On the size dependence of cumulus cloud spacing

Thirza van Laar[1] and Roel A J Neggers[2]

[1]Helmholtz-Zentrum Geesthacht
[2]University of Cologne

November 26, 2022

Abstract

In this study the spatial structure of Trade Wind shallow cumulus populations is investigated as diagnosed from large-domain high resolution cloud-resolving simulations. The main objective is to establish how inter-cloud spacing depends on cloud size, information that is crucial for understanding cloud-radiation interaction and spatial organization, and for informing grey zone parametrizations. A high-resolution cloud-resolving ICON simulation of Caribbean shallow convective cloud fields is used, based on the NARVAL South field campaign. The size statistics of the simulated cloud population are found to compare well to those derived from available satellite images. Four expressions for the nearest neighbor spacing are analyzed, including classic definitions but also novel ones. We find that the dependence of cloud spacing on cloud size strongly depends on this definition. The relation is exponential for the spacing between clouds of similar size, while it is logarithmic for the spacing between clouds of any size. Further analysis suggests that the logarithmic dependence is caused by the abundance of closely-spaced small clouds. The exponential size-dependence is argued to reflect the mesoscale dynamics driving the horizontal size of large convective cells. The implications of the obtained results are briefly discussed.

On the size dependence of cumulus cloud spacing

Thirza W. van Laar[1]*, Roel A.J. Neggers[1]

[1]University of Cologne, Institute of Geophysics and Meteorology

Key Points:

- Cloud spacing in Trade wind cumulus populations is investigated using super-large-domain high-resolution simulations
- Various definitions of cloud neighbor spacing are analyzed, including both classic and novel formulations
- Both logarithmic and exponential dependencies on cloud size are reported, reflecting differences in the spatial distribution of small and large clouds

*Current afilliation: Helmholtz-Zentrum Geesthacht, Institute of Coastal Research

Corresponding author: Thirza van Laar, `thirza.laar@hzg.de`

Abstract

In this study the spatial structure of Trade Wind shallow cumulus populations is investigated as diagnosed from large-domain high resolution cloud-resolving simulations. The main objective is to establish how inter-cloud spacing depends on cloud size, information that is crucial for understanding cloud-radiation interaction and spatial organization, and for informing grey zone parametrizations. A high-resolution cloud-resolving ICON simulation of Caribbean shallow convective cloud fields is used, based on the NARVAL South field campaign. The size statistics of the simulated cloud population are found to compare well to those derived from available satellite images. Four expressions for the nearest neighbor spacing are analyzed, including classic definitions but also novel ones. We find that the dependence of cloud spacing on cloud size strongly depends on this definition. The relation is exponential for the spacing between clouds of similar size, while it is logarithmic for the spacing between clouds of any size. Further analysis suggests that the logarithmic dependence is caused by the abundance of closely-spaced small clouds. The exponential size-dependence is argued to reflect the mesoscale dynamics driving the horizontal size of large convective cells. The implications of the obtained results are briefly discussed.

Plain Language Summary

Shallow cumulus cloud fields persistently cover large areas in the marine subtropics. These low level clouds play an important role in Earth's energy balance, because of the associated vertical transport of heat and moisture and their impact on radiation. Weather and climate models still struggle to correctly represent these cloud populations, which is partially due to our prevailing lack of understanding of their spatial structure. In this study unprecedented large-domain high resolution simulations and satellite images are used to investigate cloud spacing in more detail, revisiting classic studies that were purely based on observational data. The results show that in general cloud spacing increases with cloud size. However, the relation between size and spacing strongly depends on the way the spacing is defined: spacing between clouds of any size behaves logarithmic, while spacing between clouds of equal size shows an exponential size dependence. The results provide more insight into spatial organization of cumulus clouds, and can guide ongoing efforts to improve the representation of these clouds in circulation models.

1 Introduction

Cumuliform low-level clouds persistently cover large areas of the marine subtropics (Norris, 1998). The way these cloud fields interact with the atmospheric circulation and respond to a warming global climate, are complex scientific problems that are not completely understood yet (M. Zhang et al., 2013; Narenpitak et al., 2017). This lack of understanding is reflected in long-standing shortcomings in their representation in weather and climate models (Nam et al., 2012), and in their significant contribution to uncertainty in future climate predictions (Bony & Dufresne, 2005; Sherwood et al., 2014; Vial et al., 2016).

The spatial variability of Trade wind cumulus cloud fields, in particular the spacing between individual cumulus clouds, has been identified as a key element for understanding their role in Earth's climate system. For example, spatial aggregation is involved in the interaction between cloud fields and a changing climate (Wing & Cronin, 2015; Bretherton & Blossey, 2017; Wing, 2019). The spacing between clouds also strongly affects how they interact with solar and terrestrial radiation, in particular when the three-dimensionality of radiative fluxes is taken into account (Jakub & Mayer, 2017). Cloud spacing also plays a role in the "grey zone problem", which stands for the situation that previously unresolved convective processes are becoming (partially) resolved at the high resolutions now feasible in general circulation modeling (Wyngaard, 2004; Honnert et al., 2020). While spatial information is needed to make convection schemes scale-aware and scale-adaptive (Neggers, 2015; Brast et al., 2018), cloud spacing also affects the stochasticity in convective properties in the grey zone (Neggers et al., 2019).

Observational research of cloud spacing goes back decades. Early studies mostly relied on high-altitude photography (Plank, 1969), satellite images (Sengupta et al., 1990) and scanning radar (Ali, 1998). Joseph and Cahalan (1990) first investigated the dependence of cloud spacing on cloud size, analysing satellite snapshots of cumulus clouds at various locations on the globe. They reported a positive linear relation between cloud size and the Nearest Neighbor Spacing (NNS), suggesting that larger clouds have a bigger spacing. However, the spread in this relation was large, argued to be due to differences in meteorological and surface conditions between the snapshots. Later studies used the cumulative distribution function of NNS to quantify the spatial organization in a cloud field (Weger et al., 1992; Nair et al., 1998), yielding an organizational metric that has become frequently used (Weger et al., 1993; Tompkins & Semie, 2017). However, the size dependence in cloud spacing has not been revisited since those early days.

Large-Eddy Simulations (LES) can well be used to study cloud spacing. A key advantage of LES over satellite images is the access it provides to full four-dimensional fields at high spatial-temporal resolutions. First-generation LES studies of cumulus cloud size distributions were still severely limited by domain size, which made these numerical experiments less useful for studying cloud spacing (Neggers et al., 2003). However, ongoing advances in supercomputing are currently allowing a dramatic increase in the domain sizes that can be applied (Heinze et al., 2017; Senf et al., 2018; Vial et al., 2019). As a consequence, the simulated cloud populations also become much more complete, which is particularly important for the largest cloud sizes. These large clouds occur more frequently and abundantly in a larger domain, and are no longer dynamically constrained in an artificial way by a too small domain size. As a result, the dependence of cloud spacing on cloud size can now reliably be investigated with LES across a much broader spectrum of cloud sizes than was previously possible. While the NNS has appeared in some recent LES studies (Neggers et al., 2019), the unique new opportunities created by the use of a large domain size for studying cloud spacing have not yet been fully exploited.

In this study we revisit the classic problem of cloud spacing in Trade Wind cumulus cloud fields, now using both super-large domain LES and satellite imagery, in combination. Our prime objective is to gain more insight into the dependence of cloud spacing on cloud size. Use is made of a high-resolution cloud-resolving simulation performed with the ICON model (Zängl et al., 2014) of Caribbean shallow convective cloud fields as observed during the recent NARVAL-South campaign near Barbados (Klepp et al., 2014). These simulations were generated in the context of the HD(CP)^2 project (High Definition Clouds and Precipitation for Advancing Climate Prediction). The combination of a large domain (150×400 km^2) with a cumulus cloud-resolving horizontal resolution (150 m) allows a statistically significant investigation of cloud spacing across a broad spectrum of cloud sizes, including very large ones. The location over the ocean ensures fairly homogeneous conditions concerning the state of the atmosphere and surface characteristics. The simulated cloud populations are compared to statistics derived from MODIS satellite images at 250 m resolution. A set of four definitions of the NNS is examined, including spacing between clouds of any size, spacing between clouds of similar size, and using both center-to-center and edge-to-edge distancing. The analysis focuses on the dependence of NNS on cloud size. For reference, the obtained results are compared to i) reference NNS values reflecting purely random distributions, and ii) results from the classic observational studies on cloud spacing as mentioned above.

The data and methods used in this study are described in Section 2. Section 3 then presents the results, including an assessment of the state of the cloud field and its evolution, population statistics and their comparison to observations, and a detailed analysis of the cloud spacing. In Section 4 the main results are further interpreted, focusing on the size dependence in the cloud spacing. Section 5 then provides a brief summary of the main results and conclusions, and gives an outlook on future research inspired by this study.

2 Data and methods

2.1 20 December 2013

The NARVAL-South Campaign took place throughout December 2013 and January 2014 in the Caribbean Trade wind region upwind of Barbados, with the HALO aircraft functioning as the main instrument platform (Klepp et al., 2014). The target area of NARVAL-South is not routinely sampled by state-of-the-art meteorological instrumentation, with only a few permanent sites on islands far apart (Stevens et al. 2016, Lamer et al. 2015). Accordingly, NARVAL-South had the aim of filling the existing data gap on Atlantic Trade wind cumulus to support observational data analyses (Schnitt et al., 2017; Jacob et al., 2019) as well as high-resolution simulation efforts (Reilly et al., 2019; Naumann & Kiemle, 2020).

The day of interest for this study is 20 December 2013, on which HALO performed Research Flight 08 (RF08). The MODIS Terra satellite image shown in Figure 1a gives a good impression of the cloud field on this day, showing a cumulus cloud population featuring a broad range of cloud sizes. Such cloud patterns are typical for the Caribbean Trade wind region (Bony et al., 2020). Figure 1b zooms in on the domain of interest upwind of Barbados, indicating that in this region the cloud field was dominated by small-scale low level boundary layer cumulus with only a few larger 'flowers' present. The latter represent stratiform cumulus outflow near the Trade inversion. The MODIS reflectance is available at 250m gridspacing, which is comparable to the discretization of the LES experiment used in this study.

