcel's state throughout the vast majority of the cloud layer.

Given the importance of the entrainment process, it is worth trying to glean more about it from the LES results. To do so, we focus on the mass flux of cloudy updrafts that are undiluted, which are defined here to have a purity greater than 80%. The choice of this threshold is arbitrary aside from a desire to be consistent with Romps and Kuang (2010). As shown in Fig. 6, the undiluted flux in the CRM decays exponentially all the way up to the inversion at around 1500 m. A best-fit exponential to the undiluted flux below the inversion yields an  $e$ -folding decay distance of  $\lambda \approx 200$  m. Romps and Kuang (2010) also found an exponential decay of undiluted flux in deep convection, but with an  $e$ -folding decay distance of about 500 m.

An exponential decay with height implies that an undiluted parcel has a probability  $dz/\lambda$  of becoming diluted while traversing the distance  $dz$ . Let us assume that all entrainment events entrain a mass of environmental air at least as large as 25% of the parcel's mass. In that case, an entrainment event will knock a parcel out of the undiluted category—as defined by  $>80\%$  purity—regardless of just how much air is entrained. Therefore,

we can interpret  $dz/\lambda$  as the probability that any parcel experiences an entrainment event while ascending a distance  $dz$ . In other words, entrainment events behave like a stochastic Poisson process. The probability that a parcel travels a distance  $z$  without entraining is

$$\lim_{dz \rightarrow 0} \left(1 - \frac{dz}{\lambda}\right)^{z/dz} = e^{-z/\lambda}.$$

The probability that a parcel has its next entrainment event after a distance between  $z$  and  $z + dz$  is  $e^{-z/\lambda} - e^{-(z+dz)/\lambda} = (dz/\lambda)e^{-z/\lambda}$ . This result is similar in spirit to the treatment of entrainment events in the one-dimensional mixing model of Krueger et al. (1997), but

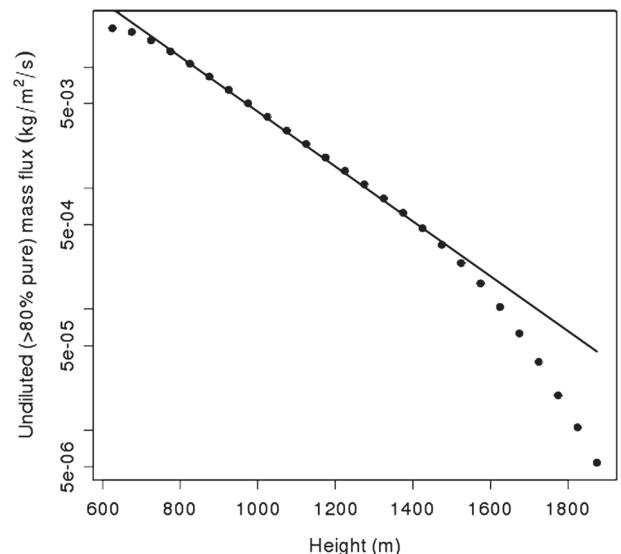


FIG. 6. The undiluted ( $>80\%$  pure) cloudy-updraft flux (circles) from the CRM and the flux-weighted best-fit exponential (line) of  $0.005 \text{ m}^{-1}$  (or  $1/200 \text{ m}$ ).

the evidence here indicates that the distance between entrainment events should be drawn from an exponential distribution (as befitting a Poisson process), not from a Poisson distribution. Using this exponential distribution, the mean distance between entrainment events is

$$\int_0^{\infty} z \frac{dz}{\lambda} e^{-z/\lambda} = \lambda.$$

Therefore,  $\lambda \approx 200$  m is the mean distance between entrainment events.

This suggests that convection may be modeled using an ensemble of parcels that are initialized identically and subjected to entrainment events using a Monte Carlo method. Different effective entrainment rates and levels of neutral buoyancy fall out of this approach naturally: lucky parcels entrain relatively little and ascend relatively high, while unlucky parcels suffer the opposite fate. A convective parameterization built around this principle would be relatively straightforward to implement, especially in the case of nonprecipitating convection where there is no parcel-to-parcel communication in the form of evaporating rain. As a step in this direction, we construct a stochastic parcel model in the next section and verify its behavior against the LES results in section 5.

