Reviewer #1:

The authors explore the sub-grid variability assumed in COSP, which many studies use to compare observations to models. Use of SPCAM at 4km resolution allows the authors to examine the impact of resolving sub-grid variability on COSP.

I really like this paper and think it is very important to get it out there to allow people to better understand the abilities of COSP and that it shouldn't be applied fecklessly to any given model. Frequently COSP is used in studies as some sort of magical talisman that bridges models and observations. This is rarely questioned as far as I can tell. As the authors point out in line 91 page 4, there are some basic resolution issues in coupling a GCM to COSP and trying to pull out something like a satellite pixel. I would almost suggest that the authors move their comments on line 91-98 into the abstract somehow so that people who just skim it will have this brought to their attention as it is critically important. However, this change is not required scientifically and may be disregarded by the authors. This paper will be a very useful reference in the COSP documentation for people trying to set their model up to run with COSP.

Thank you very much for the encouraging remarks. In our revision, we have revised our manuscript based on your helpful advices.

Line 126- convectional=convective

Ans: This correction is done.

Line 129- it is worth noting that this is still in the so-called convective grey zone, for example: Field et al. (2017). Do you think your results would change much if you doubled your grid size?

Ans: Yes, 4-km resolution is still in the so-called convective grey zone. As mentioned in Field et al. (2017), it is common practice for models operating in the convective Grey Zone to simply switch off the convection parameterization somewhere in the resolution ranging between 500 and 5km. No, we don't think our results would change much if we doubled the grid size.

Line 186 'sub-columns are'

Ans: This correction is done.

Line 262- Although not required, the authors might consider how this might contextualize results such as Nam et al. (2012).

Ans: We have added a sentence to contextualize the results from previous studies such as Nam and Quaas (2012) in our revised manuscript.

Line 374- The authors have focused on the warm rain process representation. This may be a very ignorant comment on my part, but I would be interested in how the evaluation of the first indirect effect in GCMs might be affected by the assumptions in homogeneous COSP. For example, most empirical studies of the first indirect effect utilize level 3 gridded data (McCoy et al., 2017a;Gryspeerdt et al., 2017;Bellouin et al., 2013;Quaas et al., 2008;Quaas et al., 2009), either using observed AOD/AI (Gryspeerdt et al., 2017) or reanalysis aerosol mass (McCoy et al., 2017a;McCoy et al., 2017b). These studies compare to level 3 aggregated cloud and aerosol from models and make statements regarding the ability of models to represent the first indirect effect. If the authors could comment on whether this is a valid approach that would be highly informative.

Ans: We have compared the simulated total cloud fraction by the MODIS, CALIPSO and CloudSat simulators, and the in-cloud properties by the MODIS simulator for the SPCAM5 and SPCAM5-Homogeneous simulations. As shown in the below figure (Figure S1), all the

simulated cloud properties are influenced by the sub-grid cloud variability but to different extents. The CloudSat simulation is affected most notably since the calculation of radar reflectivity is strongly sensitive to the inhomogeneous distribution of cloud droplet size. To what extent these differences will influence the aerosol-indirect effect evaluation is beyond the scope of our study, but it'd be wise to keep in mind this potential uncertainty.

Figures 2 c-d are **somewhat hard to parse.**

Ans: Figure 2c shows the distribution of large-scale (red plus signs for frac_out=1) and convective (blue plus signs for frac_out=2) cloud among the sub-columns generated by the SCOPS scheme, the variable frac_out is produced in the scops.f routine. The sub-column at certain vertical level is stratiform cloudy if frac_out =1, or connective cloudy if frac_out=2 at that vertical level. Figure 2d shows the distribution of large-scale (red plus signs for prec_frac=1), convective (blue plus signs for prec_frac=2), and mixed (green plus signs for prec_frac=3) precipitation among the sub-columns generated by the SCOPS-PREC scheme (i.e., prec_frac from prec_scops.f). We have added more detailed captions and explanations about Figure 2 in our revised manuscript to make them easy to parse. Thank you.

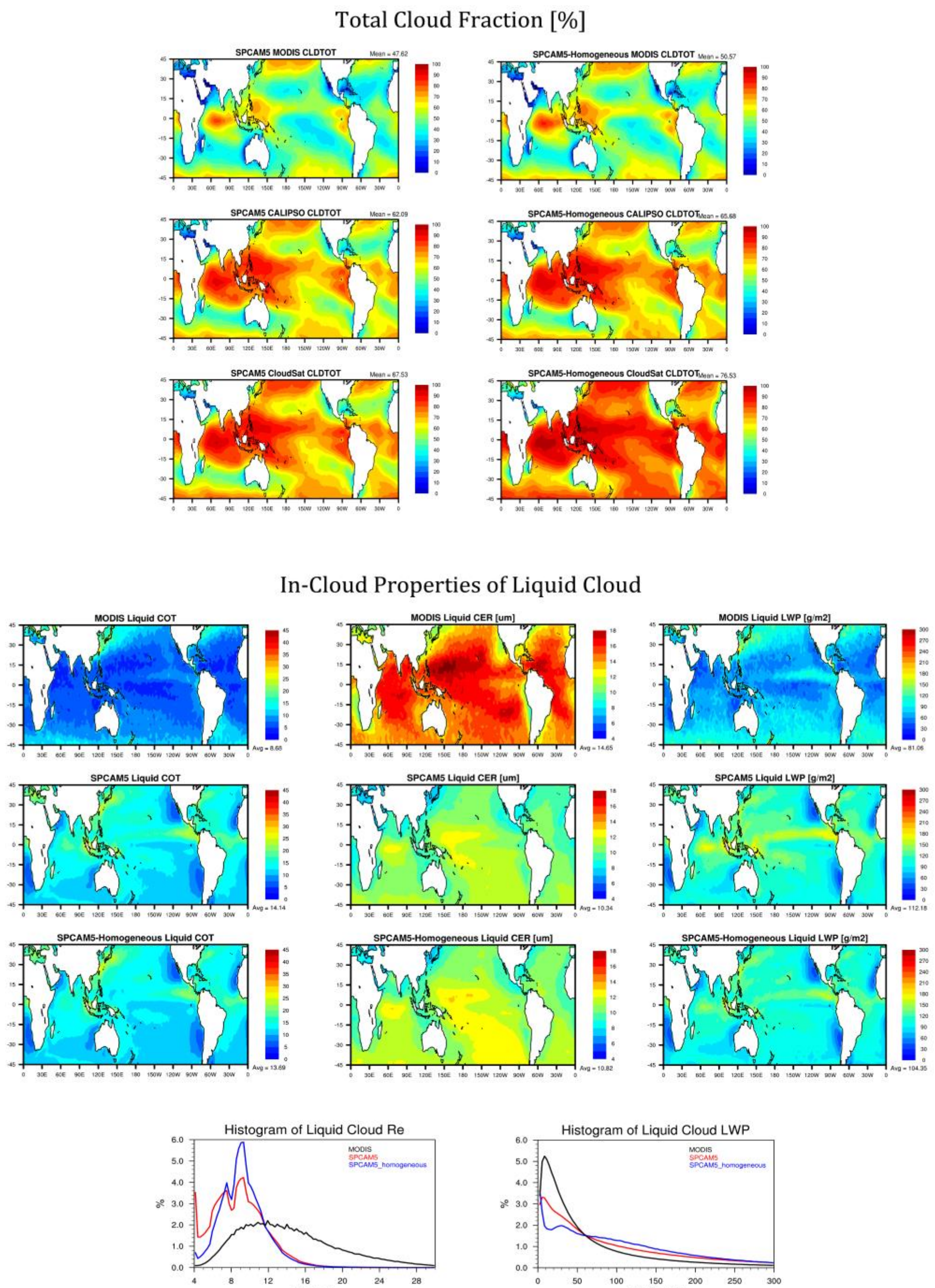


Figure S1. Top panels: Total cloud fraction from MODIS simulator, CALIPSO simulator and CloudSat simulator in SCPAM5 and SPCAM5-Homogeneous simulations. Middle panels: In-cloud properties of liquid cloud in the MODIS observation, SPCAM5 MODIS simulation, and SPCAM5-Homogeneous MODIS simulation. Low panels: Histograms of Liquid cloud effective radius and LWP over tropical ocean in the MODIS observation, SPCAM5 MODIS simulation, and SPCAM5-Homogeneous MODIS simulation.

Reviewer #2:

General Comments:

This is a well-written paper that clearly demonstrates the importance of considering the sub-grid variability of cloud and precipitation when applying the COSP MODIS and CLOUDSAT satellite simulators. The authors demonstrate that the radar reflectivities derived from the sub-grid CRM cloud and precipitation properties, versus the grid mean properties, are vastly different and excluding sub-grid variations can lead to misinterpretation of model performance (leading to the conclusion that the drizzle or rain is triggered too frequently).

I find this work to be important as its results will impact the analysis of CMIP6 model simulations, many of which will very likely be using the oversimplified COSP subcolumn generator in version 1.4.

Thank you very much for the encouraging comments. We have revised our manuscript based on your constructive advices.

Specific Comments:

Line 83: What is the pixel resolution of MODIS?

Ans: The MODIS data we used in this study is the C6 Aqua MODIS products that include the 1km geolocation products and the cloud mask product (Ackerman et al., 1998). As mentioned in Section 2.3 of our manuscript, we collocated 5 years (2006 ~ 2010) of pixel-level (i.e., level-2) MODIS and CloudSat observations using the collocation scheme developed in Cho et al. (2008). We aggregated these CloudSat and MODIS collocated level-2 data to the level-3 (gridded) data with the horizontal resolution as that in our CAM5.3, which is 1.9° latitude × 2.5° longitude.

Line 129: A more detailed description regarding clouds and microphysics in SPCAM would be appreciated. How can microphysical processes be resolved at 4km? Does SPCAM use the Morrison and Gettelman (2008) microphysical scheme mentioned?

Ans: As suggested, we have added a short paragraph to describe the physical parameterizations in SPCAM. SPCAM uses the two-moment cloud microphysics scheme of Morrison et al. (2005) to resolve microphysical processes at 4km. The Morrison and Gettelman (2008) microphysical scheme is based loosely on the approach of Morrison et al. (2005).

Fig 2 (& related Caption) - Add experiment name to plot and caption. In regards to Subplot e) Add title to columns (ie mixing ratio / eff. radius). (FYI - I like that the authors added the variable and routine 'fracout from scops.f' to the caption. This will be very helpful for other modelers).

Ans: We have modified Figure 2 as suggested in our revised manuscript.

Line 218: Consider sharing the modification to COSP to the community.

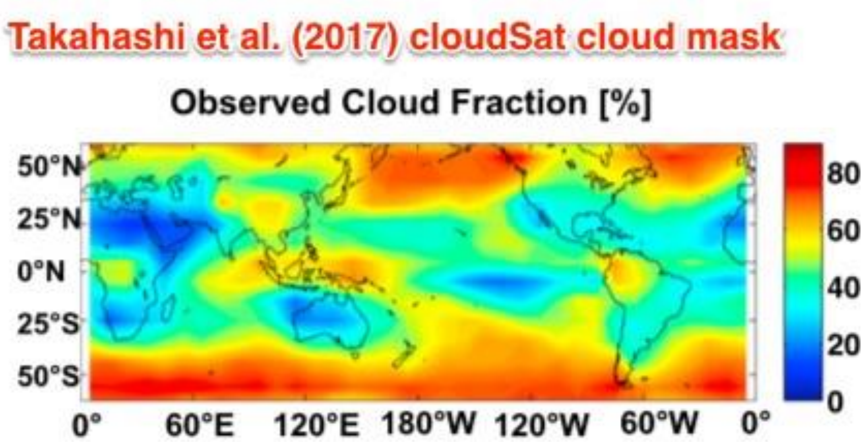
Ans: The latest version of COSP (v2.0) might have already implemented the capability for sub-column sampling. But yes, we will share our finding with the COSP to the community (through personal communication and COSP user google group https://groups.google.com/forum/#!forum/cosp-user).

