**Review of acp-2019-742**

Estimation of Cloud Condensation Nuclei number concentrations and comparison to in-situ and lidar observations during the HOPE experiments

This study compares CCN estimates from various methods. One method involves converting bulk mass measurements from both the COSMO-MUSCAT model and from observations to CCN using assumed size distributions. CCN was estimated from these size distributions and compared to CCN observations (surface CCN, lidar, and in-situ measurements on a helicopter) during the HOPE observational period in 2013. A second comparison involves comparing the 2013 estimates to estimates for the year 1985. Overall, I am unclear of the takeaways of this study, as their results seem to be primarily due to how the set up their methods, which are often not clearly stated or justified, and similar to a study that the authors were recently involved in (Hande et al. 2016). Generally, I found it somewhat difficult to assess the results and discussion and think the authors need to provide more details in several locations throughout the manuscript, including more explanations. This being said, I think there is a lot of interesting data and methods and I do think comparisons between models and parameterizations with observations can be useful. However, in the way this study was written and presented, I wasn't able to clearly discern this study's scientific contributions.

**Major Concerns:**

**The calculation of the 1985 estimates**

One of the main parts of this study is comparing results from 2013 to 1985. However, if I am understanding correctly, the manner in which the authors determine the 1985 values is by taking the 2013 simulated aerosol mass results and simply scaling them up by certain factors (Table 3). Then, the authors present the result that the CCN is higher for their 1985, but isn't this simply because they scaled up the mass concentrations in their methods. By simply scaling the 2013 model solution up and down, the authors may not be considering the different aerosol processes present in the model that may change with different emissions and concentrations within the model. For example, can the authors justify that the size distributions that are used to convert aerosol mass to the number should be the same in 2013 as in 1985. This may change their results significantly. I wonder if it would not make more sense to make assumptions about the emissions in 1985 and allows the model to run with those emissions, such that the processes are better represented with these higher concentrations are better represented.

Also in Table 3, the authors do not explain where the scaling factors (3.9, 5.3, 2) come from? They also do not explain why they assume elemental carbon should be twice as high, since they themselves state there were no emission data to support this. More details are necessary to properly assess this study, especially since these are the primary methods in terms of estimating the 1985 concentrations.

**Unclear and unjustified methods**

I appreciate that the authors' comprehensive study and comparing models to observations, which is a difficult task. However, throughout reading the manuscript, there were many instances where I either did not understand what the authors were doing and/or why they were doing it, which made it difficult to assess their results and conclusions. Although I have included some of these instances directly below, in general, I think the writing in the manuscript could be made much more clear and suggest the authors work on this.

P4, L15. The authors need to convert aerosol mass to a number size distribution, and therefore, must assume some size distribution shape. The authors end up choosing a one-modal size distribution for each species, with parameters given in Table 1. Is it a good assumption to use one modal size distribution in this region? I am not particularly familiar with aerosol size distributions in this region, but I think the authors need to justify why using a one mode size distribution is valid and better justify the parameters chosen in Table 1, as opposed to just providing references. This assumption leads to one of their main results -- that they cannot CCN at higher supersaturation doesn't compare well, because they only used one mode at accumulation mode particle sizes.

P8, L24-32: The authors clearly explain the process in which they estimate CCN. However, why do the authors not simply add up the CCN measured from from the different size selections to get the total CCN? The authors seem to currently convert the measured CCN to activated fraction, just to convert it back to CCN again? Can the authors clarify why they did this in this manner?

P5, L20: "To achieve the maximum supersaturation (as a function of vertical velocity), accounting for particle growth before and after activation, the supersaturation balance is used." I do not understand this sentence and it seems to be important for the CCN calculation. What is the supersaturation balance? This paragraph in general, I think is very important, since it goes over how the estimated aerosol size distribution is converted to CCN, and therefore, I think it needs to be made much more clear how you are doing this. In its current form, I find it difficult to follow.

P5, L25-27: I do not understand these sentences. What model are you referring to, and why would producing realistic supersaturations necessarily mean that you are also producing realistic CCN -- it also depends on if your aerosol number concentrations are realistic? Furthermore, why do the authors assume this model producing realistic supersaturations for stratiform clouds? I am confused by these sentences.

