

Supporting Information for “Investigating characteristic droplet size distributions in Large Eddy Simulations of stratocumulus clouds”

Nithin Allwayin¹ *, Daniel J. Miller^{2,3}, Kamal Kant Chandrakar⁴, Michael L. Larsen^{1,5}, Raymond A. Shaw¹

¹Department of Physics, Michigan Technological University, Houghton, MI, USA

²Climate and Radiation Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA

³GESTAR-II, University of Maryland, Baltimore County, Baltimore, MD, USA

⁴Mesoscale & Microscale Meteorology Laboratory, NSF National Center for Atmospheric Research, Boulder, CO, USA

⁵Department of Physics and Astronomy, College of Charleston, Charleston, SC, USA

Contents of this file

1. Text S1 to S4

2. Tables S1

*Current address, School of Marine and
Atmospheric Sciences, Stony Brook
University, Stony Brook, NY, USA

August 2, 2025, 2:52am

3. Figure S1

Text S1 - More details on the simulation setup

The DHARMA LES model implements bin-resolved microphysics to simulate cloud processes (Ackerman et al., 2004; Zhang et al., 2012). This case employs a domain size of $6.4 \times 6.4 \times 1.5$ km, with horizontal grid spacing of $\Delta x = \Delta y = 50$ m. The vertical grid is stretched, with a minimum spacing of 5 m at both the surface and the temperature inversion to capture small-scale turbulence at these interfaces more effectively. Additional details on the DYCOMS-II model setup and DHARMA are provided by Ackerman et al. (2009). One notable difference from these prior DHARMA simulations is that the LES output used in this study has a few more logarithmically spaced DSD bins (25 rather than 20), which helps resolve more of the DSD in transition between cloud and precipitation regimes (Miller et al., 2016).

The Cloud Model 1 (CM1) LES dynamics core (Bryan & Fritsch, 2002) with the Lagrangian super-droplet method (SDM, Shima et al., 2009) is also used in this study. The CM1-SDM modeling framework and the case setup are described in detail in Chandrakar, Grabowski, Morrison, and Bryan (2021) and Chandrakar, Morrison, and Witte (2022). This model is also configured for the DYCOMS-II setup and follows the drizzling stratocumulus intercomparison case discussed in Ackerman et al. (2009). Though it is simplified further by setting the initial winds and large-scale shear to zero. A subsidence profile is applied based on uniform large-scale horizontal divergence, and constant surface fluxes are specified and held fixed in time and space. A bulk longwave radiation scheme similar to Stevens et al. (2005) is also applied. The initial aerosol conditions follow a bi-modal

lognormal distribution consisting of ammonium sulfate, similar to Ackerman et al. (2009) except the concentration is reduced by a factor of two. A horizontal domain size of 50 km $\times$ 25 km and 1.5 km vertical domain thickness with uniform grid spacings of 100 m (horizontal) and 5 m (vertical) are used. Fine vertical grid spacing is important for representing cloud top entrainment appropriately (Mellado et al., 2018). This coarser horizontal grid spacing allows a sufficiently larger domain to capture mesoscale organization while still resolving large boundary-layer eddies of several hundred meters. A prognostic turbulent kinetic energy scheme (Deardorff, 1980) is used for subgrid-scale mixing of Eulerian prognostic model variables, and the scheme of Weil, Sullivan, and Moeng (2004) is used for subgrid-scale transport of Lagrangian particles. The Lagrangian super-particles were initialized as aerosols with 128 initial super-particles per grid box. The governing equations for Eulerian variables and Lagrangian particles are integrated with time steps of 0.6 s and 0.15 s, respectively.

The Lagrangian super-droplet scheme is uniquely suitable for investigating microphysical variability because of its detailed treatment of droplet activation, droplet size distribution evolution through condensation and coalescence, and aerosol processing. It also eliminates the problem of numerical diffusion present in Eulerian schemes that affects drop size distributions and, ultimately, precipitation development (Morrison et al., 2018; Chandrakar, Morrison, Grabowski, & Bryan, 2022).