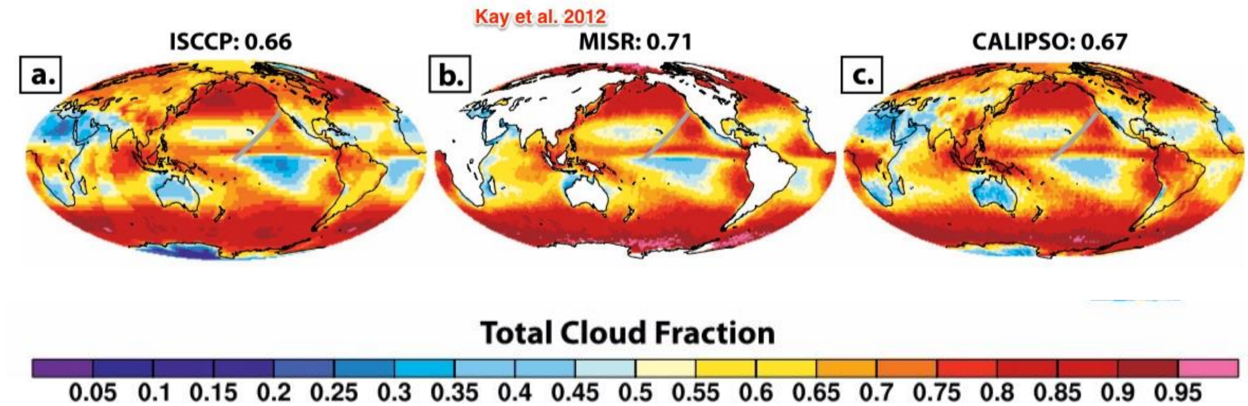
Line 274-247: The obs. pdf needs to be further analyzed. Finding that CloudSat only detects 54% of collocated warm clouds MODIS detects is a significant problem that needs to understood/explained further. Are you saying that a large chunk of the 46% of undetected clouds are too thin and can explain the sharp decline in the pdf around -40 to -25dBZ? If so, how often are warm liquid clouds too thin to be detected by CloudSat (check with CALIPSO)? Ground clutter really only influences the lowest approx. 1_km. This would imply that nearly half (or some significant fraction) of the clouds MODIS detects are within the lowest 1_km (again, check with CALIPSO). Also, is there a way of checking for frequency of attenuation (for a given altitude) in the Observations? While I understand this will very likely not change the results of this plot, it is important to note which types of clouds are being eliminated in the observations.

Ans:

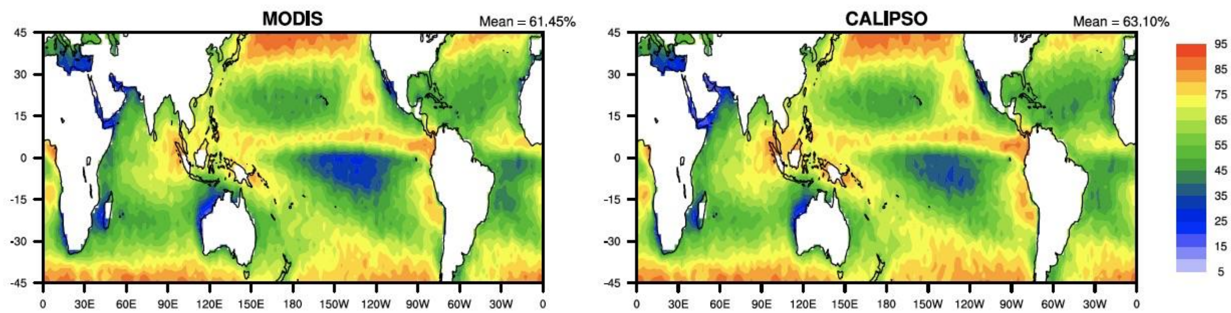
Yes, using only the CloudSat cloud mask alone (i.e., 2B-GEOPROF product) would miss significant amount of liquid-phase clouds. In addition to surface cluttering problem, some clouds are either too thin or their particle sizes are too small to generate detectable radar echo (i.e., >−30dBz), and therefore would be missed by CloudSat. Though it should be kept in mind that CloudSat is designed to detect "hydrometer" which include both cloud and more importantly precipitation. Moreover, as you pointed out, CloudSat is flying side by side with CALIPSO which is much more sensitive to thin clouds. That is why the CloudSat team developed the 2B-GEOPROF-LIDAR product which combined the CALIPSO and CloudSat for cloud detection. In our study, we mainly use CloudSat to detect drizzle and use MODIS to detect clouds.

We could not find a published reference to quantify and explain the clouds missed by CloudSat (maybe because it is well known?), but we found two papers, one by Takahashi et al. (2017) who used CloudSat only cloud mask and the other by Kay et al. (2012) who used ISCCP, MISR and CALIPSO cloud masks. Below are the cloud fractions from the two study. It is evident that the CloudSat only cloud mask detects significantly lower cloud fraction than CALIPSO or the other two passive sensors. In particular, over the stratocumulus cloud regions (e.g., SE pacific off coast Peru and NE pacific off coast of California) the cloud fraction based on CloudSat alone is only around 50% much lower than the CALIPSO values ~ 75%~85%.





One more point to note is that many studies have shown that the MODIS cloud mask agrees well with CALIPSO cloud mask. In fact, in our early paper, Song et al. (2018), we found that the total cloud fraction from MODIS is about 61% between 45S and 45N, only 2% lower than the CALIPSO cloud fraction. See Figure below.



The cloud masking product of CloudSat is beyond the scope of this study. We believe our result is robust and consistent with previous studies.

Takahashi, H., M. Lebsock, K. Suzuki, G. Stephens, and M. Wang (2017), An investigation of microphysics and subgrid-scale variability in warm-rain clouds using the A-Train observations and a multiscale modeling framework, Journal of Geophysical Research-Atmospheres, 138(669), 2151.

Kay, J. E. et al. (2012), Exposing Global Cloud Biases in the Community Atmosphere Model (CAM) Using Satellite Observations and Their Corresponding Instrument Simulators, Journal of Climate, 25(15), 5190–5207, doi:10.1175/JCLI-D-11-00469.1.

Song, H., H. Song, Z. Zhang, P.-L. Ma, S. J. Ghan, and M. Wang (2018), An Evaluation of Marine Boundary Layer Cloud Property Simulations in the Community Atmosphere Model Using Satellite Observations: Conventional Subgrid Parameterization versus CLUBB, Journal of Climate, 31(6), 2299–2320, doi:10.1175/JCLI-D-17-0277.1.

Line 339 / Section 4: Can you state which other COSP simulators, and how a few selected variables, would be influenced by the sub-grid cloud variability (and in-cloud microphysical properties)? Otherwise, I recommend changing broad statements of about the COSP simulator to more specific statements regarding the CloudSat simulator.

Ans: COSP includes simulators that are compatible with the ISCCP, PARASOL, CALIPSO, MISR, MODIS, and CloudSat observational products. In our research, we mainly focus on three COSP simulators: MODIS, CALIPSO, and CloudSat. As shown in the below figure (Figure S1), the simulated total cloud fraction by these three simulators, and the in-cloud properties by the MODIS simulator are all influenced by the sub-grid cloud variability but with different magnitudes. The CloudSat simulation is affected most obviously since the calculation of radar reflectivity is strongly sensitive to the inhomogeneous distribution of cloud droplet size.

Section 4: It needs to be emphasized that the 'sub-grid variability of mass and microphysics within each hydrometeor type' is key.

Ans: As suggested, we have added a sentence in Section 4 to emphasize the key role of sub-grid variability of mass and microphysics within each hydrometeor type.

Double check references.

Ans: We have double checked the references and made some corrections. Thank you.

Reviewer #2:

General Comments:

This is a well-written paper that clearly demonstrates the importance of considering the sub-grid variability of cloud and precipitation when applying the COSP MODIS and CLOUDSAT satellite simulators. The authors demonstrate that the radar reflectivities derived from the sub-grid CRM cloud and precipitation properties, versus the grid mean properties, are vastly different and excluding sub-grid variations can lead to misinterpretation of model performance (leading to the conclusion that the drizzle or rain is triggered too frequently).

I find this work to be important as its results will impact the analysis of CMIP6 model simulations, many of which will very likely be using the oversimplified COSP subcolumn generator in version 1.4.

Thank you very much for the encouraging comments. We have revised our manuscript based on your constructive advices.

Specific Comments:

Line 83: What is the pixel resolution of MODIS?

Ans: The MODIS data we used in this study is the C6 Aqua MODIS products that include the 1km geolocation products and the cloud mask product (Ackerman et al., 1998). As mentioned in Section 2.3 of our manuscript, we collocated 5 years (2006 ~ 2010) of pixel-level (i.e., level-2) MODIS and CloudSat observations using the collocation scheme developed in Cho et al. (2008). We aggregated these CloudSat and MODIS collocated level-2 data to the level-3 (gridded) data with the horizontal resolution as that in our CAM5.3, which is 1.9° latitude × 2.5° longitude.

Line 129: A more detailed description regarding clouds and microphysics in SPCAM would be appreciated. How can microphysical processes be resolved at 4km? Does SPCAM use the Morrison and Gettelman (2008) microphysical scheme mentioned?

Ans: As suggested, we have added a short paragraph to describe the physical parameterizations in SPCAM. SPCAM uses the two-moment cloud microphysics scheme of Morrison et al. (2005) to resolve microphysical processes at 4km. The Morrison and Gettelman (2008) microphysical scheme is based loosely on the approach of Morrison et al. (2005).

Fig 2 (& related Caption) - Add experiment name to plot and caption. In regards to

Subplot e) Add title to columns (ie mixing ratio / eff. radius). (FYI - I like that the authors added the variable and routine 'fracout from scops.f' to the caption. This will be very helpful for other modelers).

Ans: We have modified Figure 2 as suggested in our revised manuscript.

Line 218: Consider sharing the modification to COSP to the community.

Ans: The latest version of COSP (v2.0) might have already implemented the capability for sub-column sampling. But yes, we will share our finding with the COSP to the community (through personal communication and COSP user google group https://groups.google.com/forum/#!forum/cosp-user).