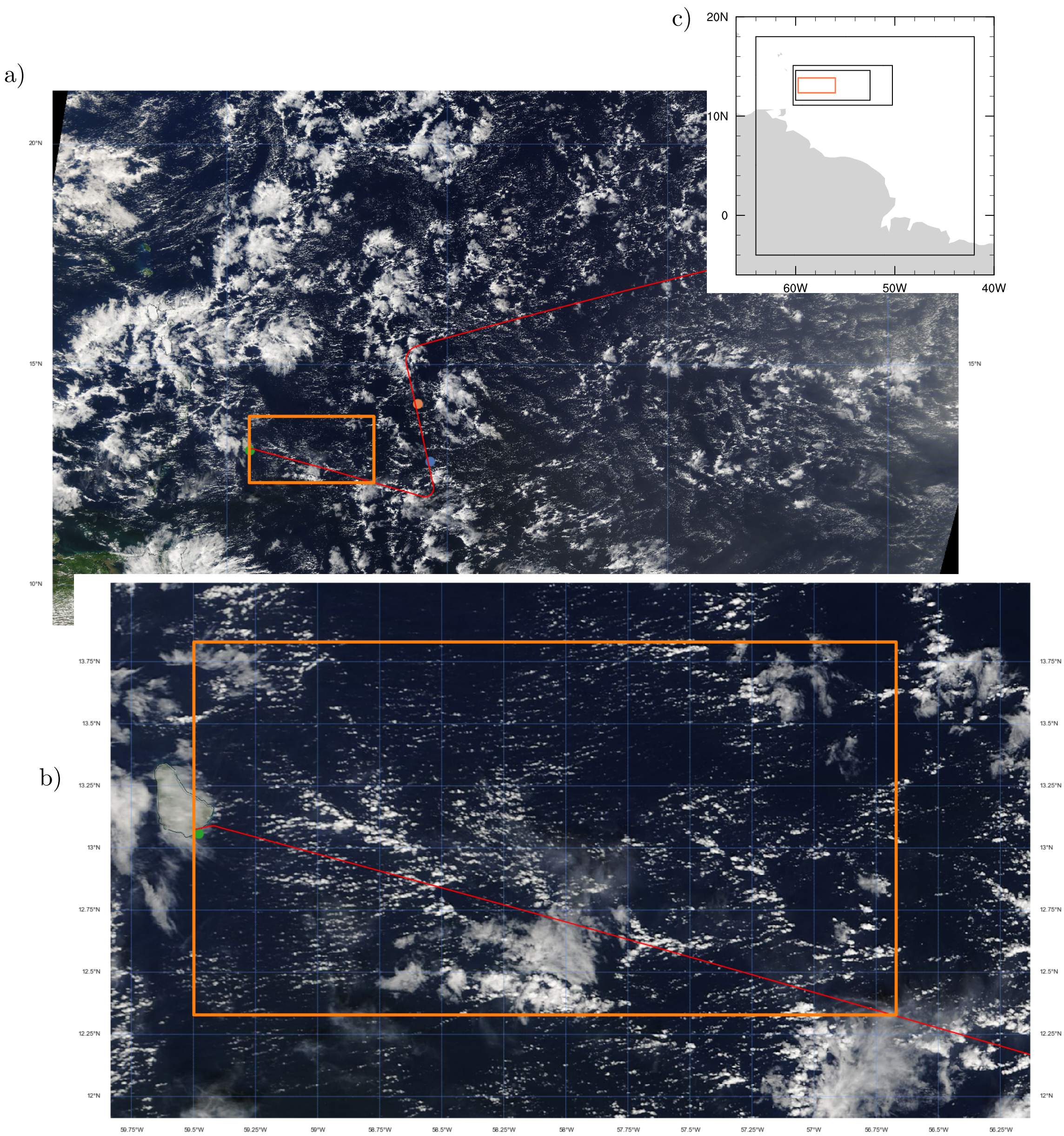

Figure 1. MODIS Terra corrected reflectance (true color) images for 20 December 2013. a) Caribbean at 1km resolution, and b) the area upwind of Barbados at 250m resolution. The Island of Barbados is visible in panel b), on the left. Panel c) gives an overview of all four ICON domains simulated with ICON LES. The orange box always indicates the inner ICON domain resolved at 150m resolution of which the results are used in this study, and within which the MODIS data is also analyzed. The HALO flight path is shown as a red line, while the locations of the first two dropsondes of HALO RF08 are indicated by the blue and orange dots. The Barbados radiosonde sounding is indicated by the green dot. Geotiff data obtained through NASA Worldview (https://worldview.earthdata.nasa.gov/).

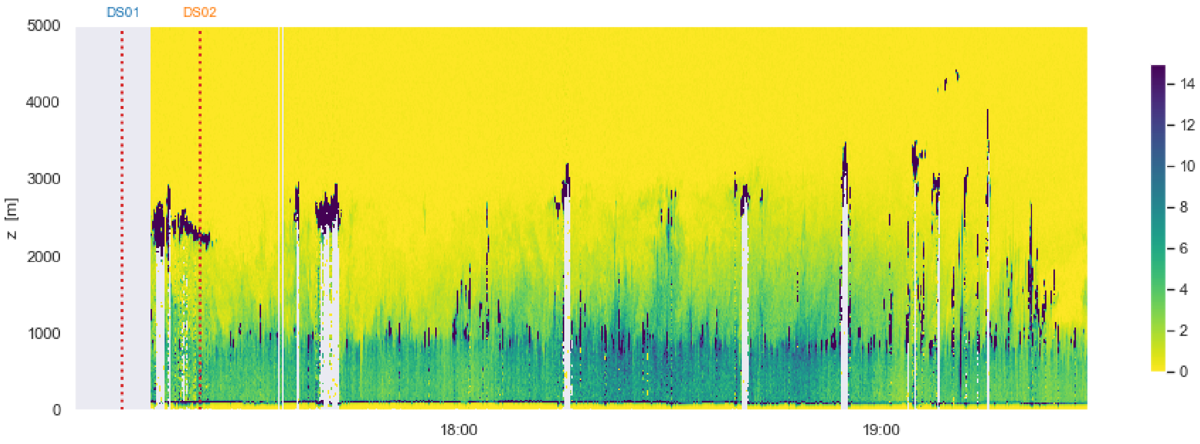

Figure 2. Time-height plot of backscatter at 532 nm as sampled by the HSRL instrument as part of the WALES system onboard HALO during RF08. Only the first part of RF08 is shown during which the observed boundary layer structure was approximately similar. Clouds show up as black areas. Dropsondes DS01 and DS02 are indicated by the dotted red lines.

More detailed information about the boundary layer cloud field in the target area is provided by the WALES instrument onboard HALO, as shown in Fig. 2. Only the first two-hour time segment of RF08 flight is shown during which defining features of the Trade wind boundary layer remained relatively unchanged, such as the cloud base and cloud top heights. After this period HALO entered a region in which the cloud structure deepened profoundly, losing these typical features. This motivated using 19 : 30 UTC as the upper time limit for the analysis of WALES cloud data. The clouds, showing up as black areas, are profoundly broken, and include many small cumulus clouds rooting in the subcloud layer as well as remnants of outflow situated below the inversion. These cloud population statistics are similar to those discussed by (Naumann & Kiemle, 2020).

Figure 3 shows observed vertical profiles sampled in or in the direct vicinity of the simulated domain. These locations are also indicated in Fig. 1a. Included are a radiosonde sounding at Barbados (at 12:00 UTC), the first two dropsondes DS01 and DS02 during HALO RF08 (launched at 17:12 and 17:23 UTC) just outside the simulated domain in upwind direction, and the cloud fraction profile as derived from the WALES data discussed above. The typical features of a shallow cumulus topped Trade wind boundary layer are evident, such as a well-mixed subcloud layer and a conditionally unstable cloud layer which is capped by an inversion layer situated between 2200-2600 m height. The

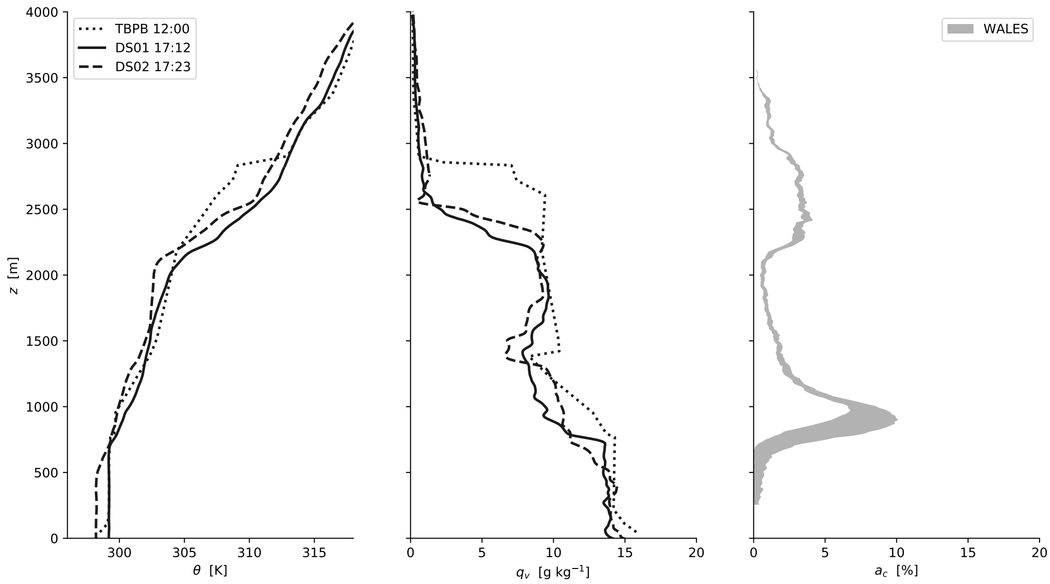

Figure 3. Observed vertical profiles of a) potential temperature θ, b) water vapor specific humidity q_v and c) cloud fraction a_c. Shown are the 12:00 UTC radiosonde sounding from the TBPB station at Barbados (dotted), the first two DropSondes (DS) from HALO RF08 launched at 17:12 (solid) and 17:23 UTC (dashed), and the WALES HSRL measurements above 250m height and time-averaged over the first two hours of RF08 (shaded grey). The horizontal range for WALES indicates the difference between the backscatter thresholds of 10 and 20 Mm^{-1} sr^{-1}.

lower free troposphere above the inversion is statically stable and very dry in all soundings, containing almost no water vapor. The cloud fraction profile shows the double peak structure typical of Trade wind cumulus as found in numerous previous studies (Stevens et al., 2001; vanZanten et al., 2011; Nuijens et al., 2014). This structure reflects the presence of cumuli above the top of the mixed-layer and cumulus outflow near the inversion as seen in Fig. 2.

All thermodynamic soundings are strikingly similar concerning the vertical structure of the boundary layer, apart from a slightly higher inversion over Barbados which might be an island effect. This good agreement between soundings that are separated quite far in both space and time suggests that the boundary layer structure was approximately in steady state as well as reasonably homogeneous across the target domain selected for simulation.

