Review of: “Droplet clustering and drop size correlations from holographic imagery suggest cloud droplet spectral broadening via entrainment-mixing”
Submitted for Consideration of Publication in Atmospheric Chemistry and Physics
John J. D’Allesandro, Robert Wood, and Peter N. Blossey
November 2nd, 2025

Summary and Overall Evaluation

This manuscript extends the existing literature that investigates sub-cm scale droplet clustering using data acquired by holographic imaging devices. The authors’ primary conclusion is that within a few cloud trensects during the CSET campaign, HOLODEC data reveals a statistical tendency for large (> 25 μm) cloud drops to be spatially isolated. In addition, this tendency seems to be related to other bulk variables (e.g. low local number concentrations and sub-saturated conditions) which the authors argue is more consistent with entrainment-mixing than other bottleneck growth hypotheses.

In general, I am pleased to see the HOLODEC data from the CSET campaign further investigated; this is a very rich important dataset that – at least in this reviewer’s opinion – has not been sufficiently discussed in the literature to date. The advantages of using instruments that detect 3-dimensional particle positions are clear and papers that extend existing techniques to more effectively leverage this information are important and needed.

However, in this reviewer’s opinion, there are significant improvements that are needed to make this work publishable in ACP.

Comment on Previous Review

Although I read this paper several times prior to writing this review, immediately preceding writing this document I also glanced at the public review supplied by Anonymous Referee #1 (posted on October 8th). Since the task of responding to critical reviewer comments often involves trying to accommodate differing (and sometimes orthogonal) opinions, let me state from the outset that I am in agreement with all elements of the general assessment and major comments (and most of the minor comments) provided by this first Anonymous Referee. I believe addressing the concerns raised by that referee are necessary for publication of this manuscript and my additional comments below are meant to augment and extend those criticisms.

Additional Major Comments

1. The authors make the claim that the approach introduced here is novel and, though this particular implementation is new, the approach builds on previously published approaches that are not yet cited. (For example, the idea of building out statistics based on separation distance from an existing droplet has striking

similarities to the conceptual starting point of the drop-size-dependent droplet clustering work using fractal and multi-fractal paradigms (see, e.g., [3, 8]). Further, the idea of using the local environmental conditions to characterize conditions under which the bottleneck problem may be better addressed was looked at in a way with strong connections to this approach in [13]. Finally, the idea of looking at averaged-local environments and combining a variety of scales rather than a scale-localized measure like $g(r)$ has been explored repeatedly including through the use of the so-called volume-averaged pair-correlation function, Fishing statistic, Clustering index, Scaled Clustering Index, and through other common second-order statistics through the correlation-fluctuation theorem (see, e.g., [1, 2, 4, 5, 7, 11, 12]).

There are already several confusions associated with these tools and how they appropriately combine scaling information prevalent, but prevailing analysis suggests that aggregate statistics corresponding to a bulk statistic measured on scale R results from an unequally-weighted contribution from all scales less than or equal to R; the nature of this scaling, however, depends on the number of dimensions related to the observation, with different weighting factors existing in 1d, 2d, and 3d systems (see, e.g., [6]). Because of this, I am glad to see that the authors do not attempt to claim their approach identifies the key spatial scales of droplet clustering, but rather just detects whether individual drops are statistically isolated.

It is true that the majority of the above materials refer to analysis of optical-array-probe or similar 1-dimensional data, and thus the previous approaches have to be adapted to the three-dimensional domain – which is one of the contributions fo this manuscript. That being said, the approach outlined here very much is related to these previous works. To avoid confusion and to provide a bridge to previous work, I would encourage the authors to go back to those manuscripts to see if the language, notation, and approach can be adapted to align with already-existing frameworks.

Indeed, the authors' stated goal of "....determin[ing] the likelihood that drops of a given size will have a significantly high number of drops surrounding them as well as likelihoods that they are significantly isolated from neighboring drops" (lines 51-53) is *very* similar to the spirit of what the clustering index, Fishing statistic, volume-averaged pair-correlation function, etc. intend to do – determine if there is net clustering associated at or below some spatial scale. As the other reviewer suggested, the degree to which this approach is fundamentally different than a size-constrained version of previously-existing measures derived from and/or related to the radial distribution function is not clear. Although the authors of this manuscript suggest that the radial distribution function characterizes volumes, it calculates derived statistical quantities based on counting particle pairs separated by specified distances and thus every component of the sum in equation (1) ultimately is taken over an individual particle location. The stated different goal of doing this centered on a "per particle" basis vs. "per volume" basis claimed/implied by the authors may be a distinction without a true difference – especially as the metric utilized is ultimately found by adapting a single term of the sum utilized in calculating the RDF.

2. This is, in spirit, a continuation of the above point but it also bears mentioning that the work presented here also has links to existing work on nearest-neighbor and/or Voronoi analysis methods – especially in cases where the closer radial annuli have greater importance in the "local concentration/clustering" metric (as seems to be the case here if I am reading sections 2 and 3 correctly). This also links back to a different way of conceiving the radial distribution function as the sum of all of the kth nearest neighbor distributions (see, e.g, [9, 10]).

