Review of the ACP manuscript “Can models robustly represent aerosol–convection interactions if their cloud microphysics is uncertain?” by White et al. The authors used the WRF model coupled with two double-moment bulk microphysics scheme to perform cloud system-resolving simulations of convection in the Congo basin, an idealized supercell, and a case of shallow cumulus convection, and tested the sensitivity of the simulated hydrometeors and precipitation to the microphysics scheme and to CDNC perturbations. The authors showed the simulations are sensitive to microphysics parameterizations much more than CDNC, which has been showed in previous studies but this study highlighted this point to imply aerosol effects are the secondary compared with the uncertainty in cloud parameterization. They further examined the shallow cumulus convection case and found that representation of autoconversion is the dominant factor that drives differences in rain production. The paper is generally written clearly but there are confusing sentences. They are some perspectives that can not be well justified physically such as saying the aerosol-cloud interactions in their study represent the upper limit.
The major problem is that their main point is based on an assumption that those twomoment schemes well represent the aerosol-cloud interactions, which is not the case based on the many past studies and a recent review paper by Khain et al 2015. Twomoment schemes have significant limitations in aerosol-cloud interaction process parameterizations such as nucleation, diffusional growth, and sedimentation, etc (detailed Khain et al. 2015). The paper did not really address the question “Can models robustly represent aerosol–convection interactions if their cloud microphysics is uncertain”, so that the title needs to be changed. More literature survey is needed, especially about those studies comparing different microphysics schemes and their responses to CCN or CDNC. Those studies need to be discussed in the introduction and the relevant places in the paper, especially about some important points on the problems with the parameterizations of some specific microphysics processes in the bulk schemes. Therefore, the paper needs major revisions to be accepted as a publication in ACP.

We thank Reviewer 2 for their thoughtful and constructive comments, which we have found useful in helping us to clarify our manuscript for the reader.

We present this study as an illustration that model uncertainty in cloud impacts arising from choice of microphysics scheme can far outweigh any aerosol effects observed within a single scheme, and that this result holds for case study simulations consisting of many cloud lifecycles, for idealised simulations with open boundaries, for idealised simulations with periodic boundaries, and (by nature of the cases used in our study) across different types of convection with their inherently different response to aerosol.

Our choice to use bulk microphysics schemes is twofold; first, they remain in wide use in the community (often for practical reasons of computational cost) and even represent state-of-the-art implementation in global models, which have historically relied on much cruder microphysics representations, and second (most importantly) there is currently only one bin microphysics scheme implemented in the public version of WRF (although it comes in two slightly different versions, a ‘full’ and ‘fast’ version which differ in the number of ice categories used) and thus our study would be impossible to perform within a single modelling framework had we opted to use bin schemes.

Specific comments:

1. The title needs to be changed. It is relevant but the authors did not conduct an unique study to really address this question. The two-moment schemes can not robustly represent aerosol–convection interactions due to the limitations in representing the most relevant processes as detailed in Khain et al. 2015. If a bin scheme is used, you might end up with similar magnitudes of aerosol indirect effects as the differences among different microphysics schemes. In addition, for specific case simulations, one can not really reveal how it is impacted by aerosols by conducting simulations with realistic aerosols/chemistry configuration. Any aerosol properties and spatial distribution change could change cloud and precipitation.

We have updated the title to emphasize our main result:
‘Uncertainty from choice of microphysics scheme in convection-permitting models significantly exceeds aerosol effects’.

Regarding the reviewer’s comments on the use of a bin scheme:
We have now performed simulations using the HUJI spectral bin microphysics scheme (SBM) implemented in WRFv3.6.1 (note that this is a newer version of WRF than used in the Morrison and Thompson simulations presented in the paper). We do not wish to further extend the paper to include discussion of the impacts of bin vs bulk microphysics in these cases (because the main aim of this particular paper is to highlight within a single paper and by using a single modelling framework the large uncertainty in many case study and GCM simulations that can result from any given combination of microphysics scheme, type of convection, and given CDNC value or prescribed CCN profile). However, we now present our bin scheme results in a supplement to our paper (Figures S2, S3 and S4) because we believe our results to be of interest. We find that the uncertainty due to choice of microphysics scheme still dominates any aerosol response within each scheme in the cases we test, regardless of whether a bin or bulk scheme is used.

