General response to both reviewers

We would like to thank you for the constructive comments and suggestions. We appreciate your time. As you will see, the manuscript has been revised follow each reviewer's suggestions. In the response, black are reviewers' comments and blue are our responses.

The major changes are:

1. Prompted by both reviewers, the cloud and rain water mixing ratios are now collocated, and the method is described in Appendix A in the revised manuscript. We combine remote sensing and adiabatic assumption to jointly estimate cloud and rain liquid water content (CLWC and RLWC) within the cloud layer. We also estimate the uncertainties in enhancement factor calculations came from our retrieval uncertainties and the results are shown in updated Figure 4.
2. Thanks for the suggestion by Reviewer 1, E_{auto} and E_{accr} in the revised manuscript are calculated at different layers of cloud to reveal the physical processes. We use averaged q_c in top five range gates to calculate E_{auto} and averaged q_c and q_r in five range gates around maximum reflectivity to calculate E_{accr}. Despite substantial changes in the data used in calculations, the trend of the new results is similar to previous one except that the values slightly increase. Thus, most of our conclusions still hold.
3. Instead of roughly assuming 10 m s^{-1} horizontal wind, we now use the mean wind speed within cloud layer from ARM merged sounding data. The terminology is changed from '2-hour…5-hour time intervals' to '60-km and 180-km model grids' as we mimic the specific model grid sizes instead of specific time intervals.
4. Mentioned by Reviewer 2, we did extensive literature reviews and rephrased sentences in both introduction and discussion sessions. Previous studies are properly cited and acknowledged.

Specific responses to Review 1

The goal of this study is to extend the results of studies such as Lebsock et al. (2013) and Boutle et al. (2014) on quantifying the effects of sub-grid scale inhomogeneity on microphysical process rates applied in GCMs from observations. The central tenet is that inhomogeneity varies with length scale and meteorological regime, thus the currently standard use of "universal" constants to characterize inhomogeneity cannot adequately describe subgrid-scale variability across a range of horizontal grid sizes or environmental conditions. The authors use a temporally extensive remote sensing dataset primarily sampling shallow convection over Graciosa Island in the Azores to develop "scale-aware" enhancement factors for the autoconversion and accretion processes (E_{auto} and E_{accr}, respectively) for several commonly used bulk microphysical parameterizations. These enhancement factors are estimated from compositing of variances and covariances of instantaneous retrievals of cloud and rain liquid water path (CLWP and RLWP, respectively) and cloud drop number concentration Nc over varying time windows, which the authors argue are roughly equivalent to a GCM horizontal grid length if a constant wind speed is assumed.

Thank you for the comments.
As stated in general response, we now use collocated q_c and q_r in the calculations and we use wind speed from merged sounding over certain periods to mimic model grid size.

I agree with the authors' basic premise that the use of constant values for E_{auto} and E_{accr} in GCM microphysics schemes is unrealistic and likely introduces precipitation biases similar (perhaps in magnitude if not sign) to assuming that grid-mean quantities (e.g. of N_c and cloud and rain liquid water mixing ratios q_c and q_r) are applicable to calculation of process rates in models with coarse grids (say horizontal grid length L greater than a kilometer or so). Furthermore, their assertion that enhancement factors should vary as a function of L as well as meteorological regime is well-stated, although they are not able to access independent information on aerosol-cloud interactions, which I suspect may be of comparable importance to the stability and LWP criteria analyzed.

Thank you for the comments.
In the revised manuscript, we keep the part of assessing enhancement factors for different grid sizes and add the uncertainties in enhancement factor calculations came from our retrieval uncertainties (Figure 4).

We agree that, within the same meteorological regime and similar LWP, aerosol-cloud-precipitation interactions will affect sub-grid cloud and precipitation variabilities. However, it is a challenge to quantitatively estimate this effect using our existing dataset, especially with large uncertainties in aerosol measurements during drizzling conditions. This is an interesting topic and worth to explore in the further.

For completeness and clarification, we add following to lines 464-466 in the revised manuscript: "The effect of aerosol-cloud-precipitation-interactions on cloud and precipitation sub-grid variabilities may be of comparable importance to meteorological regimes and precipitation status and deserves a further study."