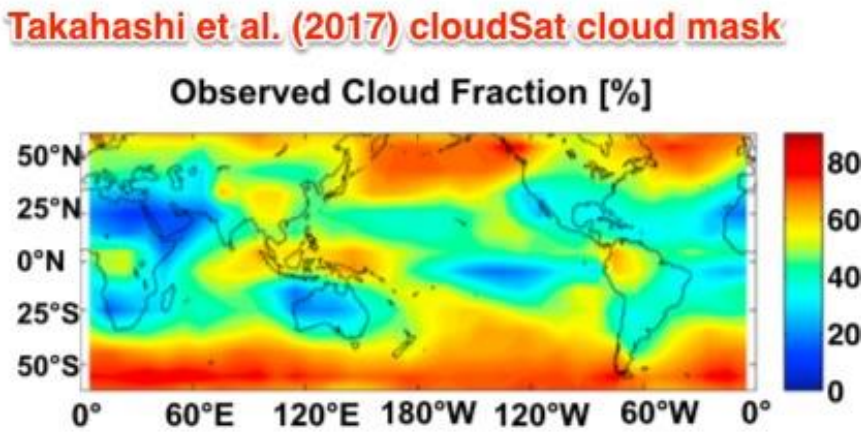
Line 274-247: The obs. pdf needs to be further analyzed. Finding that CloudSat only detects 54% of collocated warm clouds MODIS detects is a significant problem that needs to understood/explained further. Are you saying that a large chunk of the 46% of undetected clouds are too thin and can explain the sharp decline in the pdf around -40 to -25dBZ? If so, how often are warm liquid clouds too thin to be detected by CloudSat (check with CALIPSO)? Ground clutter really only influences the lowest approx. 1_km. This would imply that nearly half (or some significant fraction) of the clouds MODIS detects are within the lowest 1_km (again, check with CALIPSO). Also, is there a way of checking for frequency of attenuation (for a given altitude) in the Observations? While I understand this will very likely not change the results of this plot, it is important to note which types of clouds are being eliminated in the observations.

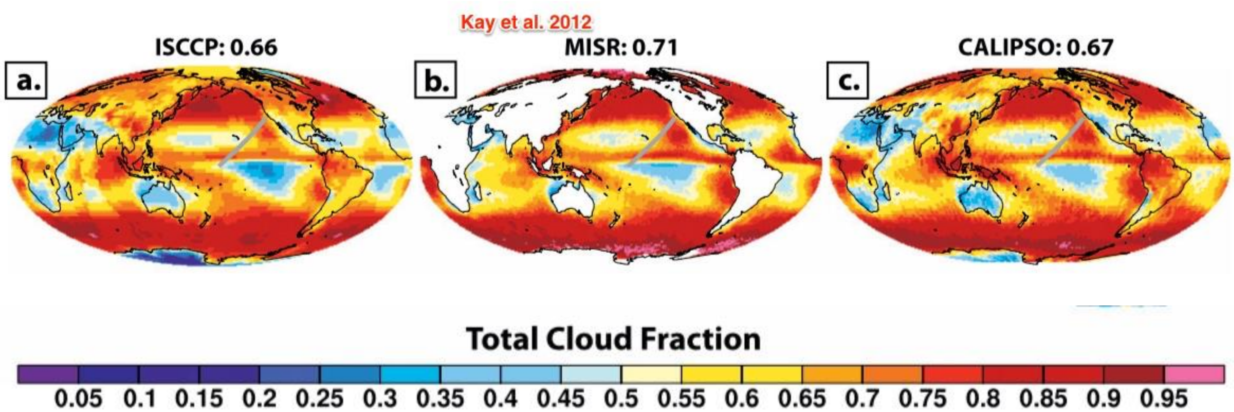
Ans:

Yes, using only the CloudSat cloud mask alone (i.e., 2B-GEOPROF product) would miss significant amount of liquid-phase clouds. In addition to surface cluttering problem, some clouds are either too thin or their particle sizes are too small to generate detectable radar echo (i.e., >−30dBz), and therefore would be missed by CloudSat. Though it should be kept in mind that CloudSat is designed to detect "hydrometer" which include both cloud and more importantly precipitation. Moreover, as you pointed out, CloudSat is flying side by side with CALIPSO which is much more sensitive to thin clouds. That is why the CloudSat team developed the 2B-GEOPROF-LIDAR product which combined the CALIPSO and CloudSat for cloud detection. In our study, we mainly use CloudSat to detect drizzle and use MODIS to detect clouds.

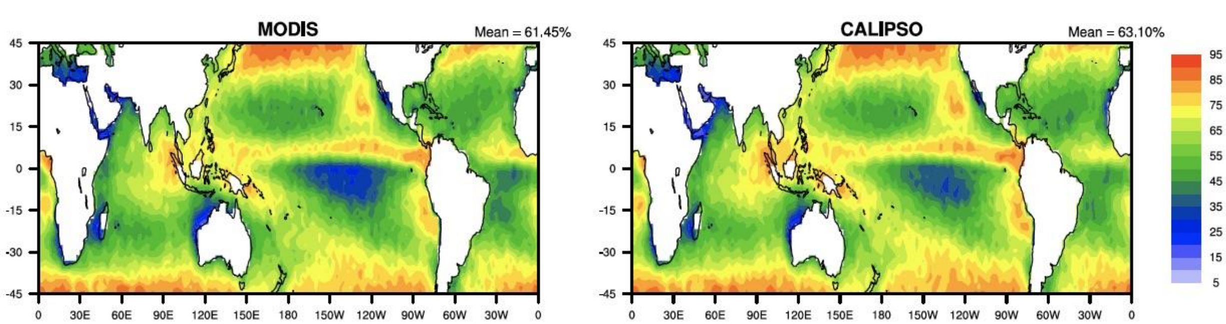
We could not find a published reference to quantify and explain the clouds missed by CloudSat (maybe because it is well known?), but we found two papers, one by Takahashi et al. (2017) who used CloudSat only cloud mask and the other by Kay et al. (2012) who used ISCCP, MISR and CALIPSO cloud masks. Below are the cloud fractions from the two study. It is evident that the CloudSat only cloud mask detects significantly lower cloud fraction than CALIPSO or the other two passive sensors. In particular, over the

stratocumulus cloud regions (e.g., SE pacific off coast Peru and NE pacific off coast of California) the cloud fraction based on CloudSat alone is only around 50% much lower than the CALIPSO values ~ 75%~85%.





One more point to note is that many studies have shown that the MODIS cloud mask agrees well with CALIPSO cloud mask. In fact, in our early paper, Song et al. (2018), we found that the total cloud fraction from MODIS is about 61% between 45S and 45N, only 2% lower than the CALIPSO cloud fraction. See Figure below.



The cloud masking product of CloudSat is beyond the scope of this study. We believe

our result is robust and consistent with previous studies.

Takahashi, H., M. Lebsock, K. Suzuki, G. Stephens, and M. Wang (2017), An investigation of microphysics and subgrid-scale variability in warm-rain clouds using the A-Train observations and a multiscale modeling framework, Journal of Geophysical Research-Atmospheres, 138(669), 2151.

Kay, J. E. et al. (2012), Exposing Global Cloud Biases in the Community Atmosphere Model (CAM) Using Satellite Observations and Their Corresponding Instrument Simulators, Journal of Climate, 25(15), 5190–5207, doi:10.1175/JCLI-D-11-00469.1.

Song, H., H. Song, Z. Zhang, P.-L. Ma, S. J. Ghan, and M. Wang (2018), An Evaluation of Marine Boundary Layer Cloud Property Simulations in the Community Atmosphere Model Using Satellite Observations: Conventional Subgrid Parameterization versus CLUBB, Journal of Climate, 31(6), 2299–2320, doi:10.1175/JCLI-D-17-0277.1.

Line 339 / Section 4: Can you state which other COSP simulators, and how a few selected variables, would be influenced by the sub-grid cloud variability (and in-cloud microphysical properties)? Otherwise, I recommend changing broad statements of about the COSP simulator to more specific statements regarding the CloudSat simulator.

Ans: COSP includes simulators that are compatible with the ISCCP, PARASOL, CALIPSO, MISR, MODIS, and CloudSat observational products. In our research, we mainly focus on three COSP simulators: MODIS, CALIPSO, and CloudSat. As shown in the below figure (Figure S1), the simulated total cloud fraction by these three simulators, and the in-cloud properties by the MODIS simulator are all influenced by the sub-grid cloud variability but with different magnitudes. The CloudSat simulation is affected most obviously since the calculation of radar reflectivity is strongly sensitive to the inhomogeneous distribution of cloud droplet size.

Section 4: It needs to be emphasized that the ’sub-grid variability of mass and microphysics within each hydrometeor type’ is key.

Ans: As suggested, we have added a sentence in Section 4 to emphasize the key role of sub-grid variability of mass and microphysics within each hydrometeor type.

Double check references.

Ans: We have double checked the references and made some corrections. Thank you.

The Importance of Considering Sub-grid Cloud Variability When Using Satellite Observations to Evaluate the Cloud and Precipitation Simulations in Climate Models

Hua Song[1], Zhibo Zhang[1, 2*], Po-Lun Ma[3], Steven Ghan[3], and Minghuai Wang[4]

1. Joint Center for Earth Systems Technology, UMBC, Baltimore, MD
2. Physics Department, UMBC, Baltimore, MD
3. Atmospheric Sciences and Global Change Division, Pacific Northwest National Laboratory, Richland, WA
4. Institute for Climate and Global Change Research & School of Atmospheric Sciences, Nanjing University, Nanjing, China

Corresponding Author:

Dr. Zhibo Zhang
Email: Zhibo.Zhang@umbc.edu
Phone: +1 (410) 455 6315

Abstract

Satellite cloud observations have become an indispensable tool for evaluating the general circulation models (GCMs). To facilitate the satellite and GCM comparisons, the CFMIP (Cloud Feedback Model Inter-comparison Project) Observation Simulator Package (COSP) has been developed and is now increasingly used in GCM evaluations. Real-world Clouds and precipitation can have significant sub-grid variations, which, however, are often ignored or oversimplified in the COSP simulation. In this study, we use COSP cloud simulations from the Super-Parameterized Community Atmosphere Model (SPCAM5) and satellite observations from the Moderate Resolution Imaging Spectroradiometer (MODIS) and CloudSat to demonstrate the importance of considering the sub-grid variability of cloud and precipitation when using the COSP to evaluate GCM simulations. We carry out two sensitivity tests: SPCAM5 COSP and SPCAM5-Homogeneous COSP. In the SPCAM5 COSP run, the sub-grid cloud and precipitation properties from the embedded cloud resolving model (CRM) of SPCAM5 are used to drive the COSP simulation, while in the SPCAM5-Homogeneous COSP run only grid mean cloud and precipitation properties (i.e., no sub-grid variations) are given to the COSP. We find that the warm rain signatures in the SPCAM5 COSP run agree with the MODIS and CloudSat observations quite well. In contrast, the SPCAM5-Homogeneous COSP run which ignores the sub-grid cloud variations, substantially overestimates the radar reflectivity and probability of precipitation compared to the satellite observations, as well as the results from the SPCAM5 COSP run. The significant differences between the two COSP runs demonstrate that it is important to take into account the sub-grid variations of cloud and precipitation when using COSP to evaluate the GCM to avoid confusing and misleading results.

1. Introduction

Marine boundary layer (MBL) cloud, as a strong modulator of the radiative energy budget of the Earth-Atmosphere system, is a major source of uncertainty in future climate change projections of the general circulation models (GCM) (Cess et al., 1996; Bony and Dufresne, 2005). Improving MBL cloud simulations in the GCMs is one of the top priorities of the climate modeling community. As the cloud parameterization schemes in the GCMs become increasingly sophisticated, there is a strong need for comprehensive global satellite cloud observations for model evaluation and improvement. However, the fundamental definitions of clouds in GCMs differ dramatically from those used for satellite remote sensing, which hampers the use of satellite products for model evaluation. In order to overcome this obstacle, the Cloud Feedback Model Inter-comparison Project (CFMIP) community has developed an integrated satellite simulator, the CFMIP Observation Simulator Package (COSP) (Zhang et al., 2010; Bodas-Salcedo et al., 2011). COSP has greatly facilitated and promoted the use of satellite data in the climate modeling community to expose and diagnose issues in GCM cloud simulations (e.g., Marchand et al., 2009; Zhang et al., 2010; Kay et al., 2012; Pincus et al., 2012; Kay et al., 2016; Song et al., 2017).

