



Statistical analysis of an 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

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

Received: 6 August 2011 – Published in Atmos. Chem. Phys. Discuss.: 17 August 2011

Revised: 3 January 2012 – Accepted: 8 January 2012 – Published: 26 January 2012

Abstract. A technique for the tracking of individual clouds in a Large Eddy Simulation (LES) is presented. We use this technique on an 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 rates display moderate correlations with cloud properties up to the inversion. Conversely, cloud base area correlates well with upper-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 pre-

condition 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, calculating 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.

While the methods of Zhao and Austin and Heus et al. were able to provide many insights into the dynamics of individual clouds in an 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 measure the effect of aerosols upon cloud lifetimes (G. 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 an shallow cumulus LES. This algorithm generates output similar to the algorithm created by Plant (2009), 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 an overview of the cloud population statistics generated by applying the algorithm to an LES shallow cumulus cloud field. In Sect. 4 we use the output generated by the algorithm to examine whether initial cloud properties at cloud base 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 Supplement 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 Rasmus-

son, 1973; Siebesma et al., 2003) setup. BOMEX simulates an idealized, non-precipitating, steady-state trade-wind cumulus cloud field based on observations made 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.

The BOMEX run was performed on a 6.4 km \times 6.4 km horizontal \times 3.2 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 Couvreux 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 regions of condensed liquid water. Tracking the resulting clouds from one time step to the next is more problematic, however. The simplest possible algorithm would identify contiguous regions containing condensed liquid water at each time step with unique ids, and then identify regions that overlap spatially in successive time steps. This information could then be used to construct a graph (in the mathematical sense) of the cloud overlaps, and connected subgraphs of this graph would represent individual tracked clouds. (For precision, we will refer to a region of condensed liquid water as a "cloud", and a connected subgraph of cloud overlaps as a "tracked cloud").

However, when this simple algorithm is applied to the BOMEX cloud field the resulting graph is dominated by a small number (less than 10) of highly connected tracked clouds that occupy over half the cloud field volume at any given time (not shown). Furthermore, the resulting tracked clouds are not spatially localised-parts of individual tracked clouds may be on opposite sides of the domain. This result occurs because this simple algorithm does not allow tracked clouds to split. Fleeting spatial connections between clouds will result in those clouds belonging to a single connected subgraph, no matter how long ago they connected or how brief the connection.

A cloud is a process, not an 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 have developed a more complex method for tracking clouds in time.

3.1 Cloud tracking algorithm

In this section we present a description of our cloud tracking algorithm. The major problem that our algorithm needs to solve is exactly how tracked clouds should merge and split. Our strategy is to split each cloud into smaller regions, which we refer to as “cloudlets”, formed around the buoyant cores of the clouds. A full implementation of our algorithm written in the Python programming language is included as Supplement to this paper, and may be illustrative to readers attempting to understand our technique.

We begin by defining three regions we use for cloud tracking (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 Couvreux 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 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 Couvreux et al., we do not require upper-level plume points to have condensed liquid water. Finally, all condensed points are also flagged as plume points regardless of their tracer concentration, so that the condensed region is always a subset of the plume.

We use the plume, condensed, and core regions to help us identify when tracked clouds merge and split. The simplest criterion would be spatial connectivity; tracked clouds merge when their condensed regions connect and split when a single condensed region separates. However, in practice this often results in tracked clouds making brief contact, merging, and then immediately splitting apart again. To prevent these transitory splitting and merging events, we only merge clouds when their core regions are connected, not simply their condensed regions. Splits still occur when the condensed region of a cloud divides into separate parts, but we additionally require newly split clouds to be connected to the surface via the plume region. This prevents cloud detritus near the inversion from splitting from its parent cloud as it evaporates. These additional constraints result in spatially localized tracked clouds which centre around actively convecting cores connected to ground level through rising plumes.

At each time step we divide each model cloud into “cloudlets” formed around contiguous, nearest-neighbour connected core regions. We identify contiguous core regions

by first selecting a core point at random and labeling it with an integer cloudlet id. Next, any core grid cells immediately above, below, north, south, east, and west of the initial core point are labeled with the same cloudlet id. This process is repeated recursively on the newly labeled points until all core points which are nearest-neighbour connected to the initial core point are labeled. Then another unlabeled core point is chosen at random, assigned a new cloudlet id, and the process repeated until all core points are assigned a cloudlet id.

Next, condensed points are labeled with cloudlet ids in a process mimicking crystal growth. We iterate through the core cloudlet ids and label condensed points that are immediately nearest-neighbour adjacent to each cloudlet with that cloudlet’s id. Condensed points adjacent to more than one cloudlet are assigned the smaller cloudlet id. Once all the cloudlets have been expanded into the condensed region around them, the iteration is repeated, causing the cloudlets to grow into the condensed region around them until all connected condensed points are assigned to a cloudlet. Remaining condensed points must be unconnected to any core points, and are assigned new cloudlet ids via the same process used to label the core points. Once all the condensed points are labeled the expansion process is repeated for the plume points until all plume points connected to a condensed region are assigned a cloudlet id. Remaining unconnected plume points are assigned to new cloudlet ids via the same process used for core and condensed cloudlets until all plume points have a cloudlet id. The entire cloudlet id assigning process is then repeated for every saved model time step.

This process of assigning plume, condensed, and core points to cloudlets results in condensed regions being divided into one or more cloudlets, each centered on contiguous regions of core points within the condensed region (Fig. 1). If a condensed region contains no core points the entire condensed region will be assigned to a single cloudlet, and if a plume region contains no condensed points the entire plume region is assigned to a single cloudlet. This creates three types of cloudlets: ones consisting of core points, surrounded by condensed points, surrounded by plume; ones consisting only of condensed points surrounded by plume; and ones consisting only of plume points.

These cloudlets are then grouped into “clusters” of related cloudlets and each cluster is assigned a unique integer tracking id. For the initial model output time step, we assign cloudlets with adjacent condensed regions to the same cluster. This causes the initial time step’s clusters to correspond to what most people would consider clouds—connected regions of condensed liquid water. At subsequent time steps, spatial overlaps between a cluster and a cloudlet in the following time step are used to associate cloudlets with previous clusters. The overlaps are examined with the cloudlet’s position corrected for advection between time steps using the mean velocity of the cloudlet’s condensed points (if condensed points are present in the cloudlet) or plume (if they are not).

