Atmos. Chem. Phys. Discuss., 11, 23231–23273, 2011 www.atmos-chem-phys-discuss.net/11/23231/2011/ doi:10.5194/acpd-11-23231-2011 © Author(s) 2011. CC Attribution 3.0 License.



This discussion paper is/has been under review for the journal Atmospheric Chemistry and Physics (ACP). Please refer to the corresponding final paper in ACP if available.

# Statistical analysis of a LES shallow cumulus cloud ensemble using a cloud tracking algorithm

# J. T. Dawe and P. H. Austin

Department of Earth and Ocean Sciences, University of British Columbia, 6339 Stores Road, Vancouver, BC, V6T 1Z4, Canada

Received: 6 August 2011 - Accepted: 12 August 2011 - Published: 17 August 2011

Correspondence to: J. T. Dawe (jdawe@eos.ubc.ca)

Published by Copernicus Publications on behalf of the European Geosciences Union.



# Abstract

A technique for the tracking of individual clouds in a Large Eddy Simulation (LES) is presented. We use this technique on a LES of a shallow cumulus cloud field based upon the Barbados Oceanographic and Meteorological Experiment (BOMEX) to calculate statistics of cloud height, lifetime, and other physical properties for individual clouds in the model. We also examine the question of nature versus nurture in shallow cumulus clouds: do properties at cloud base determine the upper-level properties of the clouds (nature), or are cloud properties determined by the environmental conditions they encounter (nurture). We find that clouds which ascend through an environment that has been pre-moistened by previous cloud activity are no more likely to reach the inversion than clouds that ascend through a drier environment. Cloud base thermodynamic properties are uncorrelated with upper-level cloud properties, while mean fractional entrainment and detrainment rate displays moderate correlations with cloud properties up to the inversion. Conversely, cloud base area correlates well with upper-

<sup>15</sup> level cloud area and maximum cloud height. We conclude that cloud thermodynamic properties are primarily influenced by entrainment and detrainment processes, cloud area and height are primarily influenced by cloud base area, and thus nature and nurture both play roles in the dynamics of BOMEX shallow cumulus clouds.

# 1 Introduction

Shallow cumulus clouds occur over large parts of the trade-wind regions (Norris, 1988), where subsiding air creates stable atmospheric conditions. They form an important part of the tropical Hadley circulation, acting to transport heat and moisture away from the surface, erode the inversion, and precondition the free troposphere for deep convection (Tiedtke et al., 1988; Neggers et al., 2007). Additionally, since short-wave radiation
 reflection from marine boundary-layer clouds is a primary component of cloud radiative effects, a proper parametrization of shallow cumulus is necessary to accurately model



the global radiation balance in General Circulation Models (GCMs; Bony and Dufresne, 2005; Medeiros et al., 2008; Wyant et al., 2009; Medeiros and Stevens, 2011).

Much effort has been devoted to understanding the dynamics of shallow cumulus, notably the work of the Global Energy and Water Cycle Experiment (GEWEX) Cloud

System Studies (GCSS; Randall et al., 2003) boundary layer cloud group. The GCSS have created several idealized test cases, based upon field campaigns, suitable for modelling via Large Eddy Simulation (LES; Siebesma and Cuijpers, 1995; Stevens et al., 2001; Brown et al., 2002; vanZanten et al., 2011). These studies have generally examined cloud fields in bulk, taking large-scale conditionally sampled mean values of quantities and fluxes. This approach provides information about the mean cloud field properties, but averages out any information concerning the dynamics of individual clouds.

A few researchers have focused instead on simulating the dynamics of individual clouds. Early work by Klaassen and Clark (1985), Bretherton and Smolarkiewicz (1989) and Grabowski and Clark (1991) performed two-dimensional simulations of a single cloud. These were quickly followed by fully three-dimensional simulations (Grabowski and Clark, 1993a,b; Carpenter et al., 1998; Blyth et al., 2005), which simulated individual clouds initiated via the application of localized heat fluxes. Unfortunately, none of these studies were able to examine how individual clouds might be affected by the presence of many other clouds in a cloud field. To address this Zhao and Austin (2005a,b) manually selected six clouds from a fully-developed LES cloud

field and analyzed their life cycles. A similar approach was taken by Heus et al. (2009), who used a virtual-reality environment to help select 79 clouds from a set of LES experiments.

<sup>25</sup> While the methods of Zhao and Austin and Heus et al. were able to provide many insights into the dynamics of individual clouds in a LES cloud field, they still suffer from two issues. First, they are time consuming, requiring human intervention to select individual clouds out of the model simulations, and second, they do not allow for a complete decomposition of a cloud field into individual clouds. Ideally one would prefer



an automated system that would be able to identify all individual clouds in an LES simulation and track them through their life history.

Two previous studies have looked at automated cloud identification and tracking in LES. Jiang et al. (2006) used a two-dimensional image registration algorithm to mea-

- <sup>5</sup> sure the effect of aerosols upon cloud lifetimes (Graham Feingold, personal communication, 2011). This method of tracking projects the cloud liquid water path of individual cumuli onto a two-dimensional surface and tracks each entity over its lifetime. It has the benefit of simplicity, but is not appropriate for clouds that overlap in the vertical. More recently, Plant (2009) presented a tracking algorithm that was able to capture the complete time evolution of each cloud in an Cloud Resolving Model (CRM). This algorithm
- operates while the CRM is running, examines cloud relationships at each time step of the CRM, and outputs diagnostics for each individual cloud.

Here we present a fully automated algorithm for the tracking of individual clouds in a LES. This algorithm generates output similar to the algorithm created by Plant (2009),

- <sup>15</sup> but with one significant difference: it can be run off-line, on pre-computed LES model fields. In Sect. 2 we describe the BOMEX LES we analyzed. Section 3 presents the cloud tracking algorithm itself, and presents an overview of the statistics generated by the algorithm. Then in Sect. 4 we use the output generated by the algorithm to address the question of whether initial cloud properties or the environment encountered by the
- cloud is more important in determining the course of the cloud's life cycle. In Sect. 5 we present our conclusions. In addition, we include the full source code for our algorithm for use by the LES modeling community, written in the Python programming language, as supplemental material to this paper.

#### 2 Model description

All LES calculations in this paper were made using version 6.7 of the System for Atmospheric Modeling (SAM; Khairoutdinov and Randall, 2003). The model run configuration was the standard GCSS Barbados Oceanographic and Meteorological Experiment



(BOMEX; Holland and Rasmusson, 1973; Siebesma et al., 2003) setup. BOMEX simulates an idealized non-precipitating steady-state trade-wind cumulus cloud field based on observations made over ocean near Barbados during June, 1969. Constant latent and sensible surface heat fluxes, and constant large-scale moisture advection, subsidence, and radiative forcings drive the model. Cloud base begins at 500 m height with maximum cloud fraction reached at 600 m, and the base of the inversion is roughly

at 1500 m.

5

10

20

The BOMEX run was performed on a  $6.4 \text{ km} \times 6.4 \text{ km}$  horizontal  $\times 3.2 \text{ km}$  vertical domain with 25 m grid resolution in all directions and a one second time step. The model was run for 6 h of model time, and the first three hours of simulation were discarded as the model was still adjusting into steady state. During the run the model steadily emits a numerical tracer from the surface which decays exponentially over time with a one minute time constant. This tracer is used to implement the conditional sampling of Couvreaux et al. (2010) in order to track cloud plumes in the dry sub-cloud layer. Model fields were output every minute for the last three hours of the simulation.

# 3 Cloud tracking

Dividing a cloud field up into individual clouds at a single moment in time is a trivial matter of finding connected cloudy regions. Tracking the resulting clouds from one time step to the next is more problematic, as a cloud is a process, not a consistent physical object. A rising parcel of moist air may condense, a parcel of air containing condensate may evaporate, and a cloud may merge with another cloud or split into multiple clouds. To handle all of these possible events, we adopt a more complex definition of what constitutes a cloud.



