Prediction for Cloud Spacing Confirmed Using Stereo Cameras

RUŞEN ÖKTEM[a,b] AND DAVID M. ROMPS[a,b]

[a] *Department of Earth and Planetary Science, University of California, Berkeley, Berkeley, California*
[b] *Climate and Ecosystem Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California*

(Manuscript received 28 January 2021, in final form 16 August 2021)

ABSTRACT: Using 3 years of the Clouds Optically Gridded by Stereo (COGS) product, the mean cloud base, cloud top, cloud width, and cloud spacing are described with respect to their seasonal and/or diurnal evolution at the Atmospheric Radiation Measurement (ARM) Southern Great Plains (SGP) site. In addition to confirming and extending prior results, the data show that the effective diameter of shallow cumuli are approximately equal to the height above ground of the lifting condensation level (LCL). Furthermore, the cloud spacing is found to closely match a prediction by Thuburn and Efstathiou for the horizontal scale of the largest unstable eddies in an unsheared convective boundary layer.

KEYWORDS: Cloud cover; Clouds; Convective clouds; Climatology; Cloud retrieval; Cumulus clouds; Surface observations

1. Introduction

Shallow cumulus (ShCu) clouds play an important role in Earth's climate, generating $\sim 5\,\mathrm{W\,m^{-2}}$ of upwelling top-of-atmosphere radiation (Chen et al. 2000), yet we lack a full understanding of the processes that determine the ShCu cloud fraction, cloud size, and intercloud spacing. The Department of Energy's Atmospheric Radiation Measurement (ARM) Southern Great Plains (SGP) site (Mather and Voyles 2013) has been collecting data on shallow clouds in Oklahoma for three decades. These measurements include cloud-base height, liquid water path, cloud-base vertical velocity, and cloud fraction. These data have been collected primarily with vertically pointing instruments, including ceilometers, cloud radars, Doppler lidars, and microwave radiometers. By virtue of their "soda straw" field of view, measurements of shallow cumulus must be collected over hours to obtain a reliable estimate of cloud fraction, during which time the cloud fraction has changed. Even data from scanning radars or lidars must be averaged over about an hour to obtain a reliable estimate of the cloud fraction (Oue et al. 2016).

Despite these challenges, it has been possible to average over many days to construct composites of the diurnal evolution of cloud fraction (Lazarus et al. 2000; Berg and Kassianov 2008; Zhang and Klein 2013), to study its sensitivity to surface fluxes and lifting condensation level (Berg and Kassianov 2008) or to the characteristics of the vertical velocities in the subcloud column (Lamer and Kollias 2015), and to probe the impact of changes in environmental conditions on the vertical cloud extent (Zhang and Klein 2013). But without the ability to measure a single day's evolution of cloud fraction—let alone cloud sizes and spacings—we are limited in what we can learn about the underlying processes.

With the aim of collecting a more complete instantaneous view of the shallow cloud field, six cameras were strategically deployed around the ARM SGP site in August 2017. These six cameras were situated along a circle centered on the SGP central facility with a radius of 6 km. Facing toward the center, these six cameras instantaneously sample tens of thousands of visually distinguishable features on the clouds within a wide field of view and, through the principles of stereo photogrammetry, allow for the reconstruction of their three-dimensional coordinates. From those reconstructed coordinates, the surfaces of the clouds are identified and stitched together to determine the regions of space that are cloudy. The result is a four-dimensional gridded map of the clouds known as the Clouds Optically Gridded by Stereo (COGS) product (Romps and Oktem 2018). COGS data are gridded binary values (cloudy or clear) spaced at 50 m in all three directions and 20 s in time within the reconstructable region of a cubic 6-km-wide domain centered on the SGP central facility (with the precise center at 36.6053°N, 97.4865°W). The term "reconstructable" here refers to the common field of view of the six cameras, which occupies more than $30\,\mathrm{km^2}$ in between an altitude of 300 m and 3 km, but reduces to $0\,\mathrm{km^2}$ at the surface and at a height of 6 km (here and throughout, altitudes are given relative to the ground). The advantage of the COGS dataset is that it enables the measurement of the geometric properties of individual clouds or fields of clouds, as well as tracking their variations through time.

In this study, we will use the COGS data to pursue two objectives. The first of these objectives is to generate a seasonal and diurnal climatology using COGS, which can be compared with earlier such climatologies made with lidar and radar. The second objective is to assess a recent theory by Thuburn and Efstathiou (2020) for the horizontal scale of eddies in the convective boundary layer. In that theory, the most unstable horizontal wavelength is predicted to be $2\sqrt{2}$ times the depth of the boundary layer. Since shallow cumulus are the condensed upper portions of those boundary layer thermals, we will check whether the shallow-cumulus cloud spacing is equal to $2\sqrt{2}$ times the height of the lifting condensation level. These objectives are pursued in sections 3 and 4, respectively. First, however, section 2 will describe the dataset.

Corresponding author: Ruşen Öktem, roktem@lbl.gov

DOI: 10.1175/JAS-D-21-0026.1

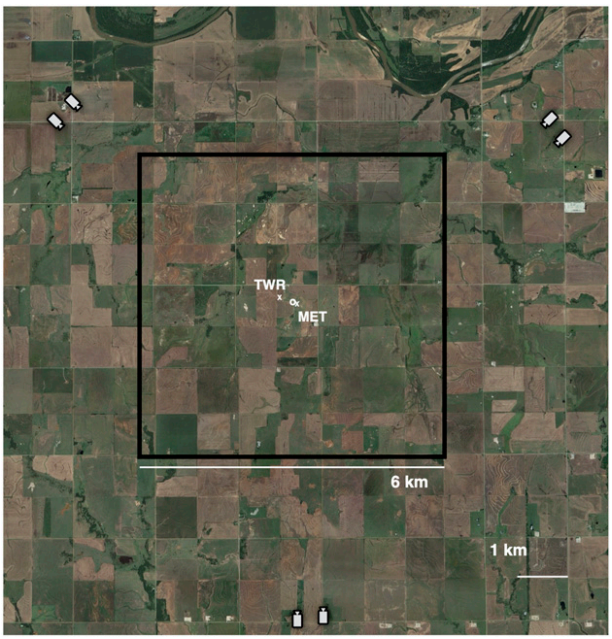

FIG. 1. Map of the stereo cameras, 60-m tower, and the MET sensor at the SGP Central Facility. The COGS domain and its center are indicated by a black rectangle and white circle, respectively. The MET sensor and tower locations are marked by crosses.