Warm rain is a unique and important feature of MBL clouds. It plays an important role in determining the macro- and micro-physical properties of MBL clouds, in particular, the cloud water budget (e.g., Stevens et al., 2005; Wood, 2005; Comstock et al., 2005). Many previous studies have investigated the warm rain simulation in GCMs using the COSP simulators. These studies reveal a common problem in the latest generation of GCMs, i.e., the drizzle in MBL clouds is too frequent in the GCM compared with satellite observations (e.g., Zhang et al. 2010; Franklin et al. 2013; Suzuki et al. 2015; Takahashi et al., 2017; Jing et al., 2017; Song et al., 2017, Bodas-Salcedo et al. 2008; Stephens et al. 2010; Bodas-Salcedo et al. 2011; Nam and Quaas 2012; Franklin et al. 2013; Jing et al., 2017). One possible reason for the excessive warm rain production in GCMs could be the model’s inaccurate representation of physical processes, such as auto-conversion and accretion that govern the precipitation efficiency in warm MBL clouds. Due to the lack of sub-grid variability of microphysical quantities in most large-scale models, the auto-conversion parameterization is overly aggressive so that the models tend to produce precipitation too quickly (Lebsock et al. 2013, Song et al. 2017).

The radar observations of warm rain from CloudSat and collocated MODIS (Moderate Resolution Imaging Spectroradiometer) cloud observations are extremely useful data for assessing and improving the GCM simulations of MBL clouds and their precipitation process. However, the dramatic spatial resolution differences between the conventional GCM (~100km) and satellite observations (~1km) become a challenging obstacle for the satellite and GCM comparisons. To overcome this obstacle, the COSP first divides the grid-level cloud and precipitation properties (e.g., grid-mean cloud water and rain water) into the so-called “sub-columns” that are conceptually similar to “pixel” in satellite observation. Then for each sub-column the COSP satellite-simulators (e.g., COSP-CloudSat and COSP-MODIS) simulate the satellite measurements (e.g., radar reflectivity) and retrievals (e.g., MODIS cloud optical depth and effective radius) which become directly comparable with satellite data. Ideally, the sub-column generation in COSP should be consistent with the sub-grid cloud parameterization scheme in the host GCM. However, in practice sub-grid variations of cloud and precipitation are often ignored or treated crudely in the COSP simulation for a number of possible reasons. First of all, the COSP is an independent package and it takes substantial efforts to implement in the COSP a sub-grid cloud generation scheme that is consistent with the host GCM. Secondly, a simple sub-column generation scheme helps alleviate the computational cost associated with the COSP simulation. Last but certainly not least, the users of the COSP might not be fully aware of the consequences of ignoring the sub-grid cloud and precipitation variability in the COSP simulations.

The current version (v1.4) of COSP provides a built-in highly simplified sub-column generator. It accounts only for the sub-grid variability of the types of hydrometeors and ignores the variability of mass and microphysics within each hydrometeor type. The water content and microphysical properties (i.e., droplet effective radius and optical thickness) of each hydrometeor are horizontally homogenous among all the sub-columns that are labeled as the same type (i.e., stratiform or convective). Here we refer to the current scheme as the "*homogenous hydrometeor scheme*". The uncertainties and potential biases caused by the *homogenous hydrometeor scheme* can be significant and should not be overlooked. A simple hypothetical example is provided in Figure 1 to illustrate the importance of accounting for the sub-grid variability of rainwater in simulating the CloudSat radar reflectivity. To be consistent with the two-moment cloud microphysics scheme (Morrison and Gettelman, 2008) that is widely used in the GCMs, we assume the sub-grid distribution of rainwater to follow the exponential distribution. In this example, the grid-mean rainwater mixing ratio ($\bar{q}$) is set to be 0.03 g/kg (dashed blue line in Figure 1a). Using the Quickbeam simulator (Haynes et al., 2007) in COSP, we simulated the corresponding 94-GHz CloudSat radar reflectivity, which is shown in Figure 1b. The grid-mean radar reflectivity based on the exponentially distributed rainwater (i.e., with sub-grid variance) is about 4 dBZ (solid red line in Figure 1b). In contrast, if the sub-grid variation of rainwater is ignored, the radar reflectivity corresponding to $\bar{q}$ = 0.03 g/kg is 13 dBZ (dashed blue line in Figure 1b). The substantial difference between the two indicates that ignoring the sub-grid variability of hydrometeors could cause significant overestimation of grid-mean radar reflectivity simulation, which in turn could complicate and even mislead the evaluation of GCMs.

The objective of this study is to investigate and demonstrate to the GCM modeling community the importance of considering the sub-grid variability of cloud and precipitation properties when evaluating the GCM simulations using COSP. Here we employ the Super-parameterized Community Atmosphere Model Version 5 (SPCAM5, Wang et al., 2015) to provide the sub-grid cloud and precipitation hydrometeor fields for a comparison study of the simulated radar reflectivity and warm rain frequencies by COSP. Fundamentally different from the convective cloud parameterization schemes in GCMs, SPCAM5 consists of a two-dimensional cloud-resolving model (CRM) embedded into each grid of a conventional CAM5 (Khairoutdinov and Randall, 2003; Wang et al., 2015). In SPCAM5, the sub-grid cloud dynamical and microphysical processes are explicitly resolved at a 4-km resolution using a two-dimensional version of the System for Atmospheric Modeling (Khairoutdinov and Randall, 2003) with the two-moment microphysics scheme (Morrison et al., 2005). We carry out two sensitivity tests: SPCAM5 COSP and SPCAM5-Homogeneous COSP. In the SPCAM5 COSP run, the sub-grid cloud and precipitation properties from the embedded CRMs of SPCAM5 are used to drive the COSP simulation. In the SPCAM5-Homogeneous COSP run, the default *homogenous hydrometeor scheme* of COSP mentioned above is used to generate the sub-grid cloud and precipitation fields for the COSP simulation. The outputs from the two runs are compared with the collocated CloudSat and MODIS observations to assess the potential problems in both runs, and also to understand the impacts of omitting sub-grid cloud variations in the COSP simulations.

The rest of the paper is organized as follows: Section 2 describes the model, COSP and satellite data used in this study. Results are represented in Section 3. Finally, Section 4 provides general conclusions and remarks.

2. Description of Model, COSP and Satellite Observations

2.1. Model

The model used in this study is SPCAM5, an application of the Multiscale Modeling Framework (MMF) (Randall et al., 2003; Khairoutdinov et al., 2005, 2008; Tao et al., 2009) to CAM5 (Neale et al., 2010), which uses the finite volume dynamical core at 1.9° latitude × 2.5° longitude resolution with 30 vertical levels and 600-s time step. The embedded 2-D CRM in each CAM5 grid cell includes 32 columns at 4 km horizontal grid spacing and 28 vertical layers coinciding with the lowest 28 CAM5 levels. The CRM runs with a 20-s time step. Details of the SPCAM5 can be found in Wang et al. (2011; 2015). The simulations are run in a "constrained meteorology" configuration (Ma et al., 2013; 2015) to facilitate model evaluation against observations, in which the model winds are nudged toward the Modern Era Reanalysis for Research Applications (MERRA) reanalysis with a relaxation timescale of 6 hours (Zhang et al., 2014). The SPCAM5 simulations are performed from September 2008 to December 2010 (28 months). The last 24 months (January 2009-December 2010) outputs of the simulations are used for analysis.

2.2. COSP

We used COSP Version 1.4, which has no scientific difference from the latest version COSP2 (Swales et al., 2018). Currently, COSP provides simulations of ISCCP (International Satellite Cloud Climatology Project), CALIPSO (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation), CloudSat, MODIS, and MISR (Multi-angle Imaging SpectroRadiometer) cloud measurements and/or retrievals (Bodas-Salcedo et al., 2011). In this study, we will focus on the MODIS and CloudSat simulators (Pincus et al., 2012; Haynes et al., 2007). COSP has three major parts, each controlling a step of the pseudo-retrieval process: (1) the *sub-column generator* of COSP first distributes the grid-mean cloud and precipitation properties from GCM into the so-called sub-columns that are conceptually similar to "pixels" in satellite remote sensing. (2) the *satellite simulators* simulate the direct measurements (e.g., CloudSat radar reflectivity and CALIOP backscatter) and retrieval products (e.g., MODIS cloud optical thickness and effective radius) for each sub-column using highly simplified radiative transfer and retrieval schemes; (3) the *aggregation scheme* averages the sub-column simulations back to grid level to obtain temporal-spatial averages that are comparable with aggregated satellite products (e.g., MODIS level-3 gridded monthly mean products).

As mentioned in the Introduction, the COSP-v1.4 has a highly simplified built-in sub-column generator based on the homogenous hydrometeor scheme. This scheme accounts only for the sub-grid variability of the types of hydrometeors and ignores the variability of mass and microphysics within each hydrometeor type. An example is provided in Figure 2 to illustrate how this default sub-column generator of COSP-v1.4 distributes the grid-mean cloud and precipitation into the sub-columns. We arbitrarily selected a grid (23°N and 150°E) with both cloud and significant precipitation from our previous CAM5 simulation (CAM5-Base simulation in Song et al., 2017). Figure 2a shows the vertical profiles of the grid-mean total (stratiform plus convective) and convective cloud fractions at the selected grid box. Figure 2b shows the vertical profiles of the grid-mean mixing ratios of each type of hydrometeors. The sub-column generator of COSP takes the grid-mean cloud fractions, hydrometeor mixing ratios and effective particle sizes (Figure 2a and Figure 2b) as inputs to generate the sub-columns for the later satellite measurement and retrieval simulation.

First, sub-columns (150 sub-columns are generated in our example) are assigned as either cloudy or clear at each model level by the Subgrid Cloud Overlap Profile Sampler (SCOPS), which was developed originally as part of the ISCCP simulator (Klein and Jakob, 1999; Webb et al., 2001). Figure 2c show the distributions of cloudy sub-columns among the 150 sub-columns at each vertical level, indicated by variable frac_out produced in the scops.f routine. The sub-column at certain vertical level is stratiform cloudy if frac_out =1, or connective cloudy if frac_out=2 at that vertical level. As illustrated in Figure 2c, the SCOPS assigns cloud to the sub-columns in a manner consistent with the model's grid box average stratiform and convective cloud amounts (Figure 2a) and its cloud overlap assumption, i.e., maximum-random overlap in this case. The next step is to determine which of the sub-columns generated by SCOPS contain precipitation hydrometeors, e.g., rain and snow. This step is necessary and critical for the COSP CloudSat radar simulator (Bodas-Salcedo et al., 2011) because radar reflectivity is highly sensitive to the precipitation hydrometeors due to their large particle size (L'Ecuyer and Stephens, 2002; Tanelli et al., 2008). The current sub-grid precipitation distribution scheme "SCOPS-PREC" is developed and described in Zhang et al. (2010). Figure 2d shows the masking of precipitation among the 150 sub-columns generated by the SCOPS-PREC for the example grid. After the cloud and precipitation are masked, the last step is to specify the mass (i.e., mixing ratio) and effective radius of hydrometeors for all the sub-columns occupied by clouds and/or precipitation. The current scheme for this step is highly simplified. As shown in Figure 2e, it assumes the mass and the microphysics of each type of hydrometeor to be horizontally homogeneous among all the sub-columns that are occupied by this type of hydrometeor at a given model level. In other words, at each model level the only difference among sub-columns is that they may be occupied by different types of hydrometeors (Zhang et al., 2010).

