A Numerical Investigation of Cumulus Thermals

DANIEL HERNANDEZ-DECKERS

Climate Change Research Centre, ARC Centre of Excellence for Climate System Science, University of New South Wales, Sydney, New South Wales, Australia, and Departamento de Geociencias, Universidad Nacional de Colombia, Bogotá, Colombia

STEVEN C. SHERWOOD

Climate Change Research Centre, ARC Centre of Excellence for Climate System Science, University of New South Wales, Sydney, New South Wales, Australia

(Manuscript received 29 December 2015, in final form 8 July 2016)

ABSTRACT

Although the steady, entraining, updraft plume is widely taken as the foundational concept of cumulus convection, past studies show that convection is typically dominated by thermals that are transient, more isotropic in shape, and possess interior vortical circulations. Here, several thousand such thermals are tracked in cloud-resolving simulations of transient growing convective events. Most tracked thermals are small (with radius \( R \), \( 300 \) m), ascend at moderate rates (\( \sim 2–4 \) m s\(^{-1}\)), maintain an approximately constant size as they rise, and have brief (4–5 min) lifetimes, although a few are much larger, faster, and/or longer lived. They show slight vertical elongation, but few, if any, would be described as plumes. As convection deepens, thermals originate higher up, are larger, and rise faster, although radius and ascent rate are only weakly correlated among individual thermals. The main force opposing buoyancy is a nonhydrostatic pressure drag, not mixing of momentum. This drag can be expressed in terms of a drag coefficient \( c_d \) that decreases as convection intensifies: deep convective thermals are less damped, with \( c_d \sim 0.2 \), while shallow convective thermals are more damped, with \( c_d \sim 0.6 \). The expected dependence of \( c_d \) based on theoretical form and wave drag coefficients for a solid sphere is inconsistent with these results, since it predicts the opposite dependence on the Froude number. Thus, a theory for drag on cumulus thermals is not straightforward. Overall, it is argued that thermals are a more realistic prototype for atmospheric deep convection than plumes, at least for the less organized convection types simulated here.

1. Introduction

Updrafts are the main dynamical component of atmospheric convection. Details about updraft characteristics strongly affect microphysical processes, which in turn strongly affect the development and water budget of the cumulus. Typical climate models do not resolve these convective updrafts but make particular assumptions about them within the convective and microphysical schemes. Other features, such as cold outflows and gravity waves, are also dynamically important for atmospheric convection. However, we leave these aside and focus only on updrafts.

Perhaps the simplest model of an updraft is that of Stommel (1947), a steady plume with increasing mass flux with height (i.e., the well-known entraining plume). This simple idea constitutes the conceptual model of updrafts in most current convection schemes (e.g., Arakawa and Schubert 1974; Tiedtke 1989; Kain and Fritsch 1990). The entraining plume transports mass, moisture, and heat from cloud base to cloud top, while it mixes with the environment. Several such entraining plumes may be considered, each having different heights and mixing properties. However, the conceptual model behind each plume remains the same.

Parameterizations based on this plume model have been successful in global atmospheric models throughout past decades. However, these parameterizations show deficiencies that are proving to be hard to fix by adjusting the existing schemes (Derbyshire et al. 2004; Del Genio 2011; Mapes and Neale 2011). A fundamental problem is that an entraining plume model cannot simultaneously simulate liquid water profile and cloud-top height (Lin and Arakawa 1997). Very low
entrainment rates are required to produce realistic cloud-top heights, at the expense of too much condensate at cloud top and little sensitivity to free-tropospheric humidity. Multiplume ensembles or interactive entrainment models (e.g., Arakawa and Schubert 1974; Mapes and Neale 2011) can mitigate these issues but often create complications by requiring additional assumptions or parameters.

Although the entraining plume model provides a simple and useful view of cumulus convection, real cumulus clouds consist primarily of transient bubbles or thermals, which cause their typical cauliflower-like appearance. This has been argued both from observations and laboratory studies since the 1950s (e.g., Scorer and Ludlam 1953; Woodward 1959; Sánchez et al. 1989; Blyth et al. 2005; Damiani et al. 2006). In spite of this, to our knowledge most current cumulus parameterizations either assume plumes or make no explicit assumption about the kinematics of the flow, except for a few recent developments (e.g., Suselj et al. 2013; Romps 2016). The “bubble theory of convection” was thought to be fully consistent and interchangeable with the entraining plume concept (Ooyama 1971). However, as Yano (2014) points out, this might not be the case, in particular because the entraining plume model does not take into account the transient nature of atmospheric convection. Regardless of whether or not explicitly representing convection in terms of thermals would offer any practical advantage, it seems worthwhile to study the dynamics of such thermals and further quantify the extent to which they are the true core entities of convection.

Updrafts—and downdrafts—have been extensively studied both through observations and simulations. Updrafts are typically sampled following LeMone and Zipser (1980), who define updraft cores as regions of at least 500 m with vertical velocity larger than 1 m s$^{-1}$. Updraft statistics can be deduced from observations or cloud-resolving models. For example, Igau et al. (1999) and Xu and Randall (2001) describe how buoyant downdraft cores and negatively buoyant updrafts occur often; Anderson et al. (2005) compare continental and oceanic updraft cores, finding almost no differences regarding average speed, size, and mass flux; Mrowiec et al. (2012) simulate one convective system during the Tropical Warm Pool–International Cloud Experiment (TWP-ICE) and describe draft properties identifying these in a similar way as LeMone and Zipser (1980). Furthermore, profilers and radar observations have also been used to study updraft properties (e.g., May and Rajopadhyaya 1999; Cifelli et al. 2002; Collis et al. 2013).

These and other studies have contributed a broad picture of the main properties of updrafts in cumulus clouds. Our approach redefines the core element as the thermal, a coherent rising structure that encompasses both the updraft and part of the surrounding fluid that is dynamically connected to it. These thermals should constitute a more physical approach that may provide new insights into the dynamics of cumulus clouds and potentially their role in climate (Sherwood et al. 2013, hereafter SHCR13).

This study makes use of high-resolution simulations of convective events to study the properties and behavior of thermals in detail. We focus on the main physical characteristics of thermals: their size, ascent rate, lifetime, distance traveled, and vertical momentum budget. Another aspect that will be crucial for developing new parameterizations is mixing. We present here only a simple analysis of mixing in relation to its impacts on thermal size with height and on the momentum budget. A more thorough and detailed analysis is left for a future paper that focuses on mixing.

In section 2 of this paper we describe the cloud-resolving model and the simulation setup, followed by a method section, where we describe the algorithm used to track and analyze thermals. Section 4 presents the results in terms of the main physical properties of the thermals and their momentum budget, and section 5 summarizes our findings.

2. Model and simulation setup

We run version 3.4.1 of the Weather Research and Forecasting (WRF) Model created by the National Center for Atmospheric Research to simulate convection explicitly in 3D. We present results from three idealized experiments running at 65-m grid length: two sea-breeze-type experiments with different island and domain sizes and one case of daytime convective development over land based on observations from the Large-Scale Biosphere–Atmosphere (LBA) field study. We denote these experiments as SB1, SB2 and LBA. All simulation data are available upon request.

Table 1 summarizes the main parameters of the domain setups. Sea-breeze simulations consist of a centered island extending along the y direction surrounded by water in the x direction. Boundary conditions are periodic in the y direction—creating an infinitely long island—and open in the x direction. Diurnal solar radiation is used.