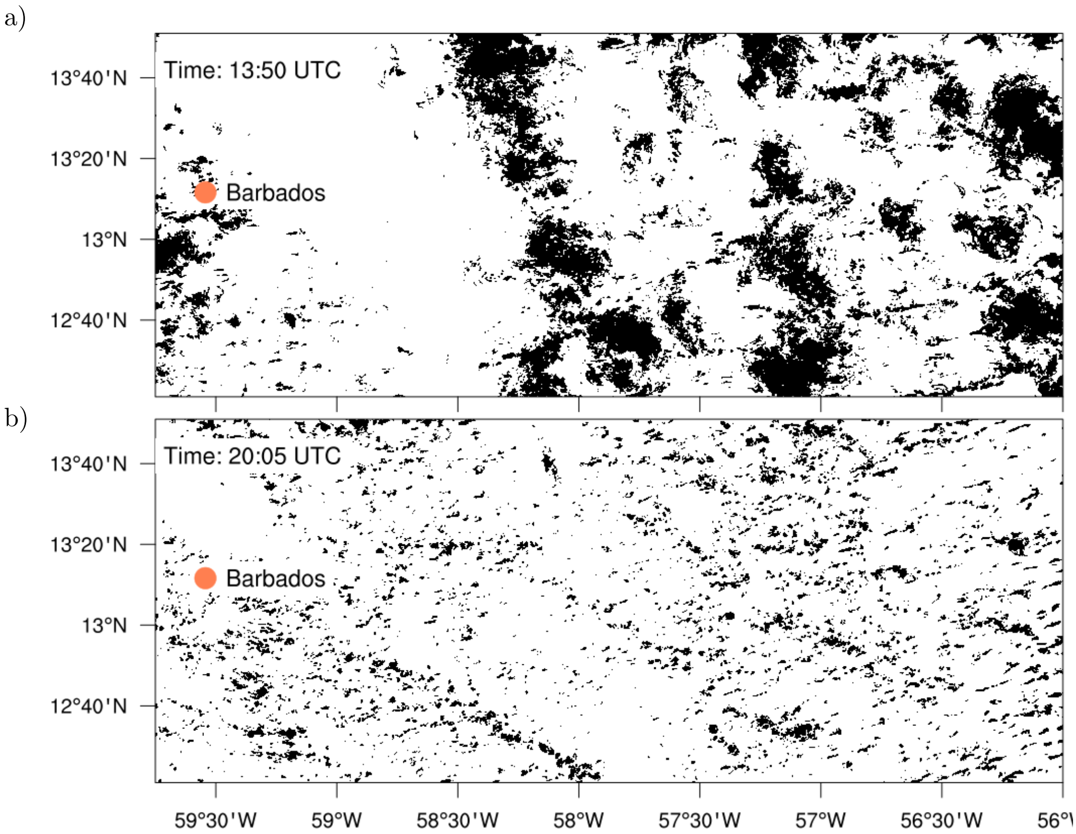

Figure 4. Snapshots of the simulated clouds on 20 December 2013 during the NARVAL-South field campaign. Shown is the cloud mask based on projected liquid water in ICON resolved at 150 m for a) 13:50 UTC and b) 20:05 UTC. Cloudy grid-points are black, while cloud-free points are white. The orange dot shows the location of the island of Barbados.

2.2 ICON simulations

The simulation data used in this study to investigate the NNS in Trade Wind cumulus were generated in the context of the HD(CP)^2 project with the Icosahedral Non-hydrostatic (ICON) model (Zängl et al., 2014; Heinze et al., 2017). At the top of the model hierarchy are the regional ICON simulations described by Klocke et al. (2017), which consist of a set of four one-way nested domains. At its boundaries the outer domain is forced by three hourly ECMWF forecast data. Their inner domain, simulated at a 1.2 km horizontal resolution, functioned as the outer domain for the higher resolution LES experiments considered in this study. The configuration of these ICON LES simulations is described in detail by the recent studies of Stevens et al. (2019), Vial et al. (2019) and Naumann and Kiemle (2020), and accordingly only a brief summary will be provided here. Three further nested domain are included, with resolutions of about 600, 300 and 150 m. This yields a total of four domains simulated with LES, as shown in Figure 1c. The inner high-resolution domain (indicated as an orange box) is used in this study for the analysis of cloud spacing. It spans approximately 150 $\times$ 400 km^2 in the horizontal and 21 km in the vertical, discretized at 150 levels (with 30 levels in the lowest 2 km). With a horizontal resolution of 150 m the resolution of the inner domain is high enough to switch off all subgrid parametrizations except the ones for the surface layer, turbulence, cloud microphysics (Baldauf et al., 2011), and radiation (Mlawer et al., 1997).

The simulation starts at 12 UTC and ends 12 hours later. Every 15 minutes, a 3D field of liquid water is available as output, which serves as input for the clustering algorithm (as described in the next subsection). Figure 4 shows snapshots of the cloud mask based on the vertically integrated liquid water at two points in time in the simulation. The domain contains numerous resolved clouds, up to approximately 4500 per snapshot. This sample size, in combination with the broad range of resolved cloud sizes, make these simulations useful for studying spacing between clouds. The surface conditions are relatively homogeneous, so that any spatial organization in the cumulus cloud field will be mainly due to large-scale effects or domain-internal dynamics. A simple comparison by eye to the satellite image in Figure 1b suggests that at the later time point the simulated cloud field agrees better with the observed cloud field, lacking the large cloud decks present in the earlier snapshot that likely reflect model spin up effects.

A thorough evaluation of LES results against measurements is of crucial importance for gaining confidence in the model and to justify its use for scientific research. This study will make simple comparisons of the simulated boundary layer clouds to the observational data discussed above. In addition, the ICON LES experiments performed for NARVAL-I have already been thoroughly confronted with available observational data in previous studies (Stevens et al., 2019; Vial et al., 2019; Naumann & Kiemle, 2020). This study builds on the encouraging results coming out of these model evaluations concerning the basic state of the trade wind boundary layer. The main focus is then to gain insight into the *two-dimensional spatial statistics* of the simulated cloud population, thus using the simulation as a virtual laboratory. Comparisons of these characteristics will be made between the LES and available satellite imagery. To this purpose a clustering algorithm is used, which is described next.

2.3 Clustering algorithm

A clustering algorithm is used to compute the cloud sizes and locations from output on the model grid. To this purpose the GRIDCLUS algorithm is applied (Schikuta, 1996), which has been used in many LES studies of cumulus cloud fields (Neggers et al., 2003, 2019; van Laar et al., 2019). The liquid water field is projected on the surface and a grid cell is considered cloudy if the integrated liquid water path is bigger than the model threshold of $1 * 10^{-8}$ kg/kg. If two cloudy cells share a cell edge, they are considered part of the same cloud. Cloud size is defined as the radius of a circle that has the same area of the cloudy grid cells belonging to the cloud (Rieck et al., 2014). The center of mass of the cloud is taken as the center of the circle, the coordinates of this point are used for determining the spacing between the clouds.

2.4 Nearest Neighbor Spacing

The first spacing considered is the distance between a cloud and its closest neighbor, regardless the size of the latter. This distance, hereafter referred to as NNS, is depicted in the schematic of Figure 5 together with the other spacings we study. NNS is calculated following the method adopted from Joseph and Cahalan (1990) and Tompkins and Semie (2017). In practice this means that for every cloud, the minimum distance is selected from the distances to all other clouds. Let $\mathcal{K}$ represent the total set of clouds, with n the total number of clouds: $\mathcal{K} = \big\{1, 2, ..., n\big\}$. NNS between cloud k and its neighbors n is defined as:

$$\mathrm{NNS}(k) = min\big\{d(n,k) \mid n \in \mathcal{K} \setminus \{k\}\big\}, \tag{1}$$

with $d(n,k)$ the great circle distance (Euclidian distance corrected for the curvature of the Earth) between the centers of cloud n and k. The second measure of cloud spacing that is considered is the distance between a cloud and the closest neighbor that has a similar size. This measure, referred to as the *equal-size* NNS (NNS_σ), only considers clouds that have a similar size (l) and belong to the same bin (σ), as determined by the clustering algorithm. This makes our set of clouds dependent on l: $\mathcal{K}_\sigma = \big\{k \in \mathcal{K} \mid l(k) = \sigma\big\}$. NNS_σ is then defined as:

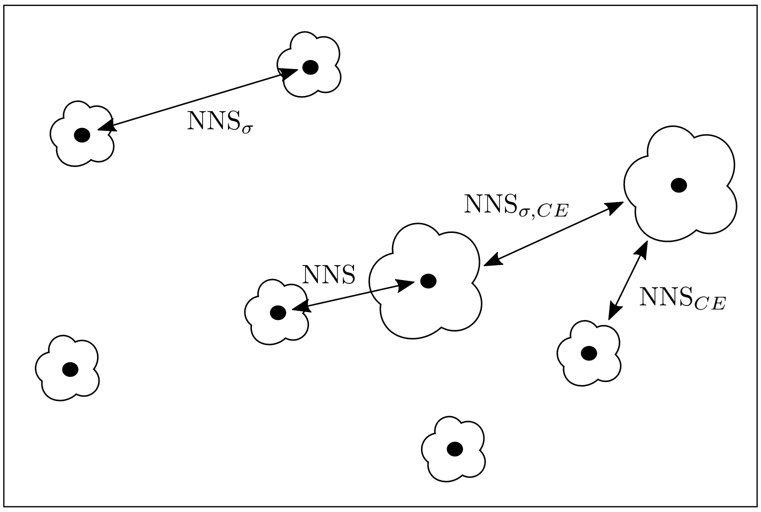

Figure 5. Example field with two cloud sizes showing the difference between NNS (distance between one cloud and its' nearest neighbor) and NNS_σ (distance between one cloud and its' nearest neighbor of a similar size).

$$\text{NNS}_\sigma(k) = min\big\{d(n,k) \mid n \in \mathcal{K}_\sigma \setminus \{k\}\big\}. \tag{2}$$

In calculating NNS and NNS_σ two different approaches were followed, yielding in total four measures of cloud spacing. First, the *cloud center spacing* is used, the distance from cloud center to cloud center. Second, the *cloud edge spacing* is the distance from cloud edge to cloud edge, computed by assuming that all clouds are perfect circles (Rieck et al., 2014; Dawe & Austin, 2013). In essence, it is the cloud center spacing minus the size (radius) of the two neighboring clouds: $d_{CE} = d(n,k) - r_n - r_k$. Then the NNS_{CE} for using cloud edge spacing is defined as:

$$\text{NNS}_{CE}(k) = min\big\{d_{CE}(n,k) \mid n \in \mathcal{K} \setminus \{k\}\big\}, \tag{3}$$

and the equal-size NNS using cloud edge spacing ($NNS_{\sigma,CE}$) as:

$$\text{NNS}_{\sigma,CE}(k) = min\big\{d_{CE}(n,k) \mid n \in \mathcal{K}_\sigma \setminus \{k\}\big\}. \tag{4}$$