Despite agreeing with the importance and timeliness of the premise of the manuscript, I have several major issues with the relevance of the observations to diagnosis of microphysical process inhomogeneity. Most importantly, the retrievals of cloud and rain/drizzle properties are not collocated; drizzle properties are only retrieved below cloud base. Cloud and drizzle properties are convolved within cloud such that what is classified as CLWP in fact includes contributions from in-cloud drizzle as well. Microphysical process rate equations assume coincident cloud and rain water mixing ratios (accretion) and coincident cloud water and drop number concentration (autoconversion), so unless it could be shown from some other dataset (LES? Aircraft observations? Maybe even a simplified 1D model?) that subcloud RLWP correlates highly with in-cloud RLWP and has similar magnitude, I have serious doubts about the physical relevance of the retrieved covariances. This may explain the apparently low ratios of cloud to rain water presented in the paper (see lines 33-34 and 291-293, Fig. 2e-f), although the authors give no "expected" value of this ratio for comparison.

Thanks for your comments and suggestions.
In the revised manuscript, collocated joint retrieval of cloud and drizzle LWC is employed to obtain q_c and q_r simultaneously. We updated the calculations accordingly, now using the variance and covariance of in-cloud mixing ratios.
In Figures 2e and 2f, we superimpose the ratio of layer-mean q_r to q_c and the ratios are both less than 15% in the two panels. This is also evident in Figure 1b that 10 times of q_r is still less than

q_c. The differences in magnitude are consistent with previous study (e.g., CloudSat and aircraft measurement presented by Boutle et al. 2014, their Figure 1a).
We add the following sentences to lines 316-318 in the revised manuscript: "In both panels, the ratios are less than 15%, which means that q_r can be one order of magnitude smaller than q_c. The differences in magnitude are consistent with previous CloudSat and aircraft results (e.g., Boutle et al. 2014)."

The use of column-integrated liquid water paths introduces further uncertainty because the partitioning of the collision-coalescence process into autoconversion and accretion sub-processes is heterogeneous in the vertical. In the shallow clouds typical of the ENA site, autoconversion will be dominant near cloud top where cloud droplets have reached a maximum size due to condensation and larger drizzle drops are rare while accretion dominates lower in cloud, where the drizzle drops initially formed at cloud top sediment and continue to grow by collecting cloud droplets. Erasing this coherent vertical variability by the use of integrated water paths may bias the results presented: in stratiform clouds, liquid water is at a maximum near cloud top (i.e. CLWP is weighted toward cloud top), such that the E_{accr} values in particular are using over-inflated liquid water values.

Thanks for your comments and suggestions.
We agree that autoconversion and accretion sub-processes dominate at different levels of cloud and it is physically reasonable to calculate them separately using different parts of the q_c and q_r profiles.

Following your suggestion, we add the followings to methodology part in lines 211-220 in the revised manuscript: "The autoconversion and accretion parameterizations partitioned from collision-coalescence process dominate at different levels in a cloud layer. Autoconversion dominates around cloud top where cloud droplets reach maximum by condensation and accretion is dominant at middle and lower parts of the cloud where drizzle drops sediment and continue to grow by collecting cloud droplets. Complying with the physical processes, we estimate autoconversion and accretion rates at different levels of a cloud layer in this study. The averaged q_c within the top five range gates (~215 m thick) are used to calculate E_{auto}. To calculate E_{accr}, we use averaged q_c and q_r within five range gates around the maximum radar reflectivity. If the maximum radar reflectivity appears at the cloud base, then five range gates above the cloud base are used."

I'm also confused about how the authors transformed liquid water paths to mixing ratios. They state that "CLWC [cloud liquid water content] values are transformed to q_c…by dividing by air density" (lines 191-192) and similar for qr (lines 194-195) but never define how they calculate CLWC or drizzle LWC. Are they dividing water path by cloud/drizzle shaft depth for an average value? Or are they applying the methods of Xie and Zhang (2015) and Wu et al. (2015) to the retrievals?

Thank you for the comments.
We first retrieve CLWC and RLWC profiles, then divided by air density vertical profiles calculated from temperature and pressure in merged sounding.

For clarification, we add the following sentences in the revised manuscript: "Using air density (ρ_{air}) profiles calculated from temperature and pressure in merged sounding, mixing ratio (q) can be calculated from LWC using $q(z) = LWC(z)/\rho_{air}(z)$." to lines 204-205 and 531-532 in methodology and Appendix A.

Is the retrieval of N_c vertically resolved? This part of the methodology is insufficiently described to understand what the authors did, and regardless, it doesn't address the issue that drizzle properties can only be retrieved below cloud using their approach.

Thank you for the comments.
In our study, N_c is not vertically resolved but is assumed to be constant in a cloud layer.
For clarification, the following is added to lines 199-200 in methodology part in the revised manuscript: "Cloud droplet number concentration (N_c) is retrieved using the methods presented in Dong et al. (1998, 2014a and 2014b) and are assumed to be constant in a cloud layer".
The drizzle properties in the revised manuscript are not from below cloud only, instead, q_r is now vertically resolved.