<table>
<thead>
<tr>
<th>Name</th>
<th>Domain size</th>
<th>Island width</th>
<th>Model top</th>
<th>$\Delta t$</th>
<th>$N_{AT}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SB1</td>
<td>50 km x 15 km</td>
<td>30 km</td>
<td>20 km</td>
<td>0.5 s</td>
<td>6</td>
</tr>
<tr>
<td>SB2</td>
<td>80 km x 20 km</td>
<td>50 km</td>
<td>20 km</td>
<td>0.5 s</td>
<td>6</td>
</tr>
<tr>
<td>LBA</td>
<td>40 km x 40 km</td>
<td>—</td>
<td>20 km</td>
<td>0.33 s</td>
<td>6</td>
</tr>
</tbody>
</table>
for a latitude of 30°N during January. We initialize these simulations at sunrise with clear-sky conditions, no wind, and temperature and humidity profiles for Darwin at 1200 UTC 11 February 2013. This sounding has a large CAPE of about 2300 J kg\(^{-1}\), with 44.8 mm of rain recorded during the following 24 h at Darwin airport. This setup was chosen to facilitate deep convection development in a short simulation time.

The LBA experiment is based on the setup described by Grabowski et al. (2006): the 23 February 1999 case during the LBA field study in the Amazon. It features monsoon-type convection with weak mesoscale organization and a strong diurnal cycle. We run this experiment from 0730 local time for 6 h, after which deep convection has developed but without mesoscale organization. Instead of specifying evolving surface fluxes and temperature tendencies due to radiation as in Grabowski et al. (2006), the model computes them interactively, assuming diurnal solar radiation at a latitude of 10°S during February and land covered by wooded wetland.

In all simulations, small random perturbations are added during the initial time steps near the surface to initiate convection. Cumulus and boundary layer schemes are switched off. We use Thompson’s microphysics scheme, a 1.5-order TKE closure, MM5 similarity scheme based on Monin–Obukhov theory for surface fluxes, and Rapid Radiative Transfer Model (RRTM) and Dudhia schemes for longwave and shortwave radiation, but with cloud radiative effects switched off to reduce the computational cost. Further details of these schemes, boundary conditions, and numerics are given by Skamarock et al. (2008).

Yamaguchi and Feingold (2012) have reported a particular issue in WRF that may impact liquid water path and cloud buoyancy because of a problem in the dynamical core. Although more recent versions of WRF have solved this issue (from version 3.7 onward), using a smaller acoustic time step also avoids this problem. We have verified that this has no impact on the results presented here by running an additional LBA simulation using 12 acoustic time steps instead of 6 and performing all the corresponding analysis (not shown here).

Figure 1 shows the evolution and intensity of convection in each simulation in terms of vertical velocity \(w\). Deep convection in the sea-breeze experiments is driven by two sea breezes that approach from opposite coasts, causing an abrupt transition from shallow to deep convection. Notice that the larger the island the more intense convection is (Robinson et al. 2011), and the later it deepens, since the sea breezes must travel farther before they converge.

From these qualitative impressions we define convective stages in each experiment, which we use for the analysis hereafter (see Fig. 1). In SB1 we define three stages that roughly describe an initial shallow stage (A), a slightly deeper one (B), and a deep stage (C). Similar stages are defined for SB2, except for B, which is now subdivided into two increasingly deeper stages: B1 and B2. Experiment LBA exhibits a more gradual deepening of convection because it is driven by surface fluxes rather than by the sudden convergence of sea breezes. For this simulation we have chosen three non-contiguous stages for tracking and analyzing thermals, which correspond to a shallow stage (A), an intermediate (congestus) stage (B), and a deep convective stage (C).

This idealized case (LBA) is often used to study the diurnal transition from shallow to deep convection over land and the processes relevant to it, such as mixing and the presence of cold pools. Although we do not pursue this direction, it is worth noting that overall our simulation has very similar features to those reported by other studies of the same case (e.g., Khairoutdinov and Randall 2006; Wu et al. 2009; Böing et al. 2012). Cloud tops rise gradually and reach \(\sim 12\) km by the end of the simulation; surface precipitation intensifies significantly just before noon, indicating the transition to deep convection (see Fig. 2). Compared to previous studies of the same case, precipitation initiates slightly earlier in our simulation, and a stratiform layer of cloud is generated around the freezing level, which is not reported in other studies. This layer might contribute to a faster transition to deep convection once updrafts reach this saturated region around 1100 local time but otherwise does not have any strong impact.

Here we use the LBA case to investigate thermals forced by a different mechanism to the convergence of sea breezes, as well as under a different thermodynamic state.
This makes our results more representative, although to fully generalize them a much larger spectrum of settings and models will need to be investigated in the future.

3. Method

a. Identification and tracking of thermals

The identification and tracking of thermals is done offline, using output data stored at intervals of 1 min. It requires the velocity, pressure, temperature, and water content fields. Our algorithm is an improved version of the method by SHCR13.

The first step is to identify thermal centers from the vertical velocity field $w$. Using the maximum filter from the multidimensional image processing module of SciPy (http://scipy.org), we identify local maxima of $w$ at each output time step with $w > 0.8 \text{ m s}^{-1}$ and at least 500 m apart from each other. We call these $w_{\text{max}}$ points, with $w_{\text{max}}$ denoting the corresponding value of $w$. The 500-m distance separation between $w_{\text{max}}$ points ensures that this simple clustering algorithm works unambiguously. Next we identify clusters of successive $w_{\text{max}}$ points that correspond to the same thermal: for each $w_{\text{max}}$ point, we identify the most probable next $w_{\text{max}}$ point and the most probable previous $w_{\text{max}}$ point; this assumes that thermals rise roughly at a speed of half $w_{\text{max}}$ (SHCR13). To account for possible accelerations, we consider next $w_{\text{max}}$ points as those that are up to a distance $1.5 dt(w_{\text{max}})$ above and previous $w_{\text{max}}$ points as those that are up to a distance $1.5 dt(w_{\text{max}})$ below, where $dt$ is the output time step (i.e., 60 s). To ensure correct tracking, two successive $w_{\text{max}}$ points are only clustered if each is the corresponding next/previous $w_{\text{max}}$ point of the other. Clusters of corresponding $w_{\text{max}}$ points with at least three points are kept, whereas those with fewer are discarded.

The initial estimates of the thermal center positions ($w_{\text{max}}$ points) is improved by smoothing their trajectories. We do this by fitting a second- or third-order polynomial (depending on the number of points in time) to the positions of the $w_{\text{max}}$ points. Note that smoothed trajectories are not fixed to the model grid, whereas the $w_{\text{max}}$ points are. These smoothed trajectories are not only better estimates of the positions of the thermal centers; they also provide an estimate of the velocity components $U(t), V(t),$ and $W(t)$ via the polynomials' first derivatives. This polynomial fit captures the acceleration of the thermals, as opposed to SHCR13’s method, which fits a straight line, thus assuming constant velocity. From here on we assume that the thermal as a whole moves with this velocity ($U, V,$ and $W$), not to be confused with the velocity at the centers. Finally, we remove any points where $W < 0.1 \text{ m s}^{-1}$ and verify that the average $W$ throughout every thermal’s tracked lifetime is at least $1 \text{ m s}^{-1}$. If not, we remove $w_{\text{max}}$ points from the beginning or the end—whichever has lower $w_{\text{max}}$—until this condition is met or the entire thermal is discarded.

Once the centers and velocity components of each thermal are known, we determine their sizes (i.e., their boundaries). Following SHCR13, we assume spherical thermals and find each radius such that the average vertical velocity of the enclosed volume matches the previously estimated vertical velocity $W$ at each output time step. This ensures physical consistency between the trajectories of the $w_{\text{max}}$ points and the vertical displacement of the fluid around them as a coherent structure.

In all experiments we track only thermals that are contained within a $12 \text{ km} \times 14 \text{ km}$ box. In the sea-breeze experiments this region is centered on the island, ensuring that both shallow and deep convection take place within this box. Deep convection in LBA is not homogeneously distributed, with some areas devoid of deep convection. Thus, in LBA the intensity of convection during stage C is sensitive to the exact location of the $12 \text{ km} \times 14 \text{ km}$ box. We therefore choose the location of this box such that it encloses a particular deep convective region during stage C.