#### 4. Stochastic parcel model

The parcel model that is used here bears close similarity to the model described by Romps and Kuang (2010). In short, the model describes the evolution of a spherical parcel of air whose properties are described by its height  $z(t)$ , volume  $V(t)$ , temperature  $T(t)$ , density  $\rho(t)$ , water vapor mass fraction  $q_v(t)$ , liquid-water mass fraction  $q_l(t)$ , and vertical velocity  $w(t)$ . A numerical model integrates the ordinary differential equations that evolve these quantities in time. To close the problem, the model is also given environmental profiles of density  $\rho_e(z)$ , water-vapor mixing ratio  $q_{ve}(z)$ , pressure  $p_e(z)$ , and temperature  $T_e(z)$ . At all times, the pressure of the parcel is set to the pressure of the environment:  $p(t) = p_e[z(t)]$ . The parcel model is run for 1 h, which is enough time for any parcel to have settled down at its level of neutral buoyancy. The appendix of Romps and Kuang (2010) provides further details of the model.

Since we are interested in modeling nonprecipitating shallow convection, we turn off precipitation in this version of the parcel model. Also, instead of using a constant entrainment rate  $\epsilon$ , we model  $\epsilon$  as a random variable that is zero except during delta-function bursts that correspond to entrainment events. For a sufficiently small  $\Delta t$ , we can model the integral of  $\epsilon$  over the time interval

$[t, t + \Delta t]$  as a random variable itself. We can write that integral as

$$\int_t^{t+\Delta t} dt' \epsilon(t') = XF,$$

where  $X$  is a random variable that takes the value of 1 if an entrainment event occurs in the interval  $[t, t + \Delta t]$  and 0 otherwise, and  $F$  is a random variable equal to the fractional mass of air entrained during the event. The probability distribution of  $X$  is defined by  $P(X = 1) = \Delta t |w| / \lambda$  and  $P(X = 0) = 1 - \Delta t |w| / \lambda$ . In practice,  $\Delta t$  is the model time step and  $F$  needs to be evaluated only when  $X = 1$ .

Although the exponential decay of undiluted flux has motivated the choice of distribution for  $X$ , there is little to inform the choice of distribution for  $F$ . Of course, for the integral of  $\epsilon$  over  $\Delta t$  to be finite,  $P(F = f)$  must go to zero as  $f$  goes to infinity, but this is not much of a constraint. For simplicity, then, we use an exponential distribution,

$$P(F = f) = \frac{1}{\sigma} e^{-f/\sigma}, \quad (1)$$

where  $\sigma$  is the mean ratio of entrained mass to parcel mass in an entrainment event. Since  $X$  and  $F$  are uncorrelated, the mean fractional entrainment per distance is  $\sigma/\lambda$ .

In practice, the entrainment rate is modeled using a Monte Carlo method. At each time step, the probability of an entrainment event is  $|w|\Delta t/\lambda$ . A random number generator is used to pick a number from a uniform distribution between 0 and 1. If the value is greater than  $|w|\Delta t/\lambda$ , then  $\epsilon$  is set to 0, and the evolution of the parcel for that time step proceeds without entrainment. Otherwise, if the value is less than  $|w|\Delta t/\lambda$ , an entrainment event is begun. In that case, a random number generator is used to select a number  $f$  from the exponential distribution in (1). Ideally, we would use a form of time splitting to entrain the entire mass of air during that one time step. In practice, we set  $\epsilon = (fM_0/M)/(N\Delta t)$  during the next  $N$  time steps, where  $N$  is chosen to be as small as possible without triggering numerical instability. Here,  $M$  is the mass of the parcel, and  $M_0$  is the mass of the parcel at the beginning of the  $N$  time steps; the ratio  $M_0/M$  guarantees that the total mass entrained is equal to  $fM_0$ . After the  $N$  time steps,  $\epsilon$  is reset to 0 until the next time that  $X = 1$ , at which point the process repeats. Since  $|w|N\Delta t$  is, in general, much smaller than  $\lambda$ , the fraction of time a parcel spends entraining is negligibly small, so this procedure does a good job of approximating instantaneous entrainment events.

To initialize the parcel model, we use cloud-base conditions saved from the CRM. During the LES, snapshots of the cloud-base layer were saved every 5 min. These snapshots include the pressure, temperature, water vapor, liquid water, and vertical velocity. Each slice contains  $256 \times 256 = 65\,536$  grid cells, which, at 5-min intervals over 5 h, provides about 4 million initial conditions. This allows us to evolve 4 million parcels, each initialized with a volume equal to the volume of a grid cell. The background values of pressure, density, temperature, and humidity through which the parcel travels are taken from profiles averaged over the 5-h LES. Each parcel is integrated for 1 h, which is a sufficient length of time for cloudy updrafts, which travel at a minimum speed of  $0.5 \text{ m s}^{-1}$ , to reach up past the inversion, which begins at 1500 m.