Text S2 - Comparison to field observations

The characteristic length scales for each distribution are calculated from both the LES simulations and the HOLODEC data (specifically, the 60 km segment from RF18 described in Allwayin, Larsen, Glienke, and Shaw (2024)). The HOLODEC data is one-dimensional in space, and we determine the lengths of contiguous segments corresponding to each characteristic distribution. The maximum segment length is taken as the characteristic length scale. To reduce noise, a gap-filling threshold of 4 points is applied. The LESs, on the other hand, are two-dimensional in space, and so we use the equivalent-circle diameters for each connected component in a 2D spatial map, applying the same gap-filling threshold. The resulting characteristic distributions are then sorted by length, and only those with lengths exceeding 500 meters are included for further analysis. The mode diameters and standard deviations of droplet sizes for these selected distributions are then estimated and presented in Table 1 of the main text.

Text S3 - Single p-value Versus ensemble KS testing

In the current implementation of the algorithm, we use the p-value from comparing two distributions using a KS test as the similarity score. Previously, for the HOLODEC data from ACE-ENA, we've used a procedure of repeatedly subsampling both distributions to a constant sample size, and computing the average pass/fail score of the KS tests as the similarity score. In these ensembles of tests, a constant p-value of 0.05 is used. Neither metric is perfect, but in our analysis, we find that the average KS pass/fail rate (your "avg-KS score") is more robust than using a single p-value. The p-value is uniformly distributed between 0 and 1 and is sensitive to the sample size. And so, even for underlying similar distributions, small differences/randomness can cause the p-value to fluctuate. In contrast, the ensemble KS testing uses a fixed sample size, and averages across large independent KS tests, helping to smooth out randomness, resulting in a more stable score. This result however is tied to the chosen sample size and the threshold p-value for the test. The ensemble tests are also much more computationally expensive. For the LES, the sample size that corresponds to the CDF height in the KS test is effectively determined by the number of diameter bins in the DSD and is a constant. The random fluctuations in the DSDs are also minimal. Hence, in this study, we use the single p-value method for computational efficiency; however, in general, the ensemble of sub-sampled KS comparisons is recommended for future applications of the algorithm when computationally feasible.

Text S4- Entrainment rate calculation for the clusters

A. Clustered Characteristics:

Entrainment rates (velocities) in this study are based on “Jump-Flux” calculations that compare mean properties of vertical wind speeds and humidity within the entrainment interfacial layer (EIL) to local deviations within said layer.

$$W_e = -\frac{\overline{w'q'_v}}{\Delta q_v} \quad (1)$$

This was performed on a column-by-column basis, which means that for a given layer of the LES, each pixel has an associated entrainment rate (as shown in our Figure 4b). When subsequently combined with the identified clusters, we obtain a population of entrainment rates associated with each cluster. From there, we experimented with different statistical representations of the entrainment rate associated with each clustered population. Many clusters were found to have relatively few members, so we decided to shy away from displaying the mean entrainment rate (though results are similar for most clusters) in favor of the more statistically robust characterization of the mode of a distribution with relatively few members – the median. The median entrainment rate is used to better capture the mode of each cluster without relatively few potential outliers driving mean statistics. This, combined with sorting the cluster label index by median entrainment rate, allows us to visually link the spatial variability in clusters to the spatial variability in entrainment rate.

B. Calculation of Entrainment Rate:

Firstly, to define entrainment rates, we need to define both the cloud top and the location of the EIL. To define the lower boundary of the EIL, we simply prescribe it based on the

location of the peak liquid water content in each given LES column – meaning that EIL top and base vary across the domain of the simulation. Within the EIL of each column, we then calculate the following deviations,

$$w'(x, y, z) = w(x, y, z) - \langle w(x, y, z) \rangle |_z, \quad (2)$$

$$q'_v(x, y, z) = q_v(x, y, z) - \langle q_v(x, y, z) \rangle |_z. \quad (3)$$

The first term in each equation represents an array of values (within the EIL), and the second term is a scalar where the angled z-brackets indicate that it is an average value (within the EIL). Then the product of these vector arrays of deviations is again averaged (overbar) to complete the numerator of the equation. Finally, the denominator expresses the jump-change in water vapor across the cloud top boundary compared to the local water vapor at the specific location and layer in question,

$$\Delta q_v = q_v(x, y, z_{\text{top}+1}) - q_v(x, y, z_i). \quad (4)$$

Where z_i indicates the layer being evaluated and $z_{\text{top}+1}$ is intended to indicate the LES layer just above cloud top.