2. P1, L21-22: aerosol can affect through aerosol radiative effects as well.
We did not intend our phrasing to imply that the only effect that aerosol has on convection is an indirect effect through cloud microphysics. We have reworded the sentence now on P1 L22 – P2 L1:
"One major way that aerosols can influence the properties of deep convection is through their effect on cloud microphysics."

3. P1, L24: Albrecht, 1989 actually showed the suppression of precipitation but for warm clouds. So, the sentence is not accurate.
We thank the reviewer for finding our omission of "warm-phase" in this sentence, and have corrected our sentence to "warm-phase precipitation", now on P2 L3.

4. P2, L29: "Until recently" should be deleted.
We replace "until recently" with "traditionally" (now on P5 L20), as it is only in recent years that schemes have started to move away from discrete ice categories (Morrison & Grabowski 2008, Harrington et al. 2013, Morrison & Milbrandt 2015), and we wish to make it clear to the reader that this is a relatively new and important development.

5. P4, L19-25, these sentences could be misleading. First, it is not clear what the authors mean by saying "the response of different microphysics schemes to perturbations in prescribed cloud parameters". Second, some past studies showed qualitatively different aerosol impact for different cloud types, with a purpose of illustrating a point that aerosol impacts depend on cloud type and dynamical and thermodynamic conditions of each case. So it is misleading to describe a study without giving information about cloud types or specific dynamical and thermodynamic environment. For example, the description about Fan et al. 2012 about aerosol reducing precipitation is not correct. The study did two different cloud cases over the eastern China – one deep convective cloud case with warm cloud base and weak wind shear and the other a winter stratiform cloud case, with aerosol increased precipitation for the former but reduced precipitation for the latter.
We thank the reviewer for pointing out that this was not clear to the reader.

We have heavily revised our introduction and included more detailed discussion on aerosol response in bin and bulk schemes, noting the convective regimes and large-scale environments used in each study. The paragraph P4 L19-25 has now been removed.

6. P5, L3: cloud and precipitation responses > cloud and precipitation responses to perturbation of CDNC
We thank the reviewer for noticing this omission and have amended our text (now on P8 L3) accordingly.

7. P5, L25: Khain and Lynn 2009 is not s study with specified CDNC. CDNC is prognostic in the bin model they used.
We thank the reviewer for this correction and have amended our text.

8. P6, first paragraph, RRTM and Goddard schemes only talk hydrometeor mass from microphysics calculation (Goddard shortwave scheme takes prognostic CDNC as well). The microphysics-radiation coupling does not account for particle size changes, which means some aerosol effects are missing in those studies.
We agree with the reviewer and we also note that the microphysics-radiation coupling is only between cloud water and ice, and none of the other frozen species. This missing aerosol effect may have an especially important impact in our Congo simulations, where the Morrison scheme develops and retains significant amounts of upper-level ice, whereas the Thompson scheme converts nearly all the ice to snow (see Figures in our response to Reviewer 1), which the radiation scheme will not see. This could have significant radiative flux and feedback impacts (Thompson 2015, Atmospheric Research), which in itself originates from the use of somewhat arbitrarily defined ice categories (e.g. if the size parameter at which cloud ice is converted to snow is changed, a bulk mass of cloud ice is removed from the radiatively-coupled ice category and moved into the non-radiatively coupled snow category).

We have included these important points in our Discussion and Conclusions of the Congo basin results, P24 L6-15.

9. Figure 3, the color scheme needs to be changed. The color difference is too little even between 0 and 100 mm.
We have changed the colour scale of this Figure to a scale which more clearly shows the large difference in accumulated rain over the 10-day period between the models and observations. The new Figure is shown below in Figure R1.