In total we tracked 2855, 3532, and 2382 thermals throughout the different stages of the SB1, SB2, and LBA experiments inside the $12 \text{ km} \times 14 \text{ km}$ box. The lower panels of Fig. 3 show how the number of tracked thermals evolves over time. In the sea-breeze experiments the number of tracked thermals slowly increases during the initial stages (A and B), decreases just before the sea breezes collide, and increase rapidly during stage C. This apparent suppression of thermals during the sea-breeze convergence is not present throughout the LBA experiment, in which after an initial increase in the number of tracked thermals during stage A, it remains...
b. Vertical momentum budget

The vertical momentum budget of a spherical thermal of mass $M$ can be approximated as follows:

$$\frac{dW}{dt} = \frac{1}{M} \int p n_z \, dS - g + F_{\text{mix}} + F_{\text{unr}} + F_{\text{entr}}.$$  

(1)

The first term on the right is the pressure gradient expressed as the surface integral of pressure times $n_z$, the vertical component of the normal vector $\hat{n}$; $F_{\text{unr}}$ denotes the unresolved momentum fluxes (per unit mass), which we neglect as they appear to be very small; and, $F_{\text{mix}}$ is the resolved convergence of vertical momentum flux, expressed as a closed integral along the thermal’s surface:

$$F_{\text{mix}} = \frac{1}{M} \int \mathbf{u'} \cdot \hat{n} \mathbf{w'} \, dS,$$  

(2)

where $\mathbf{u'} = (u', v', w')$ is the velocity vector relative to the thermal’s velocity $\mathbf{U} = (U, V, W)$. Thus, $F_{\text{mix}}$ accounts for the exchange of vertical momentum due to instantaneous entrainment and detrainment. Finally, $F_{\text{entr}}$ is the contribution due to the change in size of the thermal in time (i.e., net entrainment or detrainment):

$$F_{\text{entr}} = \frac{1}{M} \frac{dM}{dt} \overline{w'},$$  

(3)

where $\overline{w'}$ is the vertical velocity relative to the thermal averaged along the thermal’s surface. Thus, $F_{\text{entr}}$ is a drag term when the thermal grows ($dM/dt > 0$), and $\overline{w'}$ is negative. Actual thermals may also change in shape with time, which would also induce another term. However, this term does not occur because of our spherical-thermal assumption, and as long as actual thermals remain approximately spherical—which we observe qualitatively in our simulations—this effect should be much smaller than the thermal’s change in size.

Note that the pressure gradient term, together with the gravitational acceleration $g$, comprise buoyancy plus any other nonhydrostatic contributions. We can therefore approximate Eq. (1) as

$$\frac{dW}{dt} \approx B + F_{\text{nh}} + F_{\text{mix}} + F_{\text{entr}},$$  

(4)

where $F_{\text{nh}}$ is any nonhydrostatic contribution to the pressure field around the thermal, and $B$ is buoyancy, which we compute from the density field:

$$B = -\frac{1}{M} \int g \left( \frac{\rho - \overline{\rho}}{\rho} \right) \rho \, dV.$$  

(5)

Here, $\overline{\rho}$ is computed as the average buoyancy over the thermal’s environment, which we take as a volume that extends 10 times the thermal’s radius $R$ on each side, throughout a vertical layer of thickness $2R$. This choice of environment is more precise than using the entire simulation domain. For typical thermal sizes and ascent rates, a gravity wave of wavelength $R$ would travel a shorter distance than $10R$ while the thermal rises $2R$. Thus, densities farther than this distance are irrelevant for the thermal’s buoyancy. Nevertheless, we tested alternative choices for the thermal’s environment (i.e., $30R$ and the entire domain), finding no impact on the results presented here.

c. Sanity checks and algorithm performance

The vertical momentum budget can be used to test the tracking algorithm and reject thermals whenever their expected trajectories do not match their tracked trajectories. Using a thermal’s momentum budget and its initial position and velocity, we predict its final position and compare it with the tracking result. If the mismatch is too large, we remove time steps from the beginning or the end of the tracked period—whichever has lower updraft speed—until the mismatch is acceptable or until we discard approximately constant until it slightly decreases by the end of stage C.

FIG. 3. Time series of (bottom) number of tracked thermals $N$ and (top) mass flux for the three experiments. In (top), gray lines indicate the total mass flux according to $\phi_1$, and black lines indicate the thermals’ mass flux $\phi_{th}$. Numbers indicate the percentage of total mass flux captured by the thermals averaged during each stage. Vertical dotted lines indicate the tracking periods.
the entire thermal. An acceptable mismatch is 1) less than twice the thermal’s average radius or 2) less than 20% of its vertical displacement. Condition 2 is meant to avoid giving preference to shorter-lived thermals. An indication of the physical consistency of our algorithm is that very few thermals—less than 0.4% of the identified clusters of updrafts—are rejected as a result of momentum budget mismatches (m. budg.; see Table 2).

Apart from momentum budget mismatches, there are other reasons for discarding clusters of updrafts. These reasons and the corresponding percentage of discarded cases are summarized in Table 2. For example, to ensure enough contrast in the vertical velocity field so that the tracking algorithm is reliable, we reject thermals with ascent rate $W$ lower than 1 m s$^{-1}$ averaged over their lifetime (low $W$). We also reject time steps of thermals in which they are too small, i.e., with $R$ smaller than twice the horizontal grid length, so that they are adequately resolved (small $R$). Another cause for rejection is a large change in $R$ from one output time step to the next. This would suggest an incorrect estimate of $R$, so whenever this occurs we interrupt the thermal tracking and discard it if it is too short (large $\Delta R$). After visual verification of several thresholds, we opted for considering large a change of at least 80% of the smallest $R$. The last reason for rejecting clusters of updrafts is when the algorithm is unable to find a suitable radius, most probably as a result of nearby thermals and/or turbulence (invalid $R$).

Table 2 shows that invalid $R$ cases account for very few rejections, indicating that the algorithm is mostly able to find a suitable radius. The highest rejection rates because of this reason (6%–10%) occur during deep convective stages of SB1 and SB2, consistent with more thermals and more turbulence around them. Similarly, cases with large $\Delta R$ are few, with the highest numbers (15%–20%) during the deep convective stages of SB1 and SB2. Overall, one could argue that algorithm deficiencies would cause rejections because of invalid $R$, large $\Delta R$, and momentum budget mismatches. Rejections due to these three reasons combined typically account for less than 10% of rejections, except for the deepest convective stages of SB1 and SB2 (23%–33%) and LBA (15%). Overall, these rejections account for ~17% of the detected updraft clusters, which we consider satisfactory, given the simplicity of the tracking algorithm.

Most rejections are due to small $R$ or low $W$ (25.9% and 18.4% overall). There is roughly an inverse relationship between these two rejection causes: as convection deepens, fewer thermals are rejected because of small $R$, but more thermals are rejected because of low $W$. During the initial shallow stages there are large numbers of very small thermals, which cannot be tracked; as convection deepens, the number of small thermals decreases, but many of these slightly larger thermals are not fast enough to be properly tracked. Thus, the main limiting factor of our tracking algorithm is very small or very slow thermals, but these are of less interest anyway.

Overall, out of the initially detected clusters of updrafts, the algorithm tracks 38.3% of them and, during specific stages of each simulation, even more than 60%. However, the algorithm does not necessarily track thermals during their entire lifetimes, but rather during their most intense periods. To really quantify how representative these tracked thermals are of the entire convective activity, we must compare the thermals’ mass flux to the total convective mass flux. The mass flux density of an ensemble of thermals $\phi_{th}$ is given by

$$\phi_{th} = \sum_{i} \frac{M_i W_i}{\mathcal{P}},$$

where $M_i$ is the mass of the $i$th thermal, $W_i$ its ascent rate, and $\mathcal{P}$ the control volume. An equivalent
definition for total mass flux density depends on the updraft sampling criteria. We opt for a simple criterion commonly used in similar studies [e.g., Romps and Charn (2015)]: grid cells with \( w > 1 \text{ m s}^{-1} \) and cloud water content in excess of \( 10^{-5} \text{ kg kg}^{-1} \). Mass flux is then computed in the same way as it is for thermals, but summing over grid cells.

