We thank the anonymous reviewers for their useful comments and suggestions. We present the reviewers' comments in bold and our responses in standard font. For clarity, in our responses we refer to the new figure numbers, which have changed with the addition of the new Figures 1 and 2.

Given the nature of the comments from Reviewer #1 (i.e. more essay-style than point-by-point), it seems most appropriate to respond by summarising the changes we have made at the end of the comments, rather than throughout. There are also two specific points along the way, which we have responded to individually. Reviewer #2's comments are addressed point-by-point, and Reviewer #3's comments are addressed in the responses to the other two reviewers.

Anonymous Referee #1

The manuscript describes aircraft and radar observations in a line of small cumulus with bases at about +11 C and tops around -15 C. This review focuses on observations from the BAE-146 measurements in relatively new, isolated updrafts that have not been contaminated with ice from older convection.

The main point that needs to be corrected is that the manuscript consistently and persistently attributes high ice concentrations observed in the -3 to -8 C region to the Hallett-Mossop (H-M) secondary ice formation process. In several places the manuscript states that there are sufficient drops smaller than 13 microns and larger than 24 microns in the presence of graupel between -3 and -8 C, and then attributes measurements of relatively high ice concentrations to the H-M process. The manuscript goes on to make the case that the H-M process supplies ice splinters that freeze the remaining supercooled drops. However, the observations do not provide confirmatory evidence for the conclusion that the H-M process is actually operating.

The observations demonstrate the following:

1. H-M conditions are met in certain regions of clouds.
2. Relatively high ice concentrations are observed
3. Ice crystal habits are typical of growth in the temperature range from – 3 to – 8 C.
4. Large supercooled (drizzle and rain) drops appear to freeze before smaller cloud drops freeze.

From these observations the manuscript concludes that the H-M process is responsible for producing small ice splinters that then collide with and freeze the large supercooled drops. While the H-M mechanism could be responsible for the production of small ice and subsequent freezing of larger supercooled drops, it is not the only process that could be responsible. Secondary ice production in actual clouds is not well understood and not well quantified. The manuscript tends to build its case on circumstantial and sometimes facial evidence. For example, the production of small ice via drop freezing, spicule formation and ejection of small ice is another mechanism (demonstrated in the laboratory by Thomas Leisner). However, this mechanism is dismissed in the manuscript via the following discussion:

Lawson et al. (2015) stated that the presence of drops larger than 200 microns in updrafts At – 6 C was required for significant ice enhancements by droplet fragmentation.

The actual verbiage from Lawson et al. (2015) is:

"Model runs were also conducted by varying the initialized DSD and the initial ice PSD. If the DSD at -6 C does not contain drops larger than about 200 micron in diameter, the conversion to ice via drop-freezing secondary ice production and riming is greatly diminished, resulting in less ice and more supercooled water being transported higher in the cloud."

The manuscript has taken the results from a cloud model used as an aid in the interpretation of the observations and implied that the Lawson et al. (2015) stated a conclusion. This is very inappropriate and misleading.

The manuscript goes on to state:

"The mean concentrations of NRound > 200 _m in updrafts in runs 11.1, 11.4 and 13 were 0, 3.6 and 0.9 /L respectively. The average number of fragments expected from a 200 micron drop is 0.04 (Lawson et al., 2015, Fig. 12), meaning if all these drops were to freeze in the H–M zone only a minimal enhancement would be expected."

This is a misinterpretation of Fig. 12, which shows statistically the number of fragments a drop would produce based on the model results. Because this is a statistical result, the proper interpretation is that statistically, the model predicts that one in every 25 drops that are 200 microns in diameter will produce an ice particle. However, this is a cascading process, whereby this ice particle could freeze a 2avour2tre-diameter drop that could produce hundreds of ice particles, and so on, producing rapid glaciation. The statement in Lawson et al. (2015) is sharing a generalized result form the model and this not sufficient justification to eliminate one secondary ice process in 2avour of another.

It is not clear from reading Lawson et al. (2015) that this work is simply to aid the interpretation of observations- the modelling work is presented on the same footing as the observations, and was tuned to their observations. Furthermore, to our knowledge it is the only work to date to attempt to quantify the droplet freezing secondary ice production mechanism(s) in any usable fashion. Rather than being "inappropriate and misleading", the comparison was made in good faith, to give the reader an indication that the warm rain process in our study was much less active than in Lawson et al. To satisfy the reviewer's concerns, we have removed any calculations using Lawson et al, and have made the argument more qualitative, to say that based on the differences in the warm rain process, we expect drop freezing secondary ice to be less effective than in Lawson et al, but to be more quantitative you would need to use a model.