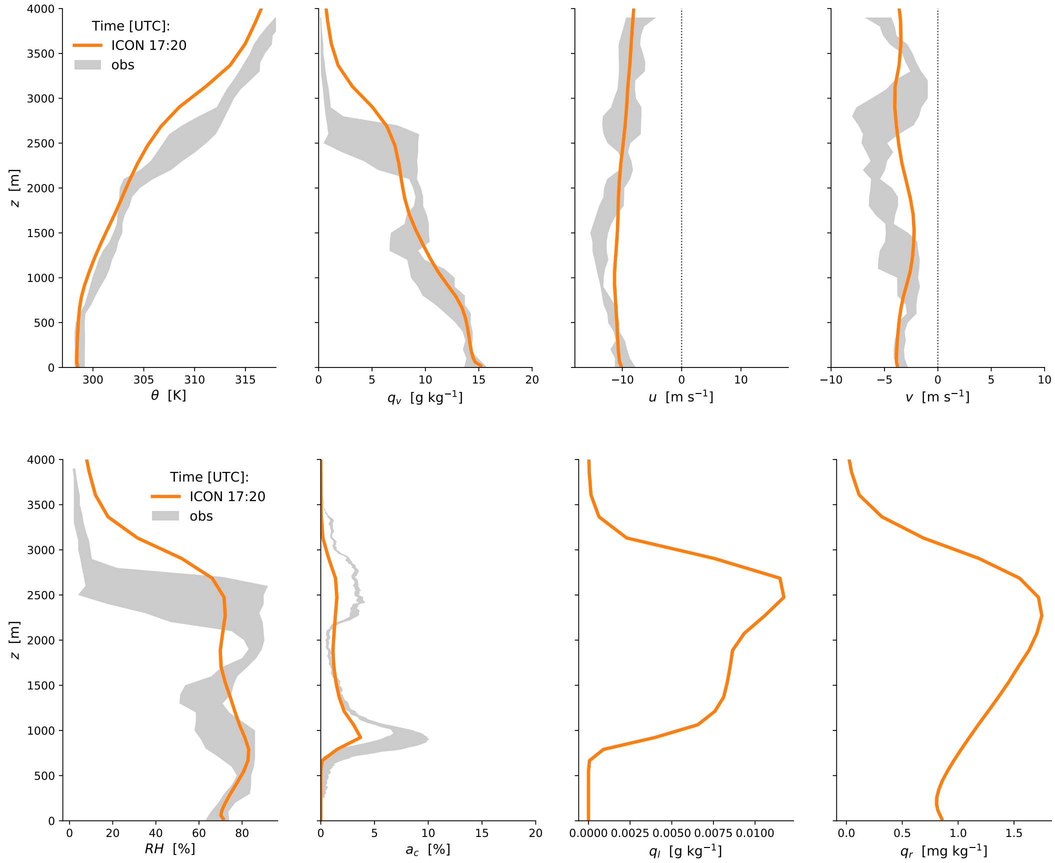

Figure 6. Vertical structure of the simulated (orange) and observed (grey) Trade wind boundary layer during RF08. Shown are a) potential temperature θ, b) water vapor specific humidity q_v, c) zonal wind speed u, d) meridional wind speed v, e) relative humidity RH, f) cloud fraction a_c, g) cloud liquid water q_c and h) rain water q_r. Sonde observations include the first two dropsondes and the Barbados radiosonde, the range indicating the minima and maxima encountered within 100m height bins. The WALES data plotted in f) is identical to those shown in Figure 3c.

2.5 Reference NNS values

Previous studies on spatial organization (Zhu et al., 1992; Nair et al., 1998) using cumulative distribution functions have yielded reference values of NNS that reflect purely randomly distributed populations. These reference values are based on a Poisson point process, which is a collection of points randomly distributed in space. The number of points can then be described by a Poisson distribution. When assuming clouds can be represented by points, the cumulative distribution of the nearest neighbor distances of cumulus clouds can be directly compared to the cumulative distribution of distances of a Poisson point process. The comparison between our simulated field and a random field can be summarized with a single value, referred to as I_{org} (Organization Index) (Tompkins & Semie, 2017). I_{org} distinguishes three regimes: randomness ($I_{org} \approx 0.5$), clustering ($I_{org} > 0.5$) and regularity ($I_{org} < 0.5$).

Based on a Poisson process and following the mathematical derivation, one can also define the mean NNS a random distribution would give (NNS_{ran}) (Weger et al., 1992):

$$NNS_{ran} = \frac{\sqrt{A}}{2\sqrt{N}}. \tag{5}$$

Here A is the domain area and N the number of clouds. As opposed to a random distribution of clouds one could also think of a regular distribution. In that case the clouds form a grid-like pattern, thereby maximizing NNS for the given amount of clouds. This NNS_{reg} could be determined as follows:

$$NNS_{reg} = \sqrt{\left(\frac{A}{N}\right)}. \tag{6}$$

3 Results

3.1 Vertical structure

Figure 6 shows vertical profiles of variables expressing the vertical structure of the simulated Trade wind boundary layer. For each variable measurements are included when available, including the first two dropsondes of HALO RF08, the Barbados radiosonde and WALES cloud fraction profile as already shown in Figure 3. The simulation data is sampled at the output timepoint (17:20 UTC) closest to the two dropsondes, and averaged over the full domain in order to optimize comparability with the sounding data which covers a similar spatial domain. The results suggest that the thermodynamic vertical structure of the cloud layer is reproduced reasonably well by the simulation, with the subcloud mixed layer and convective cloud layer situated at the right heights and featuring a similar conditional instability and humidity gradient. Slight thermodynamic biases include an overestimation in the inversion height and a small cold and moist bias in the lower free troposphere. The wind structure is realistic, including a well-defined easterly throughout the lowest 4 km featuring a small northerly component.

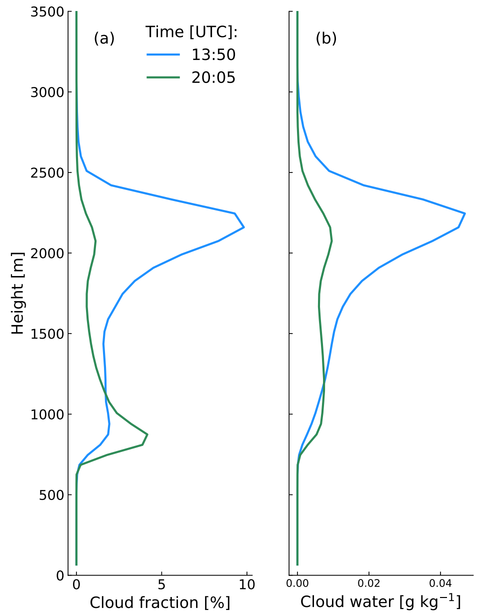

Figure 7. Simulated vertical profiles of a) cloud fraction and b) cloud water averaged over the domain at the two timepoints of the instantaneous cloud fields shown Figure 4. The line colors corresponds to the coloured time points as shown in Figure 8, for reference.

Figure 6 (e-f) focuses on the simulated cloud structure. Defining and typical Trade wind cloud features that are reproduced include the two distinct maxima in relative humidity at approximately $\sim$ 700 m and $\sim$ 2500 m height, and a concave structure situated in between. This structure and amplitude agrees well with the observations. A similar two-mode structure is evident in the cloud fraction profile, reproducing the WALES observations in this respect. The model slightly underestimates the magnitude of the profile at these two maxima, for which we speculate two reasons can exist; i) the threshold value range used to compute the observed profile from backscatter measurements, or ii) a lack of skill in LES to produce enough cloud mass. The latter would be consistent with results reported in recent studies comparing LES results to cloud observations (Y. Zhang et al., 2017). More research is required to gain insight into this question. For cloud liquid and rain water no observations are available; however, their vertical structure is similar to LES results for previous Caribbean cumulus cases (vanZanten et al., 2011).

Despite slight biases, the overall assessment is that the key features of the Trade wind boundary layer observed in the region during RF08 are reproduced to a high enough

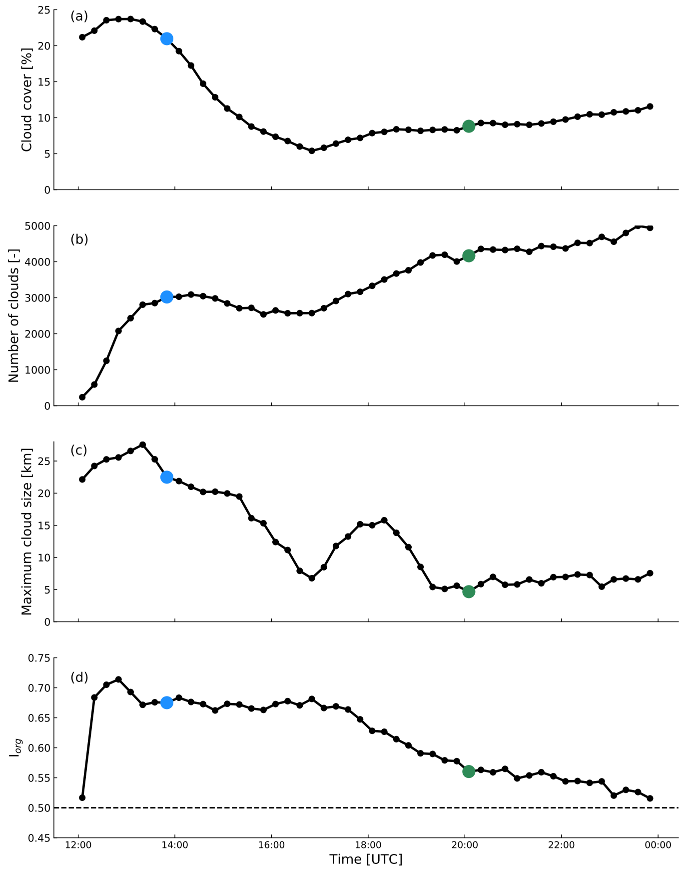

Figure 8. Timeseries of a) cloud cover, b) number of clouds, c) maximum cloud size and d) I$_{org}$. The blue and green dot correspond to the upper and lower panel of Figure 4 respectively.

degree to justify using this simulation for further investigation of cumulus cloud spacing.

3.2 Time evolution

While the good agreement with the observed vertical structure of the boundary layer during RF08 is encouraging, it is important to realize that this evaluation only applies to a brief time-window. In fact, Figure 4 already demonstrated that the spatial structure of the simulated cloud field experiences a substantial transition in time. Figure 7 further shows that the transition not only concerns the spatial structure but also the cloud vertical structure, here diagnosed at the exact two snapshots as shown in Figure 4. The first phase should be considered model spin-up, as it is still close to the initialization time of the simulation. This phase is characterized by high cloud covers and high cloud water amounts at a height slightly above 2 km, formed by large cumulus outflow clouds. The large clouds are not numerous, but do have sizes up to approximately 25 km. In contrast, later in the simulation the distribution is dominated by many more smaller clouds, and has far less large outflow cloud layers.

The transition of the cloud field is perhaps best expressed by the timeseries of cloud cover, clouds number and maximum cloud size as shown in Figure 8a-c. During the first two hours of the simulation the number of clouds rapidly increases; subsequently the increase is much more gradual. The latter phase is accompanied by a drop in cloud cover and maximum cloud size. This combination indicates that the initial large structures are gradually being replaced by smaller clouds. The increase in cloud number mainly happens at lower levels, as expressed by the cloud cover similarly increasing in that height range (Figure 7). Low-level clouds become more pronounced as time progresses, although their liquid water content is lower compared to the high-level clouds. The decrease of projected cloud cover over time (Figure 8a) is driven by the disappearance of the high-level clouds. During the second half of the simulation the maximum cloud size stays more or less constant at about 5 km, except for a modest peak around 18:00 UTC.

Apart from the transition in cloud cover, number and maximum size, their spatial distribution changes as well. Figure 4 suggests that the large cloud clusters are gradually replaced by smaller clouds that are either randomly distributed or form cloud streets. A quantification of the degree of organization is provided by I_{org}, shown in Figure 8d. The black dashed line in the figure indicates pure randomness, therefore I_{org} suggests strong organization at the beginning of the simulation. The degree of organization starts to decrease around 17:00 UTC, with values close to random at the end of the simulation.

Based on this analysis of the temporal evolution of the cloud field and the observational data showing a cloud field dominated by small-scale low level boundary layer cumulus, we decided to take the first six hours of the simulation not into account for the analysis. After this period the cloud cover, maximum size and number stay more or less constant. This behavior motivates using only the last six hours (24 time steps) for the analysis of the size distributions of cloud number and cloud spacing.