# 3.1 Cloud tracking algorithm

We begin by defining three cloud regions (Fig. 1). The first is the cloud "core", defined following Siebesma and Cuijpers (1995) as all model points containing condensed liquid water which have positive buoyancy and upward velocity. The second we refer

- to as the "condensed" region, defined simply as all model points containing condensed liquid water. Third is the cloud "plume", the region of upward moving air that is associated with the cloud. We define this region following the work of Couvreaux et al. (2010) via a numerical tracer that is emitted at the surface and subsequently decays exponentially with a one minute time constant. A point is considered to be in the plume if the tracer value of that point is larger than one standard deviation above the mean tracer
- 10 tracer value of that point is larger than one standard deviation above the mean tracer value at the current height. Additionally, the tracer value must exceed five percent of the mean of the horizontal standard deviation of the tracer values from the surface to the current height. However, unlike Couvreaux et al., we do not require upper-level plume points to have condensed liquid water. Finally, we force the condensed region
- to be a subset of the plume by setting all condensed points to also be plume points. Next, we divide the cloud field into "cloudlets" by assigning core points which are nearest-neighbour adjacent to the same cloudlet. These core cloudlets are then expanded into nearest-neighbour connected condensed points until all connected condensed points are assigned to a cloudlet. Condensed points cannot belong to two
- cloudlets, so if a cloud contains more than one contiguous volume of nearest-neighbour connected core points it will be split into separate cloudlets. Leftover condensed points are then divided into new cloudlets that have no cores. Once all the condensed points have been assigned to cloudlets, this expansion process is repeated for the plume points. This results in three types of cloudlets: ones consisting of core and condensed
- points surrounded by plume, ones consisting only of condensed points surrounded by plume, and ones consisting only of plume points. This process is then repeated for every saved model time step.



These cloudlets are then assigned to individual clouds. For the first time step, we assign cloudlets which have nearest-neighbour connected condensed regions to the same cloud. At subsequent time steps, cloudlets that spatially overlap with a cloud at the previous time step are assumed to be the same cloud. For the purpose of determining cloudlet overlaps, the cloudlet's position is corrected for advection that has occurred between snapshots using the mean velocity of the cloudlet's condensed points (if condensed points are present in the cloudlet) or plume (if they are not).

5

Four kinds of overlap are possible: condensed points in the cloudlet may overlap condensed points in a previous time step cloud (condensed  $\rightarrow$  condensed); condensed points may overlap previous plume points (plume  $\rightarrow$  condensed); plume may

- densed points may overlap previous plume points (plume → condensed); plume may overlap condensed points (condensed → plume); and plume may overlap plume (plume → plume). Several kinds of overlap may occur simultaneously, so we define a hierarchy of connection types and check each in turn. The strongest connection type is condensed → condensed, followed by plume → condensed, then plume → plume.
- Condensed → plume connections are ignored, which prevents connections between newly formed cloudlets and leftover plume from a dissipating cloud. Conversely, allowing plume → condensed connections lets us associate newly condensed clouds with plumes rising through the sub-cloud layer. Only the strongest connection type present for a given cloudlet is considered: if a cloudlet has condensed → condensed connections any condensed → plume or plume → plume connections are ignered, and if there
- tions, any condensed  $\rightarrow$  plume or plume  $\rightarrow$  plume connections are ignored, and if there are condensed  $\rightarrow$  plume connections, plume  $\rightarrow$  plume connections are ignored.

A cloudlet may connect to more than one cloud, which is used to identify merges and splits (Fig. 2). Any cloudlet which overlaps more than one cloud at the previous time step is considered to be the result of a merge between the two clouds. In this case

the cloudlet is assigned to the cloud with the largest overlap, and the other overlapping clouds are assumed to have merged with the largest overlap cloud. Next, clouds that connect to more than one cloudlet are considered for splitting. Splits occur only if a cloud contains more than one cloudlet whose plume is in contact with the ground, with the largest cloudlet being assigned to the original cloud and the smaller sub-clouds



becoming new clouds. Cloudlets which are not connected to the ground are assigned to the nearest ground-connected cloud, measured by the distance between the centroids of the cloud and cloudlet. Any cloudlets that do not overlap clouds in the previous time step are assumed to be newly formed clouds. This process of cloudlet assignment, merging, splitting, and creation is repeated for each time step until all cloudlets are assigned to a cloud.

Once the cloudlets are all assigned to a cloud, any cloud that has condensed points for less than five minutes is flagged. If any merge or split events occurred over these clouds' short lifetimes, the cloud is rejoined to the cloud it split from or merged into. This prevents decaying detritus shed from a cloud top and clouds which split then immediately re-merge with their parent from being counted as discrete cloud objects. Finally, clouds that have condensed liquid water for less than two minutes, and thus are unconnected to future or past clouds, are placed into a "cloud noise" group.

# 3.2 Cloud tracking results

10

- Applying the cloud tracking algorithm to three hours of BOMEX LES output resulted in 3171 tracked clouds. Of these, 609 (19%) were created by splitting an existing cloud, 2381 (75%) were unassociated with previous clouds, and 181 (6%) were present at the start of the output period and thus their initiation was not observed. Conversely, 261 (8%) of the clouds ended their life cycle by merging with another cloud, 2820
- (89%) ended their life cycle by decaying away, and 90 (3%) were present at the end of the output period. A total of 850 clouds either begin by splitting from or end by merging with another cloud (27%), indicating only 20 clouds both begin and end with such interactions (609 created via splitting, plus 261 destroyed via merging, minus 850, equals 20 which both split and merge). The larger number of clouds present at the start of the ended (121) then each werd (20) is the merging of the ended the ended to the split and merge.
- <sup>25</sup> of the output period (181) than at the end (90) is the result of the algorithm maintaining connections between physically separate clouds at the end of the output period.

Figure 3 displays horizontally averaged condensed region (cloud points containing condensed liquid water) properties for the longest-lived cloud in the ensemble. The



cloud's horizontal cross-sectional condensed area shows large, vertically coherent discontinuities in time due to merge events and split events. However, the majority of the merge and split events that occur to this cloud do not significantly alter the cloud condensed area, and none of the other horizontally averaged properties display significant discontinuities. Most of the cloud condensed region maintains upward velocities over 5 1 m s<sup>-1</sup> (Fig. 3b), though short-lived downdrafts are apparent in the inversion (above  $\approx$ 1.5 km) and at the end of the cloud's life. Near cloud base, the total specific water  $(q_t, \text{ units of g kg}^{-1}, \text{ Fig. 3c})$ , liquid-water potential temperature  $(\theta_t, \text{ units of K, Fig. 3d})$ , and condensed liquid water ( $q_1$ , units of g kg<sup>-1</sup>, Fig. 3e) are similar to the mean environmental properties, but the  $q_t$  excess,  $\theta_l$  deficit, and  $q_l$  of the condensed region 10 increases with height, and indications of upward propagating pulses can be seen in these fields. The condensed region buoyancy is generally positive below the inversion, with strong pulses of positive buoyancy apparent (displayed using density potential temperature  $\theta_o$ , units of K, Fig. 3f). Finally, comparing this example cloud with the example clouds presented by Heus et al. (2009, Figs. 4 and 5) shows similar magnitudes 15 and patterns in condensed region properties.

Core region (cloud points having upward velocity and positive buoyancy) properties for the same cloud (Fig. 4) display buoyant mass pulses more clearly than the condensed region. Cloud core occupies roughly 60 % the horizontal cloud area near cloud base, but essentially vanishes in the inversion. Since the condensed region properties show positive vertical velocity and negative buoyancy in the inversion, the disappearance of the core must be due to the rapid increase of environmental stability, which makes the cloud negatively buoyant. Cloud core vertical velocity increases steadily with height and  $q_t$  excess,  $\theta_l$  deficit, and  $q_l$  are all greatest at cloud top. Core buoyancy is positive by definition, and regular buoyant pulses are apparent.

#### 3.3 Tracked cloud statistics

In this section we examine the statistics of the cloud population generated by the tracking algorithm. Before calculating these statistics, we exclude any clouds which begin 23239



or end outside the model output period since we do not have access to their complete life cycles. This removes 287 clouds from the sample, leaving 2921. This filtering will tend to preferentially affect the longest-lived clouds from our sample, biasing the population slightly. However, our results show the cloud population to be overwhelmingly <sup>5</sup> composed of short-lived clouds, implying that robust statistics for the longest-lived clouds would not be possible with this three hour model run.