In this study, we have carried out two COSP simulations using the 2-year SPCAM5 CRM outputs to investigate the importance of considering the sub-grid variations of cloud and precipitation properties when evaluating the GCM simulations using COSP. The two COSP simulations are marked as SPCAM5 COSP and SPCAM5-Homogeneous COSP, respectively. For the SPCAM5 COSP simulation, we treat the sub-grid cloud and precipitation fields from the CRM of SPCAM5 outputs as sub-columns of COSP without using the COSP sub-column generator. For the SPCAM5-Homogeneous COSP simulation, we first average the sub-grid cloud and precipitation fields (including both clear and cloudy sub-grids) from the CRM of SPCAM5 to each CAM5 grid, and then input these grid-mean cloud and precipitation fields to the default COSP-v1.4 sub-column simulator described above to generate the sub-column fields. All the other processes of two COSP simulations are exactly same. The COSP simulator outputs are produced from 6-hourly calculations and the number of sub-columns used here is 32. To derive the probability of precipitation, we made some simple in-house modifications in COSP v1.4 to write out the MODIS and CloudSat simulations for every sub-column. This allows us to derive the joint statistics of COSP-MODIS and

COSP-CloudSat simulations and compare them with those derived from collocated MODIS and CloudSat level-2 products.

2.3. Satellite Data

The cloud measurements from the A-Train satellite sensors, namely MODIS and CloudSat, are used for model-to-observation comparison. The newly released collection 6 (C6) Aqua-MODIS cloud products (Platnick et al., 2017) are used to evaluate cloud fraction, cloud optical thickness and cloud droplet effective radius. For MBL cloud studies, CloudSat provides valuable information on the warm rain process that cannot be achieved by a passive sensor like MODIS. The direct measurement of CloudSat is the vertical profile of 94-GHz radar reflectivity by cloud and hydrometer particles (i.e., 2B-GEOPROF product), from which other information such as vertical distribution of clouds and precipitation can be derived. The CloudSat 2B-GEOPROF product (Marchand et al., 2008) is used for cloud vertical structure, radar reflectivity, and identification of precipitation in MBL clouds. To prepare for the comparison of joint statistics, we collocated 5 years (2006 ~ 2010) of pixel-level (i.e., level-2) MODIS and CloudSat observations using the collocation scheme developed in Cho et al. (2008). Due to the low sampling rate of CloudSat, we used 5 years (2006 ~ 2010) of observations, in comparison with the 2-year model simulation (2009 ~ 2010), to obtain enough statistics. A sensitivity study indicates that the inter-annual variability of MBL clouds is much smaller than the model-to-observation differences.

In this study, we focus on the tropical and subtropical regions between 45°S and 45°N (loosely referred to as "tropical and subtropical region"), where most stratocumulus and cumulus regimes are found. We avoid high latitudes because satellite observations, namely MODIS, may have large uncertainties to low solar zenith angles there (Kato and Marshak, 2009; Grosvenor and Wood, 2014; Cho et al., 2015).

3. Sensitivity Study: SPCAM5 COSP vs. SPCAM5-Homogeneous COSP

First, we compare the Contoured Frequency by Altitude Diagram (CFAD) of tropical clouds derived based on SPCAM5 COSP and SPCAM5-Homogeneous COSP simulations with that derived from CloudSat 2B-GEOPROF product in Figure 3. The CFAD based CloudSat observations displays a typical boomerang type shape that has been reported in many previous studies (Bodas-Salcedo et al., 2011; Zhang et al., 2010; Marchand et al., 2009). Focusing on the low clouds below 3km, we observe a rather broad distribution of radar reflectivity with a maximum occurrence frequency around −30 dBZ ~ −20 dBZ followed by a long tail extending to about 10 dBZ. As pointed out in previous studies, the peak around −30 dBZ ~ −20 dBZ is due to non-precipitating MBL clouds and the precipitating clouds with increasing rain rate give rise to the long tail. The CFAD based on two COSP simulations exhibits some characteristics similar to the CloudSat observations, but also many noticeable differences. In particular, the two COSP simulations both produce a much narrower range of radar reflectivity for low clouds, with occurrence frequency clustered mostly around −25 dBZ in SPCAM5 COSP and around 0 dBZ in SPCAM5-Homogeneous COSP. These results show that using the oversimplified COSP sub-column generator (e.g., the homogeneous hydrometeor scheme) has non-negligible influences on the simulated radar reflectivity and produces artificially high occurrences of large radar reflectivity. In consistency with many previous studies (e.g., Bodas-Salcedo et al. 2008; Stephens et al. 2010; Nam and Quaas 2012; Franklin et al. 2013; Jing et al., 2017), our results also reveal that GCMs tend to produce much larger radar reflectivity more frequently through the COSP simulator compared to the satellite observation.

The systematic biases in simulated radar reflectivity by the COSP homogeneous hydrometeor scheme might lead to the unjustified and biased evaluation of the warm rain production in GCMs, since cloud column maximum radar reflectivity (Z_{max}) is often used to distinguish precipitating from non-precipitating MBL clouds (Kubar and Hartmann, 2009; Lebsock and Su, 2014; Haynes et al., 2009).

Next we compare the simulated and observed PDFs of Z_{max} for all the sub-columns that are marked as warm liquid clouds in the domain between 45°S and 45°N. The warm liquid clouds are defined by the cloud phase and cloud top pressure derived from the MODIS simulator by the criteria that cloud phase is liquid and cloud top pressure is between 900 hPa and 500 hPa. Big differences in the PDFs of Z_{max} between the SPCAM5-Homogeneous COSP and the A-Train observations, and between SPCAM5-Homogeneous COSP and SPCAM5 COSP are shown in Figure 4. First, in the A-Train observations, about 46% of warm liquid clouds detected by the MODIS are not observed by the CloudSat. These clouds are either too thin and therefore their radar reflectivity is too weak to be detected by CloudSat, or they are too low and therefore suffer the surface clutter issue (Marchand et al., 2008). For those warm liquid clouds detected by both the MODIS and CloudSat, the PDF of Z_{max} peaks around -25 dBZ. Second, in both COSP simulations, almost all warm liquid clouds derived by the MODIS simulator have valid CloudSat radar reflectivity larger than -40 dBZ. The PDFs of Z_{max} in the SPCAM5 reasonably resemble those in the A-Train observations. However, significantly different from the other two, the distribution of Z_{max} in the SPCAM5-Homogeneous shifts to the large dBZ values and peaks around 0 dBZ. In previous studies (e.g., Takahashi et al., 2017), warm liquid clouds are categorized to three different modes by $Z_{max:}$ non-precipitating mode ($Z_{max} < -15$ dBZ), drizzle mode (-15 dBZ < Z_{max} < 0 dBZ) and rain mode (Z_{max} > 0 dBZ). The simulated and observed PDFs of Z_{max} demonstrate that a large portion of warm liquid clouds is non-precipitating in the observations and SPCAM5 COSP while most warm liquid clouds are precipitating (drizzle or rain) clouds in the SPCAM5-Homogeneous COSP. The use of the COSP homogeneous hydrometeor scheme gives us a dramatically different assessment of the warm rain production of MBL clouds in the SPCAM5 model, i.e., if we consider the sub-column variability of cloud and precipitation in the COSP simulation, we find that the SPCAM5 model can reproduce the observed warm rain production quite well. However, if we ignore the CRM sub-grid variability and use the homogeneous hydrometeor scheme, we may make the biased conclusion that the SPCAM5 model performs badly in the simulation of warm rain production.

More significant differences between the SPCAM5 COSP and SPCAM5-Homogeneous COSP simulations can be found from the spatial distributions of the probability of precipitation (POP) in MBL warm clouds (Figure 5). Here, the POP for a given grid box is defined as the fraction of liquid-phase cloud identified by MODIS observations with Z_{max} larger than a certain threshold (i.e., –15 dBZ for drizzle or rain, 0 dBZ for rain, and 10 dBZ for heavy rain, respectively) according to the collocated CloudSat observations with respect to the total population liquid-phase clouds with the cloud top pressure between 500 hPa and 900 hPa in the grid. Observations in Figure 5 suggest that roughly a third of MBL clouds observed by MODIS in the tropical and subtropical region are likely precipitating (drizzle or rain), with a domain averaged POP around 33%. The POP of drizzle plus rain has a distinct pattern: smaller (~15%) in the coastal Sc regions and increasing to ~50% in the Cu cloud regions. The observed POPs of rain and heavy rain show similar spatial patterns as those of drizzle plus rain, with much smaller domain averaged POP being about 12.5% and 3.3%, respectively.

In the same way as we define POP for observations, we define the POP for two COSP simulations as the ratio of sub-columns that have COSP-CloudSat simulated Z_{max} larger than a certain threshold with respect to the total number of liquid-phase clouds identified by COSP-MODIS. As shown in Figure 5, two COSP simulations show dramatically different spatial distributions of POPs. The SPCAM5 COSP produces the similar POP patterns as those in the observations, with the domain averaged POPs for drizzle or rain, rain and heavy rain being about 43%, 16% and 4.5%, respectively. However, the POPs in the SPCAM5-Homogeneous COSP are substantially overestimated, with the domain averaged POPs for drizzle or rain, rain and heavy rain being about 75%, 36% and 7%, respectively. Using the COSP homogeneous hydrometeor scheme will lead to the conclusion that the drizzle or rain is triggered too frequently (more than double of the observations) in the SPCAM5 model, which obviously is not a fair assessment.

Previous studies find that the warm rain production in MBL clouds is tightly related to the in-cloud microphysical properties of MBL clouds (e.g., Stevens et al., 2005; Wood, 2005; Comstock et al., 2005). Next, we check the dependence of POP on in-cloud properties liquid water path (LWP) and on liquid cloud effective radius (r_e) in both observations and two COSP simulations. Figure 6 shows the POPs of drizzle or rain (i.e., Z_{max} > -15 dBZ) as a function of in-cloud LWP and r_e overlaid by the joint PDF of LWP and r_e (white contours) in the satellite observations and two COSP simulations. The observed POPs of warm liquid clouds increase monotonically with increasing in-cloud LWP and r_e, with high POPs concentrating on the domain with large values of LWP and r_e (i.e., LWP > 200 g/m^2 and r_e > 15 μm). However, in the two COSP simulations, especially the SPCAM5-Homogeneous COSP, at each joint bin the POPs are much larger than those in the A-Train observations. When in-cloud LWP (r_e) is larger than 150 g/m^2 (17 μm), the dependence of POPs on in-cloud r_e (LWP) is small. The joint PDFs of in-cloud LWP and r_e in the observations and two COSP simulations are also quite different. There are more occurrences with large LWP and r_e in the MODIS observations than the two COSP simulations. The SPCAM5 COSP simulations have two peaks of the joint PDFs, which are converted to one occurrence peak in the SPCAM5-Homogeneous COSP simulation by using the COSP homogeneous hydrometeor scheme.

