Review of 'Revisiting particle sizing using grayscale optical array probes: evaluation using laboratory experiments and synthetic data'
O'Shea et al

The authors present a study that demonstrates that by using grayscale information available on some OAPs, uncertainties can be reduced for both sizing and counting particles that pass through the sample volume of a probe, outside of the DOF. In the conclusions, the authors present specific recommendations to users and manufacturers of this class of probes to improve the integrity of cloud particle measurements.

This study addresses an important topic that continues to be a large source of uncertainty in aircraft-based cloud measurements of hydrometeors, particularly for particles ranging in diameter from a few 10's to ~150 um. The study includes both a laboratory and modeling component, under which the conditions can be strictly controlled. They develop a methodology for determining whether to include or exclude particles in counting statistics and combine that with sizing methodologies from two well-accepted algorithms. They successfully demonstrate that these 'traditional' methods can lead to an over-counting of small (and to a much lesser extent, large) particles and develop a methodology to reduce this bias. They present data from two 'real' cases as an illustration of the impact of the new algorithm. Finally, the authors provide specific recommendations for users of OAP data sets and their conclusions are well supported by the data presented in their study.

The manuscript is well written and scientifically sound. I recommend publishing after the authors address a few specific minor comments.

Minor Comments

Page 2, line 10 – in situ measurements do not 'validate' remote sensing instrumentation, rather they can be used to (1) constrain retrievals and/or (2) validate estimates that come from these retrievals (based on remotely sensed measurements).

Section 2.1 (and more generally, section 2) – In section 2.1 the authors describe some of characteristics used in a subset of the experiments, but not all of the experiments. For example, they present threshold levels for droplet generator tests and the actual field data. However, the threshold levels used in the synthetic data set are not presented until section 2.4. I found this a bit confusing as I read through the sections and found myself often returning to previous sections because of this. I think the inclusion of a table that describes some of the key aspects for all of the experiments could help alleviate some of this confusion and would provide a single point of reference. The table could include thresholds, arm spacing used, particle transit speeds, hydrometeor sizes tested/measured, etc.

Section 2.3 – include the droplet speed in this description (I assume it is ~10 m s-1, based on page 8, line 4. But state it explicitly when describing the droplet generator).

Section 2.4 (page 7), line 12 – I find it odd that the simulations covered Zd +/- 5 cm when real CIPs used in this study had maximum probe arm distance of +/-2 and +/-3.5 cm. Further the figures 2-4 (right column) only show data out to +/-4 cm. If the analysis is limited to less than 5 cm, then it should be stated here.

Section 3.1 – Perhaps I missed it, but I did not see where the authors explicitly state what the drop sizes are for the three separate print-head tests. They present D50 at Zd=0, but this is not the independent measure provided by their camera. I'm further confused that the 90 and 120 micron print-head seem to produce drops of the same diameter?

Figures 2-4 – The dashed lines in the panels appear to match up with the synthetic data for 55, 80, and 90 um tests. Why are these same dashed lines drawn on the Left-hand side columns that show results from the drop generator? Do these sizes represent the *actual* (independently-sized) droplet diameters from the drop generator tests?

The second are third paragraph in section 3.1 describe explicitly what is shown in the middle and bottom panels of Figs. 2-4, respectively. However, no such paragraph describes explicitly the top panel. Such a description, together with interpretation, would be helpful.

Fig 5 – should be larger, including fewer particles. It is impossible to see the gray levels in the current figure.

Page 21, line 13 – change to 'mis-sized'

Page 22, line 8 – I believe the authors are referring to figure 11a

Page 22, line 26 – again, are the authors referring to figure 11? Also—I do not see error bars in the figure (they are referred to in the text).

Page 25 – the second conclusion/recommendation beginning with "Fragmented images…" I think can be removed and captured within the next two conclusions/recommendations. Why? Because, how users deal with such particles depends on the whether or not the probe is gray-scale. If gray-scale, then the authors recommend using gray-level information to remove fragmented images (3rd recommendation). However, if mono-scale, there is no such methodology for removing fragmented images. Then, what do the authors recommend? Ignoring all particles with d< 100 um? This is somewhat captured in recommendation 4, but could be made more explicit.