3. Despite rereading this paper multiple times (and being quite familiar with the general subject of droplet clustering statistics), all elements of figures 4, 5, and 6 are confusing to me. It seems like these figures provide a significant and necessary part of the central argument – but I cannot determine whether I am interpreting these results correctly. Figures 7 and 8 are slightly easier for me to comprehend, though I am still a bit stymied

by how to interpret the uncertainty envelopes. The authors *do* attempt to describe all of these figures in detail, but I am unable to follow the explanations despite multiple attempts. I strongly suggest reconsidering whether there is an alternative way to visualize and/or explain these results that is more intuitive to a fresh reader.

4. To echo and extend a comment made by the first anonymous reviewer, the information and extensive set of logic and parameters presented in table 2 and table 3 are dizzying to a reader not yet fully comfortable with this analysis technique. The authors specifically talk about the desire to avoid analysis techniques that invite "p-hacking". The introduction of *so many* parameters associated with the technique raises the question as to whether or not the conclusions or quantitative results may be sensitive to the choice of these (what seem to be) arbitrary parameters associated with the method. For example – why 7 different radii? Why these cutoffs for N for ψ and shell size (table 2), why these cutoffs for HILD determination (table 3), why the batch sizes used (and would this matter if the instrument cadence were different like new developing holographic systems)? What here might depend on the number density of observations? What would change for clouds that are not stratocumulus or for larger fields-of-view or for two-dimensional PIV images? The claim is explicitly *not* that the authors chose parameters explicitly to get a significant result, but rather the (arguably equally important) point that to apply this technique more broadly to other campaigns, instruments, systems, or environments the authors provide little guidance to the reader how to adapt their methodology more broadly.

   To be explicit, my concern here is twofold: (i) the methods, though explained in detail, are confusing to the reader. Hopefully the general approach could be explained in a simpler manner and the finer details could be moved to an appendix or supplement?, and (ii) as a reviewer who is tasked with validating the soundness of the approach as well as the veracity of the results, there are just too many unjustified and (as far as I can tell) unphysically motivated parameters associated with the method to be sure that the conclusions are not a spurious consequence of the specific parameters used/chosen.

5. One issue that may be of foundational concern – are we certain that the HOLODEC retrieval accuracy for droplets near an existing large droplet is equally effective as it would be in that region of the detector if the large drop were not present? HOLODEC reconstructions are built upon an algorithm that very easily could have lower effectiveness for other droplet detections in the vicinity of a large drop, which could mean that the entirety of this analysis could be the result of a spurious instrumental anomaly. This concern isn't merely academic – although I am not an expert on the details of in-line hologrphic reconstruction, it seems entirely reasonable that – like in other optical reconstructions – a small signal near a large signal could be missed or falsely attributed to a fluctuation related to the large signal. Before making claims about the fundamental microphysics and precipitation initiation mechanisms, it seems quite important to ensure that the instrument has a quantum efficiency of small drop detection that does not depend on the presence or absence of a large drop nearby.

   Given that it seems the statistical evidence is subtle, it bears pointing out that even if the sensitivity of small particle detection being impaired by the presence of a large particle nearby only affected part of the field of view that was retained for use in this study, that could very easily induce the entire statistical signal studied here.

   Unfortunately, I don't believe the Monte Carlo mechanisms in place throughout this paper are designed in such a way as to verify or reject this concern. At the *VERY* least, in line 358, the authors should modify the text to read "...drops within the bottleneck size range are most likely to be measured to be significantly isolated from surrounding drops". However, I do not think that merely adding this language is sufficient to ignore this overarching concern. A skeptical reader may be unconvinced by the entire message of the paper unless evidence that small particle detections near large particles are not inhibited in the holographic reconstruction.

Additional Minor Comments

1. Although this is a minor comment, I believe it is an important one. Much of the work over the last 20+ years on cloud particle clustering has been obfuscated because of the lack of a consensus and author-to-author consistency regarding terminology. Tying what has been done to previous work is particularly important (which is the foundation underlying my first two major comments), but taking care to do so in a way that doesn't further confuse things is also important. As such, I *implore* the authors to choose a different notational convention than to use the idea of "g_{drop}" (e.g. see equation (3)). Multiple papers already have been written about the conditions as to when $g(r)$ can and cannot be calculated and still contain the physical meaning ascribed to it, and many of those conditions are not easily mappable to a particle-by-particle approach as presented here. (For example, although $g(r)$ can always be calculated from any existing data-set, the underlying assumption of spatial homogeneity requires $g(r \to \infty) \to 1$ and statistical viability requires the denominator to have sufficient expectation so that shot-noise doesn't dominate the signal. Some of these concerns may not persist for individual particle detections, but assuredly if the whole point is to determine whether a particle is statistically isolated the *premise* of the approach is that there is spatial heterogeneity on the scales of investigation. Although it may seem harmless to adopt the use of the statistical metric outside its formal range of validity, we have previously seen (see, e.g, [1] and the corresponding counter-examples presented in [7] that applying radial distribution functions outside of ther range of validity for homogeneous systems results in misunderstandings.)