References

- Ackerman, A. S., Kirkpatrick, M. P., Stevens, D. E., & Toon, O. B. (2004). The impact of humidity above stratiform clouds on indirect aerosol climate forcing. *Nature*, *432*(7020), 1014–1017. Retrieved from <https://doi.org/10.1038/nature03174> doi: 10.1038/nature03174
- Ackerman, A. S., vanZanten, M. C., Stevens, B., Savic-Jovicic, V., Bretherton, C. S., Chlond, A., ... Zulauf, M. (2009). Large-eddy simulations of a drizzling, stratocumulus-topped marine boundary layer. *Monthly Weather Review*, *137*(3), 1083 - 1110. Retrieved from <https://journals.ametsoc.org/view/journals/mwre/137/3/2008mwr2582.1.xml> doi: 10.1175/2008MWR2582.1
- Allwayin, N., Larsen, M. L., Glienke, S., & Shaw, R. A. (2024). Locally narrow droplet size distributions are ubiquitous in stratocumulus clouds. *Science*, *384*(6695), 528-532. Retrieved from <https://www.science.org/doi/abs/10.1126/science.adi5550> doi: 10.1126/science.adi5550
- Bryan, G. H., & Fritsch, J. M. (2002). A benchmark simulation for moist nonhydrostatic numerical models. *Monthly Weather Review*, *130*(12), 2917–2928.
- Chandrakar, K. K., Grabowski, W. W., Morrison, H., & Bryan, G. H. (2021). Impact of entrainment mixing and turbulent fluctuations on droplet size distributions in a cumulus cloud: An investigation using lagrangian microphysics with a subgrid-scale model. *Journal of the Atmospheric Sciences*, *78*(9), 2983–3005.
- Chandrakar, K. K., Morrison, H., Grabowski, W. W., & Bryan, G. H. (2022). Comparison of lagrangian super-droplet and eulerian double-moment spectral microphysics

schemes in large-eddy simulations of an isolated cumulus-congestus cloud. *Journal of the Atmospheric Sciences*.

Chandrakar, K. K., Morrison, H., & Witte, M. (2022). Evolution of droplet size distributions during the transition of an ultraclean stratocumulus cloud system to open cell structure: An les investigation using lagrangian microphysics. *Geophysical Research Letters*, 49(17), e2022GL100511. Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2022GL100511> (e2022GL100511 2022GL100511) doi: <https://doi.org/10.1029/2022GL100511>

Deardorff, J. W. (1980). Stratocumulus-capped mixed layers derived from a three-dimensional model. *Boundary-layer meteorology*, 18(4), 495–527.

Mellado, J.-P., Bretherton, C., Stevens, B., & Wyant, M. (2018). DNS and LES for simulating stratocumulus: Better together. *Journal of Advances in Modeling Earth Systems*, 10(7), 1421–1438.

Miller, D. J., Zhang, Z., Ackerman, A. S., Platnick, S., & Baum, B. A. (2016). The impact of cloud vertical profile on liquid water path retrieval based on the bispectral method: A theoretical study based on large-eddy simulations of shallow marine boundary layer clouds. *Journal of Geophysical Research: Atmospheres*, 121(8), 4122–4141. Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015JD024322> doi: <https://doi.org/10.1002/2015JD024322>

Morrison, H., Witte, M., Bryan, G. H., Harrington, J. Y., & Lebo, Z. J. (2018). Broadening of modeled cloud droplet spectra using bin microphysics in an eulerian spatial domain. *Journal of the Atmospheric Sciences*, 75(11), 4005–4030.