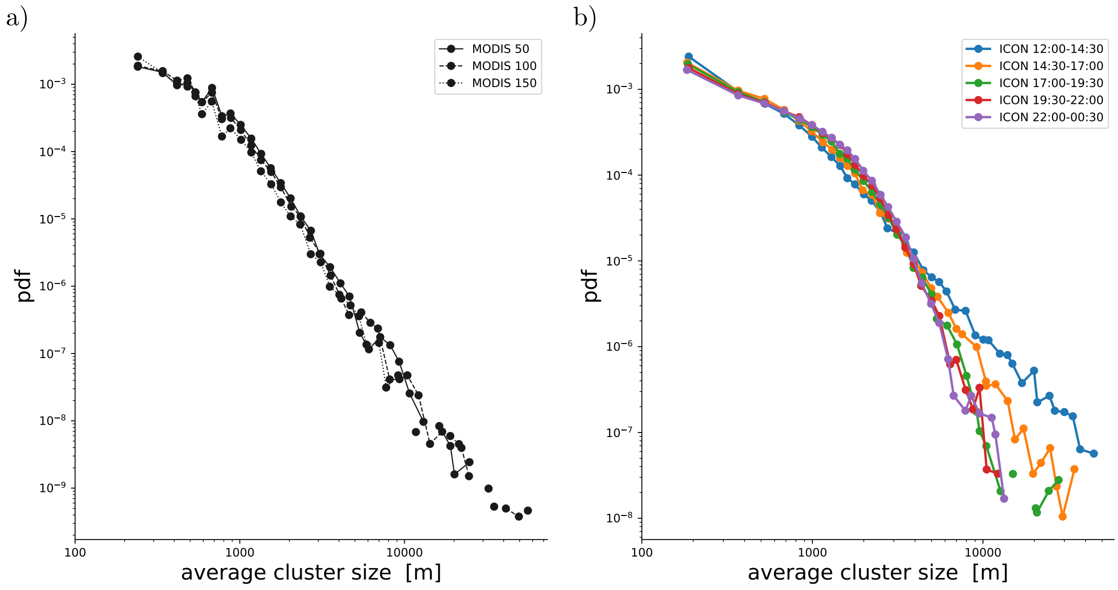

Figure 9. Size distributions of the number of clusters as derived from observations and simulations. a) Based on MODIS satellite image as shown in Figure 1b, showing results for three reflectance thresholds (50, 100 and 150) to define cloudy pixels. b) Distributions diagnosed from the integrated liquid water path in ICON for five subsequent time-periods (indicated by the UTC times in the legend). The vertical axis represents the normalized cluster number divided by the binwidth, while the horizontal axis represents the average cluster size per bin. Log-linear binning is used to calculate these histograms, as described in the text.

3.3 Size distributions

Figure 9 shows size distributions of the observed and simulated number of clouds. These CSDs (cloud size distributions) have been generated using the clustering algorithms as described in Section 2.3. The CSDs for the MODIS image as shown in Figure 1b are derived using three thresholds for reflectance in the red channel. The clusters are size-sorted using the linear-logarithmic binning as described by Quinn and Neelin (2017), which ensures that the binwidth can not be smaller than the smallest possible cluster size. To this purpose minimum binwidths of 250m and 150 m are used, which are the effective resolutions of the MODIS product and the LES simulations, respectively.

Both the simulated and observed size distributions exhibit a similar functional form in their dependence on cluster size. This shape, featuring two size-ranges with a distinctly different size-dependence, has often been reported in previous studies of cumulus convection (Neggers et al., 2003). In both the observations and the simulation the cluster

Table 1. Powerlaw exponents b resulting from least-square fits in log-log space of a single powerlaw function al^b to the size densities shown in Figure 9. Fits are applied in two ranges of cluster sizes l, including a small (0-1 km) and a large (1-10 km) range.

dataset	$0-1$ km	$1-10$ km
MODIS 50	-1.29	-3.88
MODIS 100	-1.48	-3.90
MODIS 150	-2.05	-3.85
ICON 12:00-14:30	-1.24	-2.38
ICON 14:30-17:00	-1.05	-2.82
ICON 17:00-19:30	-0.99	-3.60
ICON 19:30-22:00	-0.89	-3.71
ICON 22:00-00:30	-0.85	-3.96

size at which the dependence changes is at about 1km. Applying a single powerlaw fit in both size ranges yields powerlaw exponents as listed in Table 1. The model reproduces the distinct difference in powerlaw exponents between the two size-ranges. Note that the observed CSD shifts to the left with a higher reflectance threshold, expressing that fewer clusters are then detected, as can be expected. However, the distribution shape is still preserved. Over time the simulated CSD becomes steeper in the large size range, expressing that less and less big structures feature in the simulation.

The range of observed cluster sizes spans about three orders of magnitude, with the largest cluster size being about 60 km. Note that this maximum size does show strong dependence on the reflectance threshold, reducing to 30 km for the highest value. Accordingly, this aspect of the distribution is not very robust, also because the sample size is very small in this tail of the distribution. The simulated maximum cluster size is somewhat lower, but still significantly larger than the scales of moist boundary layer updrafts. These results reflect that while some larger stratiform outflow clouds do appear in the simulation, their total number is underestimated. Note that our main goal is to study spacing among cumuliform clouds, which sit in the left and middle part of the distribu-

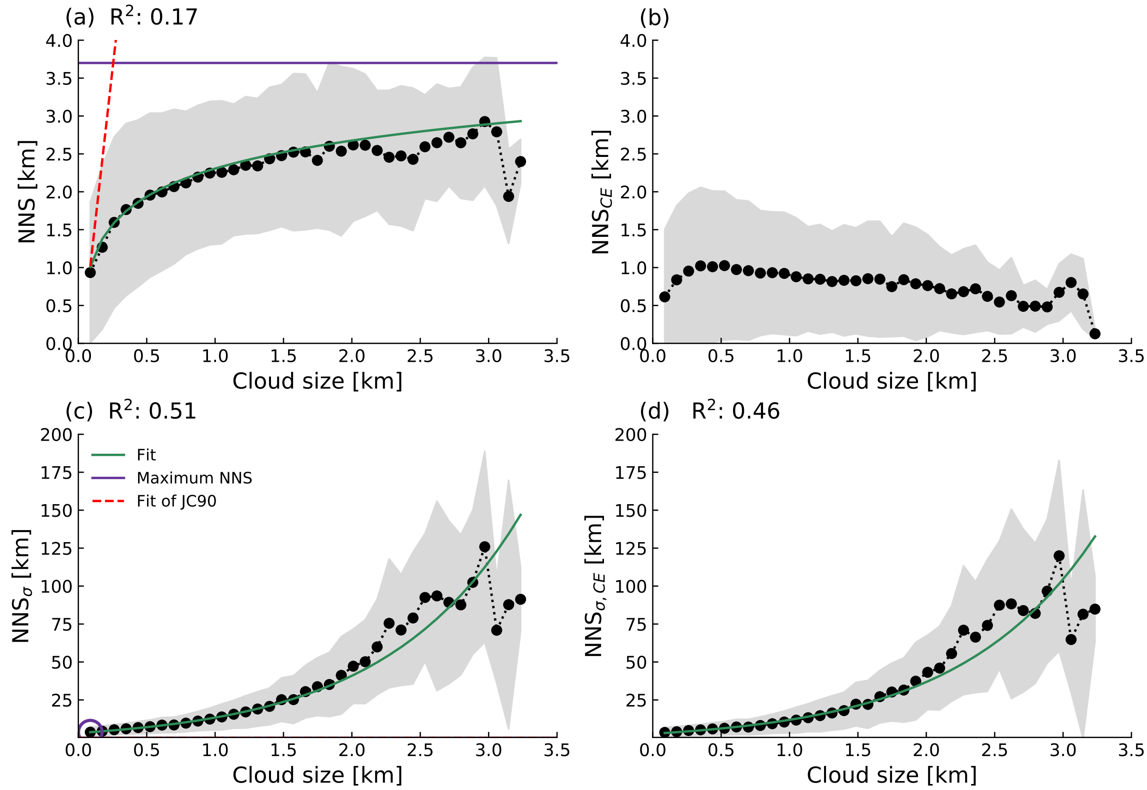

Figure 10. Cloud spacing as a function of cloud size. Shown are a) NNS, b) NNS$_{CE}$, c) NNS$_\sigma$, and d) NNS$_{\sigma,CE}$. The NNS is averaged over all analysed fields, the grey area shows the mean ± the standard deviation. The green lines show the best fits through the data with their R^2 value in the upper left corner. The purple line (a) and circle (c) indicate the maximum NNS and the Red dashed line shows the fit (Joseph & Cahalan, 1990) found.

tion. Accordingly, the underestimation of the number of large stratiform cloud decks does not harm the usefulness of the simulation.

We conclude from this comparison that the simulated cloud populations are representative of subtropical marine conditions as typically occur in the Trade wind regions, and of the cumulus cloud population as observed on 20 December 2013 in particular. This motivates using the simulation output for further analysis of cloud spacing.

3.4 Cloud spacing

The four panels of Figure 10 show the size dependence for all four definitions of the Nearest Neighbor Spacing (NNS) as defined in Section 2.4 and as applied to the ICON

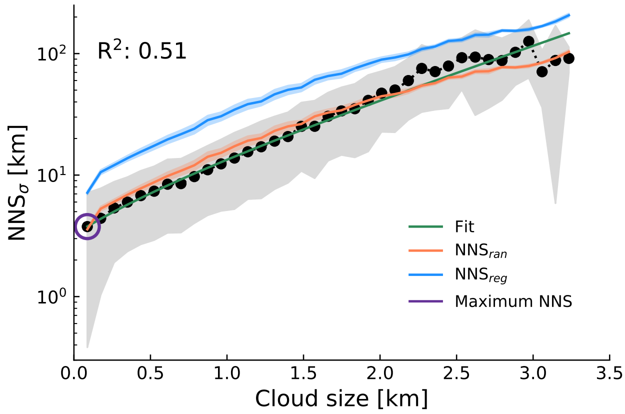

Figure 11. Same as Figure 10c but now showing NNS_σ using log-linear axes. Also included are NNS_{ran} and NNS_{reg}, for reference.

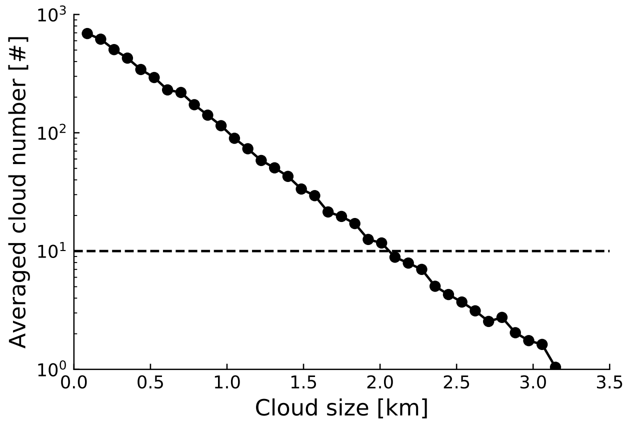

Figure 12. Number of clouds per cloud size, averaged over all snapshots used for analysis. For the calculation of this histogram a constant bin-size is used. The horizontal dashed line refers to the statistical analysis in Section 3.4.