Both cloud lifetime (defined by the length of time the tracked cloud contains condensed points, Fig. 5a) and mean condensed region mass over the cloud's lifetime (Fig. 5b) are heavily skewed toward small, short-lived clouds. The cloud lifetime distribution displays no clouds with a lifetime shorter than two minutes, due to our imposed

- <sup>10</sup> bution displays no clouds with a lifetime shorter than two minutes, due to our imposed constraint that tracked clouds be present at more than one model output time. Over 1000 of the tracked clouds, a little less than a third of the total cloud population, exist for less than 4 min, and over 2200 clouds, or about a half of the cloud population, have an average mass less than  $10^6$  kg ( $\approx$ 53 model grid cells). Conversely, there are
- four clouds that persist for longer than an hour (not shown), and 35 with mean mass larger than  $3 \times 10^7$  kg ( $\approx 1600$  model grid cells, not shown). Maximum cloud top height is much less skewed (skew = 2.0) than lifetime (3.6) and mass (10.4), but still shows many more small clouds than large clouds (Fig. 5c). Only 441 clouds, or about 15% of the population, reach a height of 1 km, and only 108 clouds (4%) reach the inversion
- <sup>20</sup> base at 1.5 km. Over 78% of the clouds have cloud base between 500 and 600 m, with the majority of the rest having cloud base values below 1 km (Fig. 5d). Thus, the overall picture of the cloud field we form is of numerous small, short-lived clouds at cloud base in a field dominated by a few large, long-lived towers that have managed to overcome convective inhibition and reach the inversion.
- As cloud size distributions have been shown to be consistent with power law scalings (Benner and Curry, 1998; Zhao and Girolamo, 2007), these results are physically plausible. However, few of these statistics have been accurately measured in real cloud fields. One cloud property that has been widely measured, however, is horizontal cloud area taken from satellite and aircraft images (Benner and Curry, 1998; Cahalan and Curry, 1998; Cahalan



Joseph, 1989; Neggers et al., 2003b,a; Zhao and Girolamo, 2007; Jiang et al., 2008). The number of clouds of a given size has been found to follow a power-law distribution of the form:

 $n(I) \propto I^{-\lambda}$ 

<sup>5</sup> where / is a horizontal length scale in meters generated by taking / = √a, a is the horizontal area of the cloud, and n(/)d/ is the number of clouds with length scales between / and / + d/. λ can be calculated by finding the slope of a line fit to a log-log plot of the relative number of clouds present at each length scale /. Often this relationship is complicated by a scale break near / = 1000 m, and so two power laws
<sup>10</sup> are fit, one for clouds with / smaller than ≈1000 m and another for clouds with / larger than ≈1000 m.

To calculate the cloud size distribution, we follow Neggers et al. (2003b) and take the projection of the clouds' condensed regions onto a horizontal plane. This simulates what a satellite directly overhead would observe, to facilitate comparison with observa-

- tions. We then take the square root of the projected cloud condensed area to generate a length scale, calculate a cloud size histogram with 10 m wide bins at each minute over the three hours of model output, then fit a line to this distribution in log-log space. We do this twice, once for cloud sizes produced by the tracking algorithm, and once using cloud condensed region areas taken directly from snapshots of the model output.
- <sup>20</sup> Comparing the size distribution generated from snapshots with the distribution generated by the tracking algorithm shows the tracking algorithm modifies the cloud size distribution by increasing the number of clouds between 100 and 1000 m length scales, and removing clouds smaller and larger than this range (Fig. 6). The algorithm removes small clouds for two reasons: first, we explicitly filter out clouds that are present in only
- one snapshot; and second, the tracking algorithm treats short-lived clouds which have recently separated from a parent as part of the parent cloud. Large clouds are underrepresented by the algorithm for two reasons: first, clouds which have connected condensed regions may still have separate cores, and so are considered separate



(1)

clouds by the algorithm; and second, the projection of the three dimensional cloud area onto the 2-D plane done for the instantaneous snapshots results in clouds which overlap in the vertical, but which are disconnected in 3-D space, inflating the numbers of large clouds seen in the snapshots. Many of the small and large clouds thus end up classified as mid- sized clouds by the tracking algorithm, inflating their numbers.

Clear scale breaks appear around 1000 m in the snapshot distribution and 900 m in the tracking algorithm distribution, and so we exclude length scales greater than 1000 m and 900 m from the line fits used to find  $\lambda$ . Additionally, a small-cloud scale break is apparent in the tracking algorithm distribution for clouds with / less than 100 m, so we exclude this length scale from the tracking algorithm line fit as well. Lines fit to

so we exclude this length scale from the tracking algorithm line fit as well. Lines fit to the distributions result in reasonably similar values of *λ*: 1.88 for the snapshots, and 1.96 for the tracking algorithm (Fig. 6). These *λ* values also compare favorably with previous estimates taken from LES (1.7, Neggers et al.,2003b; 1.9, Jiang et al., 2008), satellite observations (1.89, Cahalan and Joseph, 1989; 1.88, Zhao and Girolamo, 2007) and airplane observations (1.98, Benner and Curry, 1998; 2.3, Jiang et al., 2008). Thus, while the tracking algorithm modifies the total numbers of small, medium,

and large clouds in the distribution, the power law scaling for the medium clouds is not significantly distorted.

#### 4 Two example analyses

5

- The output of the tracking algorithm provides us with a complete decomposition of the model cloud field into individual clouds. This allows us to generate statistics of cloud behaviour and use these statistics to answer questions about cloud field dynamics. Here we present two examples of analyses that would not be possible without a large database of tracked clouds.
- <sup>25</sup> Before we turn to our examples, we briefly examine the variability of the clouds in the tracked ensemble. We examine six basic properties of the clouds' condensed liquid water regions: total specific water  $q_t$ , liquid-water potential temperature  $\theta_l$ , density



potential temperature  $\theta_{\rho}$ , vertical velocity *w*, horizontal cross-sectional area *a*, and vertical mass flux *M*. We take horizontal averages of the first four properties and horizontal sums for *a* and *M* over condensed points to generate property profiles for each cloud at each sampled time.

- <sup>5</sup> Cross-correlations between properties of each cloud sample at a given height reveal strong instantaneous relationships between the properties of the cloud condensed region (Fig. 7). Nearly identical results are found taking correlations of core properties, and so these results are applicable both to condensed and core regions. Above 800 m,  $q_t$ ,  $-\theta_l$ , and  $\theta_\rho$  are strongly correlated, to the point that they can essentially be considered a single variable; more buoyant clouds have higher  $q_t$  and lower  $\theta_l$ . Similarly, *a* and *M* have near-unity correlations from 600–1400 m, indicating that variations in horizontal cross-sectional area completely control the vertical mass flux. From this we conclude that BOMEX cloud variability between 800 m and the inversion base can be
- completely characterized by three variables:  $\theta_{\rho}$ , *w*, and *a*. Near cloud base the system becomes slightly more complicated, as *a* and *M* become uncorrelated and  $-\theta_l$  diverges from  $q_t$  and  $\theta_{\rho}$ . Below 600 m, buoyancy is completely controlled by variations in  $q_t$ , and  $-\theta_l$  becomes anti-correlated with  $q_t$  and  $\theta_{\rho}$ . Within the 700–1400 m layer cross-correlations are significant at the 99% confidence level between  $\theta_{\rho}$ , *w*, and *a*, but the magnitude of these relationships is much weaker, with a
- <sup>20</sup> correlation of  $\approx 0.75$  between *w* and  $\theta_{\rho}$ ,  $\approx 0.4$  between *a* and  $\theta_{\rho}$ , and  $\approx 0.5$  between *w* and *a*. Once the clouds reach the inversion layer, these cross-correlations between *w*,  $\theta_{\rho}$  and *a* are no longer statistically significant.