Figure 3 shows how these two mass flux estimates evolve in each simulation. Overall, tracked thermals capture about 18% of the total convective mass flux and between 15% and 23% for individual convective stages (Table 2). This captured fraction of mass flux is significantly higher than previous methods, in particular previous versions of this same method (SHCR13) and the more recent approach by Romps and Charn (2015), which captures 3% of the total mass flux. We argue that thermals tracked here are reasonably representative of overall convective activity. A clear indication is that the time series of the number of thermals and their mass flux evolve consistently with the intensity of convection (Fig. 3). Plus, we capture a similar fraction of the total mass flux throughout different stages of the experiments, despite total mass flux varying by more than two orders of magnitude in some cases. Thus, there is no indication that the representativity of the tracked thermals changes as convection strengthens or weakens.

We can also verify this qualitatively by inspecting the vertical velocity field and comparing tracked thermals with other updrafts. We do this here for SB1 (Fig. 4). Updrafts can be identified from the color scale, while circles indicate tracked thermals. During stages A and B (Figs. 4a,b) it is hard to find any updraft that does not resemble one of the tracked thermals. During stage C (Fig. 4c) the flow becomes more turbulent, and thermals are harder to identify visually. Nevertheless, updrafts that are not tracked do not look different from the tracked ones. Finally, vertical mass flux profiles (right panels of Fig. 4) suggest that, although thermals capture only a fraction of it, their mass flux approximately follows the main features of the total mass flux profile.

### d. Entrainment

We estimate instantaneous entrainment and detrainment rates in the same way as SHCR13, integrating mass flux across the thermal’s surface in the frame of reference of the moving thermal:

\[
\varepsilon = \frac{\int \mathbf{u}' \cdot \hat{n} H(\mathbf{u}' \cdot \hat{n}) \rho \, dS}{\int_V \rho w \, dV},
\]

where \( \varepsilon \) is the fractional entrainment rate \( (\text{m}^{-1}) \), \( H() \) is the step function (only air coming into the thermal counts), \( \mathbf{u}' \) is the velocity relative to the thermal, and \( \hat{n} \) is the unit vector pointing inwards from the thermal’s surface. Thus, we compute \( \varepsilon \) as the ratio of influx of mass along the thermal’s surface to the total mass transport due to the thermal’s ascent. For this, we linearly interpolate the velocity field to the thermal’s surface, which we discretize by subdividing the two relevant angles such that each surface element \( dS \) has two perpendicular sides of approximately one-fourth the model grid length. Since \( \varepsilon \) is an instantaneous estimate of entrainment, detrainment will always equal entrainment as a result of mass conservation. This estimate of \( \varepsilon \) is a relatively simple direct estimate of entrainment, compared to other direct estimates (e.g., Romps 2010; Dawe and Austin 2011). Furthermore, \( \varepsilon \) yields values for each output time step of each tracked thermal but does not reflect any net entrainment or detrainment. We compute net entrainment in between two output time steps using the thermal’s change in mass with height:
We compute this as \( \Delta M / (\bar{M} \Delta z) \), where \( \bar{M} \) is the average mass and \( \Delta M \) the change in mass while the thermal travels a distance \( \Delta z \). Thus, we obtain instantaneous entrainment (or detrainment) at output time steps and net entrainment halfway between output time steps.

### 4. Results

Results presented here have been obtained using 1-min intervals when tracking. We also tested 30-s intervals during 30 min of the LBA simulation (not shown here) and found no significant difference in any of the thermal properties analyzed here. Average quantities are computed in two ways: with equal weights for all thermals (unweighted) or weighted according to mass flux in order to focus on the most active thermals. The type of averaging is clearly stated in each case.

#### a. Shape: Thermals or plumes?

We bound thermals with spheres, but thermals are not necessarily spherical. We evaluate the average shape of thermals by inspecting a composite of the (relative) vertical velocity field of thermals scaled by their radii (Fig. 5). Mass-flux-weighted composites (right panels) are slightly noisier, since more weight is given to a few thermals, and have higher values of vertical velocity; but, overall, they have a very similar circulation to that of the unweighted composites. On average, thermals are not exactly spherical but are slightly elongated downward (i.e., ellipsoidal rather than spherical), with an eccentricity of about 1.05 and peak \( w \) slightly above the center. Given the small eccentricity, the spherical approximation still seems reasonable. To obtain a better match with the vortex ring structure, one can slightly shift downward the spheres’ centers. To do this, we first estimate the radius with the center at its original position, shift it downward a distance proportional to the estimated radius, and reestimate the radius to make it consistent with the new position. We have done this (not shown here) and found no significant impact in any of the properties analyzed here, except for a slight difference in entrainment rate, which will be described in section 4f. Otherwise, results shown here are for unshifted thermals but hold as well for shifted thermals.

The average shape of thermals does not change throughout the different stages of the simulations. However, it does change throughout the thermals’ lifetimes. To investigate this, we define a reference time for each thermal based on its \( w_{\text{max}} \). For each thermal, \( w_{\text{max}} \) typically reaches one maximum and then decreases. We set \( t = 0 \) when \( w_{\text{max}} \) reaches its peak value and construct composites for several values of \( t \) symmetric around zero.

The elongation of the thermals is largest at the beginning of their lifetime (Fig. 6) and decreases with time, suggesting that elongation is related to acceleration, usually largest for \( t = 0 \).

Apart from the average shape, we are interested in determining how much individual thermals deviate from this shape. For this we inspect the relative vertical velocity field of individual thermals. Assuming a vortex ring type of circulation such as that from Fig. 5, air rising faster than the thermal is contained within the thermal (except for other nearby thermals) and has a quasi-elliptical shape. A proxy for the shape of individual thermals is the positions of the edges of this upward-moving core with respect to the thermal’s center. Figure 7 shows the distributions of these core edges, which are the distances from the thermal’s center to where \( w' \) vanishes in each
direction along a straight line. These are smooth enough to suggest only one shape of convective elements, with some variability around that shape. Thermals’ shapes are smoothly distributed around an ellipsoid, and the deviations seem random. This rules out the possibility that the slight elongation would be due to a few very elongated plumes.

b. Characterization of the thermals

Here we describe the main characteristics of thermals: that is, their typical size, ascent rate, lifetime, starting height, and distance traveled. Table 3 shows mean values for each experiment. Unweighted average size is relatively small ($R < 300$ m) and similar throughout experiments. Mass-flux-weighted averages show that thermals that carry significant mass flux tend to be larger, and how much larger depends on convective intensity. For example, in LBA, where convection is weakest, the weighted average radius is only a factor of 2 larger than the unweighted average; but in SB2, where convection is strongest, there is a factor of almost 5.

Thermals have moderate ascent rates ($\sim 2–4$ m s$^{-1}$), with mass-flux-weighted averages more sensitive to convective intensity than unweighted averages. Average lifetime is 4–5 min and only slightly shorter when mass-flux weighted. However, our algorithm does not necessarily track a thermal’s entire lifetime. Tracked thermals represent 38% of all detected updraft clusters but capture $\sim 18\%$ of the total mass flux, so our algorithm most likely underestimates thermal lifetime. Thermals travel 500–700 m and slightly more (700–1100 m) for mass-flux-weighted averages. In general, unweighted average properties are very similar between simulations, while mass-flux-weighted averages differ considerably, indicating that differences in convective

---

**Fig. 6.** Unweighted composite of streamlines and vertical velocity relative to the thermals ($w'$), scaled by their radii at three different instants in a thermal’s lifetime ($t = -3, 0, 3$ min). Composites are constructed by setting a reference time when $w_{\text{max}}$ peaks for each thermal ($t = 0$) and averaging over corresponding instants in a thermal’s lifetime. (left) The SB1 simulation; (right) the LBA simulation.