"Modelling work by Lawson et al. (2015) suggested that the concentration of drops larger than 200 µm in updrafts at -6°C was an indicator for whether significant ice enhancements by droplet fragmentation would occur. The mean concentrations of N_{Round}>200 µm in updrafts in runs 11.1, 11.4 and 13 were 0, 3.6 and 0.9 L^{-1}respectively. Based on our observations alone we are not able to quantitatively assess the potential of droplet freezing secondary ice formation to affect ice concentrations, but these numbers will serve as a reference point for microphysical simulations which are beyond the scope of this paper. Using our observations of frozen drops and a comparison between our droplet size distributions and the observations and modelling work by Lawson et al. (2015), we are able to say that secondary ice formation associated with droplet freezing is likely to have been active in our case study, but to a lesser extent than in the tropical chimney clouds described by Lawson et al. (2015). The cooler cloud base in our case study means the drops were generally smaller when reaching freezing temperatures, and therefore less efficient at secondary ice production when freezing."

This reviewer is not saying that the H-M process was not operating in the subject clouds, that the spicule/ice production process was operating, or that some other (perhaps unknown) secondary ice process was or was not operating. The point is that there is not enough evidence to come to the conclusion that H-M was responsible for the relatively high ice concentrations. The conclusions are built on a house of cards.

For example, from the manuscript: "The crystal habits and ice concentrations of hundreds per litre make it clear that these particles were generated by the H–M process, and were lifted further up in the cloud by updrafts." This statement cannot be proved. What the observations do indicate is that ice with habits characteristic of a certain temperature range were measured in concentrations of hundreds per liter and observed in colder (higher) regions of the cloud.

The manuscript needs to be significantly modified so that the H-M process is not "promoted". Instead, the observations should be presented, the interpretation that the measured ice concentration far exceeds what is expected from primary nucleation, and that a secondary ice process may be responsible for the high ice concentrations. If desired, the manuscript could then list some candidate processes (e.g., H-M, spicule formation, drop fragmentation, crystal-crystal shattering, etc.), and also state that the basic H-M conditions are met in and near the regions where the high ice is observed. But there is not enough evidence to conclude that the H-M process is responsible. It is also inaccurate to state that the spicule ice formation process is not responsible for the high ice concentrations, because not enough is known about that process. Observational scientific papers should present the data and offer interpretations, not piece together circumstantial evidence in an effort to come up with an explanation.

There are also several minor issues that need to be addressed in the paper. For example, the Abstract overstates the certainty of existence of the H-M process:

"It is therefore clear that the freezing of supercooled drizzle drops not only provides a pathway to advance the onset of the H–M process, it also accelerates glaciation and the formation of precipitation once it has begun."

The observational evidence indicates that the measurements of ice concentration in the – 3 to – 8 C temperature range (i.e. the H-M region) greatly exceed those expected from primary nucleation. However, the manuscript states that freezing of supercooled drizzle provides a pathway to advance the onset of the H-M process, which is an implicit confirmation that the H-M process is responsible for the increased ice concentration. There is no direct evidence that the H-M process is actually operating, only circumstantial evidence, so these types of implicit confirmations of the H-M process need to be eliminated.

Another example: The manuscript cites the Harris-Hobbs and Cooper (1987) technique for estimating the number of ice particles produced by the H-M process. However, this approach,

while appropriate in its day, uses 30-year old technology, including optical probes without anti-shattering tips, and the results are not applicable today.

Although the reviewer is correct to say measurement technology has improved in the last 30 years, the approach of Harris-Hobbs and Cooper (1987) is a calculation using the ice size distribution. The calculation itself (i.e. the equation) is just as valid as it was 30 years ago. We have removed any comparison of the numbers from the older papers without anti-shatter tips. The technique is now used solely on our data, which is self-consistent.

This reviewer cannot recommend publication until the major issue, overly promoting the role of the H-M process, which is described in detail in this review, is rectified. The paper needs to focus on the observations, not the H-M process, which is mentioned 57 times in the manuscript. This reviewer recommends that the manuscript be revised to:

1) compare the observations to ice concentrations expected from primary nucleation,

2) point out that the ice concentration measurements far exceed those expected from primary nucleation,

3) show how the role of freezing of supercooled drizzle may play a role,

4) explore the possibility that there is an active secondary ice production process active, and

5) point out that the ice number concentration enhancement falls within the conditions defined by the H-M mechanism and that H-M is one of the possible candidates.

<u>General response to comments from Reviewer #1</u>

We thank the reviewer for their useful critique. We have made a number of changes to the manuscript based on Reviewer #1's comments (and some from the other two reviewers). In particular, the discussion section has been rewritten and the conclusions and abstract edited accordingly. The results and discussion regarding ice formation have been rewritten as follows:

Sections 3.2 & 3.3: Description of measurements on a run-by-run basis, as before

Section 3.4: Summary of measurements, including comparison to IN calculations from DeMott et al. (2010), highlighting the transition from primary to secondary ice and the measurement of frozen drops followed by vapour-grown ice, predominantly with columnar features. This section includes some parts that formed part of the discussion in the ACPD version.