Some of the causes of these cloud property cross-correlation patterns can be inferred from the vertical structure of the mean cloud condensed region properties

(Fig. 8). All the BOMEX clouds begin with a small range of thermodynamic properties at cloud base. Liquid-water potential temperature is remarkably uniform at cloud base, and so cloud base buoyancy is completely controlled by the amount of total specific water excess. The thermodynamic properties are then modified by mixing events as the clouds rise, resulting in accumulating variability that is strongly correlated. Conversely,



the BOMEX clouds start with a wide range of areas at cloud base, and as they rise this variation only grows a little. Finally, the clouds start with some variability in their vertical velocity at cloud base, which increases as they rise due to variability in the buoyancy forcing. Adopting the language of Romps (2010), this suggests that cloud area and mass flux results from the cloud's nature (initial conditions), and thermodynamic variables are governed by nurture (the environmental conditions they experience). We examine this question further in the next section.

# 4.1 Nature versus Nurture

5

In this section, we perform analyses on the tracking algorithm output to address the question of nature versus nurture in shallow cumulus cloud dynamics: do the properties of the air entering the cloud at cloud base, or the properties of the environment the cloud encounters as it rises, exert stronger influence on the evolution of the cloud? Romps (2010) addressed this question for individual parcels of convecting air and found the net entrainment experienced by a parcel influences the parcel's state far more than the parcel's initial properties at cloud base. Our tracking algorithm allows us to examine this question from the prospective of whole clouds, instead of sub-cloud parcels.

We use a simple one-dimensional particle tracking model to examine the causes of variability in upper-level cloud properties. For each cloud, we release a particle once a minute at some initial height and advect it vertically using the average vertical profile of

- <sup>20</sup> minute at some initial height and advect it vertically using the average vertical profile of w in the cloud. Linear interpolation is used to calculate w between model grid points and sampling times, and forward differencing is used to step the particle position in time with a time step of 1 s. The times at which the particles reach each higher model level are then recorded until the particle leaves the cloud (Fig. 9a).
- <sup>25</sup> These time-height profiles are then used to sample the mean properties of the cloud when the released particles reach each model level (Fig. 9b and c). We then combine the results from each cloud in the simulation to produce correlations between the cloud properties when the particles are released and when they reach each model height



(Fig. 9d). If a particle does not reach a given height, it is simply excluded from the correlation calculation. This allows us to determine to what extent the properties of the cloud near cloud base control the future properties of the cloud at height.

- We use the 99% significance level as a rough measure of the height at which the cloud properties are decorrelated from the properties at particle release. However, since clouds tend to experience discrete mass flux pulses on a  $\approx$ 15 min timescale and we release one particle each minute, we divide the degrees of freedom (number of particles) used to calculate the correlation significance by a factor of 15 to account for the auto-correlation of cloud properties within a pulse.
- First, we perform this calculation using particles released from a height of 300 m, in the sub-cloud layer. Since condensed liquid water does not appear in the sub-cloud layer, we use the horizontal mean properties of the cloud plume region to advect the particles and calculate property values. A total of 8012 particles are released at 300 m and tracked. For this calculation we ignore plume area and vertical mass flux, because
   the method we use to divide the sub-cloud plume up between individual clouds makes the area of the sub-cloud plume dependent on that of the cloud core above.

The  $q_t$  of the rising particles is strongly correlated with  $q_t$  at release up until the particles reach cloud base at 500 m (Fig. 10). The mean  $\theta_l$  and  $\theta_{\rho}$  of the plume are also strongly correlated with release properties. This implies that the sub-cloud plume properties propagate upward essentially unchanged by interaction with surrounding air.  $q_t$  is anti-correlated with  $\theta_l$  with a value of  $\approx 0.6$  in the sub-cloud layer, and does not show significant correlations with  $\theta_{\rho}$ . Conversely,  $\theta_{\rho}$  is strongly correlated with  $\theta_l$  but not with  $q_t$ ; thus, while in the cloud-layer buoyancy is controlled by  $q_t$ , in the

20

sub-cloud layer, it is controlled by  $\theta_l$ . The *w* of the rising particles is uncorrelated with any other properties at particle release with the exception of the velocity, but unlike the thermodynamic variables this correlation steadily declines and become insignificant 200 m above the particle release level, indicating the vertical velocity is strongly modified in the sub-cloud layer. The transition from the sub-cloud layer into cloud base fully decorrelates the *w* and  $\theta_{\rho}$  of the cloud with the sub-cloud properties at 300 m.



Only  $q_t$  and  $\theta_l$  properties survive the transit through cloud base, but the correlation with the sub-cloud properties is limited to the first 200 m above cloud base.

Next we look at correlations profiles for particles released at 600 m, the height of maximum cloud fraction in the middle of the cloud base layer. For this height we use the

- horizontal mean of the cloud core properties to advect the particles and calculate cloud 5 properties, as the cloud core definition insures the particles will always move upward. A total of 9984 particles are released and tracked. Due to the strong cross-correlations between thermodynamic variables and cloud area and mass flux, we restrict our analysis to  $\theta_{a}$ , w, and a. However, since  $\theta_{l}$  is uncorrelated with  $q_{t}$  and  $\theta_{a}$  at cloud base, we do calculate the correlation between  $\theta_1$  at 600 m and upper-level properties separately 10
- from cloud base  $\theta_{a}$ .

Cloud base  $\theta_i$  turns out to be uncorrelated with any upper-level cloud core properties (Fig. 11). Correlations between  $\theta_{\alpha}$  at cloud base and at height decline rapidly, dropping below 0.5 by 800 m and losing all statistical significance around 1000 m. Cloud base w

- and a are correlated with  $\theta_{o}$  at cloud base at levels of 0.25 and 0.5, respectively, and 15 this correlation also falls to insignificance by 1000 m. Vertical velocity shows a similar pattern, with all correlations becoming insignificant near 1000 m. Cloud core area, on the other hand, is significantly correlated with cloud base area all the way up to the inversion base at 1500 m. Cloud base w is also significantly correlated with upper-level
- cloud core area up to 1300 m, but with a much smaller correlation. In summary, upper-20 level thermodynamic properties and vertical velocity appear to be uncorrelated with cloud-base properties, but upper-level cloud core area is correlated with cloud base a and, to a lesser extent, w.

Romps (2010) explained the variation in cloud parcel properties at height as the result of the entrainment the parcel experienced as it rose from cloud base. We can 25 examine the influence of entrainment and detrainment processes on the cloud core properties by taking the mean of the total fractional cloud core entrainment  $\epsilon$  (in units of m<sup>-1</sup>) and fractional cloud core detrainment  $\delta$  (also in m<sup>-1</sup>) experienced by the particle between 600 m and the current height, and correlating these  $\bar{e}$  and  $\bar{\delta}$  values (where



the over-bar denotes the mean of the property between cloud base and the current height) with the properties of the cloud core at the current height. To do this we calculate cloud core mass entrainment and detrainment rate profiles for each individual cloud using the direct entrainment method detailed in Dawe and Austin (2011), and

- <sup>5</sup> divide these profiles by the cloud core vertical mass flux of each cloud to generate individual  $\epsilon$  and  $\delta$  values for each cloud at each particle sampling time. Additionally, we calculate the mean of the critical mixing fraction  $\chi_c$ , the fraction of environmental air in a mixture of core and environmental air needed to produce a neutrally buoyant mixture.  $\chi_c$  is used in buoyancy-sorting parametrization schemes to determine cloud core entrainment and detrainment rates (Kain and Fritsch, 1990; de Rooy and Siebesma.
- 2008). We calculate  $\chi_c$  using the mean cloud core properties and environmental properties taken as the average of all non-cloud core points at the current height within 100 m of the cloud core.
- Cloud core  $\theta_{\rho}$  is anti-correlated with  $\bar{e}$  and  $\bar{\delta}$ , and correlated with  $\bar{\chi}_c$ , with a constant <sup>15</sup> magnitude of roughly 0.5 up to 1400 m (Fig. 12). These correlations become stronger than the correlation with cloud base  $\theta_{\rho}$  roughly 200 m above particle release, and indicate that low entrainment and detrainment rates are correlated with anomalously buoyant clouds. Vertical velocity also shows stronger correlations with  $-\bar{e}$ ,  $-\bar{\delta}$ , and  $\bar{\chi}_c$ than with cloud base *w* between 800–1300 m, although in this case  $-\bar{\delta}$  has a noticeably <sup>20</sup> higher correlation with *w* than  $-\bar{e}$  does. Conversely, cloud core area is more correlated with cloud base core area than  $-\bar{e}$ ,  $-\bar{\delta}$  or  $\bar{\chi}_c$  all the way to the inversion base, at which point the  $-\bar{\delta}$  and  $\bar{\chi}_c$  values show roughly the same correlation with *a* as cloud base area.  $-\bar{e}$  is far less correlated with *a*, however, showing that cloud area variations are
  - more influenced by detrainment than entrainment.