**3: More explanations of the results**

The authors present many comparisons, but often do not fully explain why there comparisons are the same or different. I have included some of these instances below.

P 11, L15-16: The authors state that nitrate is problematic to simulate (especially in spring). However, they just discussed how it was difficult to simulate in the fall. Therefore, why especially in the spring?

P11, L15-16: The authors state that nitrate is difficult to measure (especially in the fall). Why is nitrate more difficult to measure as compared to other species? Why especially in the fall?

P12, L3: The authors state that a “too large number of CCN could result from too many large particles”. However, instead of speculating, the authors have the simulation data and the conversions to particle number size distributions and the observations to confirm whether this is the case and explain why.

P13, L5-7: This seems to be one main result of this manuscript, but it primarily based on data and results from a prior study (Hande et al., 2016). A possible contribution from this study would be to explain why their model is producing too many particles in the size range from 80-110nm?

Figure 6 and P13 L13-20: One of the main conclusions of this study is that the vertical profiles compare well with the model and observations. However, all that is compared is an average over a month time period, and the authors conclude that it compares well because it is within a factor of 2 from the observations. Given the high temporal variability (seen in Figure 2), some temporal analysis should be considered here. Furthermore, it is difficult to read Figure 6 and see what the values and differences actually are.

Figure 7: One of the main results in this study is the shape of the CCN profile in Figure 7 and the differences between the model and observations. However, the authors don’t really explain why it is different.

**4: References**

**4A: Lack of reference**

There is very little background included in this manuscript, many of which is likely very important to the authors results. On P2 L27, the authors state that comparison of modeled CCN to observations of CCN are sparse. However, since this is the entire point of this study, the authors should present the background literature here, as those results will likely be relevant to the results in this study.

**4B: Inconsistent referencing practices.**

Furthermore, the authors have many inconsistent referencing practices. For example, they include references from some instruments and projects, and do not include references for others. The authors should take some time to include relevant references throughout the manuscript. Some specific examples are listed below:
P3 L19: What is GME?
P7 L22 - P8 L2: Reference should be given for ACTRIS and EMEP.

P8 L3-4: References should be given to substantiate what HOPE is about.
P8 L9-12: References
P9, L1-2: Reference for this new instrument that the authors are using in the helicopter platform. If there is no reference, the authors should include more details about this instrument here, such as its specifications and where it was placed on the helicopter. The authors provide a lot of details in terms of ground measurements, but provide no details about the helicopter measurements.
Table 2: Is the reference for this only personal communication? This seems rather weak, and given that this is the basis of the 1985 estimates, I think the authors should supply a better reference. Who is Kevin Hausman, for whom this personal communication was with?

**Other Comments:**

**Qualitative, subjective explanations:** In general, the authors make several subjective statements instead of providing qualitative results that would be more useful to the reader. I have included a few instances below.

P11, L4: The authors state that the analyses shown in Figure 2 are in good agreement. However, this is qualitative and subjective. Can the authors provide quantitative, objective measures to describe their results here?

P11, L16: “results compare satisfactorily.” This is subjective.

P11, L20: The authors use a 1-to-1 scatter plot to compare their results, and state that the model data underestimates CCN compared to the other methods. Can the authors add more quantitative details here. Underestimate by how much? Can the authors possibly do least squares fits for their data in Figure 3 to accomplish this or some other method in order to provide some more concrete results?

P12, L5: “Figures 2 and 3 show that the CCN number concentration is in similar agreement.” Can the authors quantify this result.

P13, L17-18: The authors state that the CCN number if overestimated by less than 50% and that this is quite well. This is subjective. Can the authors put this into context, in terms of how being 50% off would impact cloud processes or better explain why they think this is quite well?

P13, L19: “decrease considerably.” How much is considerably?

**Unclear discussions**

P12, L5-8: The authors first state that the results are in similar agreement, but then state that the differences are a factor of 2, and then state the difference is about 13%. I found these statements to be quite confusing. They must be comparing different things, but I wasn’t sure what was being compared. I think the authors should be careful about making it very clear what they are referring to throughout their manuscript, since there are a lot of different data being presented in this study.

