Response to Anonymous Referee #2

We'd like to express our gratitude to the reviewer for their insightful review and we believe that the revised paper is significantly improved thanks to their comments and suggestions.

The authors compare simulations of a tropical MCS observed during a recent airborne field campaign with the in situ measurements between 0 and –40 C, where liquid water and ice could coexist (although there appears to be no liquid in the observations). There is substantial uncertainty as to how MCS updraft microphysical processes operate in nature, and improving process-level knowledge is a worthy research goal within the scope of ACP. Observations from multiple campaign flights have been reported by Leroy et al. (2015), as cited, but this appears to be the first analysis of the relationship of dynamics and microphysics observed during a flight. Overall, I am an interested reader, but I found it difficult to maintain attention on such a long paper for several reasons. First, it appears that the baseline simulation simply does not capture the event well at all, contrary to the authors' claims (in the abstract for instance), and sensitivity tests have similar errors across the board (e.g. Fig. 15). Second, several aspects of the observations appear notably odd (such as large updrafts without any additional ice content), but the authors focus on narrow elements of the observations without explaining why such apparent oddities are present. These factors combined make it difficult to be interested in nearly twenty figures comparing the simulations and observations, and even lead this reader to feel that the sensitivity tests may be futile or ill-conceived because the simulations are so far off the mark. Below I suggest the major steps that could help develop this manuscript in my estimation. Minor comments are then listed in case they are helpful.

Major comments

1. The MCS evolution in observations and simulations needs significantly more description. Highly averaged satellite data in Fig. 3 indicate that there are plenty of images that could be shown to us to see what OLR evolution looks like in the observations and the simulations. The reader needs to see these to understand if the simulated system appears far too large (in addition to being far too cold on top) compared with the observations. Is this a system coming off the ocean in observations and simulations? Is it of a similar size and duration? I would recommend showing OLR images before, during and after the aircraft sampling times used in this paper, both observed and simulated. It feels decidedly odd that these were omitted. This needs to be remedied and re-reviewed.

A timeseries of the enhanced IR imagery has been added, along with plan views of the OLR from the observations and the control simulation at 4 different times throughout the MCS lifecycle. The text describing the MCS has been expanded to read: Comparison of the modelled outgoing longwave radiation (OLR) with the satellite observations in Figure 2 show that in general, the control simulation represents the lifecycle of the MCS fairly well. The location of the mostly oceanic convective cells look reasonable, however, the modelled MCS is larger and composed of more numerous and deeper convective clouds than what was observed in the pixel level satellite OLR data and seen in the low level radar reflectivity fields shown in Figure 3. The model also produces more convection over the Tiwi Islands than what was observed at 17:30 UTC. As the MCS transitions from a developing-mature system through to a mature-decaying system, the observed reduction of deep convective cells with time is simulated, although the OLR remains significantly underestimated. During the research flight time at 23:30 UTC, the modelled MCS shows cloud positioned in a similar location to that observed with respect to the MCS structure, however, the modelled cloud is shifted somewhat to the northeast.

2. It is difficult to continue this review without understanding how the system simulated relates to the system observed in terms of overall shape and top OLR structure. Right now, it

appears to me, based on the figures shown, that the observed system is weak (Fig. 11), with low cloud top heights and surprisingly warm OLR (Fig. 3). Is this even an MCS? The simulations on the other hand do look like an MCS in terms of OLR and updraft strengths, but cloud top height seems low to me for a tropical MCS at 12.5 km with hardly any change with time. How is cloud top height defined in the observations and simulations? How do the underlying structure of cloud top heights observed and simulated compare, and what is the uncertainty in differences of definition between observations and simulations?

Apologies for making your reviewing job difficult because of these omissions. Please see the response above and note that we no longer include the cloud top height comparison due to, as you point out, difficulties in consistent definitions between satellite and models. Instead we describe the structure of the OLR as detailed in the point above. Also note that we now use the much higher resolution pixel level OLR observations, rather than the coarse resolution observations. This change shows lower observed OLR of around 120 W m^{-2}.