2. Observations and methodology

COGS data are a collection of binary values gridded on four dimensions: eastward distance, northward distance, height above ground, and time. A value of 1 (0) indicates that the corresponding grid cell is detected as cloudy (noncloudy) by the multiview stereophotogrammetric reconstruction. Stereophotogrammetric reconstruction relies on detecting the same cloud features in the synchronous shots of a camera pair to compute their real world coordinates with respect to a reference location, forming a point cloud of cloud points (PCCP). Positioning three camera pairs as shown in Fig. 1, the common field of view of the cameras are sampled from multiple views that complement each other, and the PCCP from each are stitched together to generate COGS (Romps and Oktem 2018).

The main shortcoming of COGS is its reliance on passive optical feature detection. The algorithms that generate COGS work only during daytime, only when there are distinguishable feature points on the cloud surfaces, and only when the cameras have unobstructed views of the cloud surfaces. When the window of the camera enclosure is streaked with rain, or when the clouds lack visual texture (e.g., stratus), stereo reconstruction is not possible. Shallow cumulus clouds, with their crisp features, are ideal cases for COGS so long as the cloud area fraction is not too high. When the cloud fraction is too high, parts of the domain will be obscured from the cameras' view, leading to cloudy grid cells being misidentified as clear. Our empirical analysis suggests that COGS tends to underreport cloudy volumes when the cloud fraction exceeds one-half. The COGS algorithm will also fail to register very small or translucent clouds that have too few feature points. Thus, some small, wispy clouds that are detected by a lidar may be missed by COGS. Romps and Oktem (2018) estimated that the COGS cloud boundaries are accurate to within about 100 m, which is an order of magnitude smaller than the typical width of clouds studies here (about 1 km).

We inspected the raw camera images during daytime every day over a 3-yr period spanning 1 September 2017 to 31 August 2020 and identified candidate ShCu cases that are less likely to suffer from COGS shortcomings as explained above. Our criteria during visual inspection were 1) the start and end of a ShCu event can be clearly identified (i.e., transitions from stratocumulus or stratus are eliminated), 2) precipitation is not encountered from the start to the end of the event. We did not rule out the cases with coexisting optically thin high-altitude clouds or altocumulus clouds that appear for a short duration, although the majority of the events had clear skies above the shallow cumulus. Once candidate days were identified, the COGS data were used to confirm that their cloud bases were below 3 km. We then rejected days during which the projected cloud fraction in the COGS domain exceeded 0.5 for more than 30 min in total over the entire day. Days with such persistently large cloud fractions are likely to indicate a transition to stratocumulus or a large number of larger-than-domain-size clouds, which we wish to exclude. Discarding those cases resulted in 129 days during which we can be sure that the lower troposphere transitioned cleanly from clear to shallow cumulus and back to clear.

To define a measurable and consistent start and end times for all days, a ShCu event is defined as starting when the cloud fraction exceeds 0.01 and does not fall back below 0.01 for over 30 min in the next hour. The end of the event is defined as the time when the cloud fraction falls below 0.01 and does not rise above 0.01 within the next hour. As a result, each of the ShCu cases starts and ends with a cloud fraction (CF) of around 0.01 and may contain an instance of CF at or below 0.01 in between. Of the 129 days, only one had two ShCu events according to the definition used here. For that one day, the second event was discarded for simplicity. As a result, the dataset has exactly one ShCu event for each of the 129 days.

We define the height of each $(50\,\mathrm{m})^3$ COGS grid cell to be the height of its center. For each COGS snapshot, we define the instantaneous cloud-top height (CT) and cloud-bottom height (CB) to be the 99th and the 1st percentiles of cloudy-gridcell heights. The instantaneous cloud fraction (CF) is defined as the ratio of the number of vertically projected cloudy grid cells to the total number of reconstructable grid cells at the CB level.

The instantaneous lifting condensation level (LCL) is calculated as formulated in Romps (2017) using the temperature, pressure, and relative humidity from the Surface Meteorology System (MET) sensor at 2-m elevation, and from the 25- and 60-m southeast and west sensors on the 60-m tower. The 60-m tower and the MET sensor stand approximately 200 m northwest and 60 m southeast of the center of the COGS region, respectively, as shown in Fig. 1. We calculate the LCL height separately for each of the five sensors at 1-min intervals and average them all to obtain a single time series of LCL height representative of the COGS domain.

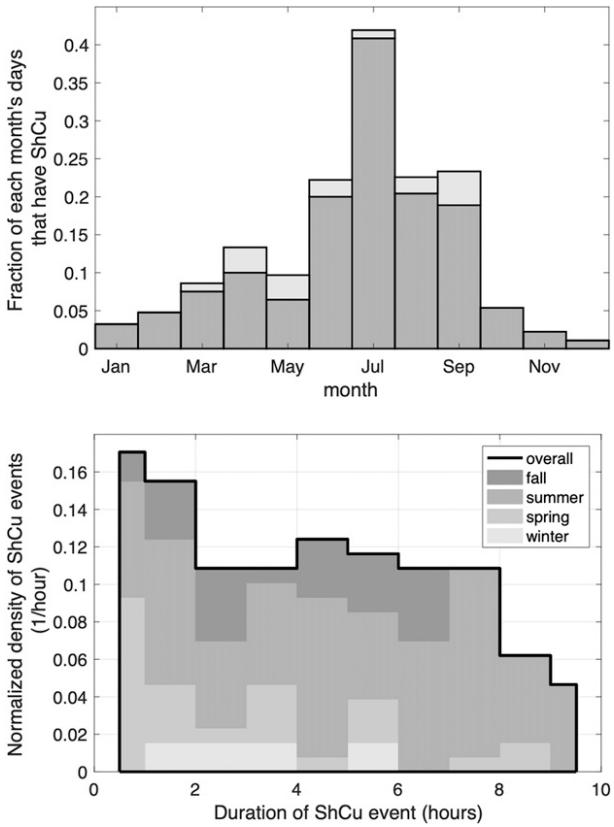

FIG. 2. (top) Fraction of the days in each month that have a ShCu event. The light-shaded parts in the top panel correspond to the days when a visual inspection confirmed the presence of ShCu but the event was discarded from the dataset due to technical problems with the cameras that caused incomplete measurements. (bottom) Normalized density of ShCu events as a function of their duration (solid black). The shading shows the contributions to the distribution from each of the seasons.