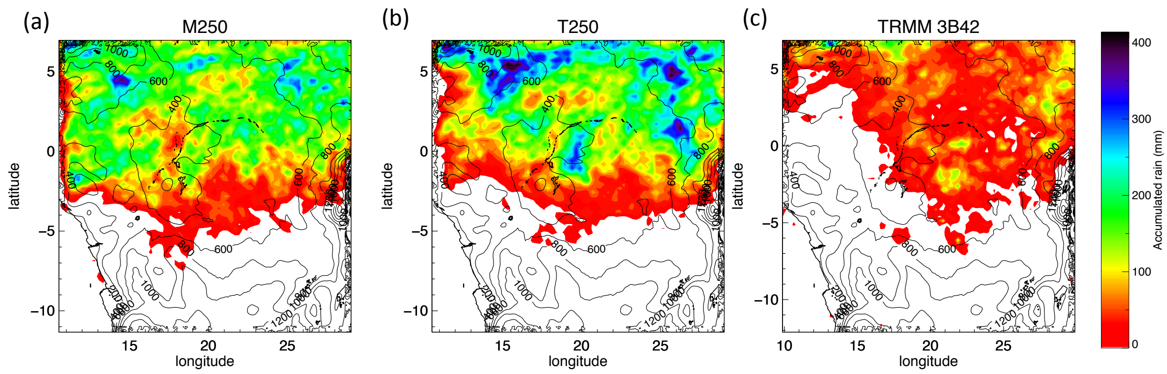

Figure R1: *Congo case: accumulated surface precipitation (mm) from 01 to 10 August 2007 in the Congo basin, showing data from (a) CONGO-M250, (b) CONGO-T250 and (c) observations from the TRMM 3B42 gridded 3-hourly mean merged precipitation product. The simulation data shown in this Figure has been coarsened to the 0.25 degree spatial resolution of the TRMM product.*

10. Figure 5, why compare with the climatology data? This is just a 10-day run, how should we expect it represent the climatology?

There were so few CloudSat overpasses during the 10-day period that the resulting histogram is very noisy compared to the model data, albeit qualitativey very similar to the climatology (see figure R2 below compared to Figure 5 in the paper). This is why we originally used the observed climatology to construct Figure 5. At Reviewer 2's suggestion, we have replaced Figure 5 in our manuscript with the histogram showing 10 days of CloudSat data, Figure R2 below.

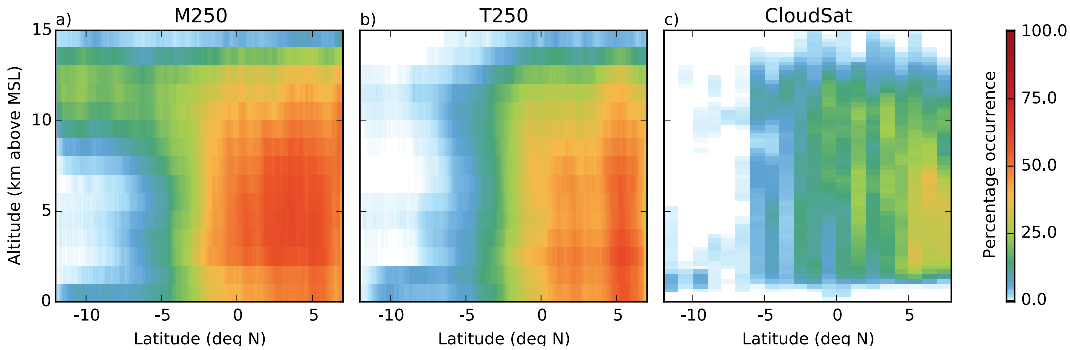

Figure R2: *Congo case: 10-day histogram for the period 1 – 10 August 2007 of model reflectivities derived from hydrometeor fields passed through the Quickbeam radar simulator, thresholded at values greater than -20 dBZ for (a) CONGO-M250, (b) CONGO-T250, and (c) the CloudSat 2B-GEOPROF product. In (a) and (b) the models have been sampled at the times of the nearest CloudSat overpasses.*

11. P. 9 L5-7, this sentence is confusing. Need to be clarified.

We have now expressed this sentence more clearly (P12 L22):

"The largest reflectivity values produced by the model occur in the convective region in the north of the domain, where the largest reflectivity values are detected by the satellite radar.'

12. P. 9, L19: domain averaged cloud water in T250 is 140 times larger than M250. Something could be wrong here. Can you plot the cloudy-point average for cloud water mass total hydrometeor mass to check if they make sense? What is the maximum cloud water mass in T250 and M250, respectively?

This is correct in the domain-average profile, because the shallow warm cloud in the M250 case precipitates out, whereas it persists in the T250 case. This leads to a significant difference in the domain-averaged cloud water mass, because there are many zero points in the M250 case which contain cloud water in T250 (see also our response, including figure, to Reviewer 2's comment 21)

Considering the cloudy-point (condensed-point) averages, we see a similar response in behaviour (increased cloud water mass in T250 compared to M250), except the magnitude is reduced because the profiles are normalized by number of points with condensed water. The condensed-point average profiles show qualitatively similar behavior in all cloud species to the domain-mean profiles (Figure R3), except cannot represent absolute changes in the number of cloudy / condensed points between the simulations.