To find the values of  $\lambda$  and  $\sigma$  that give the best fit to the CRM results, we run the parcel model 141 times using different combinations of  $\lambda$  and  $\sigma$ , each time simulating four million parcels. The mean entrainment distance  $\lambda$  was set to values from 10 m to 10 km in powers of  $2^{1/2}$ . Similarly, the mean entrainment fraction  $\sigma$  was set to values from 0.01 to 10 in powers of  $2^{1/2}$ . We seek values for  $\lambda$  and  $\sigma$  that will give the best agreement to the total flux of cloudy updrafts and what we will call the ‘‘purity flux,’’ which we define as the flux of purity tracer within cloudy updrafts. For a single cloudy updraft with a mass flux of dry air  $M$ , the purity flux will be  $\phi M$ . Note that the

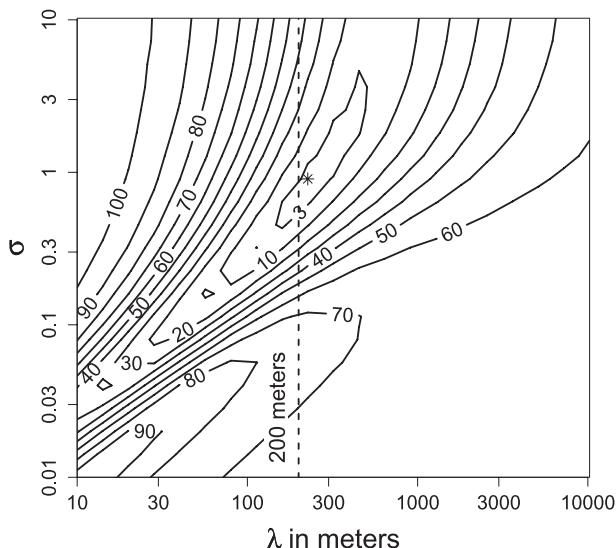


FIG. 7. The objective function in Eq. (2) normalized by its minimum value, which occurs at the location of the asterisk. The best fit has a mixing length of 226 m and an entrainment fraction of 0.91.

average purity flux,  $\overline{\phi M}$ , is not the same as the undiluted flux,  $\overline{M\mathcal{H}(\phi - 80\%)}$ , where  $\mathcal{H}$  is the Heaviside step function. To optimize  $\lambda$  and  $\sigma$ , we define the objective function to be the sum of the squared residuals on the total and purity fluxes between the parcel model and the CRM:

$$\int_{\text{cloud base}}^{\text{model top}} dz [(\overline{M}|_{\text{parcel model}} - \overline{M}|_{\text{CRM}})^2 + (\overline{\phi M}|_{\text{parcel model}} - \overline{\phi M}|_{\text{CRM}})^2], \quad (2)$$

where the bar denotes an average over cloudy updrafts. A contour plot of this objective function is shown in Fig. 7. The best fit (i.e., the minimum of the objective function) is marked by an asterisk. This best fit occurs for  $\lambda = 226 \text{ m}$  and  $\sigma = 0.91$ , which corresponds to a mean entrainment rate of  $4.0 \times 10^{-3} \text{ m}^{-1}$ . In other words, the mean distance that a parcel travels between entrainment events is 226 m. The average mass of environmental air that the parcel entrains is 0.91 times the mass of the parcel, which means that cloudy and environmental air mix in nearly a 1:1 ratio. Note that the best fit of  $\lambda = 226 \text{ m}$  agrees well with the value diagnosed from the undiluted flux in the LES. It is also reassuring to note that the mean entrainment rate obtained here is similar to the rate of  $3\text{--}4 \times 10^{-3} \text{ m}^{-1}$  diagnosed for this simulation using direct measurement (Romps 2010).