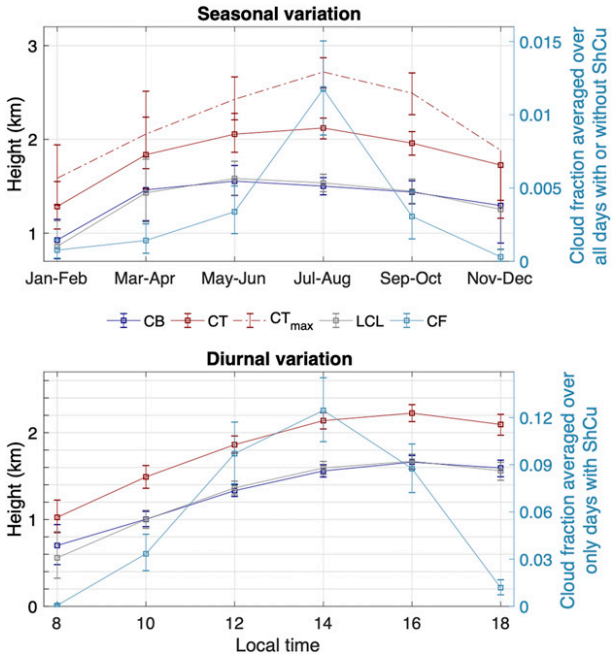

FIG. 3. (top) Seasonal variation of the time-weighted means of CB, CT, LCL height, and CF in 2-month bins. CB, CT, and the LCL are averaged over all times during ShCu events and the CF is averaged over all 24 h of all days (with or without a ShCu event). (bottom) Diurnal variation of CB, CT, LCL, and CF using 2-h bins centered on 0800, 1000 local time, and so on. CB, CT, and LCL are averaged over all days and times that have ShCu during the corresponding 2-h bin. CF is averaged over the 129 days with ShCu events, whether or not they have any cumulus clouds in the corresponding 2-h bin. The bars in both panels represent 95% confidence intervals calculated by bootstrapping ShCu events.

3. Climatology of shallow cumulus

The top panel in Fig. 2 depicts the distribution of ShCu days by month showing that ShCu are most frequent in the summer and early fall. The peak of ShCu occurs in July when nearly half of the days host shallow cumulus. The ShCu frequency abruptly drops from September to October and remains low until March. The light shaded parts of the bars represent the days that are visually identified as ShCu but are not included in our dataset due to technical problems in one or more of the cameras.

The solid black curve in the bottom panel in Fig. 2 shows the distribution of the 129 ShCu events by their duration. We see that most of the events can be found evenly distributed between 2 and 8 h, although there are events exceeding 9 h in duration and the highest density of events is at short durations less than 2 h (recall that half an hour is the shortest duration of ShCu events included in our dataset). Each shade of gray under the black curve illustrates the fractional contribution from each of the different seasons. Durations exceeding 7 h occur mainly in June and July. Short-duration events can take place in any season, but events less than 1 h in duration occur most frequently in the spring.

Previous studies have looked into the seasonal and diurnal variations of CF, CB, and CT at the SGP site, but have been constrained by the "soda straw" view of vertically pointing lidar and radar. Berg and Kassianov (2008) and Zhang and Klein (2013) focused on the diurnal cycle of clouds during the summer season only, which is when the frequency of ShCu events is greatest. Berg and Kassianov (2008) additionally inspected the year-to-year variation in the macrophysical properties of summertime clouds over five years. Lazarus et al. (2000) averaged over all days to generate a seasonal variation of shallow cloud fraction, but still unexplored are the seasonal variations in other properties of the ShCu events.

Figure 3 shows some of basic features of the seasonal and diurnal variation in ShCu as measured by COGS. In the upper panel, the light-blue curve corresponds to the projected ShCu cloud fraction averaged over all 24 h of every day in each 2-month bin. Although COGS does not report data after sunset, we assume that ShCu forms only during the day. Therefore, the light-blue curve indicates the portion of the sky covered by ShCu on average at any moment in time. In July and August, for example, an average of 1.2% of the sky is covered by ShCu, as can be seen from the scale on the right axis. This curve agrees

qualitatively with the curve of ShCu occurrence in Fig. 2, which showed ShCu occurrence maximizing in the middle of summer and minimizing in late fall and early winter, and it also agrees with the results of Lazarus et al. (2000).

Also plotted in the top panel of Fig. 3 are the averages of the LCL height (gray), CB (dark blue), and CT (solid red). These means are calculated over all times during ShCu events using time weighting (thus, long-duration events contribute more to these means than short-duration events). The bars in both panels represent the 95% confidence interval computed by bootstrapping. In bootstrapping, we treated each ShCu event as an independent sample and generated 10 000 realizations of 129 events by randomly drawing from the ShCu events with replacement. Therefore, the bars correspond to the 2.5th and 97.5th percentiles of the means calculated from the 10 000 realizations.

We see from these curves that CB tracks the LCL very closely, as expected. The ShCu clouds are thinner and the subcloud layer is significantly shallower—below 1 km—in the winter. As the boundary layer deepens starting from the spring, cloud heights also rise and we start to see considerably deeper clouds in July with cloud-base heights above 1.4 km. Looking at the difference between CB and CT, we see that the mean thickness of the ShCu cloud field is in the range of 400–600 m, with the cloud field being the thickest in July and August, when the cloud tops tend to reach the highest. Plotted in the dashed red curve is the eventwise mean of the 99th percentile of CT from each event. We see that the tallest cumulus clouds in these data tend to reach a height of 2.7 km in midsummer.

In the bottom panel of Fig. 3, the light-blue curves show the projected ShCu cloud fraction averaged over 2-h bins from all 129 days (i.e., each day is included in the average regardless of whether or not it has ShCu in that particular time interval). On days with ShCu, the ShCu CF tends to peak at 12% at 1400 local time. Also plotted in the bottom panel are CB, CT, and the LCL height averaged in each 2-h bin from all days with ShCu during that time interval. For CB, CT, and LCL, note that not all of the 129 days contribute to each time bin since the ShCu onset and end times vary. The CB and CT rise by about 1 km with a similar trend as the LCL until late afternoon, with all three peaking at around 1600 local time. Cloud thickness (CT minus CB) also peaks at 1400 local time with a mean value of 580 m. Starting from 320 m at 0800 local time, the mean cloud thickness rises to 500 m at 1000 local time, and then varies by only 80 m after then. Throughout the day, we see a strong coupling between LCL and CB, except for the time before 0900 local time. By separately inspecting each of the ShCu events that started before 0900 local time, we found that, on three of those days, the LCL height was computed to be below 300 m while the clouds formed a few hundreds of meters above the computed LCL height. We did not observe any meteorological pattern discriminating these three days from the others, so our guess is that this is a residual effect of the nocturnal boundary layer.