LES fields. The results represent averages over the last six hours of the simulation. A least squares fit is made for the size dependence in each NNS definition, adopting a functional form that yielded the largest R^2 value (proportion of total variance explained by the fit) for each definition. Unfortunately the limited availability of MODIS data for this day and area (only a single snapshot) yields a sample size too low to reliably carry out this cloud spacing analysis for the observational data. Accordingly, this is for now considered a future research topic.

Figure 10a shows the size dependence of NNS, which in the range < 600 m shows a linear relation, just like the fit reported by Joseph and Cahalan (1990). However, at larger sizes the dependence is best captured by a logarithmic relation ($y = 2.31+1.23\log_{10}(x)$), with an R^2 value of 0.17. For the larger cloud sizes, the mean falls slightly below the fit. The limited amount of data for the statistical analysis might be a reason for this; cloud number decreases strongly with cloud size, so that the largest clouds only rarely occur. This is well visible in Figure 12 b, in which a dotted line at an average of 10 clouds per bin is added for reference. The cloud size associated with this sample size is about 2 km; note that this is also the size above which the NNS starts to deviate significantly from the proposed fit (see Figure 10a). This suggests that a number of 10 clouds is the minimum sample size at which a clear functionality becomes apparent in the size dependence of the NNS.

The spacing between clouds of a similar size, NNS_σ (Eq. 2), is shown in Figure 10c. Again we find a monotonically increasing cloud spacing with cloud size; however, for this definition the relation is best captured by an exponential function ($y = -2.66+5.90exp(x)$) with an R^2 value of 0.51. Other differences with NNS include i) much larger spacing values across the spectrum and ii) an increasing spread around the mean. The larger spacing of NNS_σ in general, as compared to NNS, directly reflects that only a subset of all clouds in the population is considered when calculating the equal-size spacing; a lower density of clouds in an area is directly associated with a larger spacing. But the exponential increase with size of the NNS_σ is not so trivial, and will be further interpreted in Section 4.3.

It makes sense to compare the equal-size cloud spacing NNS_σ to the theoretical limits NNS_{ran} and NNS_{reg}, as defined in section 2.4. The results of this comparison are shown in Figure 11. NNS_σ is very similar to NNS_{ran}; for all cloud sizes NNS_{ran} stays within the spread of NNS_σ. At the same time NNS_σ has significantly lower values than NNS_{reg}. The spatial distribution of clouds of a given size is close to random, although some differences between small and big clouds can be distinguished. The equal-size spacing for small clouds is slightly smaller than what a random distribution (following a Poisson point process) would give, meaning that they are more clustered together. NNS_σ for larger clouds, on the other hand, is larger than NNS_{ran} and resembles more a regular distribution.

The impact of cloud edge spacing on the NNS is investigated in Figure 10b and d. When interpreting these results it is important to consider that the spacing for bigger clouds could be larger simply because their centers are spaced further apart, due to their

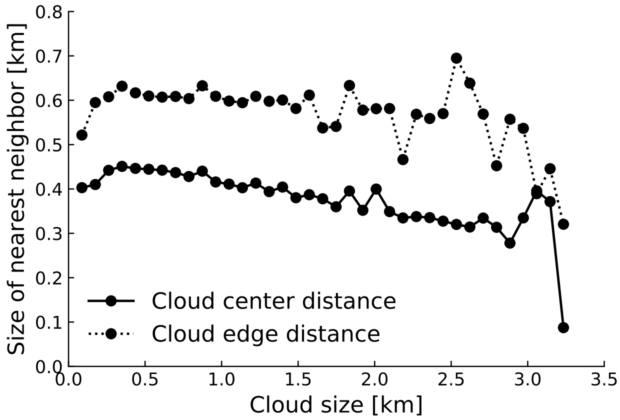

Figure 13. The averaged nearest neighbor size as a function of cloud size, using both the cloud center spacing (solid) and the cloud edge spacing (dotted).

size. Spacing definitions NNS_{CE} and $\mathrm{NNS}_{\sigma,CE}$, as defined by Equations 3 and 4, both reflect this effect. Using the cloud edge for the spacing leads to only minor differences for NNS_{σ}, preserving its functional dependence but shifting it downwards somewhat (Figure 10d). The exponential again yields the best fit ($y = -2.60 + 5.33\ exp(x)$), albeit with a slightly lower R^2 value of 0.46. In contrast, for NNS_{CE} (Figure 10b), the spacing is not only smaller, the logarithmic dependence is also lost. After a first increase of NNS_{CE} with cloud size, for clouds larger than about 400 meter a slight decrease of spacing with size is visible.

4 Interpretation

4.1 The impact of edge versus center spacing

More insight into the strong impact of adopting cloud-edge spacing versus cloud-center spacing on the size dependence of NNS is provided by considering the size of the nearest neighbors, as shown in Figure 13. The size of the nearest neighbor can be determined using both definitions of spacing. For both methods, after a slight increase for the small cloud sizes, the size of the neighboring clouds weakly decreases with cloud size. However, while both definitions share this weak size-dependence, the feature that most catches the eye is that the averaged neighbor size is universally larger when using cloud edge spacing.

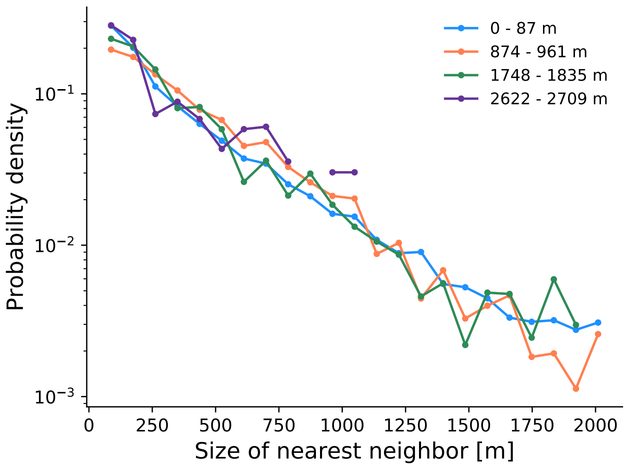

Figure 14. Normalized probability density of the size of the nearest neighbor cloud, for four different size bins.

What explains this difference in neighbor size? When edge distancing is used to determine the nearest neighbor, the radius of the clouds starts to play a role. For bigger clouds this matters more than for smaller clouds, because their edge is closer to an arbitrary cloud of interest compared to their center. As a result, the probability that a large cloud is closest is bigger when edge distancing is used. This differential impact for larger clouds also explains the strong impact of edge distancing on the functional dependence of NNS on size, as visible in Figure 10b. For larger clouds their radius makes up a larger fraction of the center-center distance; as a result, the neighbor spacing reduces more for larger clouds when switching from center-distancing (Figure 10a) to edge-distancing (Figure 10b). In effect, this counteracts the logarithmic increase in the center spacing.

4.2 Logarithmic dependence: The role of small clouds

At first glance Figure 13 seems to suggest that large clouds have smaller clouds as nearest neighbors, and vice versa. However, this dependence should be interpreted with some caution, because i) the averaging might obscure quite some spread, and ii) the dependence is weak to start with. To gain further insight the probability of having a neighbor of a certain size is investigated, as shown in Figure 14 for four different cloud sizes. For all cloud sizes considered, the probability of having a small cloud as nearest neighbor is by far the highest (note the logarithmic y-axis). Another interesting feature is that the largest clouds do not even have large clouds as nearest neighbors; smaller clouds are always closer.

In combination, these results go some way to explain why the NNS has a logarithmic size-dependence, as shown in Figure 10a. Firstly, it is important to consider that the smallest clouds are by far the most abundant in the field (Figure 9), and are also more or less randomly distributed (Figure 11). This not only means that large clouds have predominantly small clouds as nearest neighbors (Figure 14), but also that the cloud spacing that occurs most often in the domain is the equal-size distance of the smallest clouds. In this case this value is about 3.7 km, as marked by the purple circle in Figure 11. As a consequence of the abundance of this spacing, one expects that the NNS of the bigger clouds can (on average) not be much larger than this value. If this reasoning holds, then the maximum NNS would on average also be 3.7 km, thus more or less acting as a limit value. The purple line shown in Figure 10a indeed seems to act as an upper boundary.

With the equal-size spacing NNS_σ increasing exponentially with size, the picture emerges that the large clusters are swimming in a sea of small clouds. This large spacing makes it more likely that smaller clouds (with smaller spacings) are present in between the large clouds, hence the saturation for increasing cloud size. The spatial distribution of the large clouds does not play a role in this, as long as their sample size is large enough and the small clouds indeed dictate the spacing. This argumentation is summarized schematically in Figure 15.

The existence of an upper limit for the NNS would imply that the NNS will not increase anymore towards very large cloud sizes. Although the logarithmic function fits well to the data at hand, formally such saturation behaviour is not described by a logarithmic function but an asymptotic one. However, determining in a statistically significant way if the NNS actually saturates requires sampling many more large clouds, much larger than those present in these simulations. Accordingly, answering this question is for now considered future research. This could well be achieved by using an abundance of independent satellite snapshots at high resolutions covering large (ocean-covering) domains.

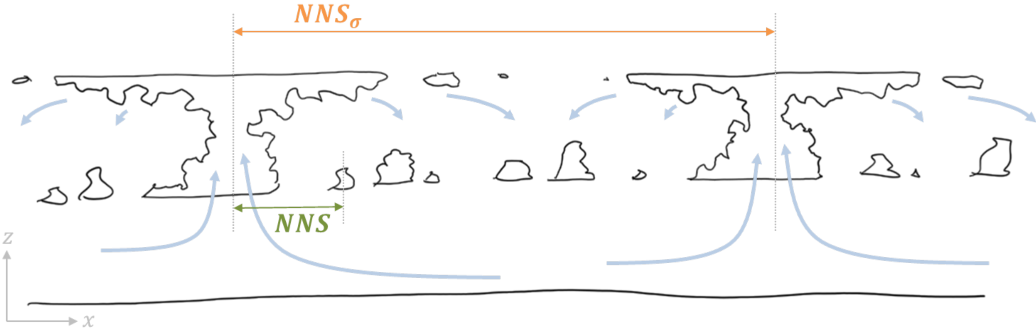

Figure 15. Schematic illustration of the difference in size dependence in the nearest neighnor spacing between clouds of any size (NNS) and between clouds of equal size (NNS$_\sigma$). In this vertical cross-section the cloud outlines are indicated in black, while the mesoscale cell circulation is shown as thick blue arrows. The two definitions of spacing that are shown apply to the largest cloud size in the domain.

4.3 Exponential dependence: Mesoscale dynamics?

The exponential dependence of the equal-size spacing NNS$_\sigma$ on cloud size is statistically significant, and also robust for the various definitions of the spacing that are considered in this study. To our knowledge it has not been reported before, but it is relevant for the representation of convection in weather and climate models. As shown in Figure 10c, in this case the spacing between clouds increases from about 10km for clouds of 1km size via 50 km for 2 km-sized clouds to 100km for clouds of 3km. These spacings are similar to the grid spacings used in global circulation models. Accordingly, they should be taken into account in the parameterization of convection in the grey zone (Wyngaard, 2004; Honnert et al., 2020), for example in the representation of stochastic effects due to subsampling (Neggers et al., 2019).