Before we move on, let us make sense of the overall shape of the objective function in Fig. 7. First, note that we expect stochastic entrainment to give the same results

as constant entrainment in the limit of small  $\lambda$  for fixed  $\sigma/\lambda$ . Therefore, in the small- $\lambda$  limit, we expect the objective function to have a relatively flat valley oriented along  $\sigma \propto \lambda$ . The flatness is expected, because the entrainment process is invariant with respect to  $\lambda$  in the small- $\lambda$  limit. The valley should be oriented along  $\sigma \propto \lambda$ , because there should be a single fractional entrainment rate that gives the best fit for the constant entrainment model, and the fractional entrainment rate is equal to  $\sigma/\lambda$ . In Fig. 7, we see that there is such a valley running from  $(\lambda, \sigma) = (10, 0.03)$  to  $(100, 0.3)$  with small fractional changes in the height of the valley floor. On the other hand, we see that the valley deviates from a straight line for  $\lambda > 100 \text{ m}$ . For example, the valley follows  $\sigma \propto \lambda^{1.6}$  in the vicinity of the best fit. This deviation from a one-to-one line tells us that stochastic entrainment is different from constant entrainment for  $\lambda \gtrsim 100 \text{ m}$ . And, in this regime, the value of  $\lambda$  matters: the valley floor drops an order of magnitude from  $\lambda = 100 \text{ m}$  to the best-fit value at  $\lambda = 226 \text{ m}$ .

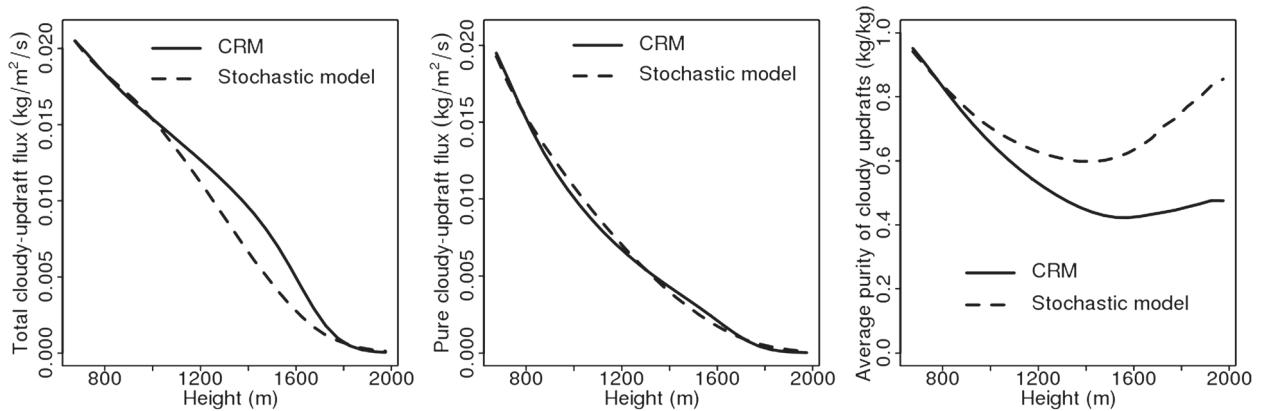


FIG. 8. The (left) total cloudy-updraft flux, (middle) cloudy-updraft purity flux, and (right) average purity of cloudy updrafts for the cloud-resolving simulation (solid) and the stochastic parcel model (dashed).

From here on, we will use the best-fit values of  $\sigma$  and  $\lambda$  in the stochastic parcel model. As shown in Fig. 8, the purity flux (middle panel) is replicated rather well by the parcel model with these values, but the total flux (left panel) is underestimated. Dividing the purity flux by the total flux gives the average purity of cloudy updrafts (right panel). Since the parcel model underestimates the total flux, the average purity is too high. This suggests that the parcel model may be having difficulty accelerating diluted parcels into the cloudy-updraft classification, which requires a vertical velocity greater than  $0.5 \text{ m s}^{-1}$ . This may be caused by using too large a drag force. It is quite plausible that individual parcels ascending within clouds do not experience the same drag force as they would ascending through still air, as we have modeled them here. Nevertheless, we have not attempted to tune the drag force because it would involve a good deal of work and it would not affect the conclusions.

## 5. Stochastic parcel results

One of the benefits of a parcel model is that it can be modified to turn off either cloud-base variability or entrainment variability to explore their relative importance. We will study three versions of the stochastic parcel model. The first is the standard run of the parcel model using the best-fit values of  $\lambda$  and  $\sigma$ , which we refer to as FullVar because it includes the full variability from both initial conditions and entrainment. In the cloud-base variability (CBVar) run, the stochastic entrainment is replaced with a constant entrainment rate, so the only source of variability comes from the initialization at the cloud base. In the entrainment variability (EntVar) run, the parcels are initialized with a fixed set of properties, so the only source of variability comes from the stochastic entrainment.