The results in Fig. 3 are mostly consistent with the previous literature. The conclusion that the mean ShCu CF peaks in the summertime agrees with data from human observers (see Fig. 9 of Lazarus et al. 2000). In the diurnal cycle, the peak of CF in the early afternoon is consistent with both human observations (see Fig. 9 of Lazarus et al. 2000) and the Active Remote Sensing of Clouds (ARSCL; Clothiaux et al. 2000) retrievals derived from vertically pointing lidar and cloud radar (see Fig. 3a of Zhang and Klein 2013). The COGS-derived CF values in Fig. 3 are much lower than the ARSCL-derived values reported by Berg and Kassianov (2008) and Zhang and Klein (2013), but those studies focused entirely on the summertime, whereas we have taken a mean over all seasons. As for the evolution of the cloud base, COGS agrees with ARSCL (Berg and Kassianov 2008; Zhang and Klein 2013) in finding the cloud-base height starting at around 1 km in the morning and roughly doubling by the late afternoon. For the cloud top, however, there is no consensus. Figure 3 shows that the cloud top tracks the cloud bottom with a fairly constant offset, in agreement with one of the ARSCL studies (see Fig. 3b of Zhang and Klein 2013), but contradicting the other (see Fig. 4a of Berg and Kassianov 2008). It is possible that the cloud-top data used by Berg and Kassianov (2008) have been contaminated by false-positive cloud detections caused by insects, which is a recognized problem (Williams et al. 2021).

A unique advantage of COGS is the ability to calculate the instantaneous vertical distribution of cloud fraction, i.e., the fraction of the horizontal domain that is cloudy at each height (not to be confused with the projected cloud fraction that has been discussed up until now). Figure 4 plots the distribution of the complete set of cloudy grid cells observed over the 129 days as a function of height and time in three different ways. In the top-left panel, the axes are simply the height above ground and the local time; the values are averages of cloud fraction over the 129 days at that height and time. The top-right panel is the same except that the height above ground has been replaced with the height above the instantaneous CB. Finally, in the bottom-left panel, the local time has been replaced with the elapsed time since the beginning of the ShCu event.

The top-right panel tells us that the cloud volume distribution is skewed toward its base and is mainly concentrated between 100 and 300 m above the cloud base. This altitude interval encompasses 58% of the total cloud volume. Furthermore, as seen in the bottom-left panel, half of the cloud volume occurs in the first 3 h of the ShCu event. The bottom-right panel of Fig. 4 shows the distribution of the length of time from cloud onset to the peak projected cloud fraction for cases with a duration longer than 2 h (76% of the cases). We see that it is most common for the projected cloud fraction to peak between 1 and 3 h after the appearance of the first shallow cumulus cloud.

4. What sets the intercloud spacing?

Thuburn and Efstathiou (2020) recently provided some insight into the origin of the horizontal length scale of the largest eddies in the convective boundary layer. Those large eddies correspond to the ascending thermals that, when the boundary extends above the LCL, manifest as shallow cumuli at their tops. If we imagine that every ascending thermal is topped by a cumulus during a ShCu event, then the cloud spacing (the

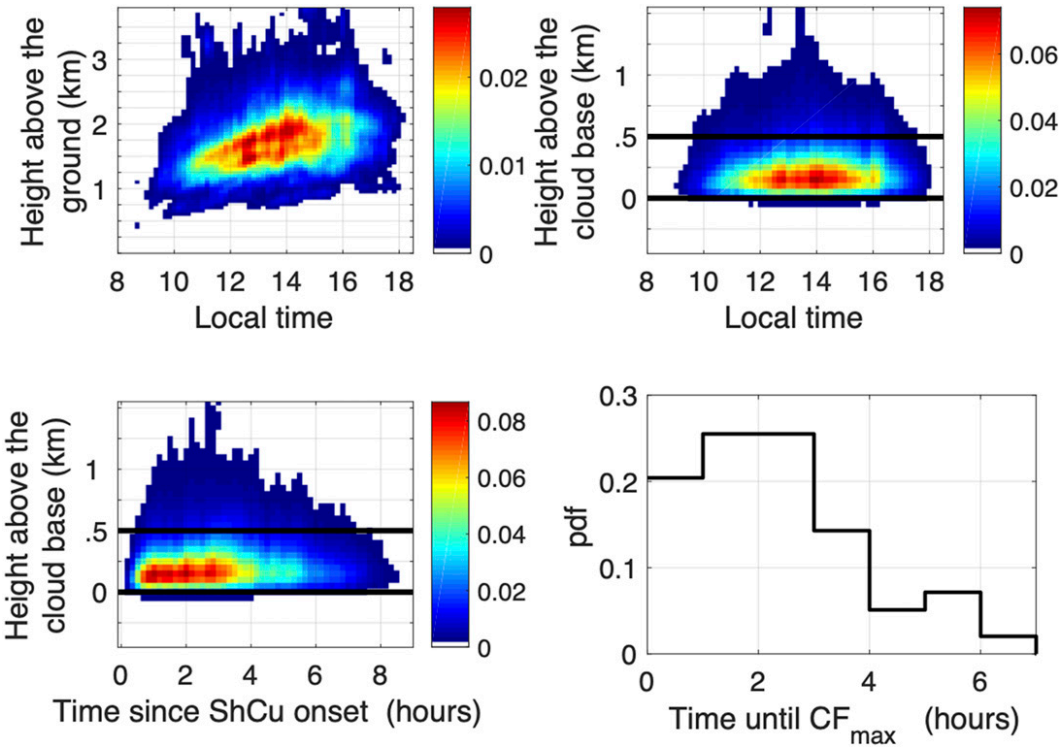

FIG. 4. (top left) Fraction of the horizontal COGS domain that is cloudy as a function of local time and height above the ground, averaged over the 129 days with ShCu. (top right) As in the top-left panel, but as a function of local time and height above the cloud base. The solid black lines indicate the level of cloud base and 500 m above the cloud base. (bottom left) As in the top-left panel, but as a function of time since ShCu onset and height above the cloud base. (bottom right) Distribution, over the days with ShCu duration longer than 2 h, of the length of time from cloud onset to the maximum vertically projected cloud fraction.

distance between adjacent clouds) should reveal the dominant eddy size.

The idea behind the derivation of Thuburn and Efstathiou (2020) is as follows. Empirically, a convective boundary layer achieves a steady state that is turbulent even though the boundary layer is well mixed, i.e., $d\theta/dz \approx 0$, where θ is the potential temperature. In other words, turbulence is maintained in the presence of near-zero instability. This motivates the marginal-stability hypothesis: that the turbulent steady state is one in which the most unstable mode has a nearly zero growth rate.