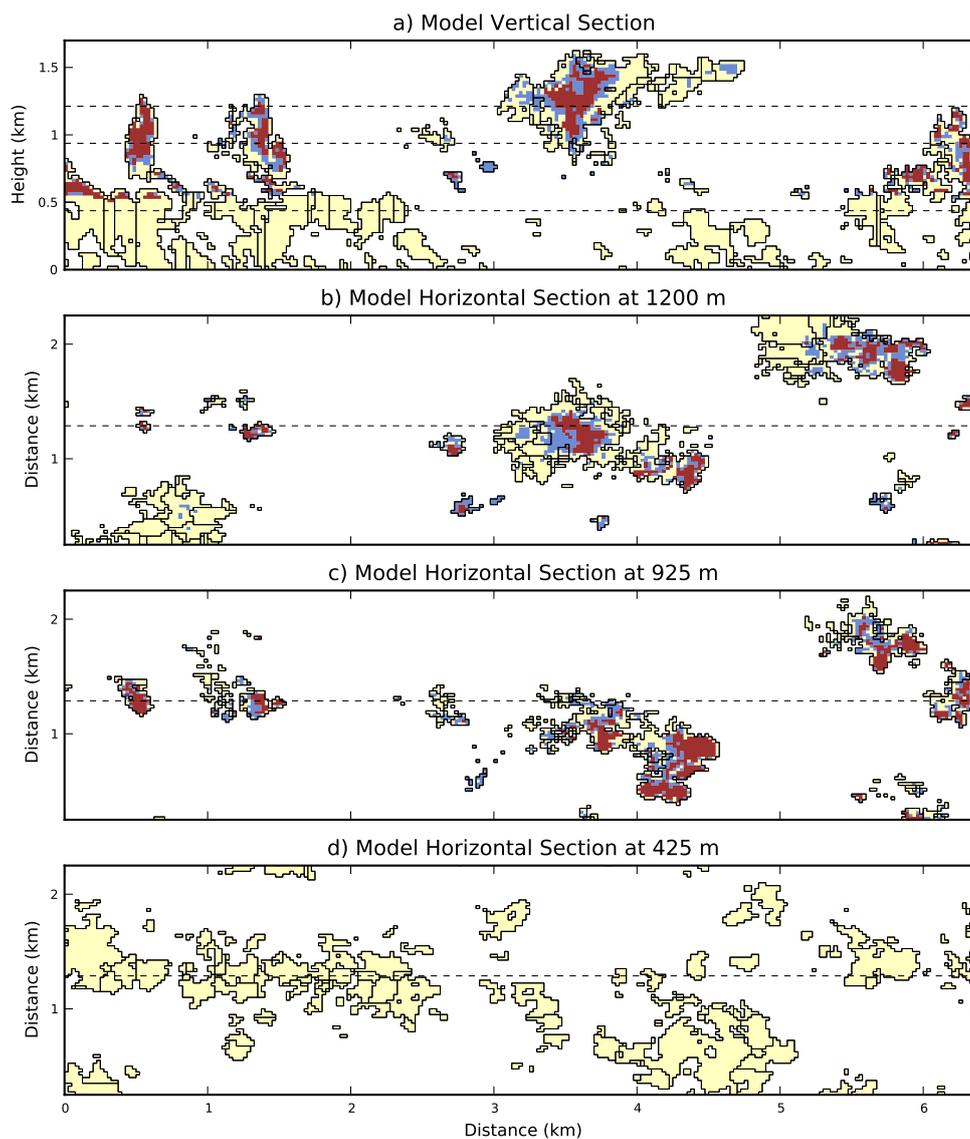


Fig. 1. Vertical and horizontal sections through the BOMEX model, showing the cloud core (dark red), condensed liquid water (blue) and the plume (light yellow) regions used by the cloud tracking algorithm. Black lines show the edges of algorithm cloudlets. **(a)** Vertical east-west section across the model domain. Dashed lines show the heights of the horizontal sections shown in **(b–d)**. **(b)** Horizontal section across the model domain at 1200 m height. Dashed line shows the position of the vertical section shown in **(a)**. Only 2 km of the 6.4 km domain are shown in the north-south direction. **(c)** Same as for **(b)**, but at 925 m height. **(d)** Same as for **(b)**, but at 425 m height.

Four kinds of overlap are possible: condensed points in the cloudlet may overlap condensed points in a previous cluster (condensed \rightarrow condensed, Fig. 2a); condensed points may overlap previous plume points (plume \rightarrow condensed, Fig. 2b); plume may overlap condensed points (condensed \rightarrow plume, Fig. 2c); and plume may overlap plume (plume \rightarrow plume, Fig. 2d). 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 \rightarrow condensed, followed by plume \rightarrow condensed, then plume \rightarrow plume. Con-

densed \rightarrow plume connections are ignored, which prevents connections between newly formed cloudlets and leftover plume from a dissipating cloud. Conversely, allowing plume \rightarrow condensed connections lets us associate newly condensed fluid 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 \rightarrow condensed connections, 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.

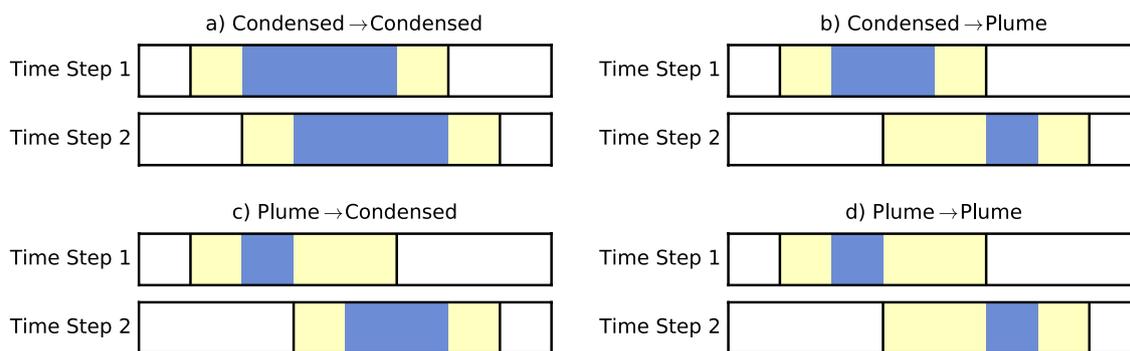


Fig. 2. The four kinds of overlap between cloudlets and tracked clouds at the previous time step. Each subfigure shows a one-dimensional representation of model points at an initial time (top) and 1 min later (bottom). Yellow indicates plume region and blue indicates condensed region. **(a)** A condensed \rightarrow condensed overlap, with the condensed region in time step 1 overlapping the condensed region in time step 2. This subfigure also contains condensed \rightarrow plume, plume \rightarrow condensed, and plume \rightarrow plume overlaps, but the condensed \rightarrow condensed overlap takes precedence. **(b)** A condensed \rightarrow plume connection, with the condensed region in time step 1 overlapping the plume region in time step 2. This diagram also contains a plume \rightarrow plume overlap, but the condensed \rightarrow plume overlap takes precedence. **(c)** A plume \rightarrow condensed connection, with the plume region in time step 1 overlapping the condensed region in time step 2. This diagram also contains a plume \rightarrow plume overlap, but the condensed \rightarrow plume overlap takes precedence. **(d)** A plume \rightarrow plume connection, with the plume region in time step 1 overlapping the plume region in time step 2.

Cloudlets that overlap a single cluster are unambiguously assigned the same tracking id (Fig. 3, cluster 1 at time step 1 overlaps cloudlet 1 at time step 2). A cloudlet which overlaps more than one cluster at the previous time step is assigned the cluster id of the largest volume overlap, and the smaller overlap cluster is flagged as possibly having merged into the larger cloud. (Fig. 3, clusters 2 and 3 at time step 1 both overlap cloudlet 2 of the figure. Clusters 2 and 3 at time step 1 are separate because their core regions have not yet connected.) Finally, clusters that contain more than one cloudlet are considered for splitting. A split occurs if a tracked cluster contains cloudlets with disconnected condensed regions, with the additional condition that those cloudlets have plume regions in contact with the ground. When a cluster splits, the cluster of connected cloudlets with the greatest condensed volume is assigned the parent cluster id, and the smaller sub-clusters are assigned new ids. Cloudlets that are not connected to the ground are assigned to the sub-cluster with the nearest centroid. (In Fig. 3, cluster 4 at time step 1 consists of two cloudlets that have connected condensed regions. At time step 2, cluster 5 splits from cluster 4 because their condensed regions are disconnected and both clusters are connected to ground through the plume. The small core cloudlet on the left side of cluster 4 does not split since it is not connected to the surface.)

Finally, any cloudlets that do not overlap clusters in the previous time step are assumed to be new clouds and are assigned new cluster ids. The process of cluster assignment, merging, splitting, and creation is repeated for each time step until all cloudlets are assigned tracking ids. At this point a graph is created, with all cloudlets having the same cluster id being connected into a tracked cloud. Any tracked cloud that has condensed points for less than five minutes is flagged,

and if any merge or split events occurred over the tracked cloud's short lifetime it is connected to the tracked cloud it split from or merged with. This further restricts 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, tracked clouds that have condensed liquid water only for a single time step are placed into a "cloud noise" group and considered separately.

3.2 Cloud tracking results

Applying the cloud tracking algorithm to three hours of BOMEX LES output with snapshots taken each minute requires 1 h and 40 min of computation time on an Intel Xeon E5645 2.4 GHz processor with 4 GB of RAM. The algorithm results 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 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 4 displays horizontally averaged condensed region (cloud points containing condensed liquid water) properties

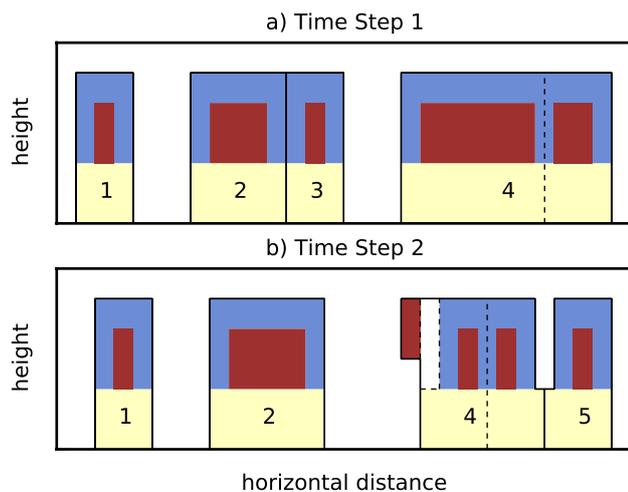


Fig. 3. Example of possible relationships between clusters and cloudlets at successive time steps. Top panel represents a vertical model section showing 4 clusters being tracked by the algorithm. Bottom panel represents the same vertical model section 1 min later. As in Fig. 1, cloud core is dark red, condensed liquid water is blue, and the tracer plume is light yellow. Solid black lines denote clusters, and dashed lines show the cloudlets that compose these clusters. The left side of the sections shows cloudlet 1 unambiguously overlapping cluster 1. The middle of the sections shows two clusters merging; in the first time step, the clusters' condensed regions have come into contact, but not until their core regions form a single contiguous area does cluster 3 merge with cluster 2. The right side of the sections shows a cluster splitting in two. Initially, cluster 4 is composed of two cloudlets with connected condensed regions. At the next time step cluster 5 splits from cluster 4 since their condensed regions have separated and cluster 5's cloudlet has a plume in contact with the ground. The disconnected core cloudlet on the left side of cluster 4 does not split since it has no plume connection to ground. The remaining two cloudlets which compose the bulk of cluster 4 do not split since their condensed regions are connected.