**Fig. 7.** Distribution of the thermal’s core edges (where $w'$ vanishes) measured along a straight line (left) in the $z$ direction and (right) in the $x$ direction for the entire population of thermals of the three simulations. The farthest distance considered is 4 times the thermal’s radius, so cases in which $w'$ does not vanish by then accumulate in the last bin. Dashed black lines indicate the mean value of the distribution.
TABLE 3. Average values of the main characteristics of thermals in each experiment. Data for each experiment includes (top) unweighted averages and (bottom) the mass-flux-weighted averages.

<table>
<thead>
<tr>
<th>Expt</th>
<th>R (m)</th>
<th>W (m s⁻¹)</th>
<th>Lifetime (min)</th>
<th>Z₀ (km)</th>
<th>ΔZ (m)</th>
<th>ε₂₀₀ (×10⁻⁵ m⁻¹)</th>
<th>ε₅₀₀ (×10⁻⁵ m⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SB1</td>
<td>276</td>
<td>3.3</td>
<td>4.5</td>
<td>3.27</td>
<td>682</td>
<td>−0.34</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td>1156</td>
<td>5.0</td>
<td>3.9</td>
<td>5.25</td>
<td>907</td>
<td>−0.14</td>
<td>0.9</td>
</tr>
<tr>
<td>SB2</td>
<td>290</td>
<td>3.3</td>
<td>4.5</td>
<td>3.11</td>
<td>689</td>
<td>−0.44</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td>1433</td>
<td>6.8</td>
<td>3.7</td>
<td>5.92</td>
<td>1122</td>
<td>0.03</td>
<td>0.7</td>
</tr>
<tr>
<td>LBA</td>
<td>247</td>
<td>2.4</td>
<td>4.6</td>
<td>2.65</td>
<td>509</td>
<td>0.03</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td>467</td>
<td>3.0</td>
<td>4.7</td>
<td>3.63</td>
<td>723</td>
<td>0.003</td>
<td>1.8</td>
</tr>
</tbody>
</table>

intensity are mostly noticeable within the most mass-flux-contributing thermals.

The distributions of thermal characteristics of each experiment (Fig. 8) reflect the similarity between SB1 and SB2 and the contrast of these relative to LBA. LBA differs the most from SB1 and SB2 on the right tails of the distributions, which is where mass-flux-weighted averages are useful. LBA has fewer large thermals (with \( R > \sim 500 \) m) but, most of all, fewer fast thermals (with \( W > \sim 3 \) m s⁻¹) than SB1 and SB2. Interestingly, despite the differences in convective intensity, thermal size, and ascent rate, thermal lifetime is almost identical in all cases. Thus, thermals travel the shortest distances in LBA. Finally, unweighted average size might be slightly overestimated. Given the large percent of rejected thermals as a result of their small size, the distribution in Fig. 8a is probably truncated at \( R = 2 \Delta z = 130 \) m. However, if the turbulent fluid is fractal, there may be no limit to the “true” number of thermals, and one must define the minimum size of interest, which we believe is appropriate here. Furthermore, since the most relevant thermals are those with considerable mass flux, the relevant sizes are clearly larger than the peaks of the distribution.

In general, we find more differences between different convective stages than between different experiments. Figure 9 shows the time evolution of several average and extreme thermal characteristics throughout SB1 and LBA. Mean \( R \) increases with convective intensity, reaching its maximum during stage C. Notice how mass-flux-weighted averages (blue lines) vary even more throughout time than unweighted averages. This is even more pronounced in terms of the largest thermals. Ascent rate \( W \) behaves similarly, with the average fastest thermals during stage C and the slowest during stage A. In fact, time series of average \( R \) and \( W \) are highly correlated, with correlation coefficients of 0.89, 0.96, and 0.79 in SB1, SB2, and LBA, respectively. This correlation decreases to 0.34, 0.49, and 0.20 for individual thermals, so although larger thermals do tend to be faster, they are not consistently faster. This weaker correlation is probably due to (i) noise in estimates, (ii) other influences on \( W \) other than drag, and/or (iii) variations in the drag for a given \( R \).

Most thermals initiate close to the boundary layer (\( Z₀ \approx 2 \) km), as expected (Fig. 8e). However, thermals do initiate higher, but only after previous thermals have reached this height. Notice that the starting height of thermals increases gradually, not suddenly or randomly (Fig. 9d). Thus, the probability of the onset of thermals at a certain level is not zero once convection has reached this level. Our algorithm may overestimate starting heights of some thermals (e.g., in some of the mismatched momentum budget or invalid \( R \) cases), but since such cases are rare, this feature is most likely real.

Thermals also travel the farthest during stage C, but still less than 1 km on average. Therefore, what makes thermals reach higher as convection deepens is not that they travel farther or live longer, but rather that they originate higher up. In fact, average lifetime is slightly shorter during stage C than during stages A and B. In general, time evolution of thermal characteristics in SB2 (not shown here) is similar to that of SB1, except that average values of \( R \) and \( W \) are slightly higher, consistent with values in Table 3.

Other features of the time evolution of convection can be seen from the distributions at different stages of each simulation. We show this for SB1 (Fig. 10). Distributions during the different stages of SB2 and LBA (not shown here) show very similar features to SB1, except that in LBA differences between stages are less pronounced than in SB1 or SB2. Regarding thermal size, during the initial stages A and B only small thermals are present; over time larger ones emerge. At stage C the distribution changes most radically: the occurrence of smaller thermals decreases by about one-third for \( R < 200 \) m, and suddenly larger thermals appear: thermals with \( R \approx 500 \) m—which are almost nonexistent during A and B—become much more frequent during C. However, it is important to note that even during the most intense convective stages, the peak of the distribution remains around small thermals (\( R \sim 200–300 \) m). The differences relevant to convective intensity are mainly visible in the tails of the distributions.
Distributions of ascent rate $W$ in SB1 (Fig. 10b) and SB2 (not shown here) evolve in such a way that, as convection deepens, the distribution grows toward larger values of $W$, keeping roughly the same number of slow thermals ($W \leq 2–3 \text{ m s}^{-1}$). For LBA (not shown here), a similar growth of the distribution toward larger values of $W$ occurs, but with a simultaneous reduction of about one-half in the occurrence of slow thermals. In terms of thermal lifetime, we find no significant differences between stages in any of the experiments. Out of hundreds of tracked thermals in each stage there are never more than a handful of thermals that live longer than 12 min, and average thermal lifetime remains approximately constant throughout all the experiments. Thus, $\Delta Z$ is mostly driven by changes in ascent rate, not in lifetime. Distributions of starting height of the thermals $Z_0$ show a clear transition between stages B and C, consistent with the intensification of convection, which is rather sudden in the sea-breeze cases and more gradual in LBA. However, it is important to keep in mind that $Z_0$ only increases after prior excursions of thermals reach that height.

Although differences are subtle, it is possible to say that during C we find the largest thermals, the fastest thermals, and the ones that travel farthest. Out of these three characteristics, $\Delta Z$ and $W$ is the only combination that is strongly correlated at the thermal level (Table 4). Thus, the faster thermals are not necessarily the larger ones, nor do the larger ones travel farther. Of course, faster thermals travel farther, as do longer-lived thermals. But because lifetime does not vary as much as $W$ between stages, $\Delta Z$ is mostly controlled by $W$. 

**Fig. 8.** Distributions of (a) thermal radius $R$, (b) mean ascent rate $W$, (c) lifetime, (d) vertical distance traveled $\Delta Z$, and (e) starting level $Z_0$ for each of the simulations. Normalization is based on the tracking time of each experiment and the area where thermals are tracked.