Section 4.3 then discusses what we may learn from the observations of:
4.3.1: Columns and mixed-habit ice: There is a secondary ice production mechanism active at temperatures where columns are the primary ice habit, and these ice crystals are also sent to different regions of the cloud.
4.3.2 Frozen drops: There is a source of small ice crystals that impact upon drizzle drops to freeze them, but are otherwise too small to see until they have grown by vapour diffusion, which takes time.
These two sections, as well as parts of the results section, invoke the concept of secondary ice production, but all discussion of the different ice multiplication mechanisms (H-M, drop freezing and

ice-ice collisions) are now contained in Section 4.3.3 in the discussion. Regarding drop-freezing secondary ice production, we have stated that it's likely to be less effective than in Lawson et al (2015) due to the colder cloud base in our study, but we can't rule it out. It is likely to have been active in some form, but full quantification is not possible without a microphysics model which includes all the processes.

The abstract and conclusions have also been rewritten to take into account the above. The main take-home message is now that there seems to be one or more source(s) of small ice crystals that freeze the drops, which then grow large and precipitate out.

Anonymous Referee #2

Review of "Observations of cloud microphysics and ice formation during COPE" by Taylor et al.

Recommendation: Requires revision before acceptance in ACP.

This study examines data collected in two lines of cumulus clouds over the Southwest Penninsula of the UK acquired during COPE. Sampling was performed first along a line of closely packed cells, followed by repeated penetrations through an isolated cell as it grew and became glaciated. The evolution of the observed cloud and aerosol properties is explained in terms of the action of the Hallett-Mossop process. The continued passes through the developing cumulus cell are especially unique, and hence should indeed be published. However, I find that some of the writing in the paper is not precise, overly speculative, and not totally justified by the presented data. In particular, the authors state that the observations show that the H-M process was initiated from the recirculated ice, with ice splinters causing the drizzle drops to freeze on contact, forming additional instant-rimers. Although it could be argued that the results presented in the paper are consistent with this explanation, there could also be other mechanisms acting that could also have explained the observations. Thus, I recommend that the writing in the paper be carefully examined to note what trends are consistent with the H-M process, rather than stating that the H-M process explains the results. There are also a number of locations in the paper where the writing could be made more precise and quantitative, with the numbers of speculative comments reduced. This is further explained below in the context of some passages of the manuscript.

We thank the reviewer for their detailed comments, which we respond to point-by-point. The question of "trends are consistent with the H-M process, rather than stating that the H-M process explains the results" has been addressed in the changes to the discussion section outlined in the responses to Reviewer #1.

A second major point relates to questions about some of the microphysical analysis that is presented in the paper. It seems that there is an overreliance on the images from the 2DC/CIP for identification of the phases of the particles, and that other information should be additional considered. The definitions of circularity and roundness are important for many of the results that are presented in the paper. With this in mind, I think that the definition of circularity should be stated in the manuscript. Also, how sensitive are your results to the definitions of low, medium and high irregularity and roundness? If a minor adjustment is made in the threshold to change these classifications are your results significantly impacted?

The definitions of circularity are listed in appendix A1. To show the sensitivity to the exact circularity thresholds, we have added the line

"Varying the circularity thresholds (in all categories concurrently) by ±0.05 caused a 5 – 20 % change in the ice and round concentrations reported by the 2DS."

More importantly, I am surprised that a lot of the shape analysis in the manuscript is based on the 2DS images rather than the CPI images which give more detailed pictures and would allow the riming to be much more easily determined.

As stated on pages 16056-7 “Quantitative hydrometeor concentrations could not be determined from the CPI during COPE” The sample volume of the CPI is much smaller than the 2DS, and it suffers from very poor counting statistics in short penetrations through mixed-phase clouds. For example, in a cloud pass where the 2DS recorded thousands of cloud drops and hundreds of ice particles, the CPI might record just 30 cloud drops and no ice. The concentrations determined by the 2DS are therefore much more quantitative. The 2DS is a well-established cloud probe and has been used successfully in many previous studies (e.g. Crosier et al., 2011; Lawson et al., 2015; Lloyd et al., 2014)

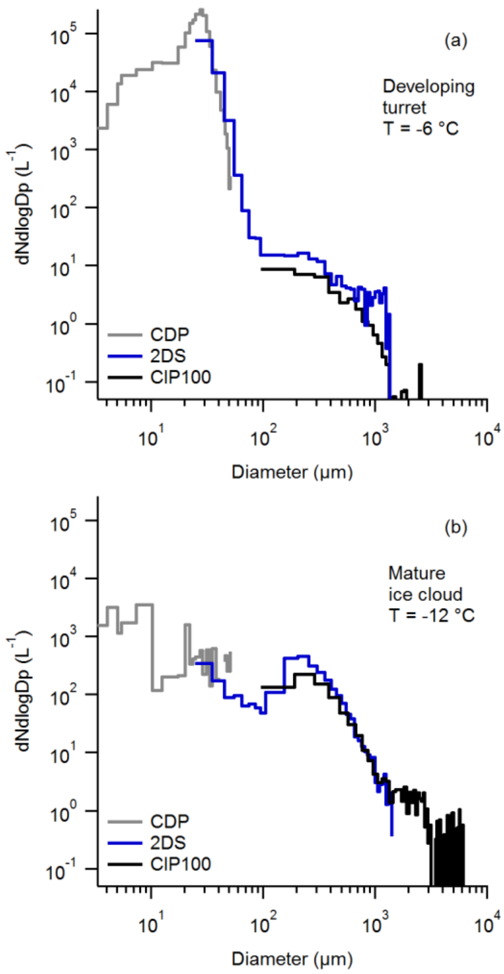
I also have difficulty in understanding why particles that occurred on the edges of the photodiode array would be any different than those occurring in other portions of the photodiode array: you are sampling the same population of particles so why would they be any less likely to be mixed-phase?