In practice, this state is achieved by modulation of the turbulent diffusivity. Before surface heating is initiated at the base of a quiescent fluid, the turbulent diffusivity κ is zero. With the application of surface heating, $d\theta/dz$ becomes negative, which causes modes to become unstable, to amplify, and to generate turbulence. That turbulence increases κ, which, in turn, tends to reduce the modes' growth rates. If κ becomes too large, it will render all modes stable, eliminating the source of energy for the turbulence, and causing κ to fall. If κ becomes too small, then many modes become unstable and grow, which spins up more turbulence and increases κ. In a steady state, κ sits at a value that allows just enough instability to drive the turbulence that maintains that value of κ. By the marginal stability hypothesis, κ can be approximated as that which makes the most unstable mode have zero growth rate.

For completeness, we will briefly replicate the derivation of Thuburn and Efstathiou (2020). We begin with the two-dimensional Boussinesq equations,

$$\frac{\partial u}{\partial t} = -\frac{\partial p}{\partial x} + \kappa\nabla^2 u, \tag{1}$$

$$\frac{\partial w}{\partial t} = -\frac{\partial p}{\partial z} + \kappa\nabla^2 w + b, \tag{2}$$

$$\frac{\partial b}{\partial t} = -wN^2 + \kappa\nabla^2 b, \tag{3}$$

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0\,, \tag{4}$$

with variables taking their usual meanings. We will linearize about a turbulent state with zero mean u and w (but not zero mean $\sqrt{u^2 + w^2}$) and interpret κ as the turbulent diffusivity. Substituting a plane-wave solution of the form $\exp[i(kx + mz - \omega t)]$, we can derive the dispersion relation for ω. Assuming the squared Brunt–Väisälä frequency N^2 is negative, the growth rate is $\mathrm{Re}(-i\omega)$, which gives

$$\text{growth rate} = \frac{\sqrt{-N^2}k}{\sqrt{k^2 + m^2}} - \kappa(k^2 + m^2)\,. \tag{5}$$

For any given k, the fastest growing mode is the one with the smallest m, which is π/H for a boundary layer of depth H. Setting $m = \pi/H$, the largest growth rate will occur when $d(\text{growth rate})/dk = 0$, which gives

$$k[k^2 + (\pi/H)^2]^{3/2} = \frac{\sqrt{-N^2}\pi^2}{2\kappa H^2}. \tag{6}$$

Using the marginal-stability hypothesis, the growth rate of the most unstable mode—i.e., of the k that solves Eq. (6)—should be zero:

$$\text{growth rate} = \frac{\sqrt{-N^2}k}{\sqrt{k^2 + (\pi/H)^2}} - \kappa[k^2 + (\pi/H)^2] = 0. \tag{7}$$

Equations (6) and (7) can be solved for κ and k, yielding

$$\kappa = \frac{2}{3^{3/2}\pi^2}\sqrt{-N^2}H^2, \tag{8}$$

$$k = \frac{\pi}{\sqrt{2}H}. \tag{9}$$

Defining the cloud spacing (CS) to be the wavelength of this marginally unstable mode, which is related to the wavenumber k by $k = 2\pi/\text{CS}$, then we get

$$\text{CS} = 2\sqrt{2}H. \tag{10}$$

Here, we see that the cloud spacing is on the order of H, as expected, but we have also derived a coefficient of proportionality: the cloud spacing is expected to be about $2\sqrt{2}$ times the LCL height.

Although this derivation is elegant and appealing, one should be wary of taking it too seriously as a model for the ShCu intercloud spacing. First, Eq. (10) gives a wavelength for two-dimensional convective rolls, but, in the absence of large-scale shear, a field of fully developed shallow cumulus does not arrange itself as convective rolls. Second, the subcloud layer does not have a constant N^2 as assumed here: instead, it varies with height. Third, the overall logic of performing a linear stability analysis on a turbulent fluid—furthermore, with the turbulence assumed to be generated as a cascade from the marginally stable linear modes—is intrinsically heuristic, and so there can be little expectation of a precise agreement with experiments or observations. Given these considerations, one might view the derivation above as little more than a confirmation that CS $\sim H$. For the sake of argument, however, we will take the factor of $2\sqrt{2}$ as a serious candidate for CS/H and compare it to observations.

To proceed, we must first decide how to calculate a cloud spacing from observations. If we imagine a snapshot of the vertically projected cloud field, there will be contiguous patches of cloud of various sizes and shapes that are somewhat randomly distributed over the domain. We are faced, then, with the question of what patches of cloudiness should count as "a cloud" and what interpatch distances should count toward the intercloud length scale. If we count every contiguous patch of cloudiness as "a cloud," then the results will depend sensitively on the presence of small, disintegrating wisps of cloud. If we try to rule out those wisps, then we are left with the challenge of defining a size cutoff, which introduces subjectivity. Even if a suitable definition of "a cloud" could be settled upon, there would still remain the issue of how to measure the distances between clouds and which pairs should contribute to the calculation of the cloud spacing. These challenges are so great, and the potential resolutions so subjective, that we will not attempt an approach along those lines.

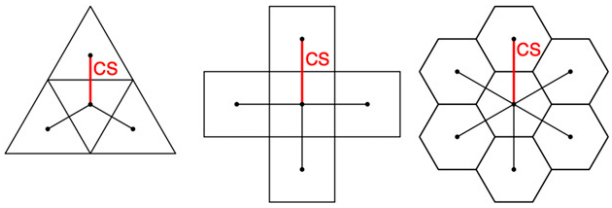

FIG. 5. The three regular tilings of a plane.

Instead, we will calculate the mean area of individual clouds and assume a certain tiling of the clouds. This still involves some subjective choices, but the choices we will make are straightforward to motivate. Let us begin by considering the issue of cloud tiling. Figure 5 shows the three regular tilings of the plane: triangular, square, and hexagonal. Let us imagine that the clouds are located at the center of these tilings so that the cloud spacing (CS) is the distance between the centers of neighboring tiles.