3. I will continue by assuming that the observed system is a small, weak system and the simulated system is a big, strong MCS, as appears to be the case from all indications in Figs. 3 and 11. Moving to the objective of this study, the title of the paper refers to phase composition, but this topic is not clearly explored. Only Figs. 13 and 14 (really one figure together) show liquid water content as a function of updraft velocity, but as far as I can tell there are no measurements of phase and no other analysis of phase.
Can the authors explain why they chose to focus on phase composition and why with this data set and this case study in particular? Also, what is a "high ice water content"? The updrafts shown here seem to have low ice water content. The authors refer to some other papers, but those seem to be focused on radar reflectivity.

In the introduction the description of the aims has been expanded to read:
The aims of this study are twofold: firstly to test different configurations of the dynamics, turbulence and microphysical formulations in the model to determine those that best represent tropical convective cloud systems and to understand the sensitivities in the modelled cloud and dynamical properties to these changes, and; secondly to determine what process control the phase composition and ice water content in the model. As mentioned previously, observations of HIWC (defined here as > 2 g m^{-3} at 1 km resolution) typically occur in glaciated conditions. However, as will be shown, the model is unable to replicate this and instead produces mixed-phase clouds under the same temperature regimes. For this reason we examine what processes control the modelled phase composition in order to understand how the model produces HIWC. This understanding will aid in improving the representation of these clouds in the model and produce a better forecasting capability.

4. The authors seem to view this modeling study as an exercise in manipulating their simplified microphysics (primarily) to better agree with the observations (unsuccessfully I would say) without investigating whether processes are actually likely to be active based on the observations. For instance, the absence of an observed bright band leads to a suggestion that particles are heavily rimed. (I think a tropical MCS should have a bright band, which to me seems another indication that the observed systems is not really an MCS. If the authors had a bright band simulator, I expect the simulated case would have one.) Later, graupel is removed from the model. What do the observed particle images look like? Do they indicate heavy rime? Is graupel observed? Leroy et al. (2015) show particle images, so I assume that they exist for this flight. Please describe what is known about the hydrometeor particles based on the flight data.

With respect to the bright band, the description of the lack of a bright band was in error. Based on the other reviewer's comment the text has been modified to read: The lack of a predominant bright band in the observations is likely due to the data being collected from volumetric scans, however, there are slightly higher reflectivities noticeable at 4 km indicating the presence of a bright band.

A discussion has been added to the section describing the MCS that reports on the presence of graupel and the observed particle images. It reads as:
There was almost no supercooled water detected during the flight, even at -10 ^{o}C, and graupel was intermittently observed. The absence of supercooled water coupled with the occasional presence of graupel is due to the system being sampled at the mature-decaying stage, where the supercooled water had been consumed in the production of graupel. Most of the time the particle images were of dense ice aggregates at flight level, except within some convective cores where graupel was observed, as also indicated by strong W-band attenuation.

5. Past literature on updraft microphysics seems to be largely ignored, as do particle size distributions themselves. The last sentence of the paper concludes that there is a need to better represent the "observed bimodal ice size distribution" but we are never shown a size distribution in the paper, either observed or simulated. How do we know that either observed or simulated are or are not bimodal and that this is important?

Based on a comment from the other reviewer, the mention of the bimodal size distribution has been deleted. Instead we retain the focus in this paper on the mass-weighted mean diameters and discuss the advantages of using a double moment microphysics scheme in representing the observed PSD variability. We have also added some discussion on updraft microphysics from other studies and note that detailed PSD studies from this campaign are currently underway. The additional text reads:
This contrasts with the lack of dependence of mean ice particle size on IWC that has been observed in earlier flights over Darwin and Cayenne in 2010 – 2012 (Fridlind et al. 2015) but agrees with more recent findings by Leroy et al. (2015). These findings show similar results to those documented by Gayet et al. (2012), with high concentrations of ice crystals occurring in regions of ice water content > 1 g m^{-3} sustained for at least 100 s at Darwin (Leroy et al. 2015) and > 0.3 g m^{-3} in the over shooting convection in the midlatitudes in Western Europe (Gayet et al. 2012). Gayet et al. (2012) proposed that the high concentration of ice crystals that appeared as chain-like aggregates of frozen drops, could be generated by strong updrafts lofting supercooled droplets that freeze homogeneously. However, using updraft parcel model simulations, Ackerman et al. (2015) showed that this process produced a smaller median mass area equivalent diameter than is observed. They proposed a number of other possible microphysical pathways to explain the observations including the Hallett-Mossop process and a large source of heterogeneous ice nuclei coupled with the shattering of water droplets when they freeze.