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 1 m s^{-1} (Fig. 4b), though short-lived downdrafts are apparent in the inversion (above $\approx 1.5 \text{ km}$) and at the end of the cloud's life. Near cloud base, the total specific water (q_t , units of g kg^{-1} , Fig. 4c), liquid-water potential temperature (θ_l , units of K, Fig. 4d), and condensed liquid water (q_l , units of g kg^{-1} , Fig. 4e) are similar to the mean environmental properties, but the q_t excess, θ_l deficit, and q_l of the condensed region 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 θ_ρ , units of K, Fig. 4f). Finally, comparing this example cloud with the example clouds presented by Heus et al. (2009, Figs. 4 and 5) shows similar magnitudes and patterns in condensed region properties.

Core region (cloud points having upward velocity and positive buoyancy) properties for the same cloud (Fig. 5) 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, θ_l deficit, and q_l are all greatest at cloud top. Core buoyancy is positive by definition, and regular buoyant pulses are apparent.

The total BOMEX cloud fraction at cloud base in our model is about 0.065, or $\approx 2.7 \times 10^6 \text{ m}^2$. The longest-lived cloud thus represents over 10% of the total cloud base area at times. As we mention at the beginning of this section, this is actually much smaller than the largest cloud tracked by a simple overlap algorithm that does not allow clouds to split and merge. The dominance of the cloud field by this cloud is a natural outcome of the power law distribution that shallow cumulus cloud areas are known to obey, since a power law implies the existence of a small number of extremely large clouds.

3.3 Tracked cloud statistics

In this section we examine the statistics of the cloud population generated by the tracking algorithm in order to sanity check the algorithm results, and to identify how and where the algorithm artificially modifies cloud field size statistics. Before calculating these statistics, we exclude any clouds which begin 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 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. 6a) and mean condensed region mass over the cloud's lifetime (Fig. 6b) 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 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,

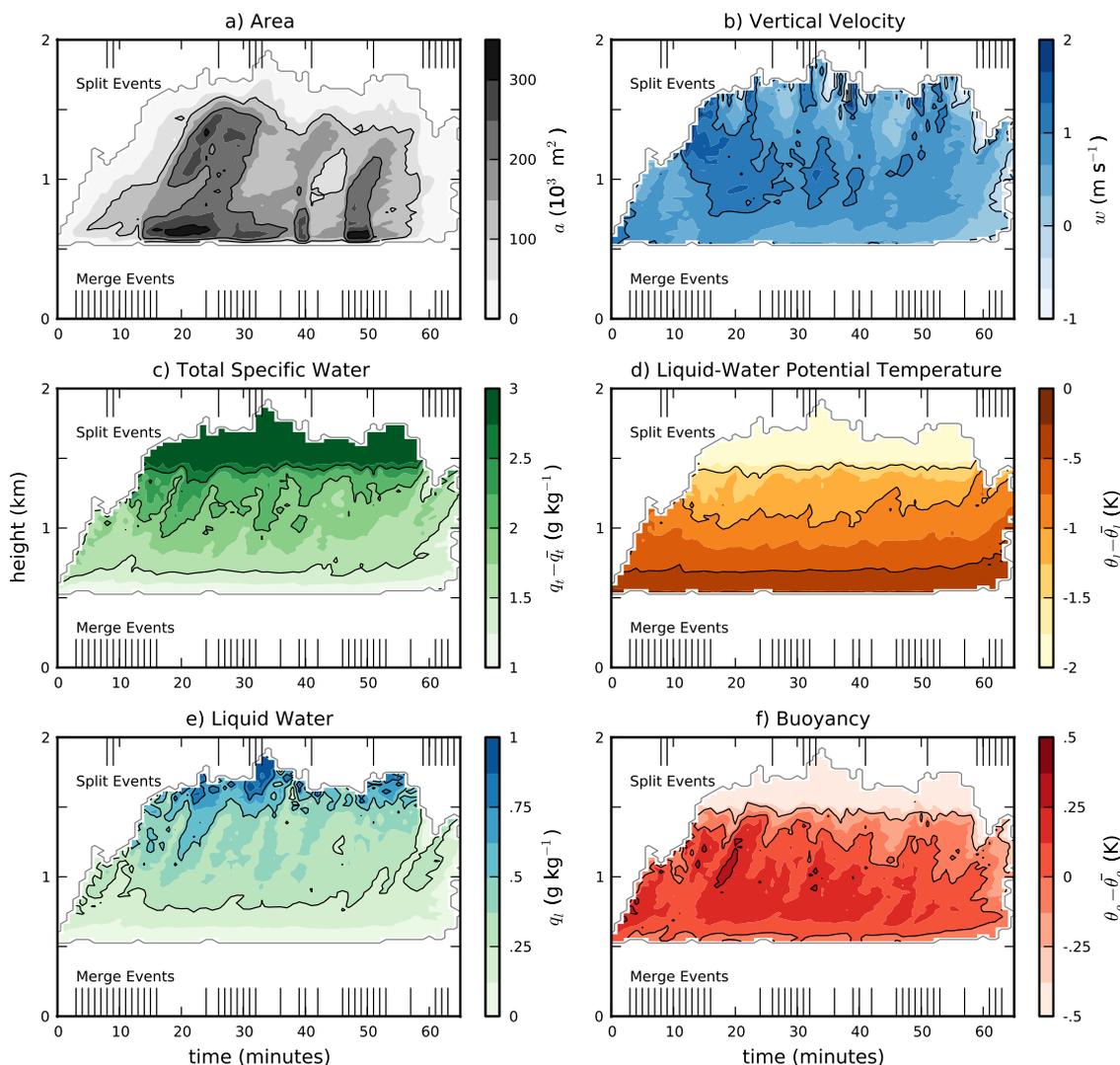


Fig. 4. Height-time profiles of (a) horizontal cross-sectional area, (b) vertical velocity, (c) total specific water surplus, (d) liquid-water potential temperature deficit, (e) specific liquid water and (f) buoyancy of the condensed liquid water region of the longest-lived tracked cloud. Surplus and 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.

and over 2200 clouds, or about a half of the cloud population, have an average mass less than 10^6 kg (≈ 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 3×10^7 kg (≈ 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. 6c). Only 441 clouds, or about 15% of the population, reach a height of 1 km, and only 108 clouds (4%) reach the inversion 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. 6d). 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 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:

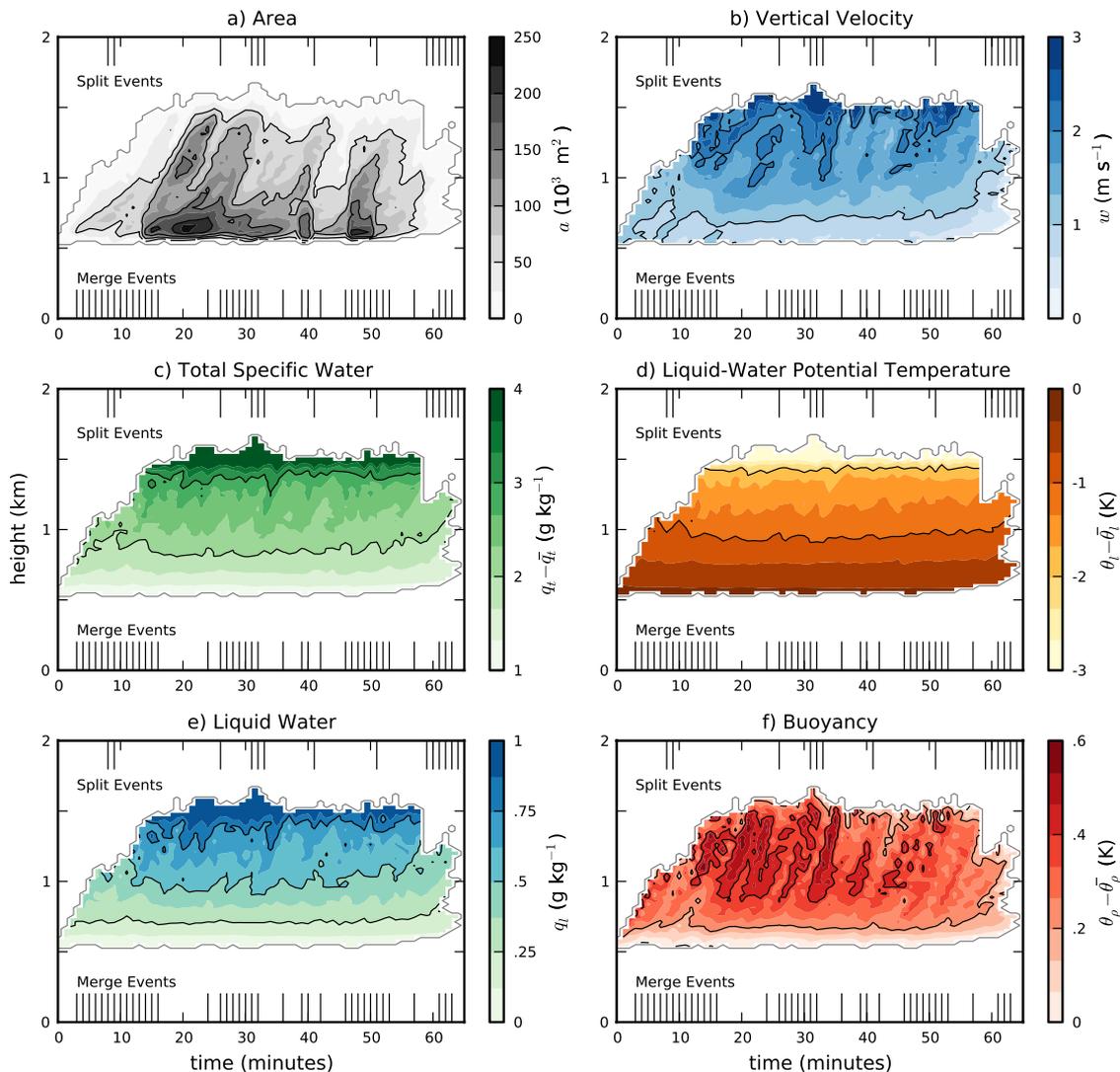


Fig. 5. As in as Fig. 4, but for the cloud core region.

$$n(l) \propto l^{-\lambda} \quad (1)$$

where l is a horizontal length scale in meters generated by taking $l = \sqrt{a}$, a is the horizontal area of the cloud, and $n(l)dl$ is the number of clouds with length scales between l and $l+dl$. 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 l . Often this relationship is complicated by a scale break near $l = 1000$ m, and so two power laws are fit, one for clouds with l smaller than ≈ 1000 m and another for clouds with l 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 observations. 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.

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. 7). 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

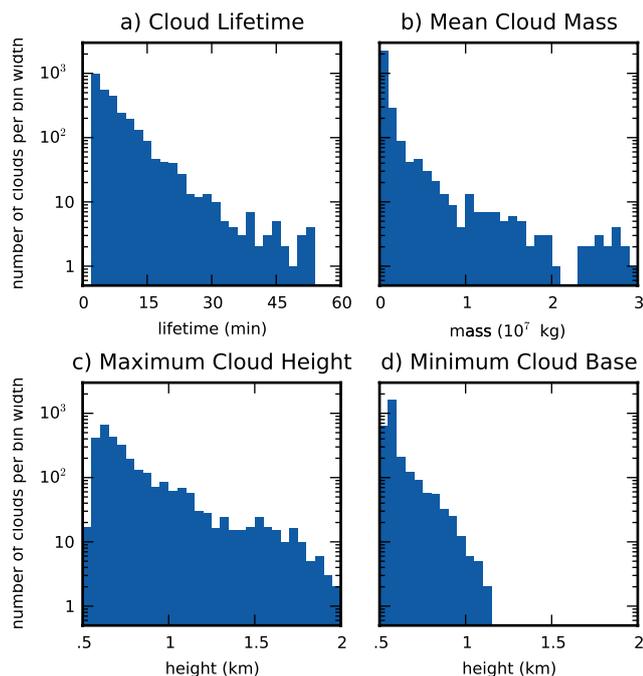


Fig. 6. Histograms of (a) cloud lifetime in minutes (with a 2 min bin width), (b) mean mass of the cloud condensed region over the cloud lifetime (10^6 kg bin width), (c) maximum height reached by the cloud top (50 m bin width), and (d) minimum height of the cloud base (50 m bin width) for tracked clouds with complete life histories.

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 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 λ . Additionally, a small-cloud scale break is apparent in the tracking algorithm distribution for clouds with l less than 100 m, 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. 7). 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

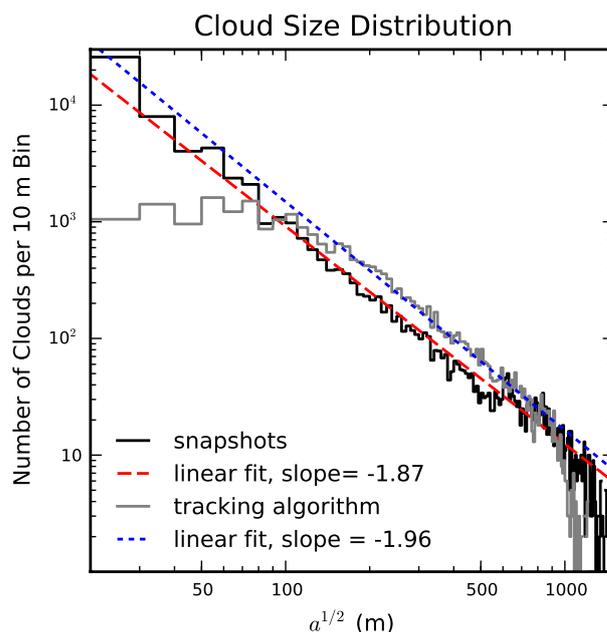


Fig. 7. Histograms of horizontal cloud length scales calculated from instantaneous cloud field snapshots (black line) and from the cloud tracking algorithm (grey line) using a 10 m bin width. The red dashed line and blue dotted line show linear best fit between 100–1000 m length scales to the instantaneous and tracked cloud histograms, respectively.

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

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.

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 θ_l , density potential temperature θ_ρ , 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.

Correlations between the horizontal mean properties of all clouds present at a given height reveal strong relationships