Since we would like to compare the Paluch diagrams among the FullVar, CBVar, and EntVar runs, it is important that the parcels in each of those runs reach the height of 1275 m. In the CBVar run with a constant entrainment rate of  $4.0 \times 10^{-3} \text{ s}^{-1}$ , none of the parcels reaches 1275 m, which should not come as a surprise. In the CRM and in the stochastic parcel model, the parcels that manage to reach 1275 m are those that have entrained at a rate far less than average. In the CBVar simulation, where we use a constant and continuous entrainment rate, we must tune the entrainment rate to get parcels to the desired height. We found that using an entrainment rate of  $\epsilon = 2.0 \times 10^{-3} \text{ s}^{-1}$  was sufficiently low to produce cloudy updrafts at 1275 m. In the EntVar

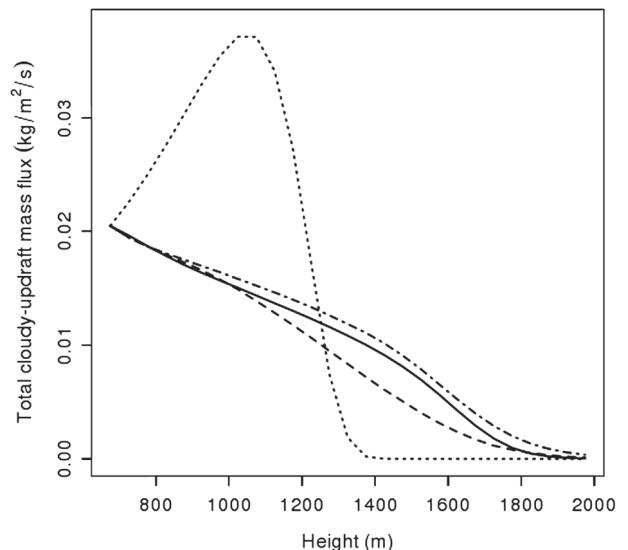


FIG. 9. The total cloudy-updraft mass flux for the CRM (solid), FullVar (dashed), EntVar (dotted-dashed), and CBVar (dotted) simulations.

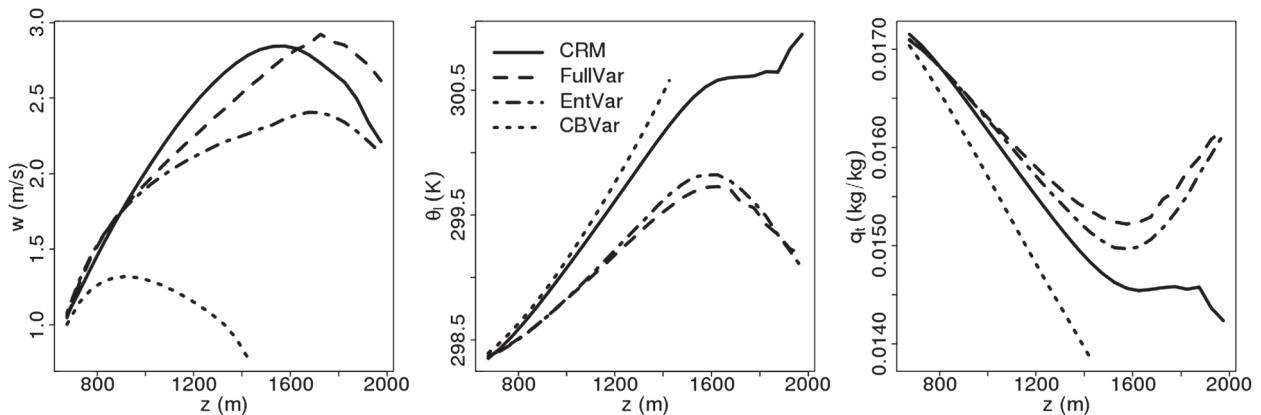


FIG. 10. The mean profiles of (left) vertical velocity, (middle) liquid-water potential temperature, and (right) total water for cloudy updrafts in the CRM (solid), FullVar (dashed), EntVar (dotted-dashed), and CBVar (dotted).

simulation, we use the same  $\sigma$  and  $\lambda$  as in FullVar, but we initialize all parcels using the mean properties (pressure, temperature, water vapor, liquid water, and vertical velocity) of the cloudy updrafts at the cloud base in the CRM. This process produces cloudy updrafts at 1275 m without any difficulty.

