We would like to thank Darrel Baumgardner, Jeff French and David Delene for their constructive comments about our manuscript. We now address their comments individually. For clarity referee comments are coloured red and responses are in black.

Darrel Baumgardner

This study summarizes laboratory studies that were conducted to evaluate the use of gray scale thresholding to better constrain the sample volume of OAPs and reduce the uncertainty in sizing of out of focus particles. This is an important contribution to the ongoing discussion of how to process cloud measurements made with imaging probes. It is written concisely and makes clear the steps that should be taken to process the measurements from gray scale OAPs. I did not identify any errors in either assumptions or interpretations.

I am attaching an annotated pdf of the manuscript with a few minor suggestions/ corrections.

The one, somewhat glaring omission that I think should be addressed concerns the lack of any discussion of doing a size correction using some combination of the gray level information instead of using the K07 corrections. The reason that using the three gray level values to derive the in-focus size could be beneficial is because application of K07 requires a table look-up that consumes processing time whereas a direct derivation from the gray levels might improve the processing speed and maybe accuracy. I think it is a reasonable request to explain why this was not done.

In order to correct the size of a particle its position in the image plane (Z_d) first needs to be determined. We investigated using several combinations of simple greyscale ratios to estimate Z_d, but none of these were found to have a monotonic dependence on Z_d across the whole depth of field (Figure 7). There is a general tendency of decreasing intensity as Z_d increases. However, this does not translate into continuous grayscale ratios when thresholds are applied. This paper does test the efficacy of RB98 which uses a single grayscale ratio to correct particle size, but it is found to perform worse than K07 and didn't fully remove the position dependence of the particle size.

Final point. Alexei Korolev has a new table with many more data points than used in the K07 paper. Was this used in this study? If not, the authors should.

Yes the updated table was used in the paper.

<u>Inline comments</u>

Page 2 Line 16 – An additional reference to Knollenberg (1970) has been added.

Page 3 Line 28 – This sentence has been reworded and the word "similarly" has been removed.

Page 4 Line 4 – The manufacturer default thresholds were 25, 50 and 75%. Other settings were used to examine a range of possible settings that might be chosen by the instrument operator.

Page 5 Line 4 – This has been reworded to say "The CDP sizes particles in the range 3 to 50 µm using the scattered light intensity from particles crossing a diode laser and assuming Mie scattering theory (Lance et al., 2010)."

Figure 10a – An additional panel has been added showing this subset of the PSD on a linear scale.

Page 24 Line 24 – Changed as suggested.

Jeff French

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).

This has been re worded to say: "In situ measurements are an important means to constrain remote sensing retrievals"

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.

A table has been added to the manuscript with the suggested information.

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).

The droplet speed was not precisely measured during these experiments. Droplet speed was estimated by varying the CIP-15 data acquisition rate in order to obtain circular images – discrepancies between the particle speed and acquisition rate manifest as elliptical images. Droplet velocity was typically of the order of 1 m/s +-0.5 m/s, and is in agreement with particle velocities determined in separate experiments using a high speed camera. This has been noted in the manuscript and is included in table 1.

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.

As stated in section 2.3, the droplet generator was mounted on a 50 mm translating state. This means that for our experiments, we only had a total range of 50 mm in the z plane. Therefore we were not able to map out the entire sample volume. However, we were able to demonstrate the symmetrical behaviour about the focal plane, and explore the degradation of image quality at high Z_d. This was done to show that the general shape of the D-Z relationship is independent of the particle size. The plots are made more general by considering a larger range of Z, while the reader can clearly see what D would be for a specific instrument. The larger range of Z doesn't impact the subsequent analysis in the paper.

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?

It was not possible to use the high-speed camera to monitor the droplet size while simultaneously scanning the CIP sample volume. The size of the printhead only provides a course control over the droplet size, since the droplet also depends on the fluid properties and the drive waveform. The median size of the CIP-15 images at Z = 0 was 60 µm, 90 µm and 90 µm for the 60, 90 and 120 µm printheads, respectively. Based on this and also by comparing with the synthetic data we estimate the true size of the droplets to be 55, 80 and 90 µm for the 60, 90 and 120 µm printheads, respectively.

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?

As described above these are the sizes that agree best with the drop generator scans and are therefore the best estimate of the true droplet diameter during the 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.

The first paragraph of this section has been extended to address this.

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

The droplets have been made larger and a smaller number are shown.

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

Done.

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

Corrected.

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).

Yes this is a reference to figure 11 and has been updated. Due to the relatively long averaging period it is only the large sizes (>700um) where the counting uncertainty is significant. For the smaller sizes they are not visible.

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.

This bullet point makes an important point that the sample volume size dependence multiplies the impact of image fragmentation and introduces the following two bullets explaining how this issue should be addressed.

Recommendation 4 has been made more explicit and now reads:

"The data from monoscale OAPs is unreliable and should not be used below approximately 100 µm due to fragmented larger particles."

David Delene

General Comments:

Very interesting work, with a mix of lab, model and field data.

Yes, it seems having an indication of at least one particle above 67 % on monoscale probes would be useful. It not clear to me, that gray scale probe is much better than a monoscole probe with such a flag.