Given CF and the mean area A of individual clouds, and assuming they are arranged as a regular tiling, then the area of an individual tile is A/CF. The area of a tile can also be written as $b\text{CS}^2$, where CS is the effective cloud spacing (the distance between the centers of adjacent tiles) and b is a constant that depends on the type of tiling ($3\sqrt{3}/4$ for a triangular tiling, 1 for a square tiling, and $\sqrt{3}/2$ for a hexagonal tiling). Combining this information, the effective intertile spacing CS can be related to A and CF as

$$\text{CS} = \begin{cases} \dfrac{2}{3^{3/4}}\sqrt{\dfrac{A}{\text{CF}}} \approx 0.88\sqrt{\dfrac{A}{\text{CF}}} & \text{triangular tiling} \\ \sqrt{\dfrac{A}{\text{CF}}} & \text{square tiling} \\ \dfrac{2^{1/2}}{3^{1/4}}\sqrt{\dfrac{A}{\text{CF}}} \approx 1.07\sqrt{\dfrac{A}{\text{CF}}} & \text{hexagonal tiling} \end{cases}. \tag{11}$$

We will define the mean area A of individual clouds as the area-weighted mean of the contiguous areas of vertically projected cloudiness. We are motivated to weight by cloud area for a few reasons. First, the contribution of a cloud to the reflection of shortwave radiation from the sun at zenith is proportional to the cloud's area, so weighting by cloud area ensures that clouds are weighted in proportion to their impact on radiation. Second, we expect that large clouds more robustly mark the tops of the largest active eddies than small ones do; area weighting ensures that small decaying cloud fragments do not contribute significantly to the calculation of

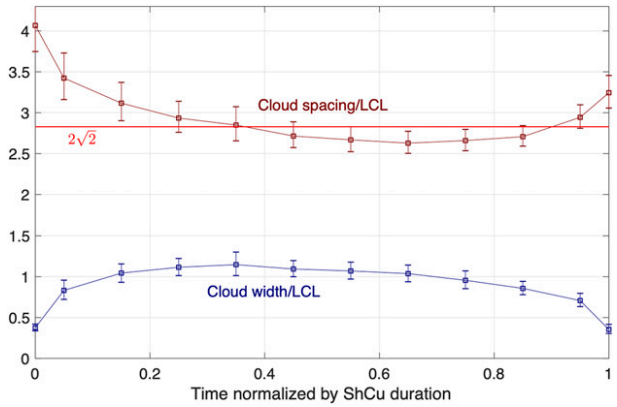

FIG. 6. Evolution of length scales through a ShCu event. Cloud width and intercloud spacing are normalized by the LCL height. The bars represent the 95% confidence interval around the mean calculated by bootstrapping. $t = 0$ and $t = 1$ mark the start and end times of the ShCu duration as defined in section 2, respectively. The red solid line shows the prediction of $2\sqrt{2}$ from Thuburn and Efstathiou (2020).

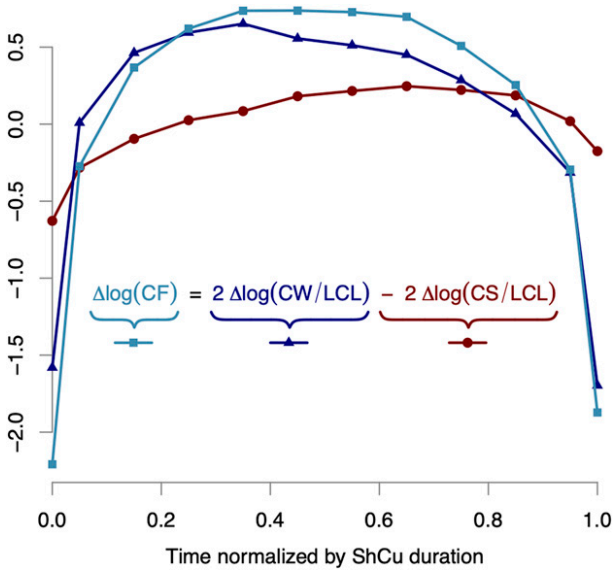

FIG. 7. Composite of $\Delta\log(\mathrm{CF})$ (light-blue squares), $2\Delta\log(\mathrm{CW/LCL})$ (dark-blue triangles), and $-2\Delta\log(\mathrm{CS/LCL})$ (dark-red circles). These are related by Eq. (15). The diurnal variation in cloud fraction is dominated by the diurnal variation in cloud width.

mean cloud area. Third, there are times when a cloud is at the edge of the COGS boundary with only part of itself inside the domain. Although it would be straightforward to exclude clouds at the edges, doing so would render ambiguous what total area we should divide by to calculate a cloud fraction; therefore, we do not attempt such an approach. Instead, weighting by area in the calculation of A deemphasizes the clouds straddling the edge. It is also worth noting that the typical cloud width (~1 km) is much smaller than the domain width (~6 km), so most of the cloudy area is from clouds that are fully contained within the COGS domain. Only 5% of the total cloud area of the whole dataset comes from clouds with effective cloud widths (defined below) less than 480 m and another 5% from clouds with effective widths greater than 3.8 km. During ShCu events, there are, on average, five clouds fully contained within the COGS domain, and those clouds constitute 56% of the cloud area.

From the COGS data, we calculate the effective cloud width CW as

$$\mathrm{CW} = 2\sqrt{A/\pi}, \tag{12}$$

which would be the diameter for a perfectly circular cloud. Since $2/\sqrt{\pi} = 1.13$, this is not much different from $\mathrm{CW} = \sqrt{A}$, which would be exact for square clouds. Since a hexagonal tiling is to be expected more often in unsheared convective boundary layers than triangular or square, we define CS as

$$\mathrm{CS} = 2^{1/2}3^{-1/4}\sqrt{A/\mathrm{CF}}\,. \tag{13}$$

In each COGS snapshot of each ShCu event, we calculate CW, CS, and LCL instantaneously; CW/LCL and CS/LCL can then be averaged across events. Note that the widths and spacings calculated in this way should be considered effective values since the observed clouds do not array themselves with a perfect hexagonal tiling.

Figure 6 plots the time series of CW/LCL and CS/LCL as functions of normalized time, averaged over the 129 ShCu events. The squares at $t = 0$ and $t = 1$ correspond to an average over the COGS snapshots at the very beginning and end of the ShCu events; recall that CF is about 0.01 at those times. Each of the other squares correspond to an average over COGS snapshots in the tenth of the event's duration centered on the square's position. The same bootstrapping method described earlier is used to calculate the 95% confidence interval around the mean. The red line marks the $2\sqrt{2}$ factor derived by Thuburn and Efstathiou (2020).

We see in Fig. 6 that, for the vast majority of the duration, the average CS remains between 2.7 and 3.1 times the LCL height (i.e., between 0.95 and 1.1 times the theoretical factor of $2\sqrt{2} \approx 2.8$). In other words, $2\sqrt{2}$ times the LCL height is a surprisingly good match for the mean intercloud spacing during ShCu events. CW, on the other hand, is well described as being equal to the LCL height itself for most of the duration of ShCu event. The cloud width begins at less than half the LCL height at onset, but quickly ramps up to match the LCL height, and the clouds maintain that size until shortly before the end.