**Fig. 9.** Time series for (left) SB1 and (right) LBA of (a) thermal radius $R$, (b) ascent rate $W$, (c) lifetime, (d) starting altitude $Z_0$, (e) distance traveled $\Delta Z$, and (f) vertical momentum budget terms, averaged over all thermals at every output time step. Gray lines in (a)–(e) indicate maximum values; blue lines indicate mass-flux-weighted average values; and solid lines in all panels indicate unweighted average values at each output time step.
The vertical profiles of mass-flux-weighted average ascent rate (Fig. 11c) depict clearly the differences in convective intensity between simulations. The profiles are qualitatively similar, but with different magnitudes. Thermal ascent rate increases almost linearly with height, although individual thermals do not accelerate much, so this increase with height must reflect a correlation between starting altitude and velocity. Instead, individual thermals accelerate early and decelerate toward the end of their lifetime (Fig. 11f), particularly in the most convective (SB2) case.

c. Thermal growth and net entrainment

Another important question is if thermals change in size as they rise. We evaluate average $R$ at each stage of the thermals’ lifetime by computing the composite thermal based on the peak $w_{\text{max}}$ as a reference. We find that average $R$ (Fig. 11d) is roughly the same at every instant of the thermal’s lifetime. However, the vertical profile of average $R$ (Fig. 11a) does show larger thermals at higher levels. These two results combined suggest that thermals have the same average size throughout their lifetime, but those that initiate higher up tend to be larger than those that initiate below.

However, the fact that thermals have similar average $R$ at each stage of their lifetime could be a misleading result because of averaging of different thermals and different numbers of thermals at each stage, just as the vertical profile of $R$ could be misleading because it reflects the average of different thermals at each vertical layer. To really assess how thermals change in size as they rise, we must compute their net entrainment rate $\varepsilon_{\text{net}}$ from Eq. (8). As expected, the mass-flux-weighted average $\varepsilon_{\text{net}}$ is smaller in magnitude than the unweighted average value (Table 3), since fractional entrainment is lower for larger thermals as a result of the lower surface to volume ratio.

Regardless of the averaging method, $\varepsilon_{\text{net}}$ is small, suggesting that, on average, thermals maintain their size as they rise. The largest value (in magnitude) of $\varepsilon_{\text{net}}$ is $-0.44 \times 10^{-3} \text{ m}^{-1}$, which corresponds to a detraining distance of 2.3 km: that is, more than 3 times the average distance traveled. Although $\varepsilon_{\text{net}}$ is always small, it evolves consistently throughout the thermals’ lifetimes: mass-flux-weighted averages indicate that thermals initially net detrain—albeit only slightly—and progressively tend toward slight net entrainment toward the end of their lifetimes (Fig. 11e). The opposite behavior is found for unweighted averages (not shown here). Thus, most thermals will initially net entrain (grow slightly in size), and then slightly shrink as they reach their final stages. However, the most relevant thermals in terms of mass flux do the opposite: they first slightly shrink as they accelerate and then slightly grow as they decelerate. As a function of height $\varepsilon_{\text{net}}$ shows no particular trend. In general it is small and slightly noisier when mass-flux weights are used (Fig. 11b) than when not (not shown here).

In the past, convective thermals have been thought to resemble self-similar thermal: that is, thermals whose surrounding environmental air velocity is proportional to their ascent rate. Such idealized thermals grow linearly in size with height, with $dR/dz \approx 0.25$. On average, we find $dR/dz = -0.03, -0.04$, and 0.002 for SB1, SB2, and LBA, respectively. Sánchez et al. (1989) showed

<table>
<thead>
<tr>
<th>$R$</th>
<th>1.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W$</td>
<td>0.34</td>
</tr>
<tr>
<td>Lifetime</td>
<td>-0.07</td>
</tr>
<tr>
<td>$\Delta Z$</td>
<td>0.18</td>
</tr>
<tr>
<td>$F_{\text{th}}$</td>
<td>-0.06</td>
</tr>
<tr>
<td>Buoy</td>
<td>0.06</td>
</tr>
</tbody>
</table>

TABLE 4. Correlation coefficients between properties of individual thermals for SB1.
that laboratory thermals traveling through a stably stratified environment require a relatively long time to become self-similar, suggesting that cumulus thermals might not reach self-similarity, and thus suffer less net entrainment. However, such thermals still entrain enough to increase exponentially their volume with distance traveled, which we do not see here. This could be because our simulated thermals are conditionally unstable with respect to the turbulent stratified environment, as opposed to the stably stratified and quiescent flow in laboratory experiments. The fact that net entrainment is strongest in the most stable simulation (LBA) supports this hypothesis.

**d. Momentum budget**

To obtain a complete description of the dynamics of thermals we must investigate the forces that act upon them. Buoyancy is the main driving force, so we could expect that thermals are always buoyant. This is true on average (Table 5), but not at all times: overall, approximately 20% of the thermals have negative buoyancy. Thus, thermals overshoot their level of neutral buoyancy often, consistent with findings by Igau et al. (1999) from aircraft observations during TOGA COARE and simulations by Xu and Randall (2001) of cumulus convection, for example.

Table 5 shows that the vertical momentum budget is dominated by a balance between buoyancy and \( F_{\text{nh}} \). This confirms our earlier results (SHCR13), those from Romps and Charn (2015), and those from de Roode et al. (2012), who found that the dominant sink term in the vertical velocity equation comes from the pressure gradient force and not from entrainment. The impact of mixing in the vertical momentum budget is represented here by \( F_{\text{mix}} \) and \( F_{\text{entr}} \), which are in fact small and often positive, implying a weak positive acceleration instead of any net drag. Only for mass-flux-weighted averages does \( F_{\text{mix}} \) become a small drag term in SB1 and SB2, suggesting that, although unimportant for most thermals, entrainment drag is present—albeit small—for flux-massive thermals during intense convection.

Time evolution of the average momentum budget terms (Fig. 9f) shows how \( F_{\text{nh}} \) follows the evolution of buoyancy, but with opposite sign (mass-flux-weighted averages—not shown here—behave similarly but are noisier). Buoyancy peaks during stage C in SB1 when \( F_{\text{nh}} \) reaches its minimum; the same holds for SB2 (not shown here). This does not imply that buoyancy causes the intense convection during stage C: the strong forcing is due to convergence of the sea breezes, and the unstable environment enhances buoyancy once the thermals are forced to rise.

On the other hand, buoyancy in LBA does not show any significant change throughout the three stages. The flux \( F_{\text{mix}} \) suffers the largest changes, decreasing progressively.
toward zero as convection deepens. Thus, instantaneous mixing is responsible for about one-third of the thermals’ acceleration during stage A of LBA, but by the end of stage C it has no significant contribution. In the sea-breeze cases it has little or no contribution throughout the three stages. Furthermore, \( F_{\text{entr}} \) has very little contribution in all experiments, with only a slight positive contribution (e.g., stage A in SB1).

The distribution of buoyancy broadens in both directions as convection deepens, and more toward positive values (Fig. 12). Consistently, \( F_{\text{nh}} \) mirrors this behavior toward negative values. The fluxes \( F_{\text{mix}} \) and \( F_{\text{entr}} \) widen their distributions symmetrically as convection deepens, remaining approximately centered around zero. The same occurs to \( dW/dt \), which broadens with convective intensity. Thus, during deeper convective stages we find more thermals with stronger positive acceleration than during shallower stages, but also more thermals with stronger negative acceleration.

Mass-flux-weighted averages of the momentum budget terms compared to nonweighted averages (Table 5) show the following features: (i) larger buoyancy; (ii) \( F_{\text{nh}} \) closer to zero in the sea-breeze cases but more negative in LBA; (iii) \( F_{\text{mix}} \) slightly negative or zero instead of slightly positive; (iv) \( F_{\text{entr}} \) closer to zero (but still positive) in SB1 and SB2 and slightly negative instead of zero in LBA; and (v) larger total acceleration in SB1 and SB2 but smaller total acceleration in LBA. These features give us hints on how thermals that contribute significantly to the total mass flux differ from the average thermals. For example, the balance between buoyancy and \( F_{\text{nh}} \)—which, on average, counteract each other in SB1 and SB2—is shifted toward buoyancy, which dominates for mass-flux-weighted averages, even though \( F_{\text{mix}} \) now aids \( F_{\text{nh}} \) in counteracting buoyancy. This shift in the balance between buoyancy and \( F_{\text{nh}} \) results in a larger total acceleration, consistent with giving more weight to more active thermals.