Finally, the authors made no attempt to quantify the uncertainty of the reported enhancement factors, such that I cannot make a determination as to whether their E_{auto} and E_{accr} are statistically distinct from the constant values introduced by Morrison and Gettelman (2008). This is particularly relevant to Figure 4.

Thanks for your comments.
To assess the uncertainty associated with the retrieved q_c and q_r, we vary q_c and q_r within their corresponding uncertainties, e.g., $(1 \pm 0.18)q_d$ and $(1 \pm 0.3)q_c$ and re-do the calculations. The mean differences are used as the boundaries of E_{auto} and E_{accr} as shown in Figure 4 in the revised manuscript.

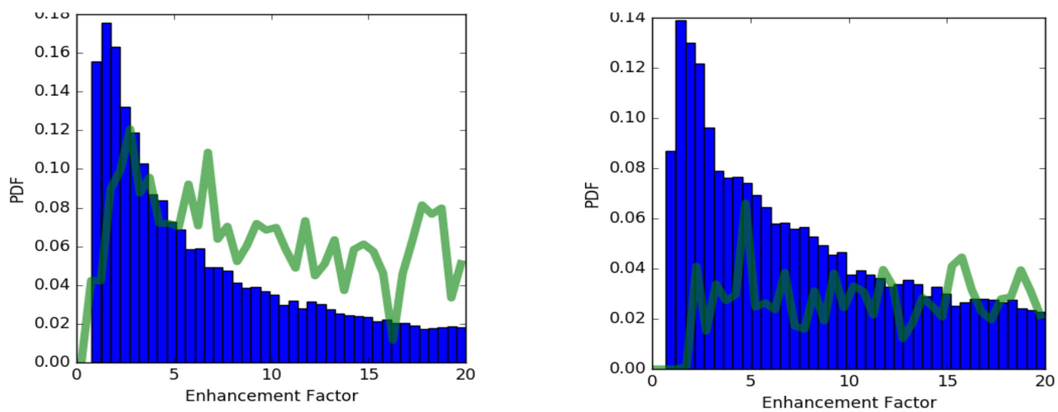
We add the following sentences to lines 206-210 to address the uncertainties of E_{auto} and E_{accr}: "The estimated uncertainties for the retrieved q_c and q_r are 30% and 18%, respectively (see Appendix A). We used the estimated uncertainties of q_r and q_c as inputs of Eqs. (4) and (7) to assess the uncertainties of E_{auto} and E_{accr}. For instance, $(1 \pm 0.3)q_c$ are used in Eq. (4) and the mean differences are then used as the uncertainty of E_{auto}. Same method is used to estimate the uncertainty for E_{accr}."

Also, in the discussion of Figure 4, we add the following sentences to lines 378-382 in the revised manuscript "The shaded areas represent the uncertainties of E_{auto} and E_{accr} associated with the uncertainties of the retrieved q_c and q_r. When model grid increases, the uncertainty slightly decreases. The prescribed E_{auto} is close to the upper boundary of uncertainties except for the 30-km grid, while the prescribed E_{accr} is significantly lower than the lower boundary."

I would also have liked to see the authors show the quantitative impact of treating qc and N_c individually with respect to calculating Eauto, as their derivation of Equation 4 assumes that the covariability of qc and N_c can be ignored. While the magnitude of Eauto is comparable for q_c or N_c individually, I don't have a good sense for what including variability of both variables implies for the predicted E_{auto} values. It's certainly a problem that CLWP and N_c are correlated in the

ARM dataset employed, but that doesn't change the fact that variability of N_c is likely substantial, especially for the longer time periods analyzed or in more cumuliform precipitation.

Thank you for the comments.
Due to N_c and LWP are highly correlated in our retrieval algorithm, we are currently unable to assess the covariance of q_c and N_c in autoconversion parameterization. In other words, the Nc is derived from LWP and other cloud variables (r_e and cloud thickness). We can use the following two figures to show why these results are artificially high: the E_{auto} calculated from the covariance of N_c and q_c for 60-km (left panel) and 180-km grid (right panel) sizes superimposed by average precipitation frequency in each bin can reach 40-50. Therefore, we only assess the individual effect of N_c as shown in Figures 2c and 2d, which are similar to the effect of q_c as shown in Figures 2a and 2b. For simplicity and clarity, only E_{auto} calculated from q_c are included in the discussions afterword.



For clarification, we add the following sentences to lines 305-308 in the revised manuscript "Because the E_{accr} values calculated from q_c and N_c are close to each other, we will focus on analyzing the results from q_c only for simplicity and clarity. The effect of q_c and N_c covariance, as stated in Section 4.1, is not presented in this study due to the intrinsic correlation in the retrieval (Dong et al., 2014a and 2014b and Appendix A of this study)."