Based on the above comparisons, we can see that the oversimplified COSP sub-column generator contributes to not only the narrow distribution of MBL cloud radar reflectivity, but also to unrealistically high POPs in the SPCAM5 model. Besides, it also changes the distribution of in-cloud microphysical properties, and the relationship between POPs and cloud microphysical properties as well.

4. Summary and Discussion

This study presents a satellite-based evaluation of the warm rain production of MBL cloud in the SPCAM5 model using two COSP simulations (SPCAM5 COSP and SPCAM5-Homogeneous COSP), with the objective to demonstrate the importance of considering the sub-grid variability of cloud and precipitation when using COSP to evaluate GCM simulations. Through the SPCAM5 COSP simulations, in which the sub-column variability of cloud and precipitation is considered, we find that the SPCAM5 model can reproduce the observed warm rain production quite well. However, in the SPCAM5-Homogeneous COSP simulation, in which we ignore the CRM sub-grid variability and use the COSP homogeneous hydrometeor scheme, the simulated radar reflectivity and POPs in the SPCAM5 are significantly overestimated compared to the observations. Therefore, use of the COSP homogeneous hydrometeor scheme gives us a significantly different assessment of warm rain production of MBL clouds in the SPCAM5 model. Our results also indicate that the sub-grid variability of mass and microphysics of each hydrometeor type is key to the realistic simulation of radar reflectivity.

The systematic and significant biases due to the limitation of current homogeneous hydrometeor scheme can mislead the evaluation of GCMs and should not be overlooked. In this regard, an improved sub-column generator needs to be developed for COSP to account for the sub-grid variances of cloud and/or hydrometer mass and microphysics. A recent study of Hillman et al. (2017) investigated the sensitivities of simulated satellite retrievals to subgrid-scale overlap and condensate heterogeneity, and demonstrated the systematic biases in the simulated MODIS cloud fraction and CloudSat radar reflectivity due to the oversimplified COSP sub-column generator. Their study also proposed a new scheme to replace the COSP current sub-column generator, and showed that the new scheme can produce much better satellite retrievals. Implementing their sub-column heterogeneous hydrometeor scheme in COSP may improve the GCM COSP simulations and give a better-justified assessment of the GCM performance in simulating warm rain processes and cloud microphysical properties.

On the other hand, since the assumptions of sub-grid variability of cloud and hydrometeors in different GCMs may be quite different, one universal sub-column hydrometeor scheme may be not applicable to all models. Based on this consideration, the latest version COSP version 2 enhances flexibility by allowing for model-specific representation of sub-grid scale cloudiness and hydrometeor condensates and encourages the users to implement the same sub-grid scheme as the host GCM for consistency (Swales et al., 2018). Nevertheless, our study also suggests that any evaluation study of warm rain production in GCMs by using COSP simulators should take this issue into account.

Code and Data Availability:

Details of SPCAM5 can be found in Wang et al. (2011, 2015). The host GCM in SPCAM5 is the Community Atmospheric Model, Version 5 (see details on the CESM website at http://www.cesm.ucar.edu/models/cesm1.1/cam/). SPCAM5 has recently been merged with CESM1.1.1 and released to the public (*Randall et al*., 2013; https://svn-ccsm-release.cgd.ucar.edu/model_development_releases/spcam2_0-cesm1_1_1). Codes of COSP V1.4 can be found in the website at https://github.com/CFMIP/COSPv1. We used the collection 6 (C6) Aqua-MODIS cloud products (Platnick et al., 2017), which can be downloaded from the NASA website at https://ladsweb.modaps.eosdis.nasa.gov/api/v1/productPage/product=MYD06_L2. The CloudSat data are distributed by the CloudSat Data Processing Center. The CloudSat 2B-GEOPROF product we used is downloaded from the website at http://www.cloudsat.cira.colostate.edu/data-products/level-2b/2b-geoprof?term=42.

Acknowledgements. This research is supported by Department of Energy (DOE), Office of Science, Biological and Environmental Research, Regional & Global Climate Modeling Program (grant #DE-SC0014641). The Pacific Northwest National Laboratory is operated for the DOE by Battelle Memorial Institute under contract DE-AC05-76RLO 1830. Minghuai Wang was supported by the Minister of Science and Technology of China (2017YFA0604001). The computations in this study were performed at the UMBC High Performance Computing Facility (HPCF). The facility is supported by the U.S. National Science Foundation through the MRI program (grant nos. CNS-0821258 and CNS-1228778) and the SCREMS program (grant no. DMS-0821311), with substantial support from UMBC. The MODIS cloud products used in this study are downloaded from NASA Level-1 and Atmosphere Archive and Distribution System (https://ladsweb.modaps.eosdis.nasa.gov/). The CloudSat products are provided by CloudSat Data Processing Center (http://www.cloudsat.cira.colostate.edu/).

Reference

Bodas-Salcedo, A., M. J. Webb, M. E. Brooks, M. A. Ringer, K. D. Williams, S. F. Milton, and D. R. Wilson, 2008: Evaluating cloud systems in the Met Office global forecast model using simulated CloudSat radar reflectivities, *J. Geophys. Res*., **113**, D00A13, doi:10.1029/2007JD009620.

Bodas-Salcedo, A., and Coauthors, 2011: COSP: Satellite simulation software for model assessment. *Bull. Am. Meteorol. Soc.,* **92**, 1023–1043, doi:10.1175/2011BAMS2856.1.

Bony, S., and J.-L. Dufresne, 2005: Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models. *Geophys. Res. Lett.*, **32**, L20806, doi:10.1029/2005GL023851.

Cess, R., and Coauthors, 1996: Cloud feedback in atmospheric general circulation models: An update. *J. Geophys. Res. Atmos.*, **101**, 12791-12794.

Cho, H. M., P. Yang, G. W. Kattawar, S. L. Nasiri, Y. Hu, P. Minnis, C. Trepte, and D. Winker, 2008: Depolarization ratio and attenuated backscatter for nine cloud types: Analyses based on collocated CALIPSO lidar and MODIS measurements, *Opt. Express*, **16**(6), 3931–3948.

Cho, H. M., and Coauthors, 2015: Frequency and causes of failed MODIS cloud property retrievals for liquid phase clouds over global oceans. *J. Geophys. Res. Atmos.*, **120**, 2015JD023161–n/a, doi:10.1002/2015JD023161.

Comstock, K. K., C. S. Bretherton, and S. E. Yuter, 2005: Mesoscale variability and drizzle in southeast Pacific stratocumulus. *J. Atmos. Sci.*, **62**, 3792–3807, doi:10.1175/JAS3567.1.

Franklin, C. N., Z. Sun, D. Bi, M. Dix, H. Yan, and A. Bodas-Salcedo, 2013: Evaluation of clouds in access using the satellite simulator package COSP: regime-sorted tropical cloud properties. *J. Geophys. Res. Atmos.*, **118**(12), 6663–6679, doi: 10.1002/jgrd.50496.

Grosvenor, D. P., and R. Wood, 2014: The effect of solar zenith angle on MODIS cloud optical and microphysical retrievals within marine liquid water clouds. Atmospheric Chemistry and Physics, 14, 7291–7321, doi:10.5194/acpd-14-303-2014.

Haynes J.M., R.T. Marchand, Z. Luo, A. Bodas-Salcedo, and G.L. Stephens, 2007: A multi-purpose radar simulation package: QuickBeam, *Bull. Am. Meteorol. Soc.*, **88**, pp. 1723-1727

Haynes, J. M., T. S. L'Ecuyer, G. L. Stephens, S. D. Miller, C. Mitrescu, N. B. Wood, and S. Tanelli, 2009: Rainfall retrieval over the ocean with spaceborne W-band radar. *J. Geophys. Res. Atmos.*, **114**, doi:10.1029/2008JD009973.

Hillman, B. R., R. T. Marchand, and T. P. Ackerman, 2017: Sensitivities of simulated satellite views of clouds to subgrid-scale overlap and condensate heterogeneity, *J. Geophys. Res. Atmos.*, under revision.

Jing, X., K. Suzuki, H. Guo, D. Goto, T. Ogura, T. Koshiro, and J. Mülmenstädt, 2017: A multimodel study on warm precipitation biases in global models compared to satellite observations. *J. Geophys. Res. Atmos.*, **122**. https://doi.org/10.1002/2017JD027310.

Kato, S., and A. Marshak, 2009: Solar zenith and viewing geometry dependent errors in satellite retrieved cloud optical thickness: Marine Sccase. *J. Geophys. Res. Atmos.*, **114**, D01202, doi:10.1029/2008JD010579.

Kay, J. E., and Coauthors, 2012: Exposing global cloud biases in the community atmosphere model (CAM) using satellite observations and their corresponding instrument simulators. *J. Climate*, **25**, 5190–5207, doi:10.1175/JCLI-D-11-00469.1.

Kay, J. E., Tristan L'Ecuyer, Helene Chepfer, Norman Loeb, Ariel Morrison, and Gregory Cesana, 2016: Recent Advances in Arctic Cloud and Climate Research. *Current Climate Change Reports,* **2**, 159-169.

Khairoutdinov, M. F., and D. A. Randall, 2003: Cloud resolving modeling of the ARM summer 1997 IOP: Model formulation, results, uncertainties, and sensitivities, *J. Atmos. Sci.*, **60**(4), 607–625.

Khairoutdinov, M., D. Randall, and C. DeMott, 2005: Simulations of the atmospheric general circulation using a cloud-resolving model as a superparameterization of physical processes, *J.* Atmos. Sci., 62(7), 2136–2154.

Klein, S. A., and C. Jakob, 1999: Validation and sensitivities of frontal clouds simulated by the ECWMF model. *Mon. Wea. Rev*., **127**, 2514-2531. doi: 10.1175/1520-0493(1999)127<2514:VASOFC>2.0.CO;2.

Kubar, T. L., and D. L. Hartmann, 2009: Understanding the importance of microphysics and macrophysics for warm rain in marine low clouds. Part I: Satellite observations. *J. Atmos. Sci.*, **66**, 2953–2972, doi:10.1175/2009JAS3071.1.

L'Ecuyer, T. S., and G. L. Stephens, 2002: An estimation-based precipitation retrieval algorithm for attenuating radars, *Journal of Applied Meteorology*, **41**(3), 272–285.

Lebsock, M., and H. Su, 2014: Application of active spaceborne remote sensing for understanding biases between passive cloud water path retrievals. *J. Geophys. Res*. Atmos., **119**, 8962–8979, doi:10.1002/2014JD021568.

Lebsock, M., H. Morrison, and A. Gettelman, 2013: Microphysical implications of cloud-precipitation covariance derived from satellite remote sensing, *J. Geophys. Res*. Atmos., **118**, 6521–6533, doi:10.1002/jgrd.50347.

Ma, P.-L., and Coauthors, 2013: The role of circulation features on black carbon transport into the Arctic in the Community Atmosphere Model version 5 (CAM5). *J. Geophys. Res. Atmos.* **118**, 4657–4669, doi:doi:10.1002/jgrd.50411.

Ma, P.-L., and Coauthors, 2015: How does increasing horizontal resolution in a global climate model improve the simulation of aerosol-cloud interactions? *Geophys. Res. Lett*., **42**(12):5058-5065. doi:10.1002/2015GL064183.

Marchand, R., G. G. Mace, T. Ackerman, and G. Stephens, 2008: Hydrometeor detection using Cloudsat — An earth-orbiting 94-GHz cloud radar. *Journal of Atmospheric and Oceanic Technology*, **25**, 519–533, doi:10.1175/2007JTECHA1006.1.