   The use of $g(r)$ on a particle-by-particle metric invites further confusion, especially if the approach introduced here gains wider traction.

2. I would find it valuable if the total number of holograms were presented for each of the flight legs in table 1 as well, so the reader can get a sense of the accepted/utilized fraction of detected holograms in each of the flight legs.

3. The paper cited on line 109 by Thiede et al. does not use HOLODEC data. (The system used in that paper has some similarities to HOLODEC, but is mounted on an aerostat, takes observational data at a much higher cadence, and is processed slightly differently).

4. Given the fairly severe aspect ratio utilized (see line 119), there are questions as to the true three-dimensionality of the domain explored here. It is notable that this has a more extreme aspect ratio than other studies that have used the HOLODEC data, which is one of the reasons that the guard-area (or, as stated here, guardrail) approach has been largely abandoned in some other work. It is not clear how that may impact the final results.

5. Line 143 – I don't believe the authors meant to say ith particles here.

6. Line 165 – "counting statistic error" – I'm not sure that "error" is the appropriate term in this context. "Uncertainty" or "bias" or "fluctuation" may be more appropriate here.

7. Line 179 – "maximum" probably should be "minimum" here.

8. Lines 208-210 – see major comment 2 related to similarities/links to Voronoi/nearest-neighbor techniques.

9. Line 222 – In what way is 0.15 cm "ideal"? Would that parameter extend to other cloud conditions/instruments/etc? This links back to major comment 4.

10. Line 243 – By choosing the same number of high and low DCFs to control for environmental conditions also seems like it introduces another scale parameter for the analysis, especially if this method were to be applied to a more heterogeneous domain in regards to number concentration or other parameters.

11. Line 246 – "randomly" – I'm not sure what is meant by this in this context.

12. Lines 470-473. The fact that holograms with $D_{\mathrm{max}} > 45\ \mu$m do not exhibit this trend is very much concerning and, to me, seems to run counter to the authors' explanation that this signal is likely induced by entrainment-mixing.

13. Lines 525-527. The fact that holograms with larger drops exhibit broader size distributions and fluctuations of other microphysical variables is more or less tautologically true, isn't it? (Especially if the instrument can only see droplets above a certain size reliably.)

14. Line 550 – at the end of the sentence including numbers, it seems to be important to specify that these are relative to the drop's *local environment* and not in absolute/cloud-averaged terms. This local environment is tied to a specific measurement scale, and that is another spatial scale of relevance when exploring this question.

15. Line 554 – "significantly isolated"; this raises the question as to whether being statistically significant is sufficient to affect localized dynamics/microphysical responses.

16. Line 570 – I believe it is important at this stage to indicate the scale (sub-cm scale) where this conclusion is likely valid.

References

[1] B. Baker and R. Lawson. Analysis of tools used to quantify droplet clustering in clouds. *Journal of the Atmospheric Sciences*, 67:3355–3367, 2010.

[2] B.A. Baker. Turbulent entrainment and mixing in clouds: A new observational approach. *Journal of the Atmospheric Sciences*, 49:387–404, 1992.

[3] Y. Knyazikhin, A. Marshak, M.L. Larsen, W.J. Wiscombe, J.V. Martonchik, and R.B. Myneni. Small-scale drop size variability: Impact on estimation of cloud optical properties. *Journal of the Atmospheric Sciences*, 62:2555–2567, 2005.

[4] A.B. Kostinski and A.R. Jameson. On the spatial distribution of cloud particles. *Journal of the Atmospheric Sciences*, 57:901–915, 2000.

[5] L. Landau and E.M. Lifshitz. *Statistical Physics*. Butterworth Heinemann, Oxford, UK, 1980.

[6] M.L. Larsen. *Studies of discrete fluctuations in atmospheric phenomena*. PhD thesis, Michigan Technological University, 2006.

[7] M.L. Larsen. Scale localization of cloud particle clustering statistics. *Journal of the Atmospheric Sciences*, 69:3277–3289, 2012.

[8] A. Marshak, Y. Knyazikhin, M.L. Larsen, and W.J.Wiscombe. Small-scale drop size variability: Empirical models for drop-size-dependent clustering in clouds. *Journal of the Atmospheric Sciences*, 62:551–558, 2005.

[9] B. Picinbono and C. Bendjaballah. Characterization of nonclassical optical fields by photodetection statistics. *Physical Review A*, A71:013812, 2005.

[10] B. Picinbono and C. Bendjaballah. Poisson processes with integrable density. *IEEE Transactions on Information Theory*, 52(12):5606–5613, 2006.

[11] R.A. Shaw. Particle-turbulence interactions in atmospheric clouds. *Annual Review of Fluid Mechanics*, 35:183–227, 2003.

[12] R.A. Shaw, A.B. Kostinski, and M.L. Larsen. Towards quantifying droplet clustering in clouds. *Quarterly Journal of the Royal Meteorological Society*, 128:1043–1057, 2002.

[13] J.D. Small and P.Y. Chuang. New observations of precipitation initiation in warm cumulus clouds. *Journal of the Atmospheric Sciences*, 65:2972–2982, 2008.