As in the CRM, we define parcels as cloudy updrafts in FullVar, EntVar, and CBVar if they have a liquid-water mixing ratio greater than  $10^{-5}$  kg kg $^{-1}$  and a vertical velocity greater than  $0.5$  m s $^{-1}$ . Therefore, parcels contribute to the cloudy-updraft averages only at heights where they satisfy those two conditions. Figure 9 shows the total cloudy-updraft mass flux for the CRM (solid), FullVar (dashed), EntVar (dotted-dashed), and the CBVar (dotted) simulations. The curves for the CRM and FullVar are the same as in Fig. 8. From the other two curves, we see that the EntVar simulation does a very good job of matching the CRM, but the CBVar simulation does a very poor job. The CBVar curve is similar to what one would expect from a set of continuously entraining parcels with similar initial conditions: first, the

parcels entrain, causing an increase in the total mass flux; then, near their level of neutral buoyancy, the mass flux quickly decreases to zero. Needless to say, it is qualitatively different from the mass flux observed in the CRM.

The flux-weighted means of three quantities— $w$ ,  $\theta_l$ , and  $q_t$ —are shown in Fig. 10 for the CRM (solid), FullVar (dashed), EntVar (dotted-dashed), and CBVar (dotted) simulations. The mean vertical velocity (left panel) is replicated well by both the FullVar and EntVar simulations. As might be expected from the very different mass-flux profile for CBVar in Fig. 9, the mean velocity for CBVar is also quite different from the CRM. Below the inversion, all three models match the mean CRM profiles of  $\theta_l$  and  $q_t$  qualitatively: as parcels get diluted with height, they entrain environmental air with a higher  $\theta_l$  and a lower  $q_t$ . As we have already seen from Fig. 8, the purity is positively biased; accordingly, in Fig. 10,  $\theta_l$  is negatively biased, and  $q_t$  is positively biased.

Since the original motivation for the stochastic entrainment was the large and unexplained heterogeneity of cloudy updrafts, we should also look at the standard

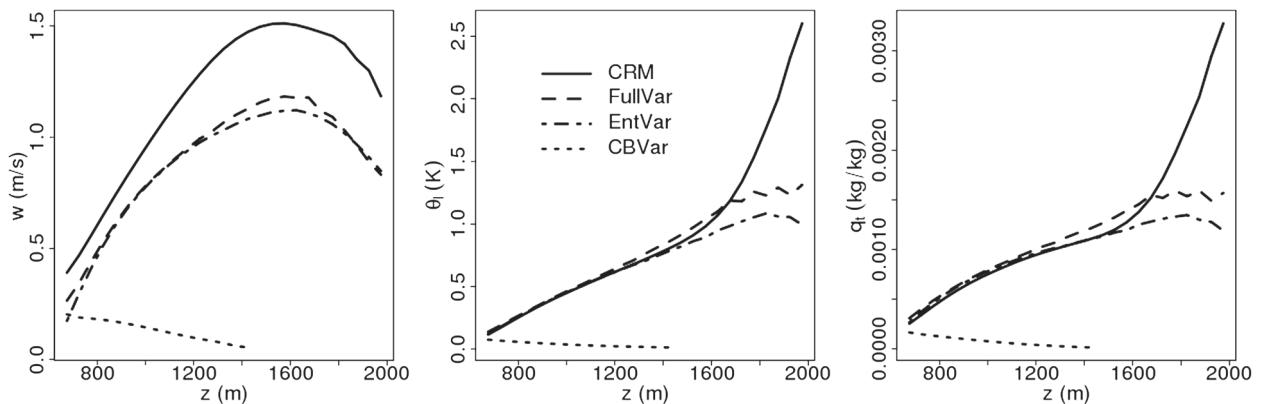


FIG. 11. The standard deviation profiles of (left) vertical velocity, (middle) liquid-water potential temperature, and (right) total water for cloudy updrafts in the CRM (solid), FullVar (dashed), EntVar (dotted-dashed), and CBVar (dotted).

