Responses to the reviewer #1:

This is an interesting and well written paper which discusses the variability of cloud water content, droplet number, and correlation between the two in stratiform clouds. The key findings seem to be the positive correlation between the water content and drop number, and the vertical dependence of this correlation, increasing towards cloud top.

The consequences of these results for autoconversion parametrizations is discussed, showing that including the correlation is crucially important for getting the enhancement factor correct, although neglecting variation in drop number entirely is a surprisingly good estimate. I recommend publication after addressing some minor comments I've listed below.

Reply: We thank this reviewer for the encouraging, insightful and constructive comments, which really help improve the manuscript significantly.

Before addressing your comments/questions below, first we would like to provide a summary of the major revisions made to the manuscript:

- We revised significantly the part about the bimodal joint distribution between qc and Nc in section 4. In particular, we pointed out that it is most likely just a coincidence that each side of the "V" shape track sampled one mode of the bimodal distribution. The along/across wind difference between the two sides is unlikely to be the cause of the bimodality.
- Three new cases that are either non-precipitating or weakly precipitating were added to the paper and they have no overall impacts on the conclusions. The flight track and radar reflectivity plots for all the cases, except for July 18, 2017, are provided in the supplementary material.
- A small bug in our code was found and fixed. This bug affects the computation of the EF based on lognormal distributions. As a result, the E_q based on the lognormal PDF agrees very well with the observation-based E_q (new Figure 6a), and the E based on the bivariate lognormal distribution agrees well with the observation-based E (new Figure 6d). Because of this, the Figure 8 was removed from the paper.
- Most figures are revised/updated per request/suggestion of the reviewers.

After these revisions, we think the paper is much improved and more focused, although the general conclusions still hold.

Minor comments

1. L45 and elsewhere - does the vertical dependence of EF really need to be accounted for, or would it suffice to simply get it right at cloud top? If the model is correctly representing the domination of accretion throughout the remainder of the cloud (e.g. as in Wood et al (2005b)), then autoconversion and the EF applied to it shouldn't matter so much here.

Reply: Indeed, the autoconversion and accretion have different relative importance at different vertical locations of the cloud. As we pointed out in the paper, most previous studies have ignored the vertical dependence of the EF and the consequential impacts on autoconversion rate simulation. In fact, this is the major motivation of this study. Based on observations, our study reveals that the EF at cloud top is significantly different (smaller) than the lower parts of the clouds and we also explain the underlying physics. Without a good understanding of the "truth" and underlying physics, how could we make sure that the model "*gets it right at cloud top*" and also "*correctly representing the domination of accretion throughout the remainder of the cloud*"?

2. L103 - it would be worth noting here or later that some CMIP6 models (e.g. Walters et al, 2019, GMD) have adopted variable enhancement factors (better options!) based on the recent work cited, so it's not quite as bad as presented.

Reply: Good point. We revised the paper to point out that the latest generation of GCMs may have adopted more advanced schemes to account for the EF than using a constant EF (around Line 141 in the revised manuscript). On the other hand, it is also important to note that Walters et al (2019) adopted the cloud-regime dependent and scale-aware schemes developed by Hill et al. (2015) and Boutle et al. (2014) to account for subgrid cloud variability. However, even these advanced schemes only consider the subgrid variability of qc only but ignore the variability of Nc and its co-variability with qc. Therefore, they also have important limitations.

3. L157 - you could also note here that this suggests an equal amount of time should be dedicated to EFs for accretion, something which certainly isn't the case in the literature!

Reply: Good point. We noted here that the vertical dependence is important for both autoconversion and accretion.

4. L271-273 - are your statistics for variance and correlation then only calculated in regions where qc>0.01g/m3, i.e. over the cloudy portion of the leg which may be <10km, or are they calculated over the entire >10km leg and include points where qc<0.01g/m3 (i.e. the zeros). As this is an important distinction, as it significantly affects the results (as shown in Witte et al, 2019, JAS) and also important for model developers to know to implement correctly in their microphysical parametrization (i.e. does the scheme work on in-cloud quantities or grid-box mean quantities).

Reply: indeed, this is important. All the analyses are based on in-cloud observations (e.g., regions with qc>0.01g/m3). We pointed this out as suggested.

5. Section 4.1 - how much do you think the results are affected by the non-stationarity of the cloud being sampled. i.e. for a model parametrization, you are interested in the variability at different heights within the cloud at the same time. But what you are sampling is the variability at different heights in the cloud at different times, up to several hours apart. The cloud is clearly evolving in this time, and how much might that evolution be affecting the results. For example, just looking at the variability in reflectivity in Figure 1 suggests there may be external factors affecting the cloud amount and drop number, and currently this variability is being attributed to the height at which the flight leg during that period happened to be at.

Reply: This is a very good question! It is certainty possible that the selected clouds in this study are "non-stationary". But it has to be noted that we observed similar vertical variations of qc and Nc in all four selected cases. It seems highly unlikely that the temporal evaluations of the clouds in all four cases conspire to confound our results in the same way. Based on this consideration, we assume that the temporal evolution of clouds is an uncertainty that could lead to random errors but not the overall conclusions. Of course, it is extremely difficult, if not impossible, to address this issue using air-borne in situ measurements alone due to their inherent limitations as you pointed out. Ground-based radars can provide a reasonable estimate of the vertical profile of qc at the temporal-resolution of 5-mintues or so. But currently there is no reliable retrieval of the vertical profile of Nc from ground radar, yet. The only useful tool in this regard is model simulation. In fact, we are currently simulating the July 18, 2017 case using a LES model which will hopefully help us understand both the spatial and temporal evolution of the clouds and thereby subgrid variability. But this is beyond the scope of this paper. Nevertheless, we pointed out this important limitation at section 4.1 and also at the end of the paper along with other limitations of this study.

6. Section 4.2 - do you have any hypothesis or theory why qc and Nc are positively correlated? Is there a physical mechanism for the correlation, or just something that happened to be the case for this study? One could imagine that areas with lower Nc will precipitate easier, thus lowering qc?

Reply: We do think there is some underlying physical processes that could lead to the positive correlation between qc and Nc at cloud top. As you pointed out, one possibility is the autoconversion process itself that converts the cloud water to rainwater and at the same time reducing the Nc. Another important possibility is the inhomogeneous mixing as a result of cloud top entrainment, which reduces the qc and Nc simultaneously leading to positive correlation between them. We pointed out these possibilities in the revised manuscript.

7. L491 - as with the comment above, do you think this really is a fortunate cancellation, or is there some underlying physical mechanism?

Reply: As mentioned above, we do believe there is some underlying physical processes that could lead to the positive correlation between qc and Nc at cloud top. As discussed in Eq. (5), because the qc and Nc is positively correlated the Ecov term is smaller than unity. In contrast, the En and Eq terms are always larger than unity. Therefore, it is expected (and explained in the paper) that the Ecov term tends to cancel the effects of En and Eq. But to what extent Ecov and En terms cancel out one another depends quantitively on the variability of Nc and its correlation with qc. But a more important point we would argue is that one should not rely on “fortunate cancellation”. It would be more robust and physically sound to take all three terms, i.e., Eq, En and Ecov, into consideration.

Typos

L41 - should say "effect that tends to make".

L60 - should probably say "e.g. Morrison and Gettleman" as theirs isn’t the first or only scheme out there.

L538 - this should say "Eq is significantly larger than E" I think.

Reply: these typos are all corrected as suggested. Thanks for pointing them out!