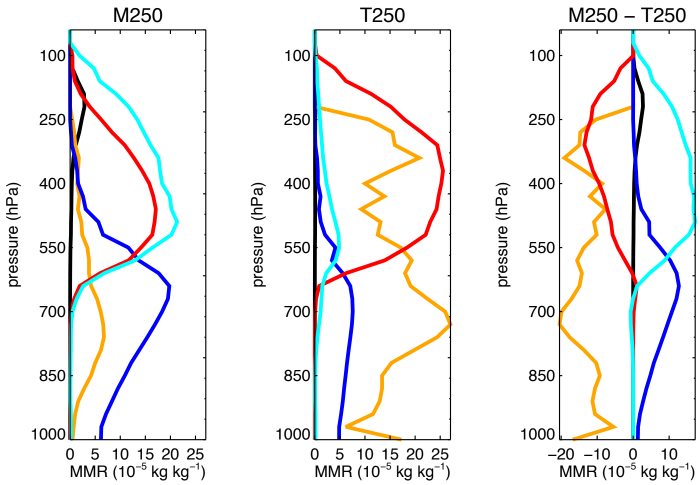

Figure R3: *Congo case: Cloud-and-condensed-point mean vertical profiles of difference in hydrometeor mass mixing ratios between the polluted and pristine cases, averaged over the 10 days of the Congo simulation for the Morrison scheme (left), Thompson scheme (centre) and difference between the two (right).*

We note also that in our WRF-SBM simulation (Figures S2 and S3 in the supplement provided with this response) the same warm cloud mass is present and has significantly greater domain-mean values than the M250 case. We believe this difference is because, as we have shown (our Figure 16), the cloud mass in the M250 case is removed through autoconversion to rain, whereas this does not occur in the T250 case or in the SBM.

13. Figure 6, this figure need to be replotted to show differences of other hydrometeor mass clearly. Right now, only cloud water differences can be seen clearly. I would use different panels for different hydrometeors.

We agree with the reviewer that this is a difficult figure to show all hydrometeor masses clearly whilst also showing the significant difference in cloud masses between the two schemes. However, we feel that using different panels for the different hydrometeors could confuse the reader, as this is not done for any other such figures. We now present the following figure, which shows all the hydrometeor classes on the same figure, as for all other such profiles in the paper, but uses a logarithmic horizontal axis even in the difference plot. This makes clear to the reader the differences of all the hydrometeor classes, not just the cloud water, whilst also showing that the difference in cloud water between the two schemes is an order of magnitude greater than between the other hydrometeor classes.