Cloud phase is often size-dependent, for example a mixed-phase cloud made up of 30µm drops and 500µm graupel. Also, larger particles are more likely to fall on the edge of the array. In practice, this means the edge particles are often either ice or liquid, rather than a mix of the two. This may not be true in other environments (such as clouds in which larger rain drops are more numerous) but it was the case in COPE. There is no perfect solution for how to deal with edge particles on which standard shape analysis cannot be performed, but this solution worked well for this dataset.

No size distributions are presented in the manuscript. How well do the 2DS and CIP distributions agree? How well does the CDP distribution agree with that of the 2DS in the overlap range? Answering these questions would help justify the robustness of the data.

We have added in the following as a new figure to the experimental section

“A comparison of example size distributions measured by the three probes is presented in Fig. 1, which shows broad agreement between the probes where their size ranges overlap.



Comparison of size distributions from the CDP, 2DS and CIP100 probes during selected cloud penetrations through (a) a developing turret and (b) a mature glaciated region. The first 2 size bins from the 2DS and CIP100 were removed, as they are subject to large uncertainty, and are not used in this analysis."

And, finally, I have some questions about the phase identification analysis. It would seem that there would be some additional information that might help better determine the phase, especially given the out of focus 2DS images that are used in some of the analysis. For example, was the shape of the CDP distributions examined? They tend to be more peaked in liquid clouds and flatter in ice clouds.

The CDP measures particles in the size range 2 – 50 µm whereas the 2DS HI,MI and LI categories are for particles larger than ~90µm. Cloud drops can coincide with larger ice particles- the clouds sampled in this analysis were predominantly mixed-phase. As shown in the figure above, the CDP distributions did tend to be monomodal in liquid clouds, but flatter and in far lower concentrations in ice cloud. This doesn't really add to the analysis though. The best measure of the cloud phase is by comparing the concentrations of cloud drops measured by the CDP and drizzle and ice measured by the 2DS.

Other questions include the following: was the H or V channel of the 2DS used, or some combination?

We have added to appendix A1:

"The 2DS has two data channels measured by vertically- and horizontally-aligned detectors. We combined images from channels to calculate an average concentration. This technique improves the counting statistics in low concentrations, as the large majority of particles are observed only on a single channel."

Were depth of field corrections applied for the small particles?

We have added to appendix A1:

"The sample volume was calculated using the measured airspeed and size-dependent sample area, as described by Heymsfield and Parrish (1978)."

It is stated that varying the inter-arrival time threshold had little impact on the derived concentrations and that a constant threshold was used: although this would be problematic in conditions of large or heavy ice, this seems reasonable for the conditions sampled. Nevertheless, it would be nice to know quantitatively what "little" effect means: 10% or a factor of 2? And, is there any way to be more quantitative about the amount of riming that is occurring?

We have added to the appendix:

"Using either no IAT threshold, or a threshold of 2×10^{-6} s, caused a 5 – 15% increase, or decrease respectively, in the ice and round concentrations reported by the 2DS."

I find the naming of the runs (e.g., 10.3.1, and 11.1) rather strange. Can the runs simply be designated by their times, or is there something more significant about these naming conventions that is used. If they are named by time or altitude, it would help give me a perspective of when/where they were.

In the archived flight logs, runs are labelled as run 10, run 11, run 12 etc so it's good practice to keep the labelling consistent with those. The numbers after the decimal points are there to signify that it is only part of one of the runs in the flight logs.

The caption to table 2 now has the sentence added:

"The run numbers correspond to full or sections of labelled runs in the archived flight logs."

To simplify things slightly, run 10.3.1 has been renamed to 10.3

I think some of the figures could be better presented to give a better perspective of where the data were obtained to help interpret the patterns within them. For example, in Figure 5 can the concentration plots be on the left hand side of the plot and the vertical velocities on the right hand side of the plot? Then the plots could be sorted vertically by altitude giving a better graphical manner to interpret the variations between the runs.

There's a lot of information plotted on Figure 7, and it went through several iterations during the drafting process, one of which was very similar to what the reviewer describes. In that configuration, the spatial link between the concentrations and the vertical wind is lost. The way it is currently plotted, the runs are still sorted by altitude, but just in 2 columns. This seemed to be the best compromise.