These composite time series of the cloud spacing and cloud width allow us to check our intuition about cloud fraction. Eliminating A from Eqs. (12) and (13), we get

$$\mathrm{CF} = \frac{\pi}{2\sqrt{3}}\frac{\mathrm{CW}^2}{\mathrm{CS}^2}\,. \tag{14}$$

Defining ΔX as the quantity X minus its time average, we can take the logarithm of this equation and subtract off the mean to get

$$\Delta\log(\mathrm{CF}) = 2\Delta\log(\mathrm{CW/LCL}) - 2\Delta\log(\mathrm{CS/LCL})\,. \tag{15}$$

Using this relation and the data points from Fig. 6, Fig. 7 plots $\Delta\log(\mathrm{CF})$, $2\Delta\log(\mathrm{CW/LCL})$, and $-2\Delta\log(\mathrm{C\,S\,/LCL})$.

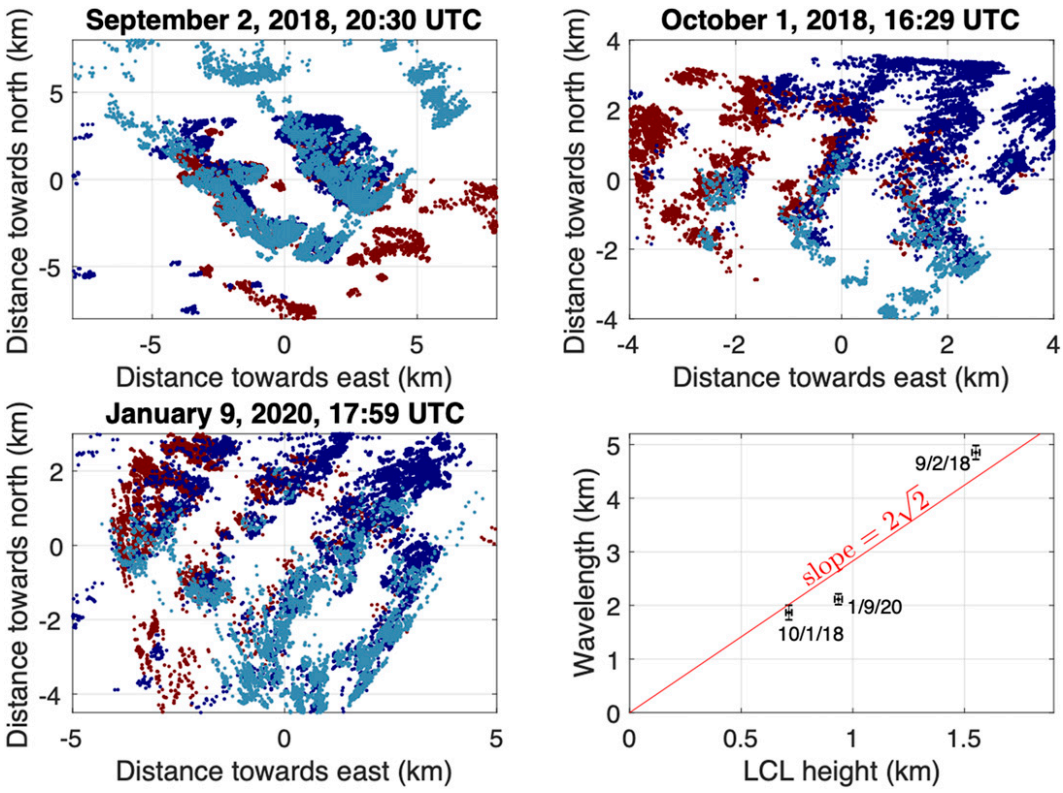

FIG. 8. (top),(bottom left) Horizontally projected PCCP feature points for representative snapshots during the occurrence of cloud streets. Data from the northwest, northeast, and southern camera pairs are shown as red, dark-blue, and light-blue dots, respectively. Note that the spatial coverage of the PCCP data changes with the altitude of the clouds since a higher cloud base gives the cameras an unobstructed view over a greater distance. (bottom right) The best-fit horizontal wavelengths plotted against the LCL height. The bars denote the standard deviations of the LCL height and wavelength over the lifetime of each convective-roll case. The straight line shows the predicted slope of $2\sqrt{2}$.

Intuitively, it is not the spacing (or, equivalently, number density) of the thermals that suddenly changes to initiate or terminate a ShCu event. To start a ShCu event, the top of the boundary layer rises above the LCL, causing the condensed portions of the tops of thermals to suddenly appear and then rapidly grow in horizontal extent. Likewise, the termination of a ShCu event occurs as the top of the boundary layer falls below the LCL, causing the condensed portions of the tops of thermals to shrink in size and disappear. Thus, the rapid growth and decay of CF at the beginning and end of a ShCu event, respectively, are driven primarily by a corresponding growth and decay in CW. Figure 7 bears this out: we see that most of the variation in CF is caused by variation in CW, with a rapid increase in cloud width to start the event and a rapid decrease in cloud width to terminate it.

We saw in Fig. 6 that $2\sqrt{2}$ closely matches the observed ratio of cloud spacing to LCL height, but we should bear in mind that the factor of $2\sqrt{2}$ in Eq. (10) was derived for two-dimensional convection or, equivalently, for three-dimensional convection with translational symmetry, which would manifest as cloud streets. Therefore, it is worth looking for times when the clouds are arranged as cloud streets to see if the $2\sqrt{2}$ prediction applies during those times. By watching time-lapse videos of the 129 ShCu events, we found three ~10-min periods of apparent cloud streets (from 2022 to 2035 UTC 2 September 2018, from 1625 to 1634 UTC 1 October 2018, and from 1756 to 1804 UTC 9 January 2020). Since the LCL height during these three periods ranged from 720 to 1550 m, these provide an opportunity to see how the horizontal wavelength of convective rolls scales with LCL height.

The top and bottom-left panels of Fig. 8 show the vertically projected PCCP feature points at representative times when cloud streets were present. The red, dark-blue, and light-blue dots are the projected cloud points detected by the northwest, northeast, and southern camera pairs, respectively. Unlike COGS, which provides regularly gridded cloudiness, the PCCP data are the coordinates of visually distinct cloud features. A downside of the PCCP data is the presence of false detections, which amount to around 1%–3% of the feature points; those false detections, which appear as noise in Fig. 8, are mostly eliminated in the COGS algorithm. The advantage of the PCCP data, however, is that the PCCP data cover a wider area than COGS because COGS only works in the intersection of the fields of view of the three camera pairs.