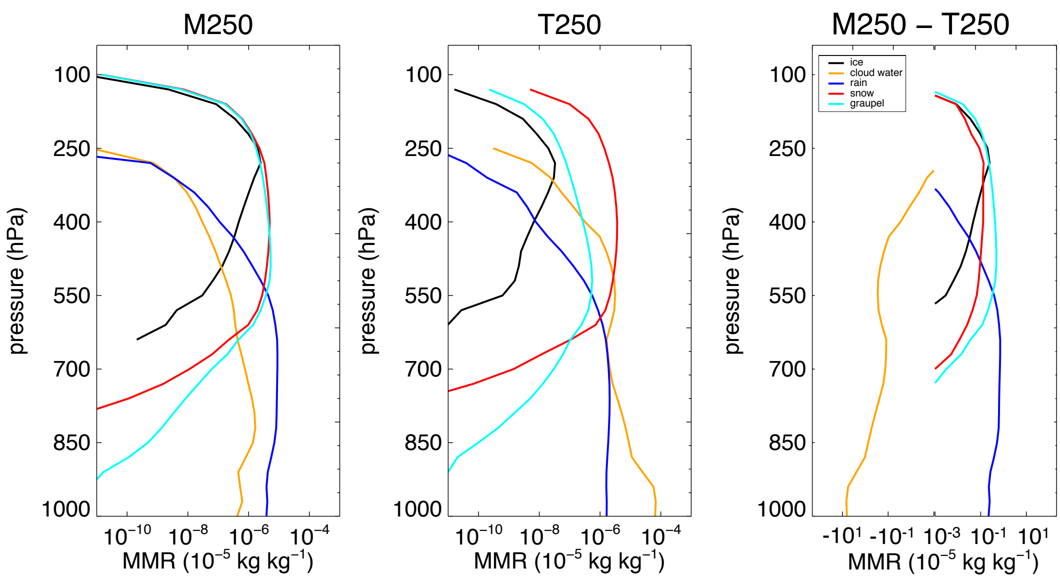

Figure R4: *Congo case: Domain-mean vertical profiles of hydrometeor mass mixing ratios (MMR) averaged over the period 1 – 10 August 2007. (a) CONGO-M250, (b) CONGO-T250 and (c) the difference in the domain-mean hydrometeor mixing ratio profiles (CONGO-M250 minus CONGO-T250). Note the logarithmic scale used on the* *horizontal axis in (a) and (b) to illustrate the large differences in the cloud water mass between the two simulations whilst also illustrating the differences in the frozen species. Note in (c) the diverging logarithmic ‘difference’ scale used on the horizontal axis, with ‘negative’ values extending from the centre to the left, and positive values extending from the centre to the right. The end of each of these axes is cut off at a value of 10^{-3}, however these axes tend towards each other towards zero.*

14. P9, last paragraph and Figure 7, the huge increase of cloud water with the increase of CDNC with the Thompson scheme seems not reasonable. How about the change of precipitation?
We agree with the reviewer that the huge increase of cloud water with CDNC in the Thompson scheme in the Congo basin simulations is surprising. Our results show that this is accompanied by a reduction in precipitation (Figure 7, also noted in P10 L2, continued paragraph from last paragraph on P9). This is consistent with warm-phase precipitation suppression (Albrecht 1989). This is also consistent with our warm-phase RICO results, which show an increase in domain-averaged cloud mass and reduction in precipitation with CDNC in the Thompson scheme (Figure 12).

We suggest that the development of the warm cloud mass is likely due to the background meteorological conditions these simulations are performed in (see also our response to Reviewer 2’s comment 21). We have shown that the autoconversion process is responsible for removing this cloud mass in the CONGO-MORR simulations. Our Figure 14 also shows that under increased CDNC, the threshold cloud water mass required for autoconversion to begin increases in the Thompson scheme. Therefore we suggest that increased levels of CDNC in the Thompson scheme in the Congo simulations even further suppress rain formation through autoconversion.

15. P. 10, L19-21: The sentence “we also see that the simulated hydrometeor classes differ between cases: the difference in the simulated hydrometeor classes in the idealised supercell configuration is different from the difference in the real-data Congo basin configuration” is not necessary. This is what it should be since they are different convective cloud types.
We have removed this sentence from our manuscript.

16. P11, first paragraph, the sentences in L5-6 and in L10-11 are repeated.
The first sentence refers to SUPER-MORR and CONGO-MORR, the second sentence to SUPER-THOM and CONGO-THOM. However, we have condensed these into a single sentence at the start of the paragraph (P14 L20):
“The SUPER-MORR and SUPER-THOM cases differs qualitatively from the CONGO-MORR and CONGO-THOM cases, respectively, both in the altitudes at which the response occurs and the sign of the response of some of the hydrometeors”

17. P11, first paragraph, the main point here should be about more significant aerosol impact on hydrometeor mass on the supercell case compared with the Congo case, not the different responses of hydrometeors between the CONGO case and the supercell case, since they should be expected for completely different cases. Many past studies have showed that aerosol impacts depend on dynamics and thermodynamics of convective clouds (e.g., Khain 2009; Fan et al. 2009).
We thought it necessary to place our results by first confirming that our results reproduce that hydrometeor response differs according to cloud type. However, we agree with Reviewer 2 that it would increase the clarity of our paper if we remove such discussion before presenting our results, and thank the Reviewer for suggesting that this can be taken as assumed knowledge. We also thank the reviewer for the suggestion of making our most important point here that the significance of aerosol impact differs between cases.

We have updated our text with these changes (P15 L1).