Abstract: I think the abstract should be considerably tightened and made more quantitative, stating what the observations were with a minimal focus on speculative comments explaining the observations.

The abstract has been largely rewritten to focus on the observations.

In addition, there is too much introductory material in the abstract that should be removed (e.g., the first two sentences are fine for an introduction, but not needed in the abstract). That would allow extra space so that it could be explicitly stated what conditions were present to justify statement that conditions of H-M process were met.

We have removed the introductory sentences, but also reference to H-M as it has been moved to later in the abstract.

In the third paragraph, the statement of "a few drizzle drops" should be made more quantitative.

We have changed it to say "drizzle concentrations increased from ~0.5 L^{-1} up to ~20 L^{-1}in around twenty minutes. Ice concentrations developed up to a few per litre, which is around the level expected of primary IN."

Also, it could be noted that graupel formed after the drizzle drops, but the freezing of the drizzle to form graupel was not explicitly observed.

It now says

"The ice images were most consistent with freezing drizzle, rather than smaller cloud drops or interstitial IN forming the first ice."

In the fourth paragraph it is stated that ice splinters were captured by supercooled drizzle drops causing them to freeze: but, again this was not a process that was observed so more quantitative comments about what was actually observed shoud be noted. In the fifth paragraph, can you stated what quantitatively a "majority of precipitation-sized particles" means. The second to last sentence in the abstract is consistent with the observations, but is not necessarily the only explanation for the observations and hence this statement should be reworded.

We have reworded these two paragraphs:

"Almost all of the initial secondary ice particles were frozen drops, while vapour-grown ice crystals were dominant in the latter stages. Our observations are consistent with the production of large numbers of small secondary ice crystals/fragments, by a mechanism such as Hallett-Mossop or droplets shattering upon freezing. Some of the small ice froze drizzle drops on contact, while others grew more slowly by vapour deposition. Graupel and columns were seen in cloud penetrations up to the -12 °C level, though many ice particles were mixed-habit due to riming and growth by vapour deposition at multiple temperatures.

Our observations demonstrate that the freezing of drizzle/raindrops is an important process that dominates the formation of large ice in the intermediate stages of cloud development. As these frozen drops were the first precipitation observed, it is clear that interactions between the warm rain and secondary ice production processes are key to determining the timing and location of precipitation."

Page 16054, line 10: Why was this particular case (3 August) chosen for analysis? A couple of sentences of explanation should be given.

We have added "…as it presented the most detailed set of observations, including repeated aircraft penetrations through cloud regions with prolific ice production"

Page 16060, line 11: Was there any evolution of the altitude of the cloud base during the height?

There may have been but we are not able to assess this with our observations. The text already states that the cloud base was similar to 1km, not exactly 1km.

Page 16061, line 3-4: Could there have been any possibility of APIPs from the King Air that could have generated ice, ultimately affecting the measurements from the BaE- 146?

This seems unlikely. Most of the time the two aircraft were sampling different clouds. Similar cloud formation and progression (i.e. rising turrets and increasing reflectivity) was observed by the ground radar in clouds that the aircraft did not fly in. If the aircraft injected significant ice into the clouds, these clouds would be expected to develop and precipitate faster than other clouds, which was not the case.

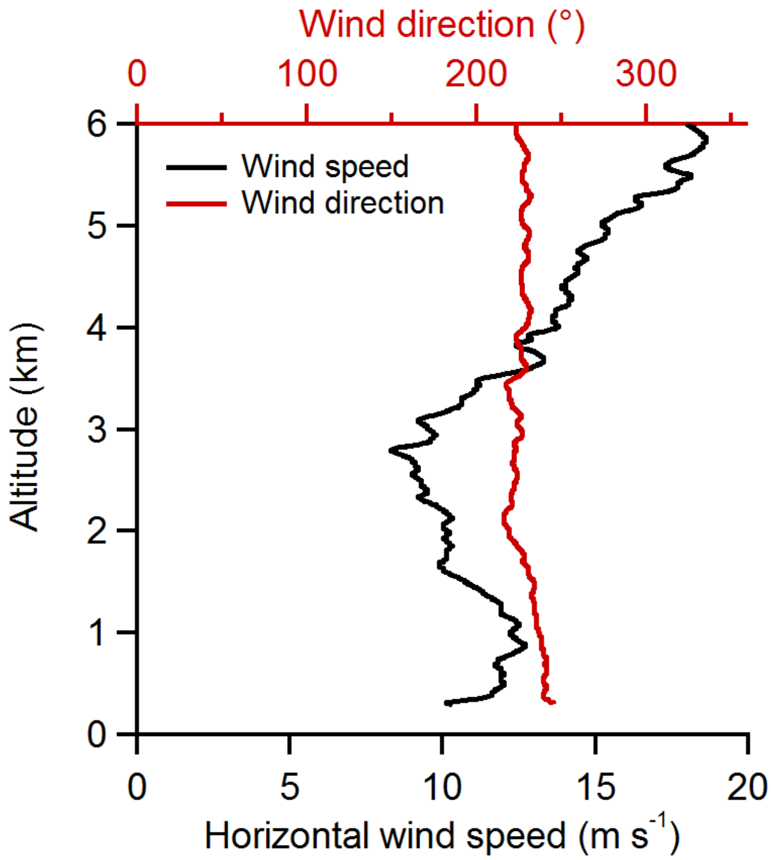
Page 16062, line 9: What do you mean by three runs? I only see a single line in Figure 3 so I am wondering how you define a run.