Marchand, R., J. Haynes, G. G. Mace, T. Ackerman, and G. Stephens, 2009: A comparison of simulated cloud radar output from the multiscale modeling framework global climate model with CloudSat cloud radar observations. *J. Geophys. Res. Atmos.*, **114**, doi:10.1029/2008JD009790.

Morrison, H., and A. Gettelman, 2008: A new two-moment bulk stratiform cloud microphysics scheme in the community atmosphere model, version 3 (CAM3). Part I: Description and numerical tests. *J. Climate*, **21**, 3642–3659.

Morrison, H., J. A. Curry, and V. I. Khvorostyanov, 2005: A new double-moment microphysics parameterization for application in cloud and climate models. Part I: Description, *J. Atmos. Sci*., **62**(6), 1665–1677.

Nam, C., and J. Quaas, 2012: Evaluation of clouds and precipitation in the ECHAM5 general circulation model using CALIPSO and CloudSat satellite data, *J. Clim*., **25**, 4975–4992, doi:10.1175/JCLI-D-11-00347.1.

Neale, R. B., and Coauthors, 2010: Description of the NCAR community atmosphere model (CAM 5.0), *Tech. Rep. TN–486+STR*, 268 pp., *Natl. Cent. for Atmos. Res*., Boulder, Colo.

Pincus, R., S. Platnick, S. A. Ackerman, R. S. Hemler, and R. J. Patrick Hofmann, 2012: Reconciling simulated and observed views of clouds: MODIS, ISCCP, and the limits of instrument simulators. *J. Climate*, **25**, 120220120058001, doi:10.1175/JCLI-D-11-00267.1.

Platnick, S., and Coauthors, 2017: The MODIS cloud optical and microphysical products: Collection 6 updates and examples from Terra and Aqua, *IEEE Transactions on Geoscience and Remote Sensing*, **55**(1), 502–525, doi:10.1109/TGRS.2016.2610522.

Randall, D., M. Khairoutdinov, A. Arakawa, and W. Grabowski, 2003: Breaking the cloud parameterization deadlock, *Bull. Am. Meteorol. Soc.*, **84**(11), 1547–1564.

Randall, D., M. Branson, M. Wang, S. Ghan, C. Craig, A. Gettelman, and J. Edwards, 2013: A community atmosphere model with superparameterized clouds, *Eos Trans. AGU*, **94**(25), 221–222.

Song, H., Z. Zhang, P.-L. Ma, S. Ghan, M. Wang, 2017: An Evaluation of Marine Boundary Layer Cloud Property Simulations in Community Atmosphere Model Using Satellite Observations: Conventional Sub-grid Parameterization vs. CLUBB. *J. Climate*, In Press.

Stephens, G. L., and Coauthors, 2002: The CloudSat Mission and the A-Train, *Bull. Am. Meteorol. Soc*., **83**(12), 1771–1790, doi:10.1175/BAMS-83-12-1771.

Stevens, B., G. Vali, K. Comstock, R. Woods, M. C. Van Zanten, P. H. Austin, C. S. Bretherton, and D. H. Lenschow, 2005: Pockets of open cells and drizzle in marine stratocumulus. *Bull. Am. Meteorol. Soc*., **86**, 51–57.

Suzuki, K., G. Stephens, A. Bodas-Salcedo, M. Wang, J.-C. Golaz, T. Yokohata, and T. Koshiro, 2015: Evaluation of the warm rain formation process in global models with satellite observations, *J. Atmos. Sci.*, **72**, 3996-4014, doi:10.1175/JAS-D-14-0265.1.

Swales, D. J., R. Pincus, and A. Bodas-Salcedo, 2018: The Cloud Feedback Model Intercomparison Project Observational Simulator Package: Version 2, *Geosci. Model Dev.*, 11, 77-81, https://doi.org/10.5194/gmd-11-77-2018.

Takahashi, H., M. Lebsock, K. Suzuki, G. Stephens, and M. Wang, 2017: An investigation of microphysics and subgrid-scale variability in warm-rain clouds using the A-Train observations and a multiscale modeling framework. *J. Geophys. Res. Atmos.*, **122**(14), 7493–7504.

Tanelli, S., S. L. Durden, E. Im, K. S. Pak, D. G. Reinke, P. Partain, J. M. Haynes, and R. T. Marchand, 2008: CloudSat's Cloud Profiling Radar After Two Years in Orbit: Performance, Calibration, and Processing, *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, **46**(11), 3560–3573.

Tao, W. K., and Coauthors, 2009: A multiscale modeling system developments, applications, and critical issues, *Bull. Am. Meteorol. Soc.*, **90**(4), 515–534.

Wang, M., V. Larson, S. Ghan, M. Ovchinnikov, D. Schanen, H. Xiao, X. Liu, Z. Guo, and P. Rasch, 2015: A multiscale modeling framework model (superparameterized CAM5) with a higher-order turbulence closure: Model description and low-cloud simulations, *Journal of Advances in Modeling Earth Systems*, **7**, 484-509, *doi:10.1002/2014MS000375*.

Wang, M., and Coauthors, 2011: The multi-scale aerosol-climate model PNNL-MMF: Model description and evaluation, *Geosci. Model. Dev.*, **4**(1), 137–168.

Webb, M., C. Senior, S. Bony, and J. Morcrette, 2001: Combining ERBE and ISCCP data to assess clouds in the Hadley Centre, ECMWF and LMD atmospheric climate models. *Climate Dyn.*, **17**, 905-922.

Wood, R., 2005: Drizzle in stratiform boundary layer clouds. Part I: Vertical and horizontal structure. *J. Atmos. Sci.*, **62**, 3011–3033.

Zhang K., and Coauthors, 2014: Technical Note: On the use of nudging for aerosol-climate model intercomparison studies. *Atmospheric Chemistry and Physics,* **14**(16): 8631-8645. doi:10.5194/acp-14-8631-2014.

Zhang, Y., S. A. Klein, J. Boyle, and G. G. Mace, 2010: Evaluation of tropical cloud and precipitation statistics of Community Atmosphere Model version 3 using CloudSat and CALIPSO data. *J. Geophys. Res. Atmos.*, **115**, D12205, doi:10.1029/2009JD012006.

List of Figures:

Figure 1. a) PDF of rainwater mixing ratio for rainwater when the horizontal variability of rainwater is assumed to follow the exponential distribution. The vertical dashed blue line indicates the mean value of rainwater mixing ratio as 0.03 g/kg. b) The corresponding PDF of the CloudSat radar reflectivity simulated by COSP assuming the Marshall and Palmer particle size distribution. The dashed blue line corresponds to the radar reflectivity based on the mean rainwater 0.03 g/kg, and the solid red line corresponds to the grid-mean radar reflectivity based on the PDF of rainwater mixing ratio.

Figure 2. At the single grid 23°N & 150°E on December 04, 2010 in the CAM5-Base simulation (Song et al., 2017): a) The grid mean total (stratiform plus convective) and convective cloud fraction. b) The grid

mean mixing ratios of cloud and precipitation hydrometeors (LS_CLIQ: large-scale (i.e., stratiform) cloud water; LS_CICE: large-scale cloud ice; LS_RAIN: large-scale rain; LS_SNOW: large-scale snow; LS_GRPL: large-scale graupel; CV_CLIQ: convective cloud water; CV_CICE: convective cloud ice; CV_RAIN: convective rain; CV_SNOW: convective snow). c) The distribution of large-scale (red plus signs for frac_out=1) and convective (blue plus signs for frac_out=2) cloud among the sub-columns generated by the SCOPS scheme (i.e., frac_out from scops.f). d) The distribution of large-scale (red plus signs for prec_frac=1), convective (blue plus signs for prec_frac=2), and mixed (green plus signs for prec_frac=3) precipitation among the sub-columns generated by the SCOPS-PREC scheme (i.e., prec_frac from prec_scops.f). e) The mixing ratio (left panels) and effective radius (right panels) of three precipitation hydrometeor types among the sub-columns.

Figure 3. Tropical averaged radar reflectivity-height histogram in the CloudSat observation (top), the SPCAM5 CloudSat simulation (bottom left) and the SPCAM5_Homogeneous CloudSat simulation (bottom right).

Figure 4. The histograms of column maximum radar reflectivity for liquid clouds over oceanic regions from 45°S to 45°N in A-Train satellite observations, SPCAM5 COSP and SPCAM5-Homogeneous COSP simulations.

Figure 5. Probability of precipitation (POP) of liquid clouds between 500hPa and 900hPa levels in the satellite observations (left panel), the SPCAM5 COSP simulation (middle panel) and the SPCAM5-Homogeneous COSP simulation (right panel). Three categories of precipitation: drizzle plus rain (column Z_{max} > -15 dBZ, top panels), rain (column Z_{max} > 0 dBZ, middle panels), and strong rain only (column Z_{max} > 10 dBZ, bottom panels). Unit of POP is %.

Figure 6. POP (drizzle or rain) of liquid clouds at each LWP and liquid cloud effective radius in the satellite observations (top), the SPCAM5 COSP simulation (bottom left) and the SPCAM5-Homogeneous COSP

simulation (bottom right). The white solid contours are joint PDF of LWP and liquid cloud effective radius. Units of POP and PDF are %.

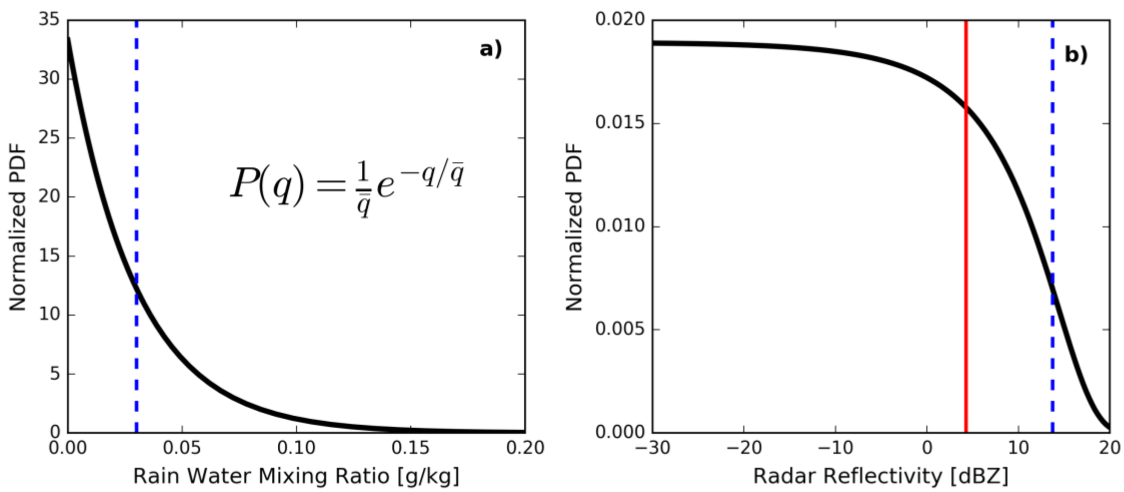


Figure 1. a) PDF of rainwater mixing ratio for rainwater when the horizontal variability of rainwater is assumed to follow the exponential distribution. The vertical dashed blue line indicates the mean value of rainwater mixing ratio

as 0.03 g/kg. b) The corresponding PDF of the CloudSat radar reflectivity simulated by COSP assuming the Marshall and Palmer particle size distribution. The dashed blue line corresponds to the radar reflectivity based on the mean rainwater 0.03 g/kg, and the solid red line corresponds to the grid-mean radar reflectivity based on the PDF of rainwater mixing ratio.

