Interactive comment on "An evaluation of clouds and radiation in a Large-Scale Atmospheric Model using a Cloud Vertical Structure classification" by Dongmin Lee et al.

Anonymous Referee #2

Received and published: 4 November 2019

The simulated cloud radiative effects (CRE) are commonly biased in most climate models. This study, with the aid of both CloudSat/CAPLIPSO retrievals and the implemented subcolumn cloud generator in NASA's GEOS-5, explores the CRE biases in association with cloud vertical structure classification. Results show that while the simulation of global CRE is in much agreement with observation, the CRE due to different cloud classification is not. Moreover, by decomposing model's CRE errors into components stemming from biases in RFO and cloudy-column CRE, the relatively good simulations of global grid-mean CRE largely benefit from compensating errors in these two terms. The method introduced in this paper can be used in other models to explore their CRE behavior, thus beneficial for the modeling community.

While I generally find the manuscript suitable for publication in GMD, further improvements are needed before the manuscript is accepted. Below I have included a list of the major comments that I think should be addressed, followed by a list of more specific comments.

Major comments:

1. The study uses 2B-FLXHR-LIDAR product as a guiding reference for CRE, which was obtained by invoking radiative transfer algorithm operating on thermodynamical fields from re-analysis and cloud properties from CloudSat/CAPLIPSO retrievals. As the authors mentioned, the SW CREs in 2B-FLXHR-LIDAR is strongly time dependent. It is worth to add CERES data as a reference as well, since a great many models are commonly tuned to resemble CERES observations.

A: The CVS class occurrences and associated radiative fluxes from 2B-FLXHR-LIDAR (obtained in the way that the reviewer describes) are defined at the scale of individual CloudSat profiles, i.e., at approximately 2 km horizontal resolution. Producing the observed global/zonal average results in our study still requires this flux-CVS type mapping at the CloudSat profile horizontal scale. That is the main reason we cannot use CERES: even at the footprint level, CERES radiative fluxes are coarser than the scale at which a CVS occurrence is defined, so such a mapping cannot be achieved. In the model, it is the subcolumn generator that allows us to define CVS occurrences within a grid cell.

2. Besides cloud overlap assumptions, the cloudiness vertical profile per se is important for the determination of CVS classification. The authors are suggested to replenish the role of layer cloud fraction when revising the paper. The CVS classification may suffer from poor representation of subgrid cloud condensation and/or overlap assumption. In addition, the vertical resolution in GCMs is typically coarser than that in CloudSat/CALIPSO retrievals, which to some extent plays an important role in calculating RFO. The authors need point out this in the paper.

A: The CVS classification of Oreopoulos et al. (2017) used in this paper does not depend on cloud fraction overlap because layer cloud fractions are either zero or one in the 2B-CLDCLASS-

LIDAR profiles. In other words, we are considering only cloud *occurrence* overlap at different vertical layers to assign cloud profiles to a CVS class. We just wanted to clarify this point once again. But we believe that what the reviewer refers to here is the quality of the model cloud profiles passed to the subcolumn generator. These may indeed affect our results. So, what we are really evaluating here is both the quality of the model's cloud scheme in terms of mean profiles, but also the quality of the subcolumn generator to translate these mean profiles into subgrid profiles. The latter issue is partially addressed in this paper by changing the overlap assumption in the subcolumn generator from generalized to maximum-random. One can of course come with other options to create the subgrid columns, and even for generalized overlap, the decorrelation length is a tunable parameter. But both versions of the generator receive as input the same mean profiles, which may be problematic. There is flexibility in changing the generator, but testing different model cloud ("moist") schemes is a bigger effort. Good point about different vertical resolutions. We have added a sentence in the discussion of the concluding section to make the point that "garbage-in" (deficient mean cloud profiles) – "garbage out" (the subcolumn generator will not rectify faulty input profiles).

3. When comparing the two overlap assumptions, the GN assumption yields more clouds than MR in almost all cloud regimes except for isolated low clouds in extratropics. Why does this occur? Is this because clouds in this region are nonadjacent separated by clear skies that have a vertical distance smaller than specified decorrelation length in GN?

A: What the reviewer suggests is certainly possible and makes sense. As shown in Oreopoulos et al. (2017), when multiple distinct cloud layers exist in a standard layer, these are still treated as a single entity in the CVS classification. So, the MR scheme sees them as separate layers (random overlap), while the generalized scheme gives a combined fraction less than random because the separation distance is small.

Specific comments:
1. The abbreviation "RFO" is not fully spelled at the first place in the main text.

A: We fixed this, thanks.

2. P9L1 regions pronounced orography -> regions with pronounced orography
P9L17 exceeds -> exceed 4. P27 As Fig. 3-> As Fig. 4.

A: Corrections applied, thanks.