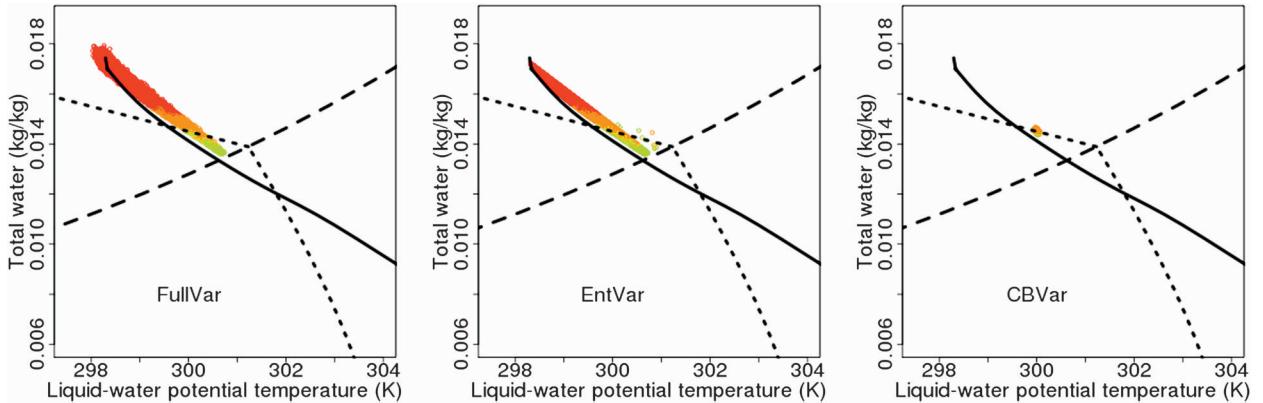


FIG. 12. The Paluch diagram of cloudy updrafts at 1275 m for (left) FullVar, (middle) EntVar, and (right) CBVar.

deviations of  $w$ ,  $\theta_l$ , and  $q_l$  among the different model runs. Figure 11 shows these standard deviations at each height in the CRM (solid), FullVar (dashed), CBVar (dotted), and EntVar (dotted–dashed) parcel models. Both the FullVar and EntVar simulations do a very good job of capturing the variability in the CRM below the inversion. The fact that EntVar does as well as FullVar further demonstrates that cloud-base variability is irrelevant to variability in the cloud layer. Indeed, the CBVar simulation replicates virtually none of the variability seen in the CRM.

Given these results, it is not surprising that FullVar and EntVar replicate well the 1275-m Paluch diagram, while CBVar does not. As shown in the left and middle panels of Fig. 12, FullVar and EntVar both reproduce the long Paluch tail. On the other hand, the Paluch diagram for CBVar, shown in the right panel of Fig. 12, consists of a tight blob of values that bears no resemblance to Fig. 2.

## 6. Conclusions

The preceding results show that nurture, not nature, is responsible for the variability among cloudy updrafts in shallow convection. We have seen that a parcel's state above the cloud base has virtually no correlation with its original state at the cloud base, which rules out theories of updraft variability that rely on an amplification of cloud-base variability through, for example, state-dependent entrainment rates. However, a parcel's state does correlate very well with its net entrainment, which suggests that stochastic turbulence is providing the observed variability. As to the form of this mixing, we have found evidence that entrainment behaves like a stochastic Poisson process.

We have found that a stochastic parcel model reproduces well the variability observed in the CRM. On

the other hand, a parcel model with a constant and continuous entrainment rate produces almost none of the observed variability. In fact, the cloud-base variability may be eliminated from the stochastic parcel model with almost no effect on the variability of parcels above the cloud base, which suggests that modeling convection as an ensemble of realizations of a single parcel with Monte Carlo entrainment may be a promising approach to convective parameterization.

*Acknowledgments.* This research was partially supported by the Office of Biological and Environmental Research of the U.S. Department of Energy under Grant DE-FG02-08ER64556 as part of the Atmospheric Radiation Measurement Program and by the National Science Foundation under Grant ATM-0754332. The numerical simulations were run on the Odyssey cluster supported by the FAS Research Computing Group.

## APPENDIX

### Definition and Derivation of $\theta_l$

To derive an expression for  $\theta_l$ , we follow, almost verbatim, the derivation for  $\theta_e$  given by Romps and Kuang (2010). We begin with the specific entropies of dry air, water vapor, liquid water, and solid water, denoted by subscripts  $a$ ,  $v$ ,  $l$ , and  $s$ , respectively:

$$\begin{aligned} s_a &= c_{pa} \log(T/T_{\text{trip}}) - R_a \log(p_a/p_{\text{trip}}), \\ s_v &= c_{pv} \log(T/T_{\text{trip}}) - R_v \log(p_v/p_{\text{trip}}) + s_{0v}, \\ s_l &= c_{vl} \log(T/T_{\text{trip}}), \quad \text{and} \\ s_s &= c_{vs} \log(T/T_{\text{trip}}) - s_{0s}. \end{aligned}$$

Here,  $c_v$  and  $c_p$  are the heat capacities at constant volume and pressure,  $p_{\text{trip}}$  and  $T_{\text{trip}}$  are the pressure and

temperature of the triple point,  $p_a$  and  $p_v$  are the partial pressures of air and vapor, and  $s_{0v}$  and  $-s_{0s}$  are the spe-

cific entropies of vapor and solid at the triple point. Using  $r$  to denote mixing ratios, the entropy per dry-air mass,

$$s_a + r_v s_v + r_l s_l + r_s s_s = (c_{pa} + r_v c_{pv} + r_l c_{vl} + r_s c_{vs}) \log(T/T_{\text{trip}}) - R_a \log(p_a/p_{\text{trip}}) - r_v R_v \log(p_v/p_{\text{trip}}) + r_v s_{0v} - r_s s_{0s},$$

is invariant under adiabatic and reversible transformations.