# **4.2** Ascending cloud top properties versus maximum cloud height

In this section, we examine the effect that the properties of the cloud top and the environment the cloud encounters as it ascends has upon the eventual height the cloud achieves. Using the cloud tracking data, we identify the earliest times that condensed



liquid water is present at each height during each cloud's ascent (Fig. 13a). We then find all condensed points in the cloud at those heights and time, and take averages of the properties of those points to generate vertical profiles of the ascending cloud top properties (Fig. 13b). At the same time, we find all environment points that are nearest-neighbour adjacent to the cloud top condensed points and generate profiles of the environment in direct contact with the rising cloud top.

From these profiles, we perform two analyses based upon the average of the property profiles between 550–750 m and between 750–1000 m. In order to not bias these averages, any cloud profiles which do not span the whole averaging range are excluded from the analysis. Clouds whose life histories begin or end outside the data period are also discarded. From the original set of 3171 clouds, this leaves 267 clouds in the 550–750 m case and 240 clouds in the 750–1000 m case.

10

The clouds that remain are then split into two categories: tall clouds which exceed a threshold height over their lifetime, and short clouds which do not. The threshold height is chosen in each case to divide the cloud sample roughly equally. For the clouds present between 550–750 m, we set the dividing height to 1125 m; for the clouds present between 750–1000 m, we set the dividing height to 1300 m. This results in 131 tall clouds and 136 short clouds in the 550–750 m case, and 127 tall and 113 short clouds in the 750–1000 m case.

<sup>20</sup> Since some of the cloud's properties, such as horizontal cross-sectional area, are not normally distributed, we use a non-parametric statistical test to determine if the properties of the tall clouds and their environment are significantly different than the properties of the short clouds. The Mann-Whitney *U* test evaluates the null hypothesis that two data samples come from the same distribution (Mann and Whitney, 1947).

It does this by summing the number of values in one sample which are greater than each value in the other sample; this sum is called U. If large samples are taken from populations with identical distributions, on average one would expect a given value from the first sample to be larger than half the values in the other sample, and U will be normally distributed with a mean value of mn/2. Thus, the further U is from mn/2, the



more likely it is that the two samples come from distributions with differing magnitudes. Note that there are actually two U values defined by this test, depending on whether the first sample is ranked against the second, or the second is ranked against the first. However, it is easy to show that  $U_1 + U_2 = mn$ , and by convention only the smaller U-value is guoted.

5

We calculate U values for total specific water  $q_t$ , liquid-water potential temperature  $\theta_l$ , density potential temperature  $\theta_{\rho}$ , and vertical velocity *w* for both cloud top and environment. Three additional quantities are calculated for cloud top alone: specific liquid water  $q_l$ , horizontal cross sectional area *a*, and vertical mass flux *M* (in units of kg s<sup>-1</sup>). All of these values are calculated for both averages over 550–750 m and 750–1000 m, resulting in a total of 22 property comparisons. For the 550–750 m samples, if both the tall and short clouds have identical distributions the mean value of *U* should be 8908, and for the 750–1000 m samples, 7175.

Since we are calculating 22 separate tests, there is a high chance that one of these
tests will show a spurious significant relationship if we test for a 95% significance
level. Several methods, such as the Bonferonni correction (Shaffer, 1995), can be
used to reduce the possibility of finding a spurious result. However, these methods
often risk erroneously excluding significant results, and much debate over their use can
be found in the scientific literature (Perneger, 1998; Nakagawa, 2004). The simplest
correction to significance for multiple statistical tests is simply to divide the p-value
sought by the number of tests performed. Thus, to achieve an overall p-value of 0.05
with 22 tests, we must achieve a p-value of 0.05/22 = 0.0023 on an individual test. This
represents a worst case scenario where only one of the 22 tests is significant, and
we must exclude the possibility this significance results from random chance; if more

than one test is significant, the p-value needed to exclude false positives is reduced. Instead of calculating these probabilities explicitly, we will simply consider any result with p < 0.05 as possibly significant and any with p < 0.0022 as definitely significant.

Our results are presented in Table 1. We find that the environmental  $\theta_1$  and  $\theta_p$  encountered by the clouds between 550–750 m are not significantly different for tall



clouds than for short clouds. Environmental  $q_t$  is on the borderline for statistical significance, with moister environments tending to encourage taller clouds. However, the effect of environmental humidity pales in comparison with the vertical velocity of the cloud environment, which is on average  $0.05 \text{ m s}^{-1}$  larger between 550–750 m for the tall clouds than the short clouds. Thermodynamic cloud top properties are likewise not significantly different between tall and short clouds. Cloud top vertical velocity is possibly significant, with faster-rising clouds slightly more likely to become tall, but much more important is the horizontal cloud cross-sectional area *a* and mass flux *M*, with extremely small p-values of  $2.6 \times 10^{-11}$  and  $5.8 \times 10^{-11}$ , respectively; clouds with larger horizontal cross-sectional areas and greater vertical mass flux are more likely to become tall clouds.

However, simply because the differences in these values are statistically significant does not mean that they are important. Examining histograms of environmental  $q_t$  and w and cloud top w and a (Fig. 14) shows that, while the differences between the tall and short clouds are unlikely to have occurred by chance, clouds tops with small areas, in relatively slowly rising environments, still have a reasonable chance of becoming tall clouds. These properties certainly have an influence on the cloud height, but it is clearly not a definitive one.

As the clouds rise above cloud base, the properties associated with short and tall clouds change (Table 1). The average environmental values of  $q_t$ ,  $\theta_l$ , and  $\theta_\rho$  encountered by the clouds between 750–1000 m are not significantly different for short clouds than for tall clouds. The vertical velocity of the environment in this height range still has an effect, but it is slightly less significant than it is closer to cloud base, although the size of the difference is about the same (0.05 m s<sup>-1</sup>). However, all tested cloud top variables show significant differences between the short and tall clouds. Clouds with large  $q_t$ ,  $q_l$ ,  $\theta_\rho$ , w, a, and M, and lower  $\theta_l$ , between 750–1000 m, are more likely to rise past 1300 m. Of these properties, cloud top area and mass flux are again the most significant differences.



#### 4.3 Discussion

5

Our results present the following picture of the dynamics of the BOMEX cloud field. The sub-cloud layer is well mixed and filled with broad regions of upward and downward motion (Fig. 1). These regions are large enough that they are relatively shielded from mixing with other air masses and property anomalies easily propagate vertically. Conversely, the upward velocity of air parcels in this region are controlled by turbulent inertia and pressure effects (but not buoyancy), and anomalies quickly dissipate.

Once the broad regions of rising air reach cloud base, they enter a different physical regime. As rising parcels reach the top of the sub-cloud layer they lose buoyancy, and must rely on inertia and pressure perturbations to allow them to continue rising. Thus, the environmental stability at cloud base acts as a filter which admits only the fastest moving parcels, or the parcels in a region of organized motion with a pressure perturbation sufficient to overcome the negative buoyancy (i.e., a cloud with a large horizontal area). Hence, larger area clouds in an upward moving environment are better able to penetrate through cloud base.

However, above the cloud base region there is a third regime. The fate of clouds above the cloud base is determined by a race between the rate the cloud moves up-

ward and the rate the cloud is mixed away into the environment. As the clouds rise, latent heating from condensation enhances their buoyancy and creates a feedback
 loop in which faster moving clouds gain more buoyancy, and more buoyant clouds move faster. At the same time, entrainment and mixing tend to wipe out this excess buoyancy. Clouds with higher moisture content and buoyancy are more likely to remain buoyant, so thermodynamic quantities become important in determining eventual cloud

height. However, cloud area dominates these effects, as larger clouds are better able to shield their cores from mixing with the environment. This process persists until the clouds reach the inversion, where the latent heating rate is incapable of keeping the clouds buoyant in the face of the strong  $\theta$  gradient, and they are mixed away.