18. P11, last paragraph: the lack of appropriate sensitivity of bulk schemes to aerosols is mainly due to the limitation of bulk scheme parameterization in nucleation, diffusional growth, and sedimentation, etc, as detailed in Khain et al. 2015. The invigoration of updrafts can not be simulated since the saturation adjustment approach for diffusional growth of droplets limit such effects. Those aspects should be considered and discussed when interpreting the results on aerosol indirect effects here. Past studies showing the limitation of bulk scheme parameterizations in representing aerosol-cloud interactions need to be surveyed and discussed.
We agree with Reviewer 2 that saturation adjustment methods can prevent important physical processes from occurring in bulk schemes. We have now made extensive note in our introduction of past studies which show the limitation of bulk schemes in representing aerosol-cloud interactions (P2 L15 through P5 L16).

We also note that some bulk schemes produce convective invigoration effects. For example, Lebo 2014 found evidence of convective invigoration under increased aerosol loading in a bulk scheme under weak shear conditions (and suppressed convection under strong shear), similar to the findings of Fan 2009 who found the same response in a bin scheme. Seifert 2006 also found higher overshooting tops and larger sizes of cumulonimbus in a weak shear environment with increased aerosol loading.

However, we emphasise that the main focus of this paper is not an investigation into aerosol and microphysical processes of convective invigoration (of which there is an extensive body of literature), but to highlight that the uncertainty due to choice of microphysics scheme can far exceed any simulated aerosol effects (even when using the WRF-SBM bin scheme, Figure S4 in the Supplement). Our Figure 10 is mainly used to illustrate that latent heating (and therefore dynamic impacts) differences due to the choice of microphysics scheme can equal those due to different levels of CDNC in a bulk scheme. (Unfortunately we cannot provide equivalent latent heating impacts for the WRF-SBM supercell simulation presented in our Supplement because we wrote the latent heating output into the Morrison and Thompson schemes for this study and have not had the chance to do this for the version of the HUJI SBM included in the public WRF release).

We agree with Reviewer 2 that our finding should be placed in the context of the literature which discusses the ability of bulk schemes to produce invigoration effects, and we have therefore included this discussion in our text.

19. P12, L5-7: Again physically it is supposed to be that for different types of convective cases. Hydrometeor differs and hydrometeor responses to CDNC are different as well.
This sentence was not supposed to convey that the hydrometeors and their response to CDNC differ in different convective cases (which, as Reviewer 2 has said, is supposed to be the case), but rather to highlight the source uncertainty due to the choice of microphysics scheme: response to CDNC varies in each scheme according to cloud / convection type (known), but the difference between the response of the two schemes to CDNC across types of convection is not systematic.

We strive to make our text as clear as possible, therefore we have rephrased our text to say 'the difference between the response of the two schemes to CDNC across types of convection is not systematic' (P16 L31-32).

20. P13, L23-25, reword the sentence. Not sure what you really want to say.
We have removed this sentence as we would need to re-run the simulations with autoconversion and accretion rates output in order to show this point.

21. P15, first paragraph, the authors showed the autoconversion process is not the significant process contributing to the large cloud mass at low –levels. Then the question is what process mainly contributes to it?
We have shown that the autoconversion process is responsible for the removal of the large cloud mass at low levels in the model configuration using the Morrison microphysics scheme. We also see that this low-level liquid cloud mass forms when we run the same simulation with the WRF-SBM implementation of the HUCM microphysics (Figures S2 and S3 in the Supplement), although to a lesser extent than in the Thompson simulations, and the warm cloud produced by the HUCM SBM produces rain. We therefore suggest that it is not the Thompson scheme per se which is responsible for producing the low-level cloud mass, but rather the meteorological conditions present in which these simulations are performed.

We now discuss this in our revised manuscript, P20 L10-16.

We provide here Figure R4 showing 3D isosurfaces of cloud water and ice mixing ratios for one snapshot in time of the WRF Congo simulations, 7 days into the simulation, at 01Z. Videos of such figures show firstly that the warm cloud mass forms mostly in the south-eastern region of the domain, over the ocean, and secondly that this cloud undergoes strong diurnal forcing.