Nonlinear optimization was used to find the wavelength and orientation of the sine wave that (when optimally translated for

each snapshot) best matched the projected PCCP points (in the sense of giving the highest sum of values of the sine curve at the points) using 33, 29, and 24 snapshots for the 2 September, 1 October, and 9 January cases, respectively. The lower-right panel in Fig. 8 displays the mean estimated wavelength plotted against the mean LCL height calculated over the set of snapshots for each case. The horizontal and vertical bars show the standard deviations of the LCL heights and wavelengths, respectively (e.g., the standard deviation of the 33 wavelengths calculated for 2 September). The close proximity of the three cases to a line of slope $2\sqrt{2}$ indicates that Eq. (10) holds as predicted for convective rolls.

5. Summary

With 3 years of video from cameras at the ARM SGP site, we have identified 129 shallow cumulus (ShCu) events in which the lower troposphere cleanly transitions from clear to shallow cumulus and back to clear. The durations of these events range from less than an hour to more than 9 h, with events occurring most frequently in the summertime. The macrophysical properties of these clouds (e.g., heights, thicknesses, and widths) have been quantified using the COGS stereo-camera product, which maps the clouds onto a 50-m grid every 20 s on a 6-km-wide cubic domain.

These data confirm various expected properties of shallow cumulus, such as the fact that the cloud base closely tracks the lifting condensation level, that the clouds are most prevalent during the summertime, and that their cloud fraction tends to peak in the early afternoon. The previous literature on ShCu at the SGP site was conflicted as to the evolution of cloud tops, with one study reporting the mean cloud-top height as being fairly constant through the day while another study found that the mean cloud-top height rises in tandem with the cloud base. Here, we find strong support for the latter: the mean height of cloud tops rises in synchrony with the cloud base, maintaining a mean cloud thickness in the range of 400–600 m. With the ability to record the evolution of individual ShCu events, it is also found here that most ShCu events that last beyond 2 h peak in cloud fraction between 1 and 3 h after onset, and most of the cloud volume is located between 100 and 300 m above the mean cloud base.

We also considered the theoretical prediction of Thuburn and Efstathiou (2020) that the horizontal wavelength of the largest eddies in an unsheared convective boundary layer should be $2\sqrt{2}$ times the depth of the boundary layer. Despite reasons to doubt the applicability of this prediction to shallow cumulus, the COGS data reveal that the mean cloud spacing matches this prediction to within ~10%. The same analysis reveals the mean effective diameters of clouds to be approximately equal to the LCL height. The suggested relation between the cloud spacing and the LCL height is further verified using three cases of short-duration cloud streets. Finally, the time series of cloud spacing and cloud width confirm that the cloud fraction grows and decays during onset and termination of a field of shallow cumulus primarily via variations in the mean cloud width, not the mean cloud spacing.

Acknowledgments. This work was supported by the U.S. Department of Energy's (DOE) Atmospheric System Research (ASR), an Office of Science, Office of Biological and Environmental Research program; Lawrence Berkeley National Laboratory is operated for the DOE by theUniversity of California under Contract DE-AC02-05CH11231. Data were obtained from the ARM user facility, a U.S. DOE Office of Science user facility managed by the Biological and Environmental Research Program.

Data availability statement. COGS, MET, and 60-m tower data used in this study are openly available from the ARM Data Center at https://doi.org/10.5439/1644322, https://doi.org/10.5439/1025341, and https://doi.org/10.5439/1025220, respectively.

REFERENCES

Berg, L. K., and E. I. Kassianov, 2008: Temporal variability of fair-weather cumulus statistics at the ACRF SGP site. *J. Climate*, **21**, 3344–3358, https://doi.org/10.1175/2007JCLI2266.1.

Chen, T., W. B. Rossow, and Y. Zhang, 2000: Radiative effects of cloud-type variations. *J. Climate*, **13**, 264–286, https://doi.org/10.1175/1520-0442(2000)013<0264:REOCTV>2.0.CO;2.

Clothiaux, E. E., T. P. Ackerman, G. G. Mace, K. P. Moran, R. T. Marchand, M. A. Miller, and B. E. Martner, 2000: Objective determination of cloud heights and radar reflectivities using a combination of active remote sensors at the ARM CART sites. *J. Appl. Meteor.*, **39**, 645–665, https://doi.org/10.1175/1520-0450(2000)039<0645:ODOCHA>2.0.CO;2.

Lamer, K., and P. Kollias, 2015: Observations of fair-weather cumuli over land: Dynamical factors controlling cloud size and cover. *Geophys. Res. Lett.*, **42**, 8693–8701, https://doi.org/10.1002/2015GL064534.

Lazarus, S. M., S. K. Krueger, and G. G. Mace, 2000: A cloud climatology of the Southern Great Plains ARM CART. *J. Climate*, **13**, 1762–1775, https://doi.org/10.1175/1520-0442(2000)013<1762:ACCOTS>2.0.CO;2.

Mather, J. H., and J. W. Voyles, 2013: The ARM Climate Research Facility: A review of structure and capabilities. *Bull. Amer. Meteor. Soc.*, **94**, 377–392, https://doi.org/10.1175/BAMS-D-11-00218.1.

Oue, M., P. Kollias, K. W. North, A. Tatarevic, S. Endo, A. M. Vogelmann, and W. I. Gustafson, 2016: Estimation of cloud fraction profile in shallow convection using a scanning cloud radar. *Geophys. Res. Lett.*, **43**, 10 998–11 006, https://doi.org/10.1002/2016GL070776.

Romps, D. M., 2017: Exact expression for the lifting condensation level. *J. Atmos. Sci.*, **74**, 3891–3900, https://doi.org/10.1175/JAS-D-17-0102.1.

——, and R. Oktem, 2018: Observing clouds in 4D with multiview stereophotogrammetry. *Bull. Amer. Meteor. Soc.*, **99**, 2575–2586, https://doi.org/10.1175/BAMS-D-18-0029.1.

Thuburn, J., and G. A. Efstathiou, 2020: Marginal stability of the convective boundary layer. *J. Atmos. Sci.*, **77**, 435–442, https://doi.org/10.1175/JAS-D-18-0222.1.

Williams, C. R., K. L. Johnson, S. E. Giangrande, J. C. Hardin, R. Öktem, and D. M. Romps, 2021: Identifying insects, clouds, and precipitation using vertically pointing polarimetric radar doppler velocity spectra. *Atmos. Meas. Tech.*, **14**, 4425–4444, https://doi.org/10.5194/amt-14-4425-2021.

Zhang, Y., and S. A. Klein, 2013: Factors controlling the vertical extent of fair-weather shallow cumulus clouds over land: Investigation of diurnal-cycle observations collected at the ARM Southern Great Plains site. *J. Atmos. Sci.*, **70**, 1297–1315, https://doi.org/10.1175/JAS-D-12-0131.1.