Using numerical tracers in a BOMEX LES, Romps and Kuang (2010) found that stochastic entrainment controlled the evolution of convecting parcels, not the thermodynamic properties of those parcels at cloud base. Our results extend this conclusion to apply to whole clouds. However, our results also show that cloud base area exerts strong controls over the upper-level area and maximum height of the cloud. Because of this we conclude that both nature, in the form of cloud base area, and nurture, in the form of entrainment and detrainment, have roles in the dynamics of BOMEX shallow cumulus clouds.

Finally, the strong effect we find that cloud area has on the eventual cloud top height loc cloud suggests a simple physical interpretation for the stochastic entrainment event parametrization proposed by Romps and Kuang (2010): entrainment events occur when a parcel reaches the edge of the cloud. This idea is supported by examining the changes in the probability distribution of  $q_t$  values at various distances from the cloud surface (Fig. 15): cloud parcels far from the cloud surface display higher mean humidities than parcels near the cloud surface, indicating that interior parcels are shielded from mixing with the environment. This hypothesis implies that the probability of a parcel experiencing an entrainment event should be related to the ratio of the total cloud

area to the area of the cloud within some mixing length of the cloud edge.

#### 5 Conclusions

We have developed an algorithm for tracking individual shallow cumulus clouds in a LES simulation which generates reasonable statistical distributions for a variety of cloud properties. The key innovations this algorithm employs are the use of non-buoyant cloud parcels as a buffer region to mediate merging and splitting of buoyant cloud plumes, and a numerical tracer which we use to evaluate if a cloud plume remains dynamically connected to the surface layer. The output of this algorithm is suitable for conducting statistical analyses of shallow cumulus cloud populations. We believe this algorithm could be easily extended to study deep convection. We include in the



supplementary material of this paper our implementation of this algorithm, written in Python, for general use by the cloud modeling community.

BOMEX cloud properties above cloud base divide into three categories of variables that are highly correlated within each category: the thermodynamic properties  $q_t$ ,  $\theta_l$ , and  $\theta_{\rho}$ ; the vertical velocity *w*; and the area *a* and vertical mass flux *M*. The thermodynamic properties show little variation at cloud base, but their variability increases steadily with height; cloud area is quite variable at cloud base and maintains its variance with height; and vertical velocity falls between these two extremes.

Analyzing the upward propagation of property anomalies shows three sets of processes operating on the clouds at different heights. Below cloud base, cloud plumes are large and homogenous and upward motion is primarily governed by turbulence. The buoyancy of plumes rapidly decreases at cloud base, which prevents all but the fastest rising plumes from penetrating into the cloud layer. If these plumes make it through the cloud base layer, dilution by the environment competes with buoyancy production from latent heating to control the ascent of the clouds to the inversion. These results suggest that nature controls the area and eventual height reached by the clouds, while nurture controls their thermodynamic properties.

# Supplementary material related to this article is available online at: http://www.atmos-chem-phys-discuss.net/11/23231/2011/

<sup>20</sup> acpd-11-23231-2011-supplement.zip.

Acknowledgements. Support for this research was provided by the Canadian Foundation for Climate and Atmospheric Science through the Cloud Aerosol Feedback and Climate network. We thank Marat Khairoutdinov for making SAM available to the cloud modeling community. Figures were generated using the matplotlib library in the Python programming language.



#### References

- Discussion ACPD 11, 23231-23273, 2011 Paper LES cloud tracking J. T. Dawe and P. H. Austin Discussion **Title Page** Paper Abstract Introduction Conclusions References **Discussion** Paper Tables **Figures** Back Full Screen / Esc **Discussion** Paper **Printer-friendly Version** Interactive Discussion
- Benner, T. C. and Curry, J. A.: Characteristics of small tropical cumulus clouds and their impact on the environment, J. Geophys. Res.-Atmos., 103, 28753-28767, 1998. 23240, 23242 Blyth, A. M., Lasher-Trapp, S. G., and Cooper, W. A.: A study of thermals in cumulus clouds, Q. J. R. Meteorol. Soc., 131, 1171–1190, 2005. 23233
- Bony, S. and Dufresne, J.: Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models, Geophys. Res. Lett., 32, L20806, doi:10.1029/2005GL023851, 2005. 23233
- Bretherton, C. S. and Smolarkiewicz, P. K.: Gravity waves, compensating subsidence and detrainment around cumulus clouds, J. Atmos. Sci., 46, 740-759, 1989. 23233
- Brown, A. R., Cederwall, R. T., Chlond, A., Duynkerke, P. G., Golaz, J.-C., Khairoutdinov, M. F., Lewellen, D. C., Lock, A. P., Macvean, M. K., Moeng, C.-H., Neggers, R. A. J., Siebesma, A. P., and Stevens, B.: Large-eddy simulation of the diurnal cycle of shallow cumulus convection over land, Q. J. R. Meteorol, Soc., 128, 1075–1093, doi:10.1256/003590002320373210. 2002. 23233
- 15

5

10

20

25

- Cahalan, R. F. and Joseph, J. H.: Fractal statistics of cloud fields, Mon. Weather Rev., 117. 261-272, 1989. 23240, 23242
- Carpenter, R. L., Droegemeier, K. K., and Blyth, A. M.: Entrainment and detrainment in numerically simulated cumulus congestus clouds. Part I: general results, J. Atmos. Sci., 55, 3417-3432, 1998. 23233
- Couvreaux, F., Hourdin, F., and Rio, C.: Resolved Versus Parametrized Boundary-Laver Plumes. Part I: A Parameterization-Oriented Conditional Sampling in Large-Eddy Simulations, Bound.-Lay. Meteorol., 134, 441-458, 2010. 23235, 23236

Dawe, J. T. and Austin, P. H.: Interpolation of LES cloud surfaces for use in direct calculations of entrainment and detrainment, Mon. Weather Rev., 139, 444–456, 2011. 23247

- de Rooy, W. C. and Siebesma, A. P.: A simple parameterization for detrainment in shallow cumulus, Mon. Weather Rev., 136, 560-576, 2008. 23247
  - Grabowski, W. W. and Clark, T. L.: Cloud-environment interface instability: rising thermal calculations in two spatial dimensions, J. Atmos. Sci., 48, 527-546, 1991. 23233
- 30 Grabowski, W. W. and Clark, T. L.: Cloud-environment interface instability. Part II: extension to three spatial dimensions, J. Atmos. Sci., 50, 527-546, 1993a. 23233 Grabowski, W. W. and Clark, T. L.: Cloud-environment interface instability. Part III: direct influ-

ence of environmental shear, J. Atmos. Sci., 50, 3821–3828, 1993b. 23233

- Heus, T., Jonker, H. J. J., den Akker, H. E. A. V., Griffith, E. J., Koutek, M., and Post, F. H.: A statistical approach to the life cycle analysis of cumulus clouds selected in a virtual reality environment, J. Geophys. Res.-Atmos., 114, D06208, doi:10.1029/2008JD010917, 2009. 23233, 23239
- Holland, J. Z. and Rasmusson, E. M.: Measurement of the atmospheric mass, energy, and momentum budgets over a 500-kilometer square of tropical ocean, Mon. Weather Rev., 101, 44-57, doi:10.1175/1520-0493(1973)101<0044:MOTAME>2.3.CO;2, 1973. 23235
  - Jiang, H., Xue, H., Teller, A., Feingold, G., and Levin, Z.: Aerosol effects on the lifetime of shal-
- low cumulus, Geophys. Res. Lett., 33, L14806, doi:10.1029/2006GL026024, 2006. 23234 10 Jiang, H., Feingold, G., Jonsson, H. H., Lu, M.-L., Chuang, P. Y., Flagan, R. C., and Seinfeld, J. H.: Statistical comparison of properties of simulated and observed cumulus clouds in the vicinity of Houston during the Gulf of Mexico Atmospheric Composition and Climate Study (GoMACCS), J. Geophys. Res.-Atmos., 113, D13205, doi:10.1029/2007JD009304, 2008. 23241.23242

