Referee Report

General assessment

This paper analyzes a valuable dataset of airborne holographic measurements by applying a droplet cluster field (dCF) method to study small-scale clustering. The approach is clearly explained and the topic is important for understanding cloud microphysics. The authors results based on their analyses are that drops having sizes within the bottleneck size range (D of 25-50 μm) are most likely to be significantly isolated from the nearest drops. Additional analyses were performed to determine under which conditions holograms are associated with this "isolated large drop" trend. They found that holograms in subsaturated conditions and having low drop concentrations are associated with this trend.

My understanding is that the main goal of the study is to detect *localized clustering* that may not be detectable using traditional system-wide statistics such as the spatial pair correlation function (or radial distribution function, $g(r)$). In a finite sample from a Poisson process (random droplet locations), one often observes patches that appear denser and others that appear sparser, simply due to statistical fluctuations. Locally, such fluctuations can look like clustering even though, globally, $g(r)$ averages to 1.

The authors therefore aim to identify clustering that varies from place to place within the sample, possibly depending on droplet size. For example, size-dependent clustering would not be captured by a global $g(r)$ unless the analysis were stratified by droplet size. My understanding is that the authors' method is essentially a localized version of $g(r)$: a droplet-centric or neighborhood-based clustering measure, rather than a single system-wide statistic.

However, it is not entirely clear why the authors did not simply compute system-wide $g(r)$ for subsets of droplets grouped by size. Such a modification of the traditional RDF approach could also reveal size-dependent clustering. Clarifying how the dCF method differs from (and improves upon) this more conventional strategy would help the reader understand the rationale for introducing a new diagnostic.

Finally, to this reader, it seems challenging for the authors' local metric to provide a robust signal based on only a few neighboring droplets around each primary droplet. A brief discussion of how statistical noise is controlled or mitigated in such local estimates would strengthen the paper.

Major comments

1. For a non-statistician such as this reviewer, sections 3 and 4 of this manuscript are mostly incomprehensible. Table 2, although presumably intended to clearly describe the classification of droplets into high or low DCFs, makes this reviewer uneasy due to what appears to be ad hoc additional classifications for high and low DCFs. This

reviewer would appreciate a demonstration of the method applied to an artificial dataset for which the clustering characteristics are specified. My concern is that achieving statistical significance based on local clustering that involves a very small number of particles could be difficult. Is the dataset large enough to overcome this difficulty? How can you demonstrate that?

Minor comments

1. lines 9-15. This could be misleading because it sounds like clustering on mm scales leads to enhanced collision rates and therefore the bottleneck solved. Therefore, you introduce a particle-by-particle neighborhood-counting method to evaluate these clustering trends. But this is not actually the same thing because the RDF at contact is the spatial factor relevant to collision rates. The neighborhood-counting method gives a different kind of clustering information (local heterogeneity), not the RDF itself. This could be revised as follows (for example): The question of how droplets rapidly grow large enough to initiate collision-coalescence has persisted for decades. Theories suggest that enhanced droplet clustering on millimeter scales can increase collision rates for droplets in the 'bottleneck' size range ($\sim$ 25-50 ?m). To investigate spatial organization in more detail, we introduce a novel droplet-centric diagnostic that evaluates local neighbor counts and proximities (i.e., clustering fields) on a particle-by-particle basis. Although not equivalent to the RDF, this approach provides complementary information about the heterogeneity of droplet environments.

2. lines 73-77: It's worth noting the confidence range for RH, since it's very sensitive and critical for distinguishing between saturated and unsaturated air. My own reading and for Rosemount probes after careful calibration is $\sim 0.2 - 0.3$ K. At best these uncertainties give $\pm 3 - 4$ % uncertainty.

3. section 2.2.2: The choice of 7 shells should be defended–why 7? Is it based on prior work, a sensitivity check, or computational limits?

4. section 2.2.2: Clarify whether edge corrections are applied for each shell separately or in a single step.

5. section 3.1: I believe Section 3.1 is analogous to analyzing spatial droplet clustering, since you're removing the dependency on droplet size. This may be worth mentioning.

6. Make it explicit in figure captions whether counts are cumulative or differential.

7. line 179: Rephrase "the maximum distance at which two drops can neighbor" since this is in fact a minimum spacing.

8. Improve figure captions so that readers can directly link each plot to the formal dCF definition.

9. lines 198-200: A sentence is duplicated here.

10. Figure 3B: It is unclear what is plotted. The caption states that the y-axis is "The number of high DCFs, low DCFs and DCFs meeting neither category." If this is the case, then the y-axis must be a number per some range of N, but this range or bin size is not stated. However, the units shown for the y-axis, as well as the text, suggest it is a number concentration. Please clarify. Furthermore, the category 'Drops used in analysis' should be explained because it does not make sense.

11. Figure 3B: Each of the upsloping patterns of 'High dCF' red points begins at one of the values for minimum ψ for high DCFs: 110, 175, 230, and 280. As a result, the patterns are dependent on the definitions listed in Table 2. What is the basis of these categories?

12. lines 285-6: The expression "meaning the DCFs will be identical for each hologram" should be clarified. I think it means that DCFs will be unchanged within a hologram, and also by extension within the entire set of holograms, because DCFs depend only on droplet spatial coordinates.

13. lines 287-8: "DCFs of the actual drop sizes" is confusing because DCFs have nothing to do with drop sizes. Please clarify.

14. line 294: Not all of us are statisticians, so please define Type I error.

15. lines 299-300: "DCF-drop size relationships are then determined for this set of holograms using the Monte Carlo-DCF methodology from Section 3.1." I don't see the need for the Monte Carlo aspect in order to determine DCF-drop size relationships.

16. line 380: "Figure 5A,C shows results separated into relatively large regions ($\sim$100 m) of subsaturated and supersaturated conditions." Please remind the reader (or explain if it was not explained already) how you identified these large regions from holograms that are only a few cm in extent. Even if you used sequences of holograms, their properties cannot be assumed to be continuous between them.

17. Figure 4: I am sorry but this figure is incomprehensible to me. The text that describes it does not help me (lines 333-343). A problem I have is evaluating the role of randomness/noise in these plots. How do we know if any of what is shown has statistical significance?

18. Figures 4, 5, 6: I suggest connecting the dots in the percentile plots to make the patterns easier to see. For me, it helped a lot to do that.

19. lines 524-7: Clarify: Do you mean "preferentially completely evaporating" not "preferentially evaporating"? It is difficult to understand your next sentence otherwise: "However, holograms experiencing this trend are also associated with broader drop size distributions (Fig. 7E) and larger drops than holograms not exhibiting this

trend (Fig. 7F)" which is the result to be expected if the most of the smaller droplets are NOT completely evaporated. Tolle and Krueger (2014, Figure 17) found that broadening by partial evaporation is common.

20. lines 524-7: It seems to this reviewer that coalescence growth could also explain the "isolated large drop trend." Especially since the "this trend is associated with portions of the cloud where precipitation/condensate reaches the lowest altitudes from the respective cloud." This may be worth mentioning if you agree.