The relevance of cloud spacing motivates gaining more insight into what processes might cause the exponential size-dependence. Convective clouds are the visible parts of a much larger convective cell, featuring a relatively narrow updraft area and a much wider area with compensating subsidence that can be either cloudy or cloud free (see Figure 15). Such cells are typically observed in many moist convective regimes, not just fair-weather cumulus (Shao & Randall, 1996; de Roode et al., 2004). Our results suggest that, on average, bigger convective clouds need more space around them to form a convective cell, in a super-linear way. In the mesoscale, the dynamics of such cells markedly changes due to the increasing occurrence of rain. For example, wider and deeper clouds are observed when cold pools occur driven by evaporation of rain (Schlemmer & Hohenegger, 2014), while precipitation has also been reported to play a role in setting the spatial scale of stratiform convective cells (Zhou et al., 2018; Zhou & Bretherton, 2019). The "flower" type of cloud patterns in the Trades is also associated with such dynamics (Stevens et al., 2019). In general, the interaction of radiation with clouds and water vapor is also thought to play a key process in convective aggregation (Bretherton et al., 2005; Muller & Held, 2012). To summarize, these known impacts on mesoscale dynamics make precipitation and radiative cooling prime candidate processes for controlling the exponentially increasing spacing with cloud size. Proving or disproving this hypothesis requires further research.

5 Summary and Conclusions

In this study Large-Eddy Simulations (LES) on super-large-domains are used to investigate how neighbor spacing in cumulus cloud populations depends on cloud size. To this purpose experiments with the ICON LES model of marine shallow cumulus cloud fields in the subtropical Atlantic as observed during the recent NARVAL South campaign were used. Cluster analyses were applied to derive size distributions of both cloud number and cloud spacing. MODIS satellite imagery is first used to test the realism and representativity of the simulated cloud fields. Despite a slight underestimation of the maximum cloud size, we find good agreement concerning the shape of the number distribution. A multitude of instantaneous snapshots from the simulation are then used to diagnose the cloud Nearest Neighbor Spacing (NNS), of which four possible definitions are considered. We find that in general the NNS increases with cloud size, a result which is in line with the findings of previous observational studies. However, the functional form of the size-dependence strongly depends on the exact definition of the NNS. Its classic definition, the spacing between clouds of any size, carries a well-defined logarithmic size-dependence. In contrast, only considering clouds of equal size yields cloud spacings that are larger but also carry a strong exponential size-dependence. Deeper investigation into this behavior reveals that the abundance of closely-spaced small clouds in the population is responsible for the logarithmic dependence. The exponential dependence is speculated to express the role of mesoscale dynamics in controlling the width of the convective cells of which the cumulus clouds are the visible parts.

The results obtained in this study are relevant for ongoing research into the spatial organization and aggregation of convection and its impact on climate (Wing, 2019). It has long been understood that cloud spacing is a key ingredient in this problem, as testified by the various metrics for the degree of spatial organization that have been proposed that are formulated in terms of the neighbor spacing (Weger et al., 1992; Tompkins & Semie, 2017). Most of these metrics depend on the spacing between clouds of any size. However, the equal-size spacing as investigated in this study could also be used to this purpose, yielding an alternative organizational metric that expresses different aspects of this phenomenon. A recent example is the B_{org} metric as proposed by Neggers et al. (2019), which exclusively relies on equal-size cloud spacing (NNS$_\sigma$) and expresses the degree of organization per cloud size. In the context of understanding cloud-climate feedbacks the exponential spacing might also be relevant, as it affects the impact of such cloud fields on radiation, in particular at low solar inclination angles.

The results of this study also have a bearing on the parameterization of convection in the grey zone (Wyngaard, 2004; Honnert et al., 2020). For example, the stochastic effects of subsampling on the cloud size distribution to be parameterized can be captured by using the neighbor spacing (Neggers et al., 2019). The functional form in the size-dependence of the cloud spacing can thus inform the further development of convection schemes based on cloud size distributions (Neggers, 2015; Sakradzija et al., 2016; Hagos et al., 2018). Through metrics relying on the neighbor spacing, constants of proportionality in such schemes could be constrained against observed and simulated cloud fields, for example using machine learning techniques.

This study has several limitations which could inspire future research efforts. Firstly, only relatively homogeneous conditions were considered, in order to focus on internal spatial organization in a cloud population. But heterogeneity in the larger-scale flow and surface can also affect the cloud spacing. Gaining insight into these impacts is needed to test the general applicability of the size dependence in cloud spacing as reported in this study. Secondly, the domain size could still artificially limit the maximum cloud size, which motivates considering even larger domain simulations. Another simplification in the experimental configuration is the use of 1D radiation in the simulation, which ignores three-dimensional effects that can change cloud alignment and spacing (Jakub & Mayer, 2017). Finally, we only applied our spacing analysis to simulated cloud fields. To seek observational support for the obtained results, an obvious next step would be to derive cloud spacing from multiple high-resolution satellite images. Such data is increasingly available, and is actively being used to investigate mesoscale spatial structures in low level cloud fields (Bony et al., 2020).

Acknowledgments

This research is supported by the SFB-TR32 'Patterns in soil-vegetation-atmosphere systems: monitoring, modelling and data assimilation', funded by the German Science Foundation (DFG). The authors want to thank Matthias Brueck from the Max-Planck-Institut für Meteorologie in Hamburg (MPI-H) who performed the ICON simulations in the scope of the HD(CP)^2 project. The NARVAL I data used in this publication is made available through the German Aerospace Center (DLR). NARVAL was funded with support of the Max Planck Society, the German Research Foundation (DFG, project HALO-SPP 1294), the European Research Council (ERC), the German Meteorological Weather Service (DWD) and DLR. The simulation data used in this study is available at the following website: https://doi.org/10.5281/zenodo.3531556.

References

Ali, K. (1998). Spatial distribution of convective clouds in north India. *Atmospheric Research*, *49*(1), 1 - 10. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0169809598000702` doi: https://doi.org/10.1016/S0169-8095(98)00070-2

Baldauf, M., Seifert, A., Förstner, J., Majewski, D., Raschendorfer, M., & Reinhardt, T. (2011). Operational convective-scale numerical weather prediction with the COSMO model: Description and sensitivities. *Monthly Weather Review*, *139*(12), 3887 - 3905. doi: 10.1175/mwr-d-10-05013.1

Bony, S., & Dufresne, J.-L. (2005). Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models. *Geophysical Research Letters*, *32*(20). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005GL023851` doi: 10.1029/2005GL023851

Bony, S., Schulz, H., Vial, J., & Stevens, B. (2020). Sugar, gravel, fish, and flowers: Dependence of mesoscale patterns of trade-wind clouds on environmental conditions. *Geophysical Research Letters*, *47*(7), e2019GL085988. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/`

10.1029/2019GL085988 (e2019GL085988 10.1029/2019GL085988) doi: 10.1029/2019GL085988

Brast, M., Schemann, V., & Neggers, R. A. J. (2018). Investigating the scale adaptivity of a size-filtered mass flux parameterization in the gray zone of shallow cumulus convection. *J. Atmos. Sci.*, *75*(4), 1195–1214. doi: 10.1175/jas-d-17-0231.1

Bretherton, C. S., & Blossey, P. N. (2017). Understanding mesoscale aggregation of shallow cumulus convection using large-eddy simulation. *J. Adv. Model Earth Syst.*, *9*(8), 2798–2821. doi: 10.1002/2017ms000981

Bretherton, C. S., Blossey, P. N., & Khairoutdinov, M. (2005). An energy-balance analysis of deep convective self-aggregation above uniform SST. *J. Atmos. Sci.*, *62*(12), 4273–4292. doi: 10.1175/jas3614.1

Dawe, J., & Austin, P. (2013). Direct entrainment and detrainment rate distributions of individual shallow cumulus clouds in an LES. *Atmos. Chem. Phys*, *13*, 7795-7811. doi: 10.5194/acp-13-7795-2013

de Roode, S. R., Duynkerke, P. G., & Jonker, H. J. J. (2004). Large-eddy simulation: How large is large enough? *J. Atmos. Sci.*, *61*(4), 403–421. doi: 10.1175/1520-0469(2004)061⟨0403:lshlil⟩2.0.co;2

Hagos, S., Feng, Z., Plant, R. S., Houze Jr., R. A., & Xiao, H. (2018). A stochastic framework for modeling the population dynamics of convective clouds. *Journal of Advances in Modeling Earth Systems*, *10*(2), 448-465. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017MS001214 doi: 10.1002/2017MS001214

Heinze, R., Dipankar, A., Henken, C. C., et al. (2017). Large-eddy simulations over Germany using ICON: a comprehensive evaluation. *Q. J. Roy. Meteorol. Soc.*, *143*(702), 69–100. doi: 10.1002/qj.2947

Honnert, R., Efstathiou, G. A., Beare, R. J., Ito, J., Lock, A., Neggers, R., … Zhou, B. (2020). The atmospheric boundary layer and the "gray zone" of turbulence: A critical review. *Journal of Geophysical Research: Atmospheres*, *n/a*(n/a), e2019JD030317. doi: 10.1029/2019JD030317

Jacob, M., Ament, F., Gutleben, M., Konow, H., Mech, M., Wirth, M., & Crewell, S. (2019). Investigating the liquid water path over the tropical atlantic with synergistic airborne measurements. *Atmospheric Measurement Techniques*

Discussions, 1–27. doi: 10.5194/amt-2019-18

Jakub, F., & Mayer, B. (2017). The role of 1-D and 3-D radiative heating in the organization of shallow cumulus convection and the formation of cloud streets. *Atmospheric Chemistry and Physics*, *17*(21), 13317–13327. doi: 10.5194/acp-17-13317-2017

Joseph, J. H., & Cahalan, R. F. (1990). Nearest neighbor spacing of fair weather cumulus clouds. *J. Appl. Meteorol.*, *29*(8), 793–805. doi: 10.1175/1520-0450(1990)029⟨0793:NNSOFW⟩2.0.CO;2

Klepp, C., Ament, F., Bakan, S., Hirsch, L., & Stevens, B. (2014). The NARVAL campaign report. *Berichte zur Erdsystemforschung / Max-Planck-Institut für Meteorologie*, *164*.