6. The concern of the authors with model dynamics is likely well founded. Some discussion of past model resolution studies would be helpful. Question: why bother with this exercise if the resolution of this model is too coarse to properly represent the updrafts observed, given that such updrafts are the only location where phase composition is interesting? If the authors do believe that the updrafts are grossly misrepresented dynamically, why spend so much time examining details of what occurs within them microphysically? Do the authors have evidence that this model is adequate to sufficiently represent uprafts being compared with observations? Why should I not conclude that this is the wrong tool to study phase composition in a tropical MCS?

The discussion on past studies of model resolution and the effect on updrafts has been expanded. It reads: …These values are well outside the range of maximum vertical velocities presented for oceanic convection by Heymsfield et al. (2010) and agree with other studies showing excessive tropical vertical velocities simulated by convection permitting models. Hanley et al. (2014) demonstrated that the UM with a grid length of 1.5 km simulated convective cells that were too intense and were initiated too early, as was also shown by Varble et al. (2014a), suggesting that convection is under resolved at grid lengths of order 1 km. Improved initiation time was shown by Hanley et al. (2014) to occur when the grid length was reduced to 500 and 200 m. However, the intensity of the convective cells was not necessarily improved, with the results being case-dependent. Varble et al. (2014a) also showed that in the tropics the intensity of the updrafts remained overestimated even at the 100 m grid length. Both of these studies suggest that there are missing processes in the model and/or the interactions between convective dynamics and microphysics are incorrectly represented.

We also note that recent cloud-resolving model intercomparison studies of tropical convection use a similar horizontal grid length to what is used in this study (e.g. Fridlind et al. 2012; Varble et al. 2014a,b). Some of these recent studies focus on convective updraft properties, which as described in the introduction, is important because these models are used to develop convection parameterisations for coarser resolution models. Therefore, a detailed understanding of how these models represent convective updraft processes is necessary.

7. Throughout the abstract, broad claims are made that are not clearly limited to these particular simulations. For instance, the last sentence of the abstract states that "... the entrainment and buoyancy of the air parcels is controlled by the ice particle sizes, demonstrating the importance of the microphysical processes on the convective dynamics." I think the authors mean in this particular system simulated by their particular model, which does not appear to resemble the system observed as far as I can tell.
The statements made in the abstract need to be more carefully delineated to refer to their particular model with coarse resolution and one-moment microphysics, especially given the apparently poor resemblance of results to observations in almost every way shown (e.g., updrafts, reflectivities, OLR, ice mean size, ice water content, and ice water content versus updraft strength). I credit the authors with showing these myriad flaws of their simulations (that is truly useful), but I would be more interested to see conclusions related to what model factors need to be changed to improve the simulations rather than conclusions about whether ice size controls updraft strength, given the unrealistic nature of the simulations.

The abstract has been revised to read:

Simulations of tropical convection from an operational numerical weather prediction model are evaluated with the focus on the model's ability to simulate the observed high ice water contents associated with the outflow of deep convection, and to investigate the modelled processes that control the phase composition of tropical convective clouds. The 1 km horizontal grid length model that uses a single moment microphysics scheme simulates the intensification and decay of convective strength across the mesoscale convective system. However, deep convection is produced too early, the OLR is underestimated and the areas with reflectivities > 30 dBZ are overestimated due to too much rain above the freezing level, stronger updrafts and larger particle sizes in the model. The inclusion of a heterogeneous rain freezing parameterisation and the use of different ice size distributions show better agreement with the observed reflectivity distributions, however, this simulation still produces a broader profile with many high reflectivity outliers demonstrating the greater occurrence of convective cells in the simulations. Examining the phase composition shows that the amount of liquid and ice in the modelled convective updrafts is controlled by: the size of the ice particles, with larger particles growing more efficiently through riming, producing larger IWC; the efficiency of the warm rain process, with greater cloud water contents being available to support larger ice growth rates, and; exclusion or limitation of graupel growth, with more mass contained in slower falling snow particles resulting in an increase of in-cloud residence times and more efficient removal of LWC. In this simulated case using a 1 km grid length model, horizontal mass divergence in the mixed-phase regions of convective updrafts is most sensitive to the turbulence formulation. Greater mixing of environmental air into cloudy updrafts in the region of -30 to 0 degrees Celsius produces more mass divergence indicative of greater entrainment, which generates a larger stratiform rain area. Above these levels in the purely ice region of the simulated updrafts, the convective updraft buoyancy is controlled by the ice particle sizes, demonstrating the importance of the microphysical processes on the convective dynamics in this simulated case study using a single moment microphysics scheme. The single moment microphysics scheme in the model is unable to simulate the observed reduction of mean mass-weighted ice diameter as the ice water content increases. The inability of the model to represent the observed variability of the ice size distribution would be improved with the use of a double moment microphysics scheme.