The line referring to this figure (now Figure 5) says "Figure 5 shows in situ microphysical data recorded during one of these runs" so it is clear that only one run is presented. Data from the other two longitudinal runs did not add anything further to the analysis so are not presented in the manuscript.

Page 16062, line 10: There were several times in the paper I was wanting to see a vertical profile of the wind. Could you show one?

We have added the following plot to the start of the results section

"Figure 2 shows the vertical profile of wind speed and direction. The wind direction was southwesterly and showed little variation with altitude. The wind speed was 8 – 13 ms^{-1} up to 3 km, then increased with height to reach 19 ms^{-1} by 5.8 km.



Vertical profile of horizontal wind speed and direction measured from the 1500 UTC radiosonde launched from Davidstow, the same location as the NCAS precipitation radar."

Page 16062, first paragraph in Section 3.2: Is it possible to also include a plot of the liquid fraction? I think that would be very useful for many of the analysis fields that you presented.

For the liquid fraction, there is no single answer. For example, both the liquid and ice number and mass fractions are very size-dependent. The best way to tell is to compare the CDP concentration and/or LWC to the ice and drizzle concentrations measured by the 2DS, which are already plotted.

Are the components of ice included or not included in the calculation of effective radius? It would seem important to exclude them for calculating the effective radius.

It was included, we have clarified by redefining it as R_{32}

"The variable R_{32} in Fig. 5 is defined as the ratio of the third to the second moment of the hydrometeor size distribution generated by combining all-accept data from the CDP, 2DS and CIP100. For spherical drops, this is the same as the effective radius; when used in clouds containing nonspherical particles it is a useful indicator of the average particle size."

This definition is sufficient as R_{32} is not used quantitatively (e.g. to compare to remote sensing measurements), but as an indicator of particle size.

Page 16062, line 26 (also several other points in paper): I have trouble seeing where you are getting the information about the spatial scale of the vertical motion from. Was some sort of FFT analysis applied? Or what else was done to determine the spatial scale of the vertical motion?

It's just from looking at figure 5c and seeing how wide the up/downdrafts are.

Page 16063, line 11: Could look at shape of CDP to verify that the cloud was almost exclusively composed of liquid drops?

The text refers to a measured CDP concentration of 200 cm^{-3}. If this was all ice, the concentration would be three orders of magnitude higher than typical ice concentrations in mature ice clouds.

Page 16063, line 15: It would be interesting to show the size distributions to better visualize these comments on the contributions of drops with different diameters.

The relevant part of the new discussions section (in 4.3.3) now refers to the new figure 1 which shows a size distribution

Page 16064, paragraph on Region II: I'm not convinced that I see evidence of liquid water from the images that are presented for Region II. It is claimed that this is a more mature cloud because of the presence of the ice crystals.

We have stated in the experimental section that the LWC is measured by the CDP, and

"…agreed with the onboard Johnson–Williams hotwire probe within ~12% below ~0.9 gm^{-3} (on the CDP), above which the hotwire probe began to saturate. The reported LWC values are from the CDP, as the hotwire probe suffered from wetting and saturation artefacts"

Additionally, the measured CDP number concentration was 60/cm3, and there is no known mechanism to generate such a high concentration of small ice in atmospheric clouds.

Is there some way that is more objective that can be used to estimate the age of the cloud to make the analysis more objective?

We've added in

"Regions I – III were parts of the line that had emerged 30 mins earlier, but was still active dynamically, with new cells emerging on the edges of pre-existing cloud regions. Regions IV and V were further downwind in a region that originated around an hour earlier, passed through a dynamically active phase and was now quiescent and stratiform."

And also about region V

"Following the aircraft penetration, these new updrafts developed into a reinvigorated dynamic cloud system."

Given the spatial and temporal resolution of the radar, it's very difficult to be more precise than this.

Page 16064, line 13: Are some of the donut type crystals seen in region III assumed to be liquid particles? Out of focus ice particles can appear as round donuts in 2DS/CIP images, so it is not necessary that these particles are liquid.

No the liquid was measured by the CDP, the figure shows the 2DS only measured a few per litre in region III.

Page 16064, line 17: See my comments on microphysics analysis. How well do you really know what the concentrations of graupel or round particles are? What is their uncertainty baed on the shape analysis?

This is now addressed in appendix A1, see previous comment about varying the circularity thresholds

Page 16064, discussion of Region III: Inevitably, you may be seeing some differences in the regions from what part of the cell was penetrated (edges or cores). This is certainly acknowledged in the text, but I am left wondering to what degree some of your conclusions can be affected by these differences.