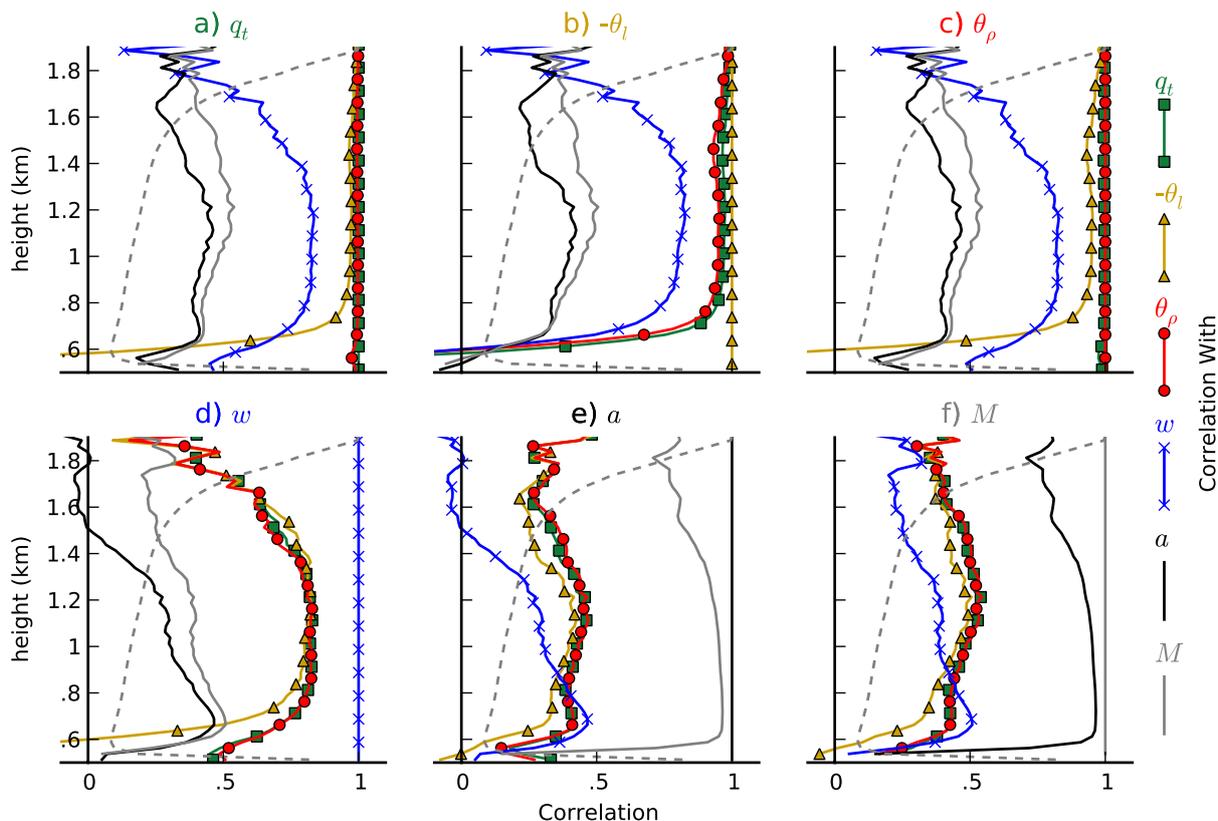


Fig. 8. Vertical profiles of cross-correlations between condensed region (a) mean total specific water q_t , (b) mean liquid-water potential temperature θ_l , (c) mean density potential temperature θ_ρ , (d) mean vertical velocity w , (e) horizontal cross-sectional area a and (f) vertical mass flux M . Green squares denote correlations with q_t , yellow triangles correlations with $-\theta_l$, red circles correlations with θ_ρ , 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.

between the mean cloud properties (Fig. 8). 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 θ_ρ are strongly correlated, to the point that they can essentially be considered a single variable; more buoyant clouds have higher q_t and lower θ_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: θ_ρ , 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 θ_ρ . Below 600 m, buoyancy is controlled by variations in q_t , and $-\theta_l$ becomes anti-correlated with both q_t and θ_ρ , reaching a correlation of ≈ -0.9 at 500 m (not shown). Within the 700–1400 m layer cross-correlations are significant at the 99 % confidence level between θ_ρ , w , and a , but the magnitude of these relationships is much weaker, with a correlation of ≈ 0.75 between w and θ_ρ , ≈ 0.4 between a and θ_ρ , and ≈ 0.4 between w and a . Once the clouds reach the

inversion layer, these correlations between w , θ_ρ 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. 9). All the BOMEX clouds begin with a small range of thermodynamic properties (relative to their mean values) at cloud base. 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.

Although M is related to a and w via the relation $M = \rho wa$, where ρ is the air density in kg m^{-3} , M is only weakly correlated with w , despite being strongly correlated with a . This unintuitive result arises due to the relative contributions of a and w to the variance of M . Changes in M can be expressed in terms of changes in a and w as

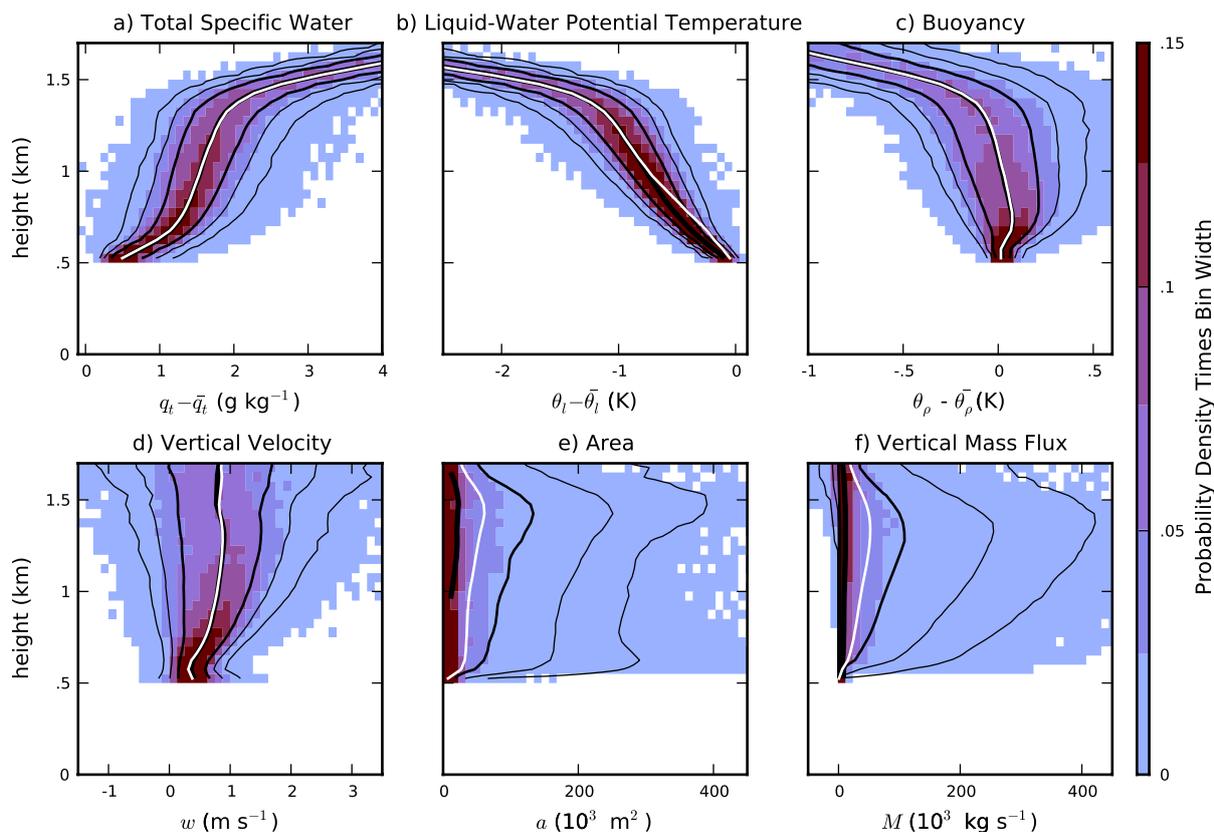


Fig. 9. Variation with height of the probabilities in range Δx ($P(x)\Delta x$) of the condensed liquid-water regions of the tracked clouds having a certain **(a)** mean total specific water excess, **(b)** mean liquid-water potential temperature deficit, **(c)** mean buoyancy, **(d)** mean vertical velocity, **(e)** horizontal cross-sectional area, and **(f)** vertical mass flux. Black lines show contours of the cumulative distribution functions of each variable at the 0.01, 0.05, 0.16, 0.5, 0.84, 0.95, and 0.99 levels, from left to right. White line denotes the mean values for each variable.

$$dM = \rho w \partial a + \rho a \partial w \quad (2)$$

Choosing representative cloud layer values into Eq. (2) is complicated by the power-law distribution that governs cloud areas. The median values of w and a are roughly 0.5 m s^{-1} and $10\,000 \text{ m}^2$ (Fig. 9), while 66% of w and a fall between roughly $0.1\text{--}1.0 \text{ m s}^{-1}$ and $1000\text{--}100\,000 \text{ m}^2$, respectively. Differences in cloud vertical velocity will thus result in mass flux values of approximately $1000\text{--}10\,000 \text{ kg s}^{-1}$, while differences in cloud cross-sectional area will result in mass flux values between $500\text{--}50\,000 \text{ kg s}^{-1}$, a range an order of magnitude larger. Thus, cloud mass fluxes are primarily controlled by the area of the cloud, producing near unity correlations between a and M .

Adopting the language of Romps and Kuang (2010), these results suggest 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

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 and Kuang (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 perspective 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 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. 10a).