Vertical profiles of the momentum budget terms (Figs. 13a–c) support the fact that buoyancy and \( F_{\text{nh}} \) are anticorrelated, while \( F_{\text{mix}} \) and \( F_{\text{entr}} \) are small. The main difference between unweighted average profiles (not shown here) and mass-flux-weighted average profiles is that \( F_{\text{nh}} \)—particularly in SB2—is closer to zero, while \( F_{\text{mix}} \) becomes slightly more negative. The same feature is visible in the thermals’ lifetime composite (Figs. 13d–f), suggesting that, for deeper convection, the flux-massive thermals suffer less drag due to \( F_{\text{nh}} \) but slightly more due to entrainment (\( F_{\text{mix}} \)).

**e. Drag law**

An important question for parameterizing convection is what controls the strength of \( F_{\text{nh}} \), the main source of drag. In most of the foregoing results, \( F_{\text{nh}} \) appears to vary closely with buoyancy \( B \), but on physical grounds we may expect any draglike force to depend on \( W \) rather than \( B \). Furthermore, in the thermal’s lifetime composite (Figs. 13d–f), \( F_{\text{nh}} \) peaks slightly later than \( B \) (at \( t = 0 \), when \( W \) peaks). Lag correlations between \( F_{\text{nh}} \) and \( B \) confirm this (not shown here). Thus, our results do not unequivocally distinguish whether \( F_{\text{nh}} \) is more directly related to \( W \) or to \( B \) with a time lag, but physical reasoning favors \( W \) rather than \( B \).

In standard dissipative drag formulation (or as evident from dimensional analysis if the only scales are a
velocity $W$ and radius $R$), drag is proportional to $W^2R^{-1}$. Following Romps and Öktem (2015), we can write drag (per unit mass) on a sphere as $(3/8)cdW^2R^{-1}$, where $cd$ is the drag coefficient. We first estimate $cd$ by fitting our data to $-F_{nh} = (3/8)cdW^2R^{-1}$. To do this, we obtain the average values of $-F_{nh}$ and $(3/8)W^2R^{-1}$ for each thermal throughout its lifetime and then use mass-flux-weighted least squares (mf-WLS) to obtain $cd$ for each convective stage of each simulation (Fig. 14). The use of the mf-WLS method is justified since we are interested in the thermals that carry most of the mass flux. We have seen in the previous section that mass-flux-weighted average values of the momentum budget terms (and the relevant balances) change compared to unweighted averages. Since these flux-massive thermals are outnumbered by small and weak thermals, we must give them more weight when fitting the drag coefficient.

We find that $cd$ is not constant and decreases for increasing convective intensity (Fig. 14), typically within the range 0.2–0.5. We also perform the linear fit for separate vertical layers during entire simulations or during each stage of each simulation to obtain vertical profiles of $cd$ (Fig. 15). These profiles show that $cd$ is largest around $Z = 2$ km during the shallow convective stages, and it acquires a relatively uniform value upward as convection deepens. The value of $cd$ decreases abruptly below $\sim 2$ km, even reaching negative values at the lowest levels. The rapid drop of $cd$ toward zero at the lowest levels is most likely because $F_{nh}$ does not consist of drag only. The quantity $F_{nh}$ may include other nonhydrostatic forces, such as those produced by density currents near the surface, for example. The lowest tracked thermals are slightly larger and slower than average, which makes drag on these thermals particularly small. Thus, any other body force may have a large relative impact on $F_{nh}$ there. Indeed, the lowest tracked thermals show mostly positive values of $F_{nh}$ (not shown here), whereas elsewhere the distribution of $F_{nh}$ is close to that from Fig. 12. This is not surprising, given that it is at these low levels where most thermals initiate as a result of density currents, random perturbations because of surface fluxes, etc.

Recent estimates by Romps and Charn (2015) suggest values of $cd \approx 0.6$, which is toward the upper end of our $cd$ value for shallow convection. They argue that thermals are “sticky” because of the dominant balance between buoyancy and drag, which we also find here. Acceleration is only a small fraction of buoyancy; thus, thermals are strongly damped. However, expected drag would be much larger (at least 8 times) if computed for updraft cores rather than for coherent thermals, since the former have at least twice the thermal’s vertical velocity ($w_{max}$ is typically 2$W$) and about half its radius. Furthermore, we confirm the absence of drag from mixing reported by SHCR13, which under standard assumptions would be significant in a highly entraining thermal. Finally, we also find that drag on thermals is weaker as convection intensifies, particularly when giving more weight to the more flux-massive thermals. In other words, the most relevant thermals for deep convection are less damped, with $cd \approx 0.2$, far from the sticky regime ($cd \approx 1$) advocated by Romps and Öktem (2015).

Furthermore, although drag follows this relationship ($\sim W^2R^{-1}$) on average, there is considerable spread. Some thermals suffer less drag and travel farther than expected; others suffer more drag and travel less distance than expected (Fig. 16). This dispersion could have a nontrivial net impact on convective growth given the nonlinearities and local moisture feedbacks involved.
Neglecting entrainment drag, $c_d$ should include the form and wave drag coefficients: $c_d = c_{d,\text{form}} + c_{d,\text{wave}}$ [see Romps and Öktem (2015) and references therein]. For comparison, $c_{d,\text{form}}$ for a solid sphere moving through a quiescent fluid at high Reynolds numbers is between 0.2 and 0.4 (Batchelor 1967, section 5.11). Its value for cumulus thermals is unknown, but we do expect high Reynolds numbers. For wave drag on a solid sphere moving through a stratified fluid, an analytical expression that is a nonlinear function of the Froude number $Fr = W/(N R)$ can be obtained [see Eq. (6) from Romps and Öktem (2015)], where $N = \sqrt{(g/\theta)(d\theta/dZ)}$ is the buoyancy frequency. To test this $c_d(Fr)$ relationship, we plot values of $c_d$ and $Fr$ for individual thermals (Fig. 17a). We compute $c_{d,\text{form}}$ and $Fr$ from each thermal’s $F_{\text{nh}}$, $W$, and $R$ (instead of using the mf-WLS method), and for $Fr$ we obtain $N(z)$ for each stage of each simulation. The large spread suggests that much of the variance in $c_d$ is not captured by this relationship, and/or our data are noisy. To reduce the noise, we bin the data by $Fr$ (Fig. 17b) and use the standard deviation of $c_d$ as an estimate of uncertainty. Because of the large spread, we cannot completely discard the $c_d(Fr)$ relationship, but our results do question its validity, particularly at small values of $Fr$. Wave drag appears to be negligible, and form drag, rather than being constant, decreases for $Fr < \sim 1.5$. Clearly, a theory for drag on cumulus thermals is not straightforward and requires further investigation.

f. Why does mixing not cause drag?

We found little net entrainment or detrainment taking place, which explains why $F_{\text{entr}}$ does not cause any significant drag. However, instantaneous entrainment and detrainment could still take place, potentially causing drag via $F_{\text{mix}}$. Since $s_{\text{net}}$ is small, thermals must either be undiluted or entrain as much as they detrain. Our estimates of $\varepsilon$ are not small: from 0.7 to $2.6 \times 10^{-3} \text{m}^{-1}$, depending on the averaging method and convective intensity (Table 3). So thermals entrain and detrain at similar rates rather than being undiluted but without this being a cause of drag.

To understand this, notice that the composite thermal’s surface is mostly outside of the upward-moving core (see Fig. 5). That is, the relative vertical velocity at the surface is mostly negative. So entrainment is a negative contribution to $F_{\text{mix}}$, but detrainment is a positive contribution to $F_{\text{mix}}$: entrainment causes drag, but detrainment accelerates the thermal. Since both occur at nearly the same rate, the net effect is small, which is what we find. The precise balance between entrainment drag and detrainment acceleration is determined by the exact distribution of $\varepsilon$ and $\delta$ along the thermal’s surface.
However, this distribution is sensitive to the location of the thermal’s boundaries. As pointed out earlier, we can shift the thermals’ centers so that the spheres match the vortex ring structure better. This results in a slight change in the distribution of mixing rates along the surface (not shown here), but it does not change the average values of $F_{\text{mix}}$, suggesting that the integrated balance between entrainment drag and detrainment acceleration is not sensitive to this shift.