Subtracting  $r_l s_{0v}$ , dividing by  $c_{pa}$ , exponentiating, and multiplying by  $T_{\text{trip}} (p_0/p_{\text{trip}})^{R_a/c_{pa}}$  produces another in-

variant, which we call the liquid-water potential temperature:

$$\theta_l \equiv T \left( \frac{p_0}{p_a} \right)^{R_a/c_{pa}} \left( \frac{T}{T_{\text{trip}}} \right)^{(r_v c_{pv} + r_l c_{vl} + r_s c_{vs})/c_{pa}} \left( \frac{p_{\text{trip}}}{p_v} \right)^{r_v R_v/c_{pa}} \exp\{[-r_l s_{0v} - r_s (s_{0v} + s_{0s})]/c_{pa}\}.$$

This formulation has the benefit of taking a simple form and reducing to the potential temperature for dry air. If  $r_s = 0$ , as is the case in the simulations of shallow convection presented here, then

$$\theta_l \equiv T \left( \frac{p_0}{p_a} \right)^{R_a/c_{pa}} \left( \frac{T}{T_{\text{trip}}} \right)^{(r_v c_{pv} + r_l c_{vl})/c_{pa}} \left( \frac{p_{\text{trip}}}{p_v} \right)^{r_v R_v/c_{pa}} e^{-r_l s_{0v}/c_{pa}}.$$

#### REFERENCES

- Holland, J., and E. Rasmusson, 1973: Measurements of the atmospheric mass, energy, and momentum budgets over a 500-kilometer square of tropical ocean. *Mon. Wea. Rev.*, **101**, 44–55.
- Jonas, P., 1990: Observations of cumulus cloud entrainment. *Atmos. Res.*, **25**, 105–127.
- Krueger, S., Q. Fu, K. Liou, and H. Chin, 1995: Improvements of an ice-phase microphysics parameterization for use in numerical simulations of tropical convection. *J. Appl. Meteor.*, **34**, 281–287.
- , C.-W. Su, and P. McMurry, 1997: Modeling entrainment and finescale mixing in cumulus clouds. *J. Atmos. Sci.*, **54**, 2697–2712.
- Lin, Y.-L., R. Farley, and H. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteor.*, **22**, 1065–1092.
- Lord, S., H. Willoughby, and J. Piotrowicz, 1984: Role of a parameterized ice-phase microphysics in an axisymmetric, nonhydrostatic tropical cyclone model. *J. Atmos. Sci.*, **41**, 2836–2848.
- Neggers, R. A. J., A. P. Siebesma, and H. J. J. Jonker, 2002: A multiparcel model for shallow cumulus convection. *J. Atmos. Sci.*, **59**, 1655–1668.
- , M. Köhler, and A. C. M. Beljaars, 2009: A dual mass flux framework for boundary layer convection. Part I: Transport. *J. Atmos. Sci.*, **66**, 1465–1487.
- Raymond, D., and A. Blyth, 1986: A stochastic mixing model for non-precipitating cumulus clouds. *J. Atmos. Sci.*, **43**, 2708–2718.
- Romps, D. M., 2008: The dry-entropy budget of a moist atmosphere. *J. Atmos. Sci.*, **65**, 3779–3799.
- , 2010: A direct measure of entrainment. *J. Atmos. Sci.*, in press.
- , and Z. Kuang, 2010: Do undiluted convective plumes exist in the upper tropical troposphere? *J. Atmos. Sci.*, **67**, 468–484.
- Siebesma, A., and J. Cuijpers, 1995: Evaluation of parametric assumptions for shallow cumulus convection. *J. Atmos. Sci.*, **52**, 650–666.
- , and Coauthors, 2003: A large-eddy simulation intercomparison study of shallow cumulus convection. *J. Atmos. Sci.*, **60**, 1201–1219.
- Tiedtke, M., 1989: A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Mon. Wea. Rev.*, **117**, 1779–1800.
- Warner, J., 1977: Time variation of updraft and water content in small cumulus clouds. *J. Atmos. Sci.*, **34**, 1306–1312.