The text already states

"in some of the cloud penetrations, particularly the later penetrations where the cloud was more developed and reflectivity was higher, the LWC and updraft strength may be biased low compared to the updraft cores."

which is fairly clear. Also, near the start of section 3.3.2 we have added

"Some variation in maturity of cloud regions, and concentrations within those regions, is to be expected due to penetrations through different parts of the cloud, but the general progression from young liquid cloud to mature ice cloud was still clear, as was the timescale over which this transition took place."

Page 16065, line 8: I would recommend removing the last sentence of this paragraph unless something more concrete can be said rather than the speculative statement.

Done

Page 16065, line 11: Is it differences in ages, or do what extend could some of these differences be associated with where in the cells the penetrations are made? Do you see any gradual maturation of the cells on the radar that can be quantified?

We have added to the start of section 3.2

"The ground-based precipitation radar observed newly-formed cloud regions increasing in height and reflectivity as they matured and moved downwind. The timescale from initiation, development of precipitation and dissipation was of the order of one hour, though mixing between different cloud regions meant the precise age of any one region was often difficult to determine."

Page 16065, line 25-26: I think it would be more fair to say that your results are consistent with the action of the H-M process: it does not really infer that this process is going on.

We have phrased the rewritten discussion section in this way

Page 16066, line 3: Where does this 4 km come from? It seems to be more than 4 km away from the other cells on this figure.

Reading off figure 5 it's 4.7 km, so it has been changed to ~5 km

Page 16066, line 14: can you quantify what you mean by relatively low?

We've added that it was below -5 dBZ

Page 16066, line 16: Are these numbers correct? The 0.5 m^-3 of 1 mm drops seems a little high for this reflectivity.

The numbers are correct. To clarify and give a better feel of the numbers, we have changed the text to

"For comparison, the peak reflectivity of −3 dBZ is equivalent to a concentration of 0.5 $m^{-3}$of 1 mm raindrops , 500 $L^{-1}$of 100 µm drops, or 2000 $cm^{-3}$of 25 µm drops"

Page 16067, line 17: Can you say something more quantitative than young clouds?

The text refers to figure 5 showing Run 11.2. We have added to the discussion in section 3.2:

"Regions I – III were parts of the line that had emerged 30 mins earlier, but was still active dynamically, with new cells emerging on the edges of pre-existing cloud regions. Regions IV and V were further downwind in a region that originated around an hour earlier, passed through a dynamically active phase but was now quiescent and stratiform"

Page 16068, line 10: I find this hard to justify with presented data because so many of the images are out of focus.

The text says

"As a general trend, in the earlier, lower altitude runs, the cloud was composed almost entirely of liquid cloud drops, and the peak LWC increased with altitude. This trend proceeded until significant precipitation was observed, at which point the LWC began to decrease, likely by scavenging and entrainment. At increasing altitude/time, the cloud shifted to mixed-phase and finally was nearly glaciated in the final run."

The liquid drops were mainly cloud drops measured by the CDP, which does not have such ambiguities.

Page 16068, line 18: Can something more quantitative be stated to show that the downwind updraft region was more turbulent?

This now reads

"In Run 11.1, the downwind updraft region was more turbulent, with vertical velocity varying on the scale of tens of metres, where the upwind updraft was 300m wide"

Page 16069, line 10 and after: Can you refer to which particular figure you are seeing this information about the particle characteristics? I'm not convinced that you really have the resolution in the particle images to observe this. So many particles on the CIP/2DS don't really have sufficient resolution. Can't you be using the CPI and showing specific images to show that this is occurring?

We have already discussed the fact the CPI did not record very many particles due to its low sample volume. Example images we are referring to are now highlighted in figures 8 and 9 with red arrows

Page 16069. line 22: What is the basis of saying that the downwind side is more turbulent?

The upwind side consists of two broad updraft sections, 800 and 1000m across, whereas the downwind side has several up/downdrafts each 50-250m across. This is visible on figure 7

Page 16070, line 8: You can say frozen drops but not recently frozen drops. There is nothing in the observations that says the time at which the drops were frozen.

The text now reads

"…a mixture of small columns, recently frozen drops and rimed ice. The frozen drops had some riming but were still recognisable as frozen drops, so are likely to have frozen fairly recently."

Page 16070, line 18: Remember that you are not necessarily sampling the exact same locations in the cloud. Plus with the evolution and movement of particles, you can't necessarily equate one part of the cloud to the other. Thus, some of the discussion should be adjusted accordingly.

Please see earlier comment regarding this question

Page 16071, line 17: reword "following cloud upwind". I think I know what you mean, but this is worded awkwardly.

The subsection title is now "Ice in the next cloud upwind" and the paragraph says "…the BAe-146 made three runs near the top of the next cloud to the southwest along line CD (i.e. upwind)"

Page 16072, paragraph beginning line 3: A lot of the comments in this paragraph are overly speculative (suggesting, likely, may have, etc.). I think it would be better to say the data are consistent with these processes. Are there any more processes that the data are also consistent with?