Klocke, D., Brueck, M., Hohenegger, C., & Stevens, B. (2017). Rediscovery of the doldrums in storm-resolving simulations over the tropical Atlantic. *Nature Geosci.*, *10*(12), 891–896. doi: 10.1038/s41561-017-0005-4

Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., & Clough, S. A. (1997). Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *Journal of Geophysical Research: Atmospheres*, *102*(D14), 16663 - 16682. doi: 10.1029/97jd00237

Muller, C. J., & Held, I. M. (2012). Detailed Investigation of the Self-Aggregation of Convection in Cloud-Resolving Simulations. *Journal of the Atmospheric Sciences*, *69*(8), 2551-2565. Retrieved from `https://doi.org/10.1175/JAS-D-11-0257.1` doi: 10.1175/JAS-D-11-0257.1

Nair, U. S., Weger, R. C., Kuo, K. S., & Welch, R. M. (1998). Clustering, randomness, and regularity in cloud fields: 5. the nature of regular cumulus cloud fields. *J. Geophys. Res.: Atmospheres*, *103*(D10), 11363–11380. doi: 10.1029/98jd00088

Nam, C., Bony, S., Dufresne, J.-L., & Chepfer, H. (2012). The 'too few, too bright' tropical low-cloud problem in CMIP5 models. *Geophysical Research Letters*, *39*(21). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012GL053421` doi: 10.1029/2012GL053421

Narenpitak, P., Bretherton, C. S., & Khairoutdinov, M. F. (2017). Cloud and circulation feedbacks in a near-global aquaplanet cloud-resolving model. *Journal of Advances in Modeling Earth Systems*, *9*(2), 1069-1090. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016MS000872 doi: 10.1002/2016MS000872

Naumann, A. K., & Kiemle, C. (2020). The vertical Structure and spatial Variability of lower tropospheric Water Vapor and Clouds in the Trades. *Atmospheric Chemistry and Physics*. doi: 10.5194/acp-20-6129-2020

Neggers, R. A. J. (2015). Exploring bin-macrophysics models for moist convective transport and clouds. *J. Adv. Model Earth Syst.*, *7*(4), 2079–2104. doi: 10.1002/2015MS000502

Neggers, R. A. J., Griewank, P. J., & Heus, T. (2019). Powerlaw scaling in the internal variability of cumulus cloud size distributions due to subsampling and spatial organization. *J. Atmos. Sci.*. doi: 10.1175/jas-d-18-0194.1

Neggers, R. A. J., Jonker, H. J. J., & Siebesma, A. P. (2003). Size statistics of cumulus cloud populations in large-eddy simulations. *J. Atmos. Sci.*, *60*(8), 1060–1074. doi: 10.1175/1520-0469(2003)60⟨1060:SSOCCP⟩2.0.CO;2

Norris, J. R. (1998). Low Cloud Type over the Ocean from Surface Observations. Part II: Geographical and Seasonal Variations. *Journal of Climate*, *11*(3), 383-403. Retrieved from https://doi.org/10.1175/1520-0442(1998)011<0383:LCTOTO>2.0.CO;2 doi: 10.1175/1520-0442(1998)011⟨0383:LCTOTO⟩2.0.CO;2

Nuijens, L., Serikov, I., Hirsch, L., Lonitz, K., & Stevens, B. (2014). The distribution and variability of low-level cloud in the north atlantic trades. *Q. J. Roy. Meteorol. Soc.*, *140*(684), 2364–2374. doi: 10.1002/qj.2307

Plank, V. G. (1969). The size distribution of cumulus clouds in representative Florida populations. *J. Appl. Meteor.*, *8*(1), 46–67. doi: 10.1175/1520-0450(1969)008⟨0046:TSDOCC⟩2.0.CO;2

Quinn, K. M., & Neelin, J. D. (2017). Distributions of Tropical Precipitation Cluster Power and Their Changes under Global Warming. Part I: Observational Baseline and Comparison to a High-Resolution Atmospheric Model. *Journal of Climate*, *30*(20), 8033-8044. Retrieved from https://doi.org/10.1175/JCLI-D-16-0683.1 doi: 10.1175/JCLI-D-16-0683.1

Reilly, S., Gesso, S. D., & Neggers, R. (2019). Configuring LES in sparsely sampled areas in the subtropical atlantic. *Journal of Applied Meteorology and Climatology*, *submitted*.

Rieck, M., Hohenegger, C., & van Heerwaarden, C. C. (2014). The influence of land surface heterogeneities on cloud size development. *Mon. Wea. Rev.*, *142*(10), 3830–3846. doi: 10.1175/MWR-D-13-00354.1

Sakradzija, M., Seifert, A., & Dipankar, A. (2016). A stochastic scale-aware parameterization of shallow cumulus convection across the convective gray zone. *J. Adv. Model Earth Syst.*, *8*(2), 786–812. doi: 10.1002/2016ms000634

Schikuta, E. (1996). Grid-clustering: an efficient hierarchical clustering method for very large data sets. *Proceedings of the 13th International Conference on Pattern Recognition*, *2*, 101-105.

Schlemmer, L., & Hohenegger, C. (2014). The Formation of Wider and Deeper Clouds as a Result of Cold-Pool Dynamics. *Journal of the Atmospheric Sciences*, *71*(8), 2842-2858. Retrieved from `https://doi.org/10.1175/JAS-D-13-0170.1` doi: 10.1175/JAS-D-13-0170.1

Schnitt, S., Orlandi, E., Mech, M., Ehrlich, A., & Crewell, S. (2017). Characterization of Water Vapor and Clouds During the Next-Generation Aircraft Remote Sensing for Validation (NARVAL) South Studies. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(7), 3114–3124. doi: 10.1109/jstars.2017.2687943

Senf, F., Klocke, D., & Brueck, M. (2018). Size-resolved evaluation of simulated deep tropical convection. *Mon. Wea. Rev.*, *146*(7), 2161–2182. doi: 10.1175/mwr-d-17-0378.1

Sengupta, S., Welch, R., Navar, M., Berendes, T., & Chen, D. (1990). Cumulus cloud field morphology and spatial patterns derived from high spatial resolution Landsat imagery. *J. Appl. Meteorol.*, *29*(1245-1267). doi: 10.1175/1520-0450(1990)029⟨1245:CCFMAS⟩2.0.CO;2

Shao, Q., & Randall, D. A. (1996). Closed Mesoscale Cellular Convection Driven by Cloud-Top Radiative Cooling. *Journal of the Atmospheric Sciences*, *53*(15), 2144-2165. Retrieved from `https://doi.org/10.1175/1520-0469(1996)053<2144:CMCCDB>2.0.CO;2` doi: 10.1175/1520-0469(1996)053⟨2144:CMCCDB⟩2.0.CO;2

Sherwood, S., Bony, S., & Dufresne, J. (2014). Spread in model climate sensitivity traced to atmospheric convective mixing. *Nature*, *505*, 37-42. doi: 10.1038/nature12829

Stevens, B., Ackerman, A. S., Albrecht, B. A., Brown, A. R., Chlond, A., Cuxart, J., … Stevens, D. E. (2001). Simulations of trade wind cumuli under a strong inversion. *Journal of the Atmospheric Sciences*, *58*(14), 1870 - 1891. Retrieved from `https://journals.ametsoc.org/view/journals/atsc/58/14/1520-0469_2001_058_1870_sotwcu_2.0.co_2.xml` doi: 10.1175/1520-0469(2001)058⟨1870:SOTWCU⟩2.0.CO;2

Stevens, B., Ament, F., Bony, S., Crewell, S., Ewald, F., Gross, S., … Zinner, T. (2019). A High-Altitude Long-Range Aircraft Configured as a Cloud Observatory: The NARVAL Expeditions. *Bulletin of the American Meteorological Society*, *100*(6), 1061-1077. Retrieved from `https://doi.org/10.1175/BAMS-D-18-0198.1` doi: 10.1175/BAMS-D-18-0198.1

Tompkins, A. M., & Semie, A. G. (2017). Organization of tropical convection in low vertical wind shears: Role of updraft entrainment. *J. Adv. Model Earth Syst.*, *9*(2), 1046–1068. doi: 10.1002/2016ms000802

van Laar, T. W., Schemann, V., & Neggers, R. A. J. (2019). Investigating the diurnal evolution of the cloud size distribution of continental cumulus convection using multiday LES. *J. Atmos. Sci.*, *76*(3), 729–747. doi: 10.1175/jas-d-18-0084.1

vanZanten, M. C., Stevens, B., Nuijens, L., Siebesma, A. P., Ackerman, A. S., Burnet, F., … Wyszogrodzki, A. (2011). Controls on precipitation and cloudiness in simulations of trade-wind cumulus as observed during rico. *Journal of Advances in Modeling Earth Systems*, *3*(2). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011MS000056` doi: https://doi.org/10.1029/2011MS000056

Vial, J., Bony, S., Dufresne, J.-L., & Roehrig, R. (2016). Coupling between lower-tropospheric convective mixing and low-level clouds: Physical mechanisms and dependence on convection scheme. *J. Adv. Model Earth Syst.*, *8*(4), 1892–1911. doi: 10.1002/2016MS000740

Vial, J., Vogel, R., Bony, S., Stevens, B., Winker, D. M., Cai, X., … Brogniez, H. (2019). A new look at the daily cycle of trade wind cumuli. *Journal of Advances in Modeling Earth Systems*, *11*(10), 3148-3166. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001746` doi: 10.1029/2019MS001746

Weger, R. C., Lee, J., & Welch, R. M. (1993). Clustering, randomness, and regularity in cloud fields: 3. the nature and distribution of clusters. *J. Geophys. Res.*, *98*(D10), 18449-18463.

Weger, R. C., Lee, J., Zhu, T., & Welch, R. M. (1992). Clustering, randomness and regularity in cloud fields: 1. theoretical considerations. *J. Geophys. Res.*, *97*(D18), 20519. doi: 10.1029/92jd02038

Wing, A. A. (2019). Self-aggregation of deep convection and its implications for climate. *Current Climate Change Reports*, *5*(1), 1–11. doi: 10.1007/s40641-019-00120-3

Wing, A. A., & Cronin, T. W. (2015). Self-aggregation of convection in long channel geometry. *Q. J. Roy. Meteorol. Soc.*, *142*(694), 1–15. doi: 10.1002/qj.2628

Wyngaard, J. C. (2004). Toward numerical modeling in the 'Terra Incognita'. *J. Atmos. Sci.*, *61*(14), 1816–1826. doi: 10.1175/1520-0469(2004)061⟨1816:TNMITT⟩2.0.CO;2

Zängl, G., Reinert, D., Rípodas, P., & Baldauf, M. (2014). The ICON (ICOsahedral non-hydrostatic) modelling framework of DWD and MPI-m: Description of the non-hydrostatic dynamical core. *Q. J. Roy. Meteorol. Soc.*, *141*(687), 563–579. doi: 10.1002/qj.2378

Zhang, M., Bretherton, C. S., Blossey, P. N., et al. (2013). CGILS: Results from the first phase of an international project to understand the physical mechanisms of low cloud feedbacks in single column models. *J. Adv. Model Earth Syst.*, *5*(4), 826–842. doi: 10.1002/2013MS000246

Zhang, Y., Klein, S. A., Fan, J., et al. (2017). Large-eddy simulation of shallow cumulus over land: A composite case based on arm long-term observations at its southern great plains site. *J. Atmos. Sci.*, *74*. doi: https://doi.org/10.1175/JAS-D-16-0317.1

Zhou, X., Ackerman, A. S., Fridlind, A. M., & Kollias, P. (2018). Simulation of Mesoscale Cellular Convection in Marine Stratocumulus. Part I: Drizzling Conditions. *Journal of the Atmospheric Sciences*, *75*(1), 257-274. Retrieved from `https://doi.org/10.1175/JAS-D-17-0070.1` doi: 10.1175/JAS-D-17-0070.1

Zhou, X., & Bretherton, C. S. (2019). Simulation of mesoscale cellular convection in marine stratocumulus: 2. nondrizzling conditions. *Journal of Ad-*

vances in Modeling Earth Systems, *11*(1), 3-18. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018MS001448 doi: 10.1029/2018MS001448

Zhu, T., Lee, J., Weger, R. C., & Welch, R. M. (1992). Clustering, randomness, and regularity in cloud fields: 2. cumulus cloud fields. *J. Geophys. Res.*, *97*(D18), 20537. doi: 10.1029/92jd02022