P13, L17-18: The authors state that the ccn number if overestimated by less than 50%, but that seems to only be the case of the lidar measurements but not the ACTOS measurements. Can the authors clarify? A much different picture is present in the ACTOS measurements, which are not really discussed.

P11, L25-32: One large uncertainty in this analysis and comparison revolves around the size distributions being used in the model conversion from mass to number and whether these are representative for this site. As such, I think this should be included here.

P14, L1-2. The authors state that there is increased variability in CCN number conc. In the free troposphere, which I think is based on the increased 25%-75% quartile range in Figure 5. However, the authors state this is "mainly an expression of the considerably increased detection uncertainty". However, the same trend is present in the model data, which does not have this detection uncertainty. Can the authority clarify why they believe this to be the case?

**Additional Specific Comments**

P11, L22-24: Since SOA is so important to this region, why wasn't it included in the COSMO-MUSCAT simulations?

P2 L2: "For a realistic simulation of cloud adjustment..." What is cloud adjustment?

P2 L12-14: I do not understand what the authors are describing with these two sentences.

P2 L32: Why 1985? Can the authors provide background information here as to why one would be interested in the year 1985?

P3 L2: "This implies...". Can the authors clarify what this refers to? I am unclear and generally confused by this sentence?

P3 L4-9: The authors state they will be comparing 5 CCN estimates, but they seem to be estimates of different things. However, it is unclear why they are comparing these 5 items? As this is the introduction to the authors study, the authors should make this clear. What is the ultimate goal of this comparison and study?

P3 L18: What is a composition cycle?

P3 L19: The authors state that the simulations will be reinitialized every 48 hours, and provide other details, without yet describing the basics of the simulation. Therefore, it is difficult to understand why a 48 initialization time is reasonable. It is suggested that the authors provide more details at the beginning of this section to make this section more readable.

P4 L8: Why do the authors only use model data up to 8km? The authors should provide a reason to justify or explain why they are not using the full model data?

Figure 1: Why is only Melpitz shown, when Julich is also mentioned? Were measurements taken at the different city, Julich? I was confused about this throughout the manuscript.

P4 L2: "in order to be considered" for what?

P4 L3: The authors state the model simulations were run. Are these ICON-LES simulations, which were just mentioned in the previous sentence or COSMO-MUSCAT simulations? I think it is the COSMO-MUSCAT simulations, but can the authors make this more clear?

P4 L17-18. It seems that the authors imply that total mass concentrations is converted to total number concentrations before the log-normal size distribution parameters are set? However, these size distribution would be necessary for the initial conversion to mass to number. So I am ultimately confused by what the authors are actually doing?

P5 L7: Which of the values are not according to Hande et al. (2016) and why?

P5, L14: The authors mentioned that ammonium sulfate has a kappa of 0.6, but then use 0.51 in their table. Why?

P5, L33-34 -- this should be included in the introduction.

P5, L7: The authors reference Table 1, but mean to reference Table 2.

Table 4: How were the measured aerosol mass concentrations obtained? Can the authors include this information in the caption, similar to how they mention that the modeled concentrations came from the COSMO-MUSCAT simulations.

P10, L16-19: The authors discuss why the measured dust was larger than the modeled dust. Can the authors make these statements more clear. For example, why is it OK to assume the difference in total mass from the measured species must be from dust?

P11, L8: The authors state the nitrate was overestimated by a factor of 2. Can the authors present this information more clearly, possible on the same figure to make this more clear.

P11, L8-16: A majority of the discussion is focused on a 3 day period, which is very difficult to see in Figure 2. Can the authors create a new figure that zooms in on this period, to allow the reader to more easily follow along in the figure.

Figure 6: Why isn't the fall period shown here?

P14, L8. The authors state that for this calculation a vertical velocity of 1 m/s is used. I don't understand what this is referring too.

Figure 7: What supersaturation is used for this analysis?

Can the authors changes their units to $cm^3$ from $m^3$, such that it is easier to comprehend the values presented?