We have rewritten the discussion section to consider other processes that may explain our results.

Page 16072, line 23: Emphasize that the results are consistent with some processes, state if there are any other processes that should also explain the results, and note that there really is nothing that proves what is stated.

This has been taken into account when rewriting the discussion section

Page 16073, Section 4.1: Is this section really needed? I think the earlier reference in the paper to the 2004 floods should be sufficient.

The section is useful as a comparison of the similarities and differences between this case study and the Boscastle event, which was one of the main motivations for the project. They were similar, but not the same, and it is useful to explicitly point this out. It helps provide some context e.g. for modellers also looking at this case study

Page 16074, line 13: It would be nice to show the wind gradient or vertical shear in the paper somewhere.

We've added a reference to the new figure 2 showing the wind vertical profile

Page 16076, line 6: It should be emphasized that it takes time for any ice crystal produced at cloud top to fall to the measurement level, so the approach discussed here should provide a maximum estimate of the ice crystal concentration.

We've added that to the previous paragraph

Page 16077, line 11: Can you be more quantitative rather than use words such as minimal?

This has been removed in the restructuring/rewriting of the discussion section

Page 16078, line 2: I would argue that this is a possible explanation or that the results are consistent with this explanation.

The comment refers to ice seeding from pre-existing cloud layers aloft. We have already ruled this out in the section on primary ice

"the cloud region that was the focus of this analysis was initially isolated from neighbouring clouds by several kilometres, and no aerosol layers or regions containing outflow from previous clouds were detected in the immediate vicinity."

Page 16078, line 3: classified rather than classed

This has been removed in the restructuring/rewriting of the discussion section

Page 16078, line 13: Can there be some quantitative analysis presented to justify that ice in the downdraft became more rimed.

It's clear from the images in Figures 8 and 9

Page 16078, line 21: It is not clear these particles were generated by the H-M process, but the results are consistent with their generation by the H-M process.

We have used language like this when rewriting the discussion section

Page 16079, line 8: Remove word "recently"

Done

Page 16808, line 2: I think you determined that the H-M process was consistent with the observations, not necessarily that it was responsible for generating the ice crystal concentrations up to several hundred per liter.

We have taken this into account when rewriting the discussion section

Page 16084, lines 5-12: I don't think this paragraph is needed. I agree that it would be a good subject for a subsequent paper.

It is not strictly necessary for our observations but it helps to place them into the context of the COPE project as a whole (i.e. modelling of this and similar case studies is being undertaken)

Table 2: Is there some significance to the nomenclature that is used to label the runs?

Yes, as discussed above it's the same as the archived flight logs

Figure 1: Can you also show quartiles of the distributions on the plots as well as the individual points in order to give a better idea on what are the trends with respect to altitude?

Quartile statistics vs altitude isn't really an appropriate way to display the trends in the data, which are more temporal than vertical (i.e. cells glaciating as they mature, rather than ice or liquid being in particular levels). For example, consider Run 11.2, shown in Figure 5. This is a run at one altitude, but with very distinct and clear boundaries between liquid clouds (containing very little ice) and ice clouds (containing very little liquid). Averaging over the entire run therefore gives a skewed picture of the data- the average cloud becomes an average of liquid and ice clouds, but doesn't actually look like any real cloud measured.

Figure 3: Shouldn't the upper panel be labeled a, and the other panels subsequently reordered?

Done, and the same for Figure 12.

Figure 4: Should there be a vertical wind field for panel (b) for the first penetration?

Done. The left/right axes of panel (b) have also been swapped as it makes more sense now

Anonymous Referee #3

After careful reading of the manuscript, I saw that already two comments on the paper were posted. I read both of them and found that all questions and other points I would have liked to say are covered by the two reviews, in particular by the second referee. Specifically, I think the authors should take into account his comments with respect to the H-M process and the identification of particle habits using the scheme of Korolev and Sussmann (2000).

So I see no need to repeat the points already made, but like to emphasize my general impression of the manuscript: the topic of the study is of high importance and the measurements (instrumentation and measurement strategy) are impressive. I liked the introductory material a lot since it conveys the relevance of this detailed case study. Further, the data analysis, presentation and discussion are straightforward, the article is well organized and reads smoothly. The figures are clear and appropriate and no redundant material is presented.

In summary, the paper should be published after the revisions recommended above.

We thank the reviewer for their comments. The issues regarding H-M have been addressed in the changes to the discussion, summarised in the response to Reviewer #1.

Regarding the habit classification we have addressed the uncertainties raised by Reviewer #2, adding to the appendix A1

"Varying the circularity thresholds (in all categories concurrently) by ±0.05 caused a 5 – 20 % change in the ice and round concentrations reported by the 2DS."

and

"Using either no IAT threshold, or a threshold of 2×10^{-6} s, caused a 5 – 15% increase, or decrease respectively, in the ice and round concentrations reported by the 2DS."