These time-height profiles are then used to sample the mean properties of the cloud when the released particles reach each model level (Fig. 10b 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. 10d). 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 ≈ 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. 11). The mean θ_1 and θ_ρ 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 θ_1 with a value of ≈ 0.6 in the sub-cloud layer, and does not show significant correlations with θ_ρ . Conversely, θ_ρ is strongly correlated with θ_1 but not with q_t ; thus, while in the cloud-layer buoyancy is controlled by q_t , in the sub-cloud layer, it is controlled by θ_1 . 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 θ_ρ of the cloud with the sub-cloud properties at 300 m. Only q_t and θ_1 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 correlation 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 hor-

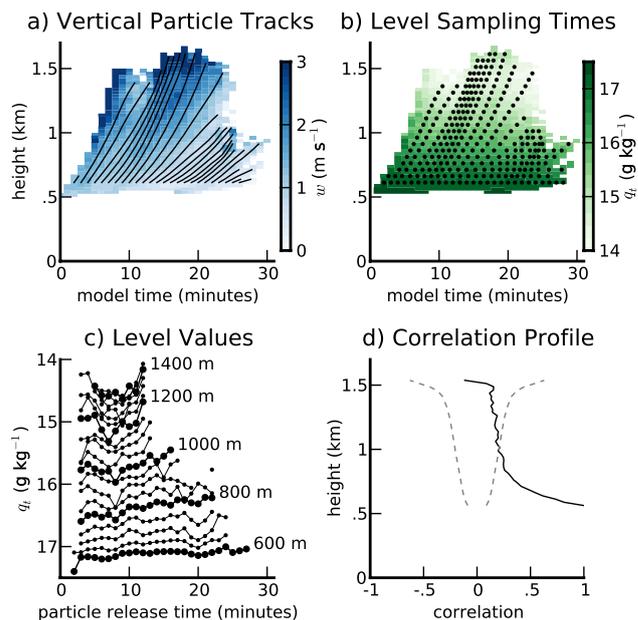


Fig. 10. Method used to determine correlations between lower- and upper-level cloud properties. (a) Numerical particles are released once per minute from an initial level in the cloud and advected vertically with the mean vertical velocity of the cloud until the particle leaves the cloud. (Lines show the time-height trajectories of the numerical particles and colours show the cloud's vertical velocity.) (b) The times at which particles reach each model level are then identified and the cloud properties at those times are recorded. (Dots show the time each particle reaches each model level and colours show the cloud's vertical velocity. Only half the model levels have been plotted for clarity.) (c) The properties encountered by the particles at a given height are then arranged by the time each particle was released, forming a set of pseudo-time series at each height. (Dotted lines show the total specific water values of the cloud at the time each particle reached a given height. The 600 m, 800 m, 1000 m, 1200 m, and 1400 m height particle values are highlighted and labeled. Only half the model levels have been plotted.) (d) Correlations are then taken between the properties of the particles at release and the properties at higher levels to calculate correlation profiles. (Solid line shows the correlation between total specific humidity of the particles at release and the total specific humidity of the cloud at various heights. Dotted lines show the 99% confidence level for a correlation to be significantly different than zero.)

izontal mean of the cloud core properties to advect the particles and calculate cloud 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 θ_ρ , w , and a . However, since θ_1 is uncorrelated with q_t and θ_ρ at cloud base, we do calculate the correlation between θ_1 at 600 m and upper-level properties separately from cloud base θ_ρ .

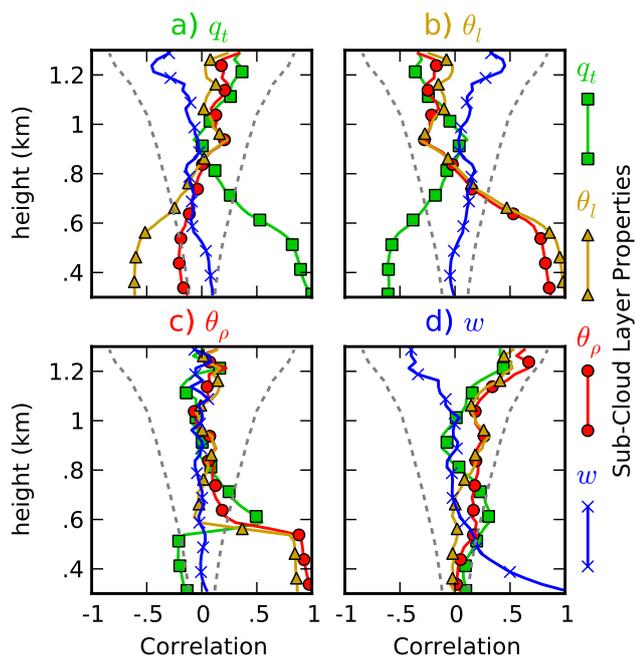


Fig. 11. Correlation profiles between cloud plume properties near the middle of the sub-cloud layer at 300 m height and at higher levels for (a) total specific water q_t , (b) liquid-water potential temperature θ_l , (c) density potential temperature θ_ρ , and (d) vertical velocity w . Green squares show correlations with 300 m level q_t , yellow triangles correlations with 300 m θ_l , red circles correlations with 300 m θ_ρ , 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.

Cloud base θ_l turns out to be uncorrelated with any upper-level cloud core properties (Fig. 12). Correlations between θ_ρ 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 θ_ρ at cloud base at levels of 0.25 and 0.5, respectively, and 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-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 and Kuang (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 examine the influence of entrainment and detrainment processes on the cloud core properties by taking the mean of the total fractional cloud core entrainment ϵ (in units of m^{-1})

and fractional cloud core detrainment δ (also in m^{-1}) experienced by the particle between 600 m and the current height, and correlating these $\bar{\epsilon}$ 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 divide these profiles by the cloud core vertical mass flux of each cloud to generate individual ϵ and δ values for each cloud at each particle sampling time. Additionally, we calculate the mean of the critical mixing fraction χ_c , the fraction of environmental air in a mixture of core and environmental air needed to produce a neutrally buoyant mixture. χ_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 χ_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 θ_ρ is anti-correlated with $\bar{\epsilon}$ and $\bar{\delta}$, and correlated with $\bar{\chi}_c$, with a constant magnitude of roughly 0.5 up to 1400 m (Fig. 13). These correlations become stronger than the correlation with cloud base θ_ρ 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 $-\bar{\epsilon}$, $-\bar{\delta}$, and $\bar{\chi}_c$ than with cloud base w between 800–1300 m, although in this case $-\bar{\delta}$ has a noticeably higher correlation with w than $-\bar{\epsilon}$ does. Conversely, cloud core area is more correlated with cloud base core area than $-\bar{\epsilon}$, $-\bar{\delta}$ 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. $-\bar{\epsilon}$ is far less correlated with a , however, showing that cloud area variations are more influenced by detrainment than entrainment.

Our results largely agree with the results of Romps and Kuang (2010): upper-level cloud properties are governed by the entrainment and detrainment experienced by the cloud as it rises, and cloud base properties have little influence on upper-level cloud properties, suggesting that nurture is more important than nature in determining shallow cumulus cloud properties. The exception to this is cloud area, which is correlated with cloud base area and which Romps and Kuang were not able to examine with their parcel model. Nevertheless, cloud base area and entrainment/detrainment rates still exert roughly equal influence over upper-level cloud area.

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.

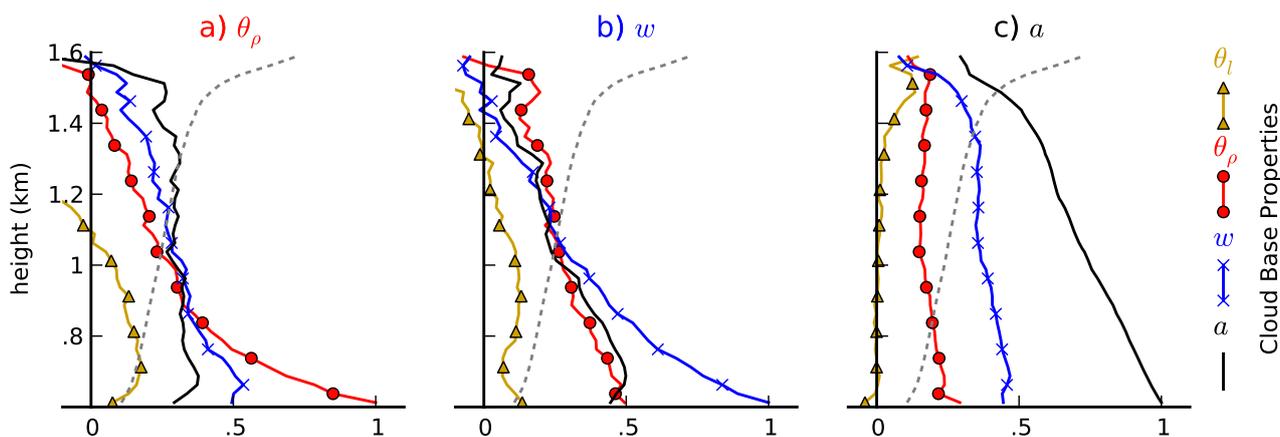


Fig. 12. Correlation profiles between cloud base properties at 600 m and at higher levels for (a) density potential temperature θ_ρ , (b) vertical velocity w , and (c) horizontal cross-sectional area a . Yellow triangles show correlations with cloud base θ_l , red circles show correlations with cloud base θ_ρ , 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.