On the other hand, the average value of $\varepsilon$ does change when shifting the centers down. Direct estimates of entrainment are sensitive to the exact placement of the boundaries (e.g., Dawe and Austin 2011). In our case, average $\varepsilon$ reaches a minimum value when shifting the centers by 20%–30% of the initially estimated radii, and this minimum is 5%–10% lower than the values shown in Table 3. Of course, we cannot tell which is the ‘right’ position of the thermals’ boundaries, but since none of the other properties analyzed here show any sensitivity to this shift, we do not investigate this further. Only for more detailed studies related to mixing will this be relevant.

5. Summary

We track a large number of thermals in 3D idealized simulations of convection using WRF at 65-m grid length. Two experiments simulate diurnal heating over an island so that convergence of the sea breezes creates a deep convective event reaching above 14 km, preceded by shallow and congestus regimes. A third experiment simulates diurnal evolution of convection over land, with a smooth transition from shallow to deep convection reaching up to 12 km. Our tracking method is based on a simple algorithm that tracks maxima of vertical velocity and sets a sphere around these points such that the average ascent rate $W$ of the enclosed volume matches the tracked points’ $W$. Although it cannot track all thermals, nor track them through their entire lifetime, we find that this simple method successfully identifies a large number of thermals, capturing features such as their acceleration and change in size throughout time. Tracked thermals contribute only $\approx 18\%$ of the total mass flux, but we find indications that these thermals are representative of most of the convective activity. Most updrafts that are not captured by our method typically correspond to either very small or very slow thermals. We do not find any coherent upright plumes.

Most thermals are small ($R < 300$ m) and typically ascend at moderate rates ($W \approx 2$–4 m s$^{-1}$). They travel short distances ($\Delta Z \approx 500$–700 m), since their lifetime—which appears to be insensitive to simulation type and convective intensity—is short ($\approx 4$–5 min on average). This highlights the transient nature of thermals, as opposed to the traditional view of steady plumes spanning the convective layer, which do not occur in these simulations. Furthermore, thermals experience very little net entrainment or detrainment and thus maintain an approximately constant size as they rise.

Average values quoted here are representative of the large majority of thermals and do not vary much.

---

**FIG. 15.** Vertical profiles of $c_d$ obtained for (left) the entire set of simulations and for each stage of (center left) SB1, (center right) SB2, and (right) LBA.

**FIG. 16.** Actual vs expected vertical distances traveled by thermals during stage C of SB1 and LBA. The expected distance is calculated starting from the first tracked time step of each thermal and from there using a simple kinematic equation with constant acceleration in between each pair of output time steps. This acceleration is given by Eq. (4), the momentum budget equation, in which all terms are kept as computed for each thermal at each time step, except for $F_{\text{nh}}$, which is replaced by $(3/8)c_dW^2R^{-1}$, with $c_d$ set to the fitted value from Fig. 14. Points above the $y = x$ line correspond to thermals that suffer less drag than expected, and vice versa.
between simulations. However, the tails of the distributions (i.e., the largest or the fastest thermals) do vary between experiments and with time, mostly reflecting changes in convective intensity. Since thermals carrying more mass flux are likely more important, we also computed mass-flux-weighted averages, which reveal differences between strong and weakly convective situations that are not reflected in the more numerous, small thermals.

The circulation in and around thermals clearly resembles the vortex ring structure described by Woodward (1959). These vortex rings have a slight vertical elongation and thus are more like ellipsoids with eccentricity of roughly 1.05. This elongation is most pronounced during the early stages of a thermal’s lifetime, suggesting that it is related to its vertical acceleration. Importantly, however, we do not find any thermals sufficiently elongated to be reasonably called a plume.

The main change in thermal properties as convection deepens is their starting heights. The vast majority of thermals originate below 3–4 km. As convection intensifies, thermals develop at higher levels, but always at levels that have been reached by previous thermals. Thus, what makes convection deeper is not thermals traveling farther, but rather originating at higher altitudes, at least within the limits of our tracking procedure. A realistic parameterization of convection may need to “dispatch” thermals at different altitudes, not only from the boundary layer as usually expected (e.g., Ooyama 1971). Our results reinforce the picture from previous observational studies (e.g., Yuter and Houze 1995) that convection consists more of a local mixing process, less characterized by coherent troposphere-spanning plumes, than is often believed (see also Sherwood and Risi 2012; Allen and Landuyt 2014).

Past studies have often assumed that drag due to entrainment or mixing has significant impact on thermals, implying that larger thermals will be less affected and will rise faster. While we do find that larger thermals rise somewhat faster on average, it is not for this reason. First, consistent with previous findings (e.g., de Roode et al. 2012), the main source of drag is not linked to mixing, but rather to the nonhydrostatic pressure term $F_{\text{nh}}$. Moreover, while average size and ascent rate of thermals each increase as convection becomes deeper, these two properties are fairly weakly correlated at the thermal scale. The net force acting on a thermal is typically a small fraction of buoyancy, with buoyancy nearly counteracted by the nonhydrostatic force as also pointed out recently by Romps and Charn (2015). On the other hand, mixing does not cause drag because thermals detrain as much as they entrain, and although entrainment causes drag, detrainment does the opposite, so the net effect is small.

We investigate how $F_{\text{nh}}$ relates to standard drag formulation, $(3/8) c_d W^2 R^{-1}$, where $c_d$ is the total drag coefficient that incorporates both form and wave drag ($c_d = c_{d,\text{form}} + c_{d,\text{wave}}$). Using the mass-flux-weighted least squares method, we obtain different $c_d$ values for different convective stages and vertical layers. We find that $c_d$ is mostly sensitive to convective intensity and not to altitude once above $\sim 2$ km. For shallow convection, $c_d$ has high values that reach up to 0.8, and as convection deepens $c_d$ decreases down to values of 0.1–0.2 throughout the entire column. Thus, deep convective thermals tend to be more “slippery” than shallow convective thermals.
We also compare our results to theoretical expressions for drag on a sphere. On the one hand, $c_d\text{form}$ for a solid sphere traveling through quiescent fluid at high Reynolds numbers is expected to be between 0.2 and 0.4. On the other hand, $c_d\text{wave}$ for a solid sphere moving through a stratified fluid is expected to depend nonlinearly on the Froude number and can be computed analytically. It grows exponentially for low Froude numbers and tends to zero for high Froude numbers. Our results show a different behavior, with an approximately constant drag coefficient at high Froude numbers (>1.5), but decreasing drag at low Froude numbers (<1.5). This is likely because drag on a solid sphere arises in a high-shear boundary layer, which does not exist on the boundary of a vortex ring; drag in the latter case must arise from other mechanisms, and these appear to vary significantly within convective systems. Thus, our simulations suggest that drag on cumulus thermals behaves in a different way to drag on a solid sphere.

In any case, drag is not weak, counteracts most of the buoyancy, and is unrelated to mixing. Furthermore, a stochastic description of drag may be useful, given that some thermals suffer more and others less drag than expected by the standard drag formulation, most likely because of complexities in the flow that are not accounted for in this framework.

Acknowledgments. We thank David Fuchs for his generous help in developing and optimizing the Python scripts used to track and analyze thermals. We also thank David Romps and four other anonymous reviewers for helpful comments and suggestions. WRF runs were carried out at the National Computational Infrastructure (NCI) at the ANU, Australia. This work was funded by the Australian Research Council Centre of Excellence for Climate System Science (ARCSS), and the first author was also supported through Agreement FP44842-134-2015 with the “Patrimonio Autónomo Fondo Nacional de Financiamiento para la Ciencia, la Tecnología y la Innovación, Francisco José de Caldas” of Colciencias and the Universidad Nacional de Colombia (UNAL). We also thank Nancy L. Villegas for her kind support at UNAL.

REFERENCES