15

20

5

- Kain, J. S. and Fritsch, J. M.: A one-dimensional entraining/detraining plume model and its application in convective parameterization, J. Atmos. Sci., 47, 2784–2802, 1990. 23247 Khairoutdinov, M. F. and Randall, D. A.: Cloud resolving modeling of the ARM summer 1997 IOP: model formulation, results, uncertainties, and sensitivities, J. Atmos. Sci., 60, 607–625, 2003. 23234
- Klaassen, G. P. and Clark, T. L.: Dynamics of the cloud-environment interfact and entrainment in small cumuli: two-dimensional simulations in the absence of ambient shear, J. Atmos. Sci., 42, 2621–2642, 1985. 23233

Mann, H. B. and Whitney, D. R.: On a Test of Whether one of Two Random Variables

is Stochastically Larger than the Other, Annals of Mathematical Statistics, 18, 50-60, 25 doi:10.1214/aoms/1177730491.1947.23248

Medeiros, B. and Stevens, B.: Revealing differences in GCM representations of low clouds, Clim. Dynam., 36, 385–399, 2011. 23233

Medeiros, B., Stevens, B., Held, I. M., Zhao, M., Williamson, D. L., Olson, J. G., and Bretherton,

- C. S.: Aquaplanets, climate sensitivity, and low clouds, J. Climate, 21, 4974–4991, 2008. 30 23233
  - Nakagawa, S.: A farewell to Bonferroni: the problems of low statistical power and publication bias, Behavioral Ecology, 15, 1044-1045, doi:10.1093/beheco/arh107, 2004. 23249



Neggers, R. A. J., Duynkerke, P. G., and Rodts, S. M. A.: Shallow cumulus convection: a validation of large-eddy simulation against aircraft and Landsat observations, Q. J. R. Meteorol. Soc., 129, 2671–2696, 2003a. 23241

Neggers, R. A. J., Jonker, H. J. J., and Siebesma, A. P.: Size statistics of cumulus cloud

- populations in large-eddy simulations, J. Atmos. Sci., 60, 1060–1074, 2003b. 23241, 23242 Neggers, R. A. J., Neelin, J. D., and Stevens, B.: Impact mechanisms of shallow cumulus convection on tropical climate dynamics, J. Climate, 20, 2623–2642, 2007. 23232
  - Norris, J. R.: Low cloud type over the ocean from surface observations. Part II: geographical and seasonal variations, J. Climate, 11, 383–403, 1988. 23232
- Perneger, T. V.: What's wrong with Bonferroni adjustments, The British Medical Journal, 316, p. 1236, 1998. 23249
  - Plant, R. S.: Statistical properties of cloud lifecycles in cloud-resolving models, Atmos. Chem. Phys., 9, 2195–2205, doi:10.5194/acp-9-2195-2009, 2009. 23234

Randall, D., Krueger, S., Bretherton, C., Curry, J., Duynkerke, P., Moncrieff, M., Ryan, B., Starr,

D., Miller, M., Rossow, W., Tselioudis, G., and Wielicki, B.: Confronting models with data: the GEWEX cloud systems study, B. Am. Meteorol. Soc., 84, 455–469, doi:10.1175/BAMS-84-4-455, http://journals.ametsoc.org/doi/abs/10.1175/BAMS-84-4-455, 2003. 23233
 D.M. A direct measure of entrainment 1. Attemps 26: 67, 1000, 1007, 2010.

Romps, D. M.: A direct measure of entrainment, J. Atmos. Sci., 67, 1908–1927, 2010. 23244, 23246

20 Romps, D. M. and Kuang, Z.: Nature versus Nurture in Shallow Convection, J. Atmos. Sci., 67, 1655–1666, 2010. 23252

Shaffer, J. P.: Multiple Hypothesis Testing, Annual Reviews of Psychology, 46, 561–584, 1995. 23249

Siebesma, A. P. and Cuijpers, J. W. M.: Evaluation of parametric assumptions for shallow cumulus convection, J. Atmos. Sci., 52, 650–666, 1995. 23233, 23236

25

- Siebesma, A. P., Bretherton, C. S., Brown, A., Chlond, A., Cuxart, J., Duynkerke, P. G., Jiang, H., Khairoutdinov, M., Lewellen, D., Moeng, C.-H., Sanchez, E., Stevens, B., and Stevens, D. E.: A large eddy simulation intercomparison study of shallow cumulus convection, J. Atmos. Sci., 60, 1201–1219, doi:10.1175/1520-0469(2003)60j1201:ALESIS¿2.0.CO;2, 2003.
  23235
  - Stevens, B., Ackerman, A. S., Albrecht, B. A., Brown, A. R., Chlond, A., Cuxart, J., Duynkerke, P. G., Lewellen, D. C., Macvean, M. K., Neggers, R. A. J., Sanchez, E., Siebesma, A. P., and Stevens, D. E.: Simulations of trade wind cumuli under a strong inversion, J. Atmos. Sci., 58,



1870-1891, 2001. 23233

- Tiedtke, M., Heckley, W. A., and Slingo, J.: Tropical forecasting at ECMWF: The influence of physical parametrization on the mean structure of forecasts and analyses, Q. J. R. Meteorol. Soc., 114, 639–664, doi:10.1002/qj.49711448106, 1988. 23232
- vanZanten, M. C., Stevens, B., Nuijens, L., Siebesma, A. P., Ackerman, A. S., Burnet, F., Cheng, A., Couvreaux, F., Jiang, H., Khairoutdinov, M., Kogan, Y., Lewellen, D. C., Mechem, D., Nakamura, K., Noda, A., Shipway, B. J., Slawinska, J., Wang, S., and Wyszogrodzki, A.: Controls on precipitation and cloudiness in simulations of trade-wind cumulus as observed during RICO, Journal of Advances in Modeling Earth Systems, 3, M06001, doi:10.1029/2011MS000056, 2011. 23233
- Wyant, M., Bretherton, C., and Blossey, P.: Subtropical low cloud response to a warmer climate in an superparameterized climate model: part I. regime sorting and physical mechanisms, Journal of Advances in Modeling Earth Systems, 1, doi:10.3894/JAMES.2009.1.7, http://dx. doi.org/10.3894/JAMES.2009.1.7, 2009. 23233
- <sup>15</sup> Zhao, G. and Girolamo, L. D.: Statistics on the macrophysical properties of trade wind cumuli over the tropical western Atlantic, J. Geophys. Res.-Atmos., 112, D10204, doi:10.1029/2006JD007371, 2007. 23240, 23241, 23242
  - Zhao, M. and Austin, P. H.: Life cycle of numerically simulated shallow cumulus clouds. Part I: transport, J. Atmos. Sci., 62, 1269–1290, doi:10.1175/JAS3414.1, 2005a. 23233
- Zhao, M. and Austin, P. H.: Life cycle of numerically simulated shallow cumulus clouds. Part II: mixing dynamics., J. Atmos. Sci., 62, 1291–1310, doi:10.1175/JAS3415.1, 2005b. 23233



**Table 1.** Mann-Whitney U-test results comparing differences between encountered environment and rising cloud top properties for tall vs short clouds. For average properties between 550–750 m, tall clouds are defined as clouds that reach heights over 1050 m; for average properties between 750–1000 m, tall clouds are those that reach heights over 1300 m. Shown are differences in the means (tall minus short), the differences in the means normalized by the standard deviation, the U-value, and the p-value that this U-value would arise for two samples taken from the same distribution.