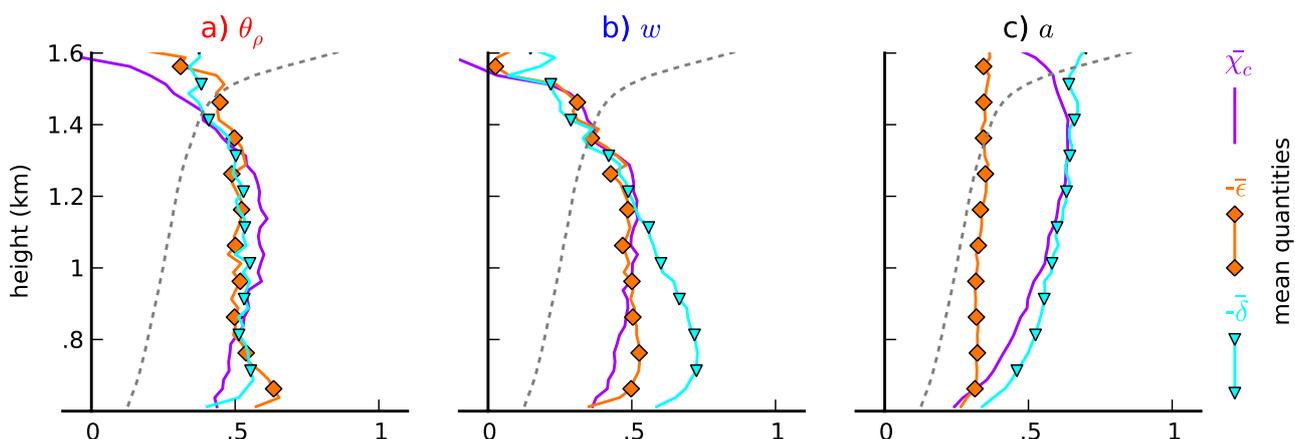


Fig. 13. Correlation profiles between cloud properties at each level and the mean of the entrainment/detrainment variables from cloud base at 600 m to the current level for (a) density potential temperature θ_ρ , (b) vertical velocity w , and (c) horizontal cross-sectional area a . Yellow triangles show correlations with mean critical mixing fraction $\bar{\chi}_c$, orange diamonds show correlations with the negative of the mean fractional entrainment rate $\bar{\epsilon}$, and cyan triangles show correlations with the negative of the mean fractional detrainment rate $\bar{\delta}$. Dotted lines show the 99 % confidence level for a correlation to be significantly different than zero.

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. 14a). We then find all condensed points within 100 m of 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. 14b). Additionally, we find all environment points that are nearest-neighbour adjacent to the cloud top condensed points between 0 m and 50 m below cloud top and extend the topmost nearest-neighbour points 50 m above cloud top to generate profiles of the environment encountered by 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 257 clouds in the 550–750 m case and 240 clouds in the 750–1000 m case.

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.

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 quoted.

We calculate U values for total specific water q_t , liquid-water potential temperature θ_l , density potential temperature θ_ρ , 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^{-1}). 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 $p < 0.0023$ as definitely significant.

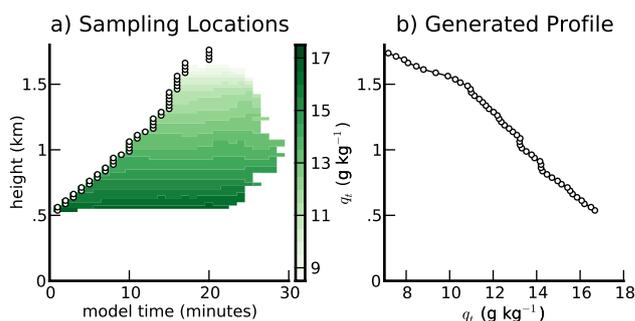


Fig. 14. Method used to generate profiles of environmental properties encountered by a cloud during its initial ascent. **(a)** At each height, the earliest time the cloud is present is used to sample the mean properties of environmental points directly adjacent to the cloud, and the mean cloud points. **(b)** These samples are used to generate a profile of the initial properties encountered by the cloud top.

Our results are presented in Table 1. We find that the environmental q_t , θ_l and θ_ρ encountered by the clouds between 550–750 m are not significantly different for tall clouds than for short clouds. Environmental w is on the borderline for statistical significance, with velocities on average 0.05 m s^{-1} larger between 550–750 m for the tall clouds than the short clouds. On the other hand, cloud top properties show strong differences between tall and short clouds, with only θ_l showing no differences. Cloud top θ_ρ is possibly significant, with more buoyant clouds slightly more likely to become tall, but much more important are the cloud humidity, vertical velocity, and especially horizontal cloud cross-sectional area and mass flux, which have extremely small p-values of 2.2×10^{-16} and 6.8×10^{-21} , respectively; moist clouds, rising quickly, 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 w and cloud top θ_ρ , w , and a (Fig. 15) 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, similar results appear for the properties associated with short and tall clouds change (Table 1). None of the environmental properties encountered by the clouds between 750–1000 m are significantly different for short clouds than for tall clouds. However, all tested cloud top variables show significant differences between the short and tall clouds. Clouds with large q_t , q_l , θ_ρ , w , a , and M , and lower θ_l , between 750–1000 m, are more likely to rise

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. Variables with p-values <0.05 (“possibly significant”) are shaded in light grey, and variables with p-values < 2.3×10^{-3} (“definitely significant”) are shaded in darker grey.

	Variable	Number of tall clouds	Number of short clouds	Mean difference (tall minus short)	Mean difference divided by variable standard deviation	Mann-Whitney U-value	Mann-Whitney p-value
environment 550–750 m		131	126				
	q_t			0.03 g kg ⁻¹	0.14	7391	0.15
	θ_l			-0.0015 K	-0.02	8167	0.89
	θ_ρ			0.004 K	0.09	7603.5	0.28
	w			0.05 m s ⁻¹	0.38	6450	2.5×10^{-3}
cloud 550–750 m							
	q_t			0.031 g kg ⁻¹	0.43	6044	2.1×10^{-4}
	q_l			6.7×10^{-3} g kg ⁻¹	0.41	6386	1.7×10^{-3}
	θ_l			-1.1×10^{-3} K	-0.05	8006	0.68
	θ_ρ			0.018 K	0.37	6450	2.5×10^{-3}
	w			0.08 m s ⁻¹	0.47	6039	2.0×10^{-4}
	a			7748 m ²	0.95	3383	3.0×10^{-16}
	M			6193 kg s ⁻¹	1.04	2667	6.8×10^{-21}
environment 750–1000 m		127	113				
	q_t			0.018 g kg ⁻¹	0.07	6893	0.60
	θ_l			-0.008 K	-0.07	6932	0.65
	θ_ρ			-3.5×10^{-4} K	-0.006	7028	0.78
	w			0.03 m s ⁻¹	0.17	6659	0.34
cloud 750–1000 m							
	q_t			0.17 g kg ⁻¹	1.10	2544	6.3×10^{-18}
	q_l			0.06 g kg ⁻¹	1.05	2836	6.3×10^{-16}
	θ_l			-0.05 K	-1.00	2921.5	2.1×10^{-15}
	θ_ρ			0.10 K	1.02	2918	2.1×10^{-15}
	w			0.15 m s ⁻¹	0.70	4212	3.4×10^{-8}
	a			8100 m ²	1.07	2405	6.3×10^{-19}
	M			11 656 kg s ⁻¹	1.11	1716	2.7×10^{-24}

past 1300 m. Of these properties, cloud top area and mass flux are again the most significant differences.

4.3 Discussion

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 mean upward velocity of plumes in this region is uncorrelated with the plume’s mean buoyancy and plume velocity anomalies dissipate quickly as the plumes rise, suggesting the sub-cloud plume dynamics are dominated by inertia and pressure effects.