In light of these concerns, I must recommend that this manuscript be **rejected** in its current form. A revised version of the manuscript only addressing autoconversion would be more feasible and would also be very useful to the parameterization development community, although as mentioned above, I would ask that the authors address the question of whether ignoring covariability of qc and N_c is a reasonable assumption.

Thank you for the comments.
In fact, we found very high covariance between the two variables, which is a result of our retrieval method in which N_c is derived from LWP and other cloud variables. As stated in the response to last comment, the results using N_c and q_c covariance could result in large variations of E_{auto} that are artificially high. To address this issue, independent retrieval methods for N_c and q_c are needed, that is what we plan to explore in the future.

Thanks for suggesting to use the jointly retrieved q_c and q_r, we think it is reasonable to keep accretion part in the manuscript.

I would be happy to review a revised and refocused manuscript. Until remote sensing datasets can unambiguously partition in-cloud condensed water into cloud and drizzle components, analysis of cloud-rain covariance from the present spatially disjoint cloud and rain retrievals cannot be used to inform accretion parameterizations.

Thank you for the comments.
Please see above responses that we tried to retrieve q_c and q_r profiles in the cloud and re-do the calculations.

A technique like that of Luke and Kollias (2013; doi:10.1175/JTECHD-11-00195.1) that uses skewness of the Doppler spectrum to differentiate between cloud and drizzle could be combined with a method similar to Frisch et al. (1998; doi:10.1029/98JD01827) to retrieve vertically-resolved profiles of cloud and rain water, albeit likely only in stratiform clouds. If such an approach could be developed, the analysis performed in this manuscript would be more tractable although it would likely need to be validated before application to the GCM cloud inhomogeneity problem given the amount of technical work necessary to provide confidence in the retrievals.

Thank you for the suggestions.
We used an alternative way as presented in Appendix A to retrieve CLWC and RLWC and then calculate q_c and q_r. The uncertainties of the retrieval are difficult to quantify without aircraft *in situ* data or other retrieval results. In the uncertainty analysis part, we used 18% as uncertainty for RLWC (rain LWC) from drizzle properties in Wu et al. (2015) and 30% for CLWC (cloud LWC) from cloud properties in Dong et al. (2014a and 2014b). The actual uncertainties may vary depend on the accuracy of merged sounding data and WACR detectability near cloud base.

In Appendix A, we add the following sentences to address the retrieval uncertainties to lines 517-526: "It is difficult to quantitatively estimate the retrieval uncertainties without aircraft in situ measurements. For the proposed retrieval method, 18% should be used as uncertainty for RLWC from drizzle properties in Wu et al. (2015) and 30% for CLWC from cloud properties in Dong et al. (2014a and 2014b). The actual uncertainty depends on the accuracy of merged sounding data, the detectability of WACR near cloud base and the effect of entrainment on cloud adiabaticity during drizzling. In the recent aircraft field campaign, the Aerosol and Cloud Experiments in Eastern North Atlantic (ACE-ENA) was conducted during 2017-2018 with a total of 39 flights over the Azores, near the ARM ENA site on Graciosa Island. These aircraft in situ measurements will be used to validate the ground-based retrievals and quantitatively estimate their uncertainties in the future."

References:
Boutle, I. A., Abel, S. J., Hill, P. G., and Morcrette, C. J.: Spatial variability of liquid cloud and rain: Observations and microphysical effects. Quart. J. Roy. Meteor. Soc., 140, 583–594, doi:10.1002/qj.2140, 2014.

Dong X., Ackerman, T. P., and Clothiaux, E. E.: Parameterizations of Microphysical and Radiative Properties of Boundary Layer Stratus from Ground-based measurements, J. Geophys. Res., 102, 31,681-31,393, 1998.

Dong, X., Xi, B., Kennedy, A., Minnis, P. and Wood, R.: A 19-month Marine Aerosol-Cloud_Radiation Properties derived from DOE ARM AMF deployment at the Azores: Part I: Cloud Fraction and Single-layered MBL cloud Properties, J. Clim., 27, doi:10.1175/JCLI-D-13-00553.1, 2014a.
Dong, X., Xi, B., and Wu, P.: Investigation of Diurnal Variation of MBL Cloud Microphysical Properties at the Azores, J. Clim., 27, 8827-8835, 2014b.
Wu, P., Dong, X. and Xi, B.: Marine boundary layer drizzle properties and their impact on cloud property retrieval, Atmos. Meas. Tech., 8, 3555–3562. doi: 10.5194/amt-8-3555-2015, 2015.