As is shown in the paper having a 67% greyscale threshold effectively removes fragmented spherical particles and restricts the sample volume to Zd < 4.8. The paper also examines how different greyscale thresholds that are available on an increasing number of commercially available probes can also be used to filter data.

While the paper states that monoscale OAPs, the ice particles less than 100 um are not usable. The gray scale probes, while better, also may not be that usable. The comparison between the CIP gray and the holographic imager don't agree.

This paper suggests that a combination of K07 and the presented greyscale relationships can be used to determine particle size distributions for spherical particles less than 100 um. As discussed in Sect 3.4.2 neither of these are directly applicable to non-spherical particles and are therefore subject to greater uncertainty. Further work would be needed to derive greyscale Z relationships for specific particle habits.

There are many places where comma is missing for introduction phases. This should be correct throughout the manuscript. One example is on page 22 , line 13.While the papers is typically consistent with the use of past and present tense. The paper's use of present tense for items published and hence "known", and past tense for items in the paper, make things harder to understand in my option. I like past tense for published work and present tense for what is presented in the papers. An example of switching between tense is on page 22 line 5-8, where one sentence is past, then present and then past. In the same paragraph. I would encourage the authors to consider a different tense usage to make things easier to understand of the reader.

We believe that the manuscripts use of tense is clear. The collection of the data is described in the past tense and its interpretation in the present tense. We will follow the editor's advice about whether this needs changing. The manuscript has been checked for use of commas.

Specific Comments:

Page 5 Line 20: Can it be stated what exact droplet generator was used?

Part numbers for the printheads have been added to the manuscript.

Page 6 Line 16: What is the pixel size for your magnification, or range of pixel sizes?

1.4 µm per pixel. This has been added to the manuscript.

Page 8 Line 8: How are these box and whisker plots, look like a histogram plot?

These are box whisker plots showing the 10, 25, 50, 75 and 90th diameter percentiles. Since the resolution of the probe is 15 µm some of percentiles are the same. A description of the percentiles has been added to the caption.

Page 9 Line 3 and 4: How is the droplet diameter estimated. Not just the print head size since Figure 3 and 4 and different. Figure caption should state how the estimate is done.

Please see our response to Jeff French which addresses this point.

Page 15 Line 2: What is the difference between the left and right rows? How about a label across the top?

These plots show different combinations of greyscale ratios and are labelled on the y-axis.

Page 16 Line 15: What type of particles? Assume it is liquid spherical droplets? Important to state that these are liquid droplets, right?

Yes this has been added.

Page 17 Figure 9. Can the printhead size be include as label or title in the plots themselves?

These have been added to the plot.

Page 17 Figure 9. Can the figure caption state that this is for a 15 um CIP measurement?

This has been added.

Page 18 Line 6: “Similar” to what? Figure 8 and 9? Please state.

This refers to equation 3 and has been clarified in the manuscript.

Page 21 Line 2: Figure captions should provide all details necessary to understand what is presented;

Hence, should not include “this period” but directly state the time period and date. Move details about the figures out of the text and into the figure caption. Many figure captions do not provide the necessary details.

These details have been added to the caption and also to fig 11.

Page 21 Figure 10. For this time period. Are there out of cloud measurements? At is the minimum concentration for 1 Hz measurements for the time period? The CDP spectrum does not seem smooth like I would expect, at 30 um the concentration doubles from one channel to the next. Why such a change in concentration from one channel to the next?

Page 21 Figure 10: I don’t understand having such a large range of values. Hard to see the CDP/CIP overlap, which is what is important. Can the plot not be done from 10 to 100 um diameter and on a linear scale?

A panel has been added showing a subset of the PSD on a linear scale. These measurements are in cloud, however the CDP concentration was very low (1 cm^{-3}) and a significant proportion of the droplet spectrum was larger than the CDP size range, which may explain the anomalously high concentrations in the largest CDP bin.

Page 22 Line 8: Do you mean figure 11a instead of figure 10a? Again details need to be in the figure caption, not the text.

Corrected. These details have been added.

Technical Corrections:

Abstract: Page 2 Lines 1-3: Abstract states, “We make .. and We also raise...” Can the main specific recommendation about grey scale OAP and bias in OAP without Greyscale be simply stated in the abstract?

A full list of recommendations have been made in the conclusions. We believe that this would make the abstract overly long if included here as well.

Page 6 Line 8: Need space between number and unit, dwelling for 3 s.

Done.

Page 6 Line 12: Use complete sentence, no verb.

Done.

Page 11 Figure 5. Can you put the detector intensity on the left side of these images, adjust figure caption? Also, this is for CIP with 15 um pixels, right? Figure caption should include this information.

Done.

Page 11 Line 6 The “middle palnel” of what figure?

A reference to figures 2-4 has been added.

Page 14 Line 7: Need comma, “position, we” Likewise Line 8 needs comma.

Done.

Page 16 Line 1: Be direct and state equation. “Equation 3 allows ...”

Done.

Page 19 Line 4: Another example of not having a comma for the introduction clause. Please check full manuscript for this issue. “concentration, the” comma required. Another example, page 19, line 17.

Done.

Page 20 Line 2: Space between value and unit. 14 0C. This issue occurs several places in manuscript so check all units/values.

Done.