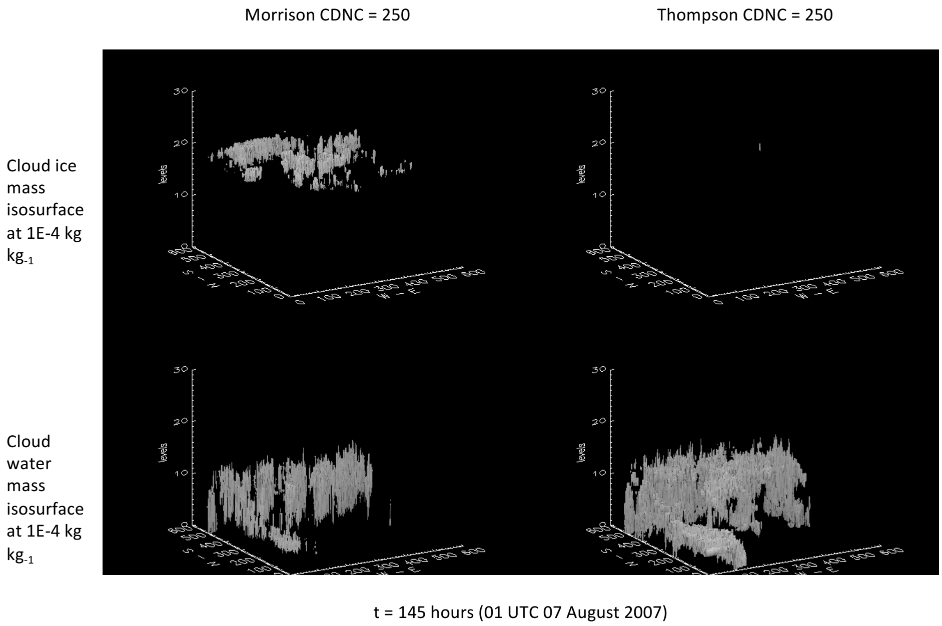

Figure R4: *Three-dimensional isosurfaces of cloud ice mass (top row) and cloud water mass (bottom row) for the M250 (left column) and T250 (right column) Congo case. The isosurface shown is the 1.E^{-4} kg kg^{-1} surface.*

However, regardless of the reasons for the development of the mass of warm cloud at low levels, the importance of our results is that under identical initial and lateral boundary meteorological conditions, the choice of microphysics representation on cloud development equals or exceeds that of cloud response to CDNC within each scheme.

22. P15, L19-20, again, the limitation of two-moment bulk schemes in representing aerosol impacts on microphysics processes should be discussed.
Bulk schemes have been shown to be limited in their ability to produce convective invigoration. Other studies using bulk and bin-bulk schemes have identified aerosol impacts on precipitation of up to about 15% (e.g. Kalina et al. 2014, Morrison 2012, Morrison 2011, Lebo et al. 2012, Lee & Feingold 2010, Lee & Feingold 2013, Lee 2011, van den Heever et al. 2006). Indeed, even studies using bin schemes have been shown to have little impact on total precipitation, although inducing a shift in rainfall rates (Fan et al. 2013).

We note that global modelling studies of aerosol indirect effects use bulk microphysics representations (e.g. Ghan et al. 2016, Zhang et al. 2016), and therefore our results have important implications for such studies.

We emphasise that our main result is to show that the variability due to, and within, schemes dominates any aerosol impacts on microphysics. Our results using the WRF-SBM in the idealised supercell case show that aerosol impacts in the bin scheme are of equal magnitude to those in the bulk schemes (Figure S4 in the Supplement).

We have included such a discussion in this section of the paper (P22 L24 through P23 L1) and also in the introduction (P2 L15 through P5 L16).

23. P17, L21-23, this sentence appears in a few places throughout the study, but the point can not be well justified even for aerosol indirect effects. First, you not know what the reality of aerosol look like in composition and spatial variability, many studies showed that aerosol spatial distribution could significant change storm location such as urban aerosols impact significantly on the precipitation in the downwind area of cities through aerosol indirect effects. Second, since two-moment bulk schemes even can not represent aerosol-cloud interaction processes physically, then how do you justify the aerosol impact here represent the upper limit?
We have removed these statements from our manuscript.

24. P18, L29, this is definitely not the first study to consider two and more cloud cases. A thorough literature search would give you those past studies.
We have removed this entire paragraph to make the conclusions more concise.

25. P18, L30-31, again, it has been a basic understanding that hydrometeors and their responses to CCN or CDNC vary with different cloud types and convective cases.
We agree with Reviewer 2 that this is basic understanding. As Reviewer 2 feels that this can be taken as assumed knowledge, we have increased the clarity of our paper by removing all such contextual discussion.