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 upward 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

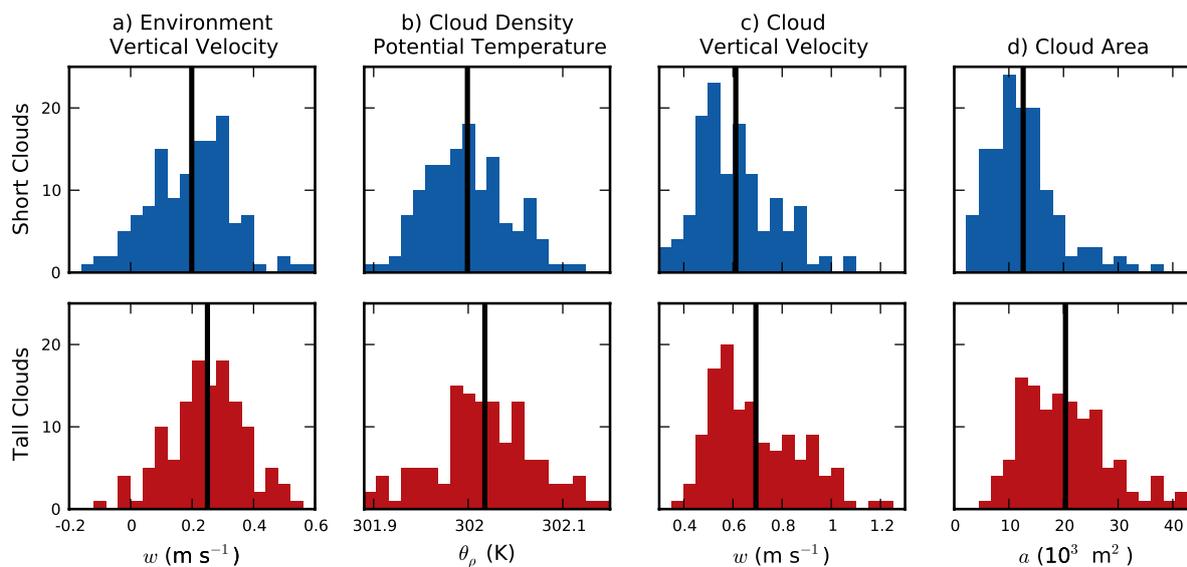


Fig. 15. Probabilities in range Δx ($P(x)\Delta x$) of the mean properties of cloud tops between 550–750 m having a certain (a) environmental vertical velocity, (b) cloud density potential temperature, (c) cloud vertical velocity, and (d) horizontal cloud cross-sectional area. Top row, blue histograms 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.

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 θ gradient, and they are mixed away.

Neggiers et al. (2002) construct a theory in which fractional entrainment rate is inversely proportional to vertical velocity and cloud parcels enter cloud base with a wide range of properties. The variations in parcel properties then set the entrainment rate and thus control the future evolution of the parcel properties. Our results do not support the assumption of Neggers et al. that parcels have a range of initial thermodynamic conditions. We have not directly examined the dependence of entrainment rate on vertical velocity; however, the strong relationship between eventual height reached by the clouds and the cloud area suggests that fractional entrainment is more likely dependent on cloud area and any relationship between vertical velocity and entrainment is due to larger area clouds shielding their cores from entrainment, producing higher buoyancies and vertical velocities.

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 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. 16): 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 an 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

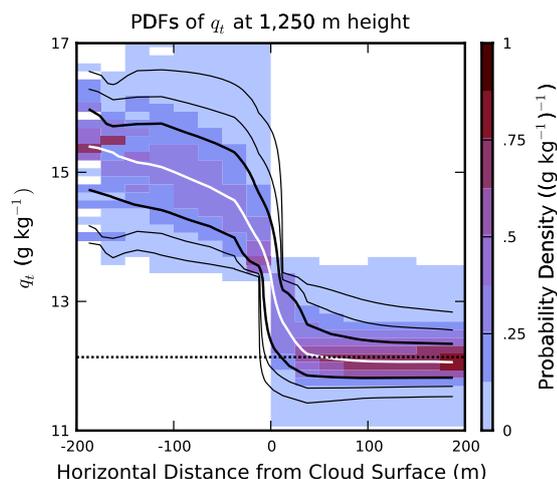


Fig. 16. Variation in the probability density function of total specific water q_t at 1250 m height with distance from the nearest cloud surface. Probability density functions are taken at 25 m distance intervals, to match the model grid spacing. Negative distances represent points inside the cloud. Black lines show contours of the cumulative distribution functions of each variables at 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. Horizontal dotted line shows the domain-average specific humidity.

analyses of shallow cumulus cloud populations. We believe this algorithm could be easily extended to study deep convection. We include in the Supplement 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 , θ_l , and θ_ρ ; 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.net/12/1101/2012/acp-12-1101-2012-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.

Edited by: B. Stevens

References

- 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.
- Blyth, A. M., Lasher-Trapp, S. G., and Cooper, W. A.: A study of thermals in cumulus clouds, *Q. J. Roy. Meteorol. Soc.*, 131, 1171–1190, 2005.
- 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.
- Bretherton, C. S. and Smolarkiewicz, P. K.: Gravity waves, compensating subsidence and detrainment around cumulus clouds, *J. Atmos. Sci.*, 46, 740–759, 1989.
- Brown, A. R., Cederwall, R. T., Chlond, A., Duynkerke, P. G., Gollaz, 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. Roy. Meteorol. Soc.*, 128, 1075–1093, doi:10.1256/003590002320373210, 2002.
- Cahalan, R. F. and Joseph, J. H.: Fractal statistics of cloud fields, *Mon. Weather Rev.*, 117, 261–272, 1989.
- 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.
- Couvreur, F., Hourdin, F., and Rio, C.: Resolved Versus Parametrized Boundary-Layer Plumes. Part I: A Parameterization-Oriented Conditional Sampling in Large-Eddy Simulations, *Bound.-Lay. Meteorology*, 134, 441–458, 2010.
- 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.
- de Rooy, W. C. and Siebesma, A. P.: A simple parameterization for detrainment in shallow cumulus, *Mon. Weather Rev.*, 136, 560–576, 2008.
- 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.

- 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.
- Grabowski, W. W. and Clark, T. L.: Cloud-environment interface instability. Part III: direct influence of environmental shear, *J. Atmos. Sci.*, 50, 3821–3828, 1993b.
- 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.
- 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.
- Jiang, H., Xue, H., Teller, A., Feingold, G., and Levin, Z.: Aerosol effects on the lifetime of shallow cumulus, *Geophys. Res. Lett.*, 33, L14806, doi:10.1029/2006GL026024, 2006.
- 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.
- 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.
- 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.
- 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.
- 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, doi:10.1214/aoms/1177730491, 1947.
- Medeiros, B. and Stevens, B.: Revealing differences in GCM representations of low clouds, *Clim. Dyn.*, 36, 385–399, 2011.
- 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.
- 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.
- Neggers, R. A. J., Siebesma, A. P., and Jonker, H. J. J.: A Multiparcel Model for Shallow Cumulus Convection, *J. Atmos. Sci.*, 59, 1655–1668, 2002.
- 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. Roy. Meteorol. Soc.*, 129, 2671–2696, 2003a.
- 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.
- 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.
- Norris, J. R.: Low cloud type over the ocean from surface observations. Part II: geographical and seasonal variations, *J. Climate*, 11, 383–403, 1988.
- Perneger, T. V.: What's wrong with Bonferroni adjustments, *The British Medical Journal*, 316, 1236, 1998.
- 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.
- 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, 2003.
- Romps, D. M. and Kuang, Z.: Nature versus Nurture in Shallow Convection, *J. Atmos. Sci.*, 67, 1655–1666, 2010.
- Shaffer, J. P.: Multiple Hypothesis Testing, *Annual Reviews of Psychology*, 46, 561–584, 1995.
- Siebesma, A. P. and Cuijpers, J. W. M.: Evaluation of parametric assumptions for shallow cumulus convection, *J. Atmos. Sci.*, 52, 650–666, 1995.
- 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)60<1201:ALESIS>2.0.CO;2, 2003.
- 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.
- 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. Roy. Meteorol. Soc.*, 114, 639–664, doi:10.1002/qj.49711448106, 1988.
- 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.
- 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, 7, doi:10.3894/JAMES.2009.1.7, 2009.
- 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.
- 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.
- 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.