- Shima, S.-i., Kusano, K., Kawano, A., Sugiyama, T., & Kawahara, S. (2009). The super-droplet method for the numerical simulation of clouds and precipitation: A particle-based and probabilistic microphysics model coupled with a non-hydrostatic model. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 135(642), 1307–1320.
- Stevens, B., Moeng, C.-H., Ackerman, A. S., Bretherton, C. S., Chlond, A., de Roode, S., ... others (2005). Evaluation of large-eddy simulations via observations of nocturnal marine stratocumulus. *Monthly weather review*, 133(6), 1443–1462.
- Weil, J. C., Sullivan, P. P., & Moeng, C.-H. (2004). The use of large-eddy simulations in lagrangian particle dispersion models. *Journal of the atmospheric sciences*, 61(23), 2877–2887.
- Zhang, Z., Ackerman, A. S., Feingold, G., Platnick, S., Pincus, R., & Xue, H. (2012). Effects of cloud horizontal inhomogeneity and drizzle on remote sensing of cloud droplet effective radius: Case studies based on large-eddy simulations. *Journal of Geophysical Research: Atmospheres*, 117(D19). Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012JD017655> doi: <https://doi.org/10.1029/2012JD017655>

Table S1. Details of the horizontal layers from the LES simulations used in the analysis.

Multiple layers, corresponding to different altitudes, are analyzed for each time step.

LES Model	Simulation Time(s)	Vertical Level	Altitude (m)
DHARMA LES- bin-resolved microphysics	2304	33, 41, 50, 54	546, 695, 781, 802
	4637	33, 41, 50, 54	546, 695, 781, 802
	8036	33, 41, 50, 54	546, 695, 781, 802
Cloud Model 1- Lagrangian super- droplet method	11001	150	747.5
	15802	135, 150, 157	672.5, 747.5, 782.5

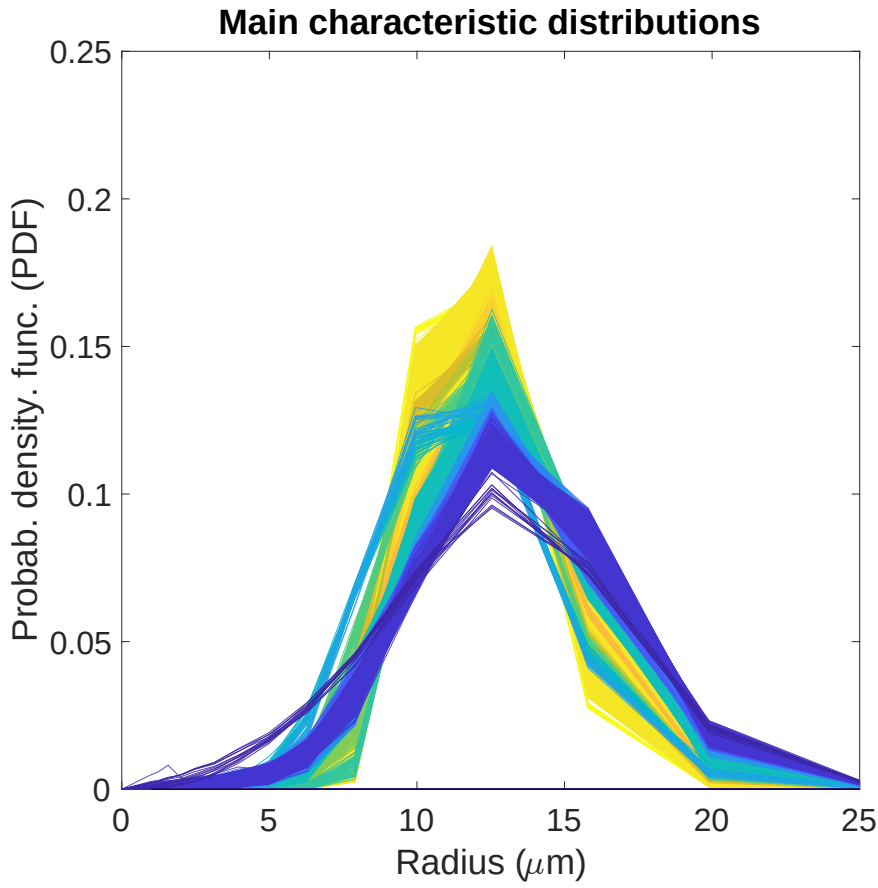


Figure S1. The primary characteristic distributions with length scales exceeding 500 m, as presented in Table 1 of the main text for the LES bin microphysics case. The distributions broaden as the mean diameter increases.