	Variable	N tall	N short	Δ mean	∆ mean/ st. dev.	U-value	p-value
environment 550–750 m		131	136				
	$q_t$			0.03 g kg <sup>-1</sup>	0.24	6803	0.015
	$\theta_{I}$			-0.008 K	-0.19	7173.5	0.07
	$ heta_ ho$			-0.002 K	-0.09	7767	0.42
	W			$0.05{\rm ms}^{-1}$	0.45	6113	$3.3 \times 10^{-4}$
cloud 550–750 m							
	$q_t$			$0.018\mathrm{gkg^{-1}}$	0.19	7358	0.13
	$q_I$			$4.9 \times 10^{-3} \mathrm{g  kg^{-1}}$	0.16	7533	0.23
	$\theta_{l}$			−8.6 × 10 <sup>−5</sup> K	-0.004	8083.5	0.78
	$\theta_{ ho}$			0.013 K	0.20	7350	0.13
	W			$0.05{\rm ms}^{-1}$	0.28	6911	0.024
	а			2138 m <sup>2</sup>	0.78	4282.5	$2.6 \times 10^{-11}$
	М			2440 kg s <sup>-1</sup>	0.77	4351	5.8 × 10 <sup>-11</sup>
environment 750–1000 m		127	113				
	$q_t$			0.03 g kg <sup>-1</sup>	0.14	6574	0.26
	$\theta_{I}$			–0.007 K	-0.12	6533	0.23
	$ heta_ ho$			-0.003 K	-0.08	6888	0.59
	W			$0.05{\rm ms^{-1}}$	0.45	5145	$1.6 \times 10^{-4}$
cloud 750–1000 m							_
	$q_t$			0.1 g kg <sup>-1</sup>	0.70	4368	$1.7 \times 10^{-7}$
	$q_I$			$0.04{ m gkg^{-1}}$	0.66	4468	$4.6 \times 10^{-7}$
	$\theta_{I}$			–0.03 K	-0.51	4974	$4.1 \times 10^{-5}$
	$\theta_{\rho}$			0.07 K	0.71	4290	7.7 × 10 <sup>-8</sup>
	Ŵ			0.11 m s <sup>-1</sup>	0.46	5225	$2.8 \times 10^{-4}$
	а			2661 m <sup>2</sup>	0.89	3333.5	$8.3 \times 10^{-13}$
	М			$4711  \text{kg s}^{-1}$	0.92	2938	$2.9 \times 10^{-15}$











horizontal distance

**Fig. 2.** Example of possible relationships between clouds at two adjacent times. Top panel represents a vertical model section showing 4 clouds being tracked by the algorithm. Bottom panel represents the same vertical model section 1 minute later. As in Fig. 1, cloud core is dark red, condensed liquid water is blue, and the tracer plume is light yellow. The left side of the sections shows cloud 1 tracked between time steps. The middle of the sections shows two clouds merging; in the first time step, the clouds' condensed regions have come into contact, but not until their cores contact does cloud 3 merge into cloud 2. The right side of the sections shows a cloud splitting in two; cloud 5 splits from cloud 4 since their condensed regions have separated, while the disconnected core region on the left side of cloud 4 has not separated since it is not connected to ground level by the cloud plume.





Fig. 3. Height-time profiles of (a) horizontal cross-sectional area, (b) vertical velocity, (c) total specific water surplus, (d) liquid-water potential temperature deficit, (e) specific liquid water and (f) buoyancy of the condensed liquid water region of the longest-lived tracked cloud. Surplus and deficit values in (c) and (d) are relative to the horizontal mean properties of the model. The black lines in (b) and (f) denote the zero contours. The longer line markers at the top and bottom of the plots denote the times at which clouds split from and merge into this cloud, respectively.







Fig. 4. As in as Fig. 3, but for the cloud core region.





**Fig. 5.** Bin-width normalized number densities (n(x)dx) of **(a)** cloud lifetime in minutes, **(b)** mean mass of the cloud condensed region over the cloud lifetime, **(c)** maximum height reached by the cloud top, and **(d)** minimum height of the cloud base for tracked clouds with complete life histories.









**Fig. 7.** Vertical profiles of cross-correlations between condensed region (a) mean total specific water  $q_t$ , (b) mean liquid-water potential temperature  $\theta_l$ , (c) mean density potential temperature  $\theta_{\rho}$ , (d) mean vertical velocity w, (e) horizontal cross-sectional area a and (d) vertical mass flux M. Green squares denote correlations with  $q_t$ , yellow triangles correlations with  $-\theta_l$ , red circles correlations with  $\theta_{\rho}$ , blue x markers correlations with w, unmarked black lines correlations with a, and unmarked grey lines correlations with M. The dashed line denotes the 99% confidence level for the correlation to be significantly different than zero.







**Fig. 8.** Variation with height of the bin-width normalized probability density functions ( $P(x)\Delta x$ ) of (a) mean total specific water excess, (b) mean liquid-water potential temperature deficit, (c) mean buoyancy, (d) mean vertical velocity, (e) horizontal cross-sectional area, and (f) vertical mass flux of the condensed liquid-water regions of the tracked clouds. Black lines show contours of the cumulative distribution functions of each variable at the 0.01, 0.05, 0.16, 0.84, 0.95, and 0.99 levels, from left to right. White line denotes the mean values for each variable.



**Fig. 9.** Method used to determine correlations between lower- and upper-level cloud properties. **(a)** Dotted lines show the location of numerical particles released each minute from an initial level in the cloud and advected with the mean vertical velocity of the cloud. Colors represent cloud vertical velocities. **(b)** The time each particle reaches each model height is recorded (dotted lines) and the mean profiles of cloud properties are used to determine **(c)** the properties of the cloud at each height at the time when the particle reaches that height. **(d)** These property profiles are then used to determine correlation profiles (solid line) between the properties of the cloud at the release height and each height reached by the rising particles. The dotted lines show the 99 % confidence level for a correlation to be significantly different than zero.





**Fig. 10.** Correlation profiles between cloud plume properties near the middle of the sub-cloud layer at 300 m height and at higher levels for (a) total specific water  $q_t$ , (b) liquid-water potential temperature  $\theta_{l}$ , (c) density potential temperature  $\theta_{\rho}$ , and d) vertical velocity *w*. Green squares show correlations with 300 m level  $q_t$ , yellow triangles correlations with 300 m  $\theta_{l}$ , red circles correlations with 300 m  $\theta_{\rho}$ , and blue x markers correlations with 300 m *w*. Dotted lines show the 99 % confidence level for a correlation to be significantly different than zero.





**Fig. 11.** Correlation profiles between cloud base properties at 600 m and at higher levels for (a) density potential temperature  $\theta_{\rho}$ , (b) vertical velocity w, and (c) horizontal cross-sectional area a. Yellow triangles show correlations with cloud base  $\theta_{l}$ , red circles show correlations with cloud base  $\theta_{\rho}$ , blue x markers correlations with cloud base w, and the black line show correlations with cloud base a. Dotted lines show the 99% confidence level for a correlation to be significantly different than zero.





**Fig. 12.** Correlation profiles between cloud properties at each level and the mean of the entrainment/detrainment variables from cloud base at 600 m to the current level for (a) density potential temperature  $\theta_{\rho}$ , (b) vertical velocity *w*, and (c) horizontal cross-sectional area *a*. Yellow triangles show correlations with mean critical mixing fraction  $\chi_c$ , orange diamonds show correlations with the negative of the mean fractional entrainment rate  $\bar{e}$ , and cyan triangles show the 99% confidence level for a correlation to be significantly different than zero.





**Fig. 13.** Method used to generate profiles of environmental properties encountered by a cloud during its initial ascent. **(a)** At each height, the earliest time the cloud is present is used to sample the mean properties of environmental points directly adjacent to the cloud, and the mean cloud points. **(b)** These samples are used to generate a profile of the initial properties encountered by the cloud top.





**Fig. 14.** Bin-width normalized probability density functions (P(x)dx) of mean properties between 550–750 m of **(a)** environmental total specific water, **(b)** environmental vertical velocity, **(c)** cloud vertical velocity, and **(d)** horizontal cloud cross-sectional area. Top row, blue histograms show clouds with maximum height less than 1150 m; bottom row, red histograms show clouds with maximum height greater than 1150 m. Black lines show mean value of each histogram.





**Fig. 15.** Variation in the probability density function of total specific water  $q_t$  at 1250 m height with distance from the nearest cloud surface. Probability density functions are taken at 25 m distance intervals, to match the model grid spacing. Negative distances represent points inside the cloud. Black lines show contours of the cumulative distribution functions of each variables at at the 0.01, 0.05, 0.16, 0.84, 0.95, and 0.99 levels, from bottom to top; white line shows the 0.5 level of the cumulative distribution function.