Minor comments

1. Page 8, line 14: How well are cloud bases observed by satellite? Cloud base throughout this system is at 3 km? That seems quite high to me for a tropical MCS.
Over ocean?

This paragraph has been deleted based on comments from the other reviewer.

2. Page 9, line 32: CloudSat IWP uncertainty is less than 25%?

This sentence refers to a comparison that was made between the tropical IWP derived from VISST and that from CloudSat. In the cited study, the comparison showed that VISST derived IWP was underestimated compared to the CloudSat derived IWP by 25%. But we take the point that CloudSat has its own uncertainties and have modified the text to read:
The observed IWP is only valid for the daytime from about 22:30 UTC or 8 am local time, and while the simulations with the generic PSD parameterisation compare well with the satellite derived value, the comparison of VISST IWP with CloudSat in tropical regions was shown by Waliser et al. (2009) to be underestimated by 25%, likely due to the maximum retrieved optical depth being limited to 128. Together with the CloudSat uncertainties (30% bias, 80% root mean square error; Heymsfield et al. 2008), this suggests that the modelled domain mean IWP may be underestimated from 22:30 – 23:30 UTC.

3. Page 11, first paragraph: There is a lot of discussion of divergence and convergence here, but to me the peaks above 15 km in Fig. 5 look like oscillatory gravity waves.
What evidence do the authors have that the peaks in motion above 12 km are not dominated by oscillatory motions?

Analysing vertical velocity profiles of the convective cells shows a smooth profile up to about 16 km, with oscillatory motions above this height. This finding also fits with the PDF of cloud top heights at this time that shows a distinct change in the distribution at 16 km. We note this in the revised manuscript.

4. Page 16, line 1: Both rain and ice appear bimodal to me; could they be related to one another?

Thank you for pointing this out. The text has been revised to state that the PDF is bimodal. Looking at the observed PDF distribution at heights in between 6 and 2.5 km shows that the bimodality does not persist throughout the vertical and, therefore, they do not appear to be related.

5. Figure 15: These observations need some explanation. There is a 10 m/s updraft with less than 90% RHI between –20 and –30 degrees C? Is there a problem with the observations? Fig. 12 shows IWC remaining low to 15 m/s at 0 to –5 degrees C. I think a section should be devoted to noting and explaining such features when these observations are first shown. Are

they somehow atypical? Is this strange strong updraft(s) associated with some aspects of the chaotic and odd diameter trends shown in Fig. 17?

Based on this comment we analysed the RH observations from all of the Darwin flights. This analysis confirms that there are erroneous observations and, therefore, this figure and discussion have been removed.

Most of the flight time was at temperatures colder than -10 °C and the limited number of samples affects the results for this temperature range. We now include the results for all of the Darwin flights to increase the sample size. However, there are still not a great deal of observations within this warmest temperature regime and the figure only includes the results of the compositing when there are more than 5 samples. The effect of this is to eliminate the chaotic trends. Additional text has been added to the beginning of this section that reads:
Due to the small sample size of observations from the single research flight on 18/02/2014, the observations from 18 of the Darwin HIWC flights have been used to allow for a more robust comparison of the model to the observations (Fig. 12 and 14). The majority of the flight time for these cases was in clouds with temperatures < -10 °C and vertical motions within the range of -2 to 2 m s^{-1}. Therefore, when comparing the model to the aircraft observations the focus is on this subset of cloud conditions as there are limited observational samples outside of these ranges.

The text describing the comparison of the simulations to the aircraft observations has been modified accordingly, but we note that apart from the increasing IWC in the downdrafts, the main conclusions have not changed.

6. I found it difficult to follow and maintain interest after the jump from Fig. 12 to Fig. 16 on page 20.

This section has been significantly revised. Figures 12 and 14 are now represented by a single figure and Figures 13 and 15 have been removed. The text has been streamlined throughout to focus more on the key points.