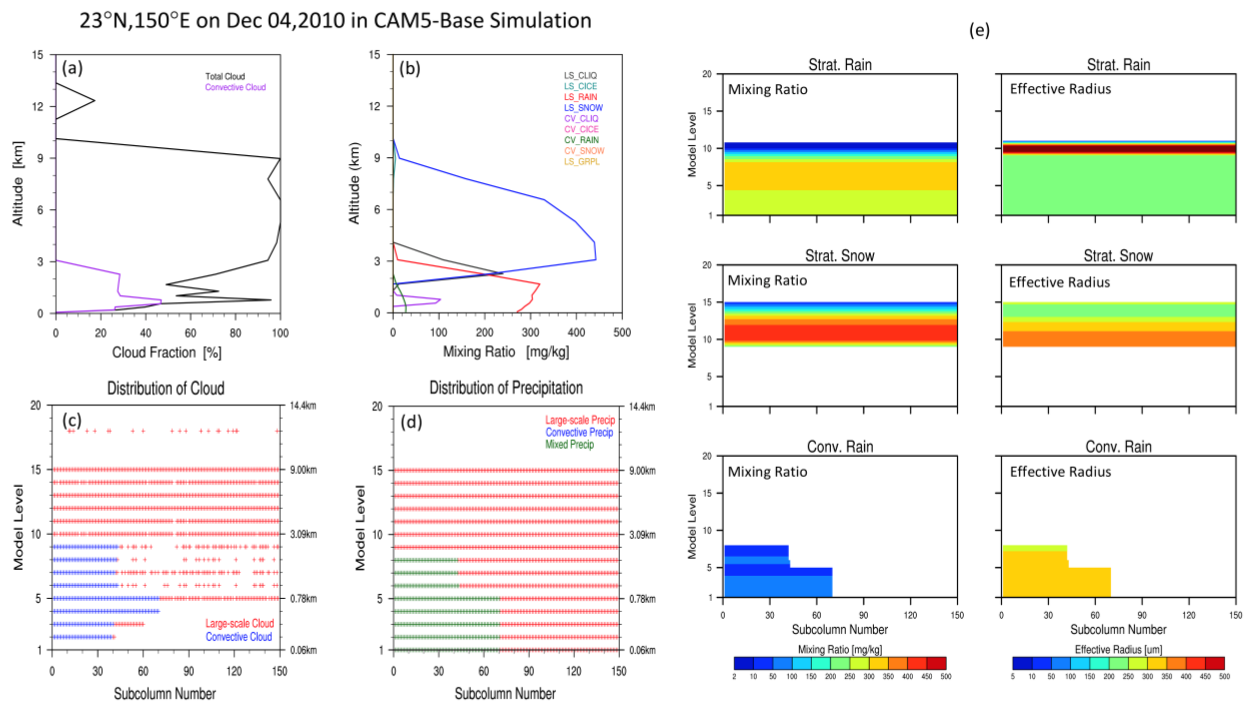


Figure 2. At the single grid 23°N & 150°E on December 04, 2010 in the CAM5-Base simulation (Song et al., 2017): a) The grid mean total (stratiform plus convective) and convective cloud fraction. b) The grid mean mixing ratios of cloud and precipitation hydrometeors (LS_CLIQ: large-scale (i.e., stratiform) cloud water; LS_CICE: large-scale cloud ice; LS_RAIN: large-scale rain; LS_SNOW: large-scale snow; LS_GRPL: large-scale graupel; CV_CLIQ: convective cloud water; CV_CICE: convective cloud ice; CV_RAIN: convective rain; CV_SNOW: convective snow). c) The distribution of large-scale (red plus signs for frac_out=1) and convective (blue plus signs for frac_out=2) cloud among the sub-columns generated by the SCOPS scheme (i.e., frac_out from scops.f). d) The distribution of large-scale (red plus signs for prec_frac=1), convective (blue plus signs for prec_frac=2), and mixed (green plus signs for prec_frac=3) precipitation among the sub-columns generated by the SCOPS-PREC scheme (i.e., prec_frac from prec_scops.f). e) The mixing ratio (left panels) and effective radius (right panels) of three precipitation hydrometeor types among the sub-columns.

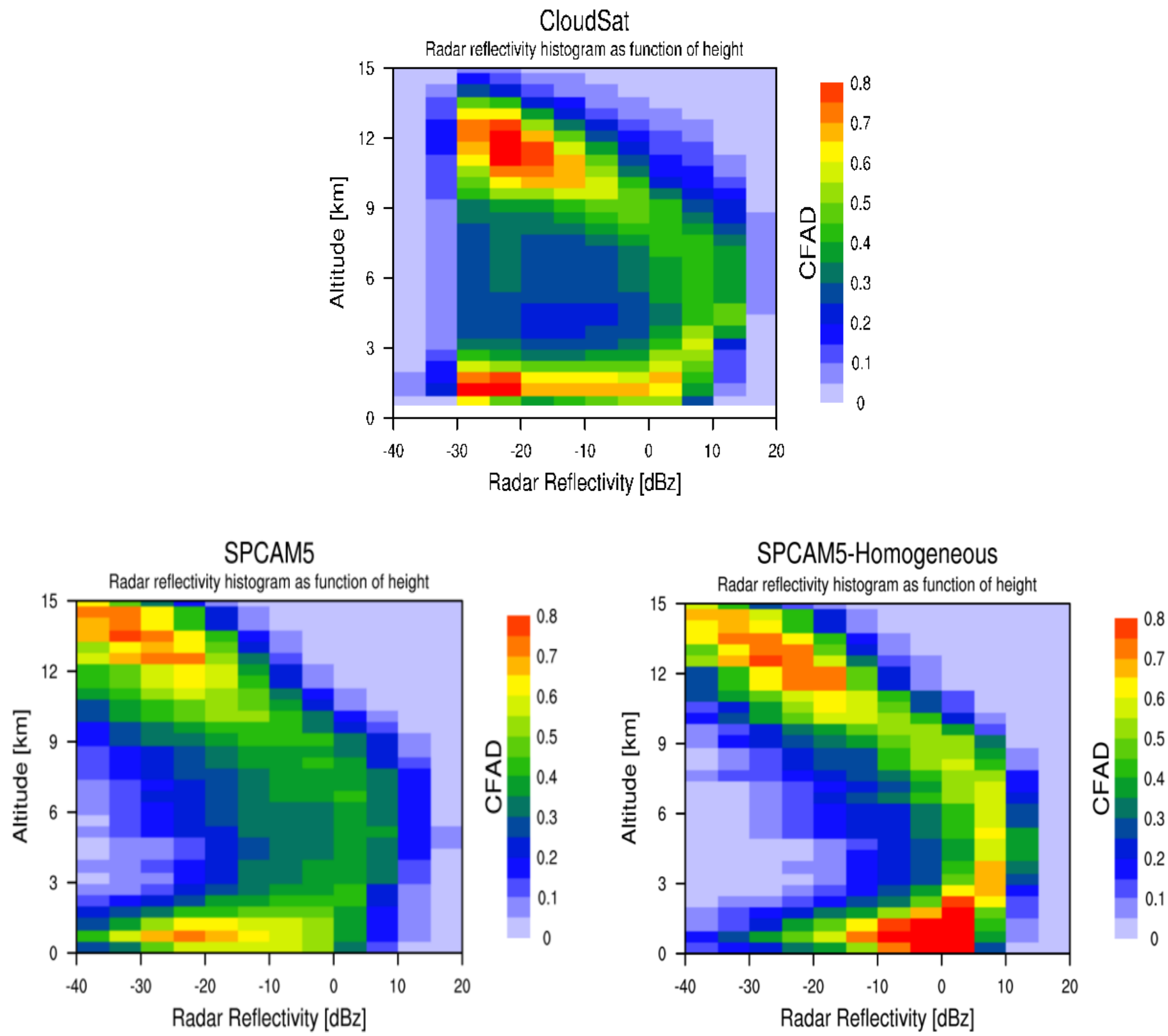


Figure 3. Tropical averaged radar reflectivity-height histogram in the CloudSat observation (top), the SPCAM5 COSP simulation (bottom left) and the SPCAM5-Homogeneous COSP simulation (bottom right).

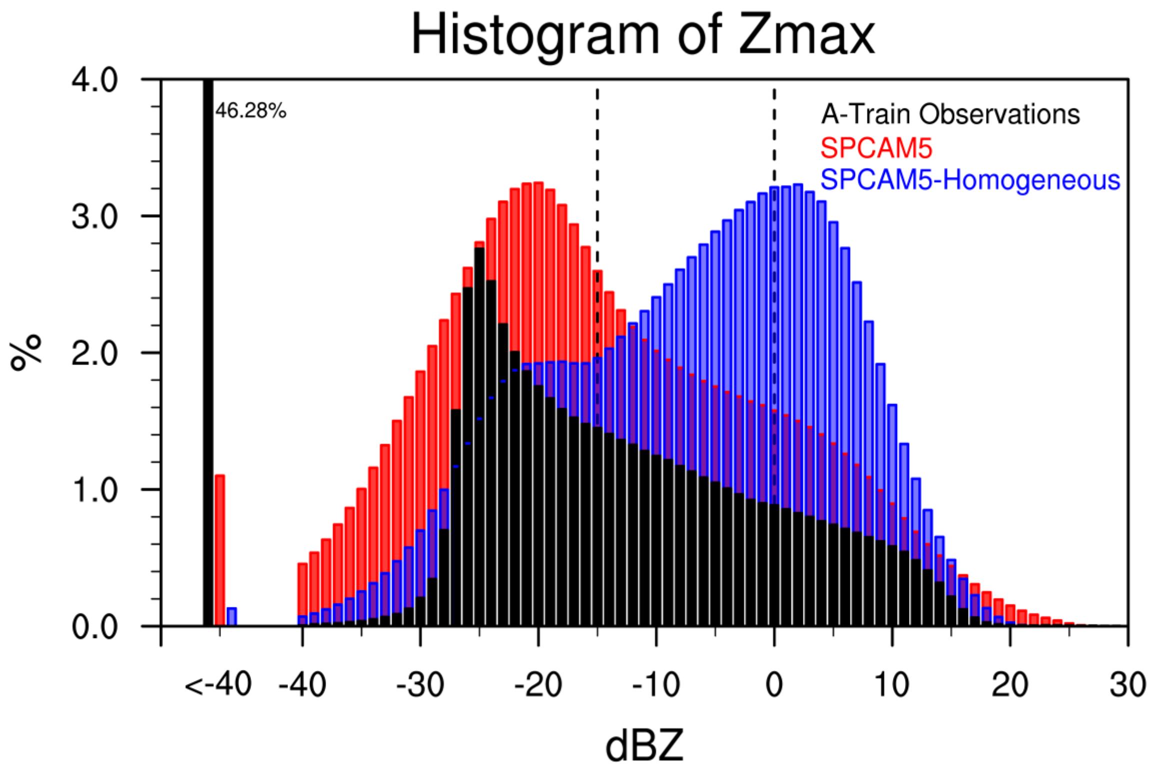


Figure 4. The histograms of column maximum radar reflectivity for liquid clouds over oceanic regions from 45°S to 45°N in A-Train satellite observations, SPCAM5 COSP and SPCAM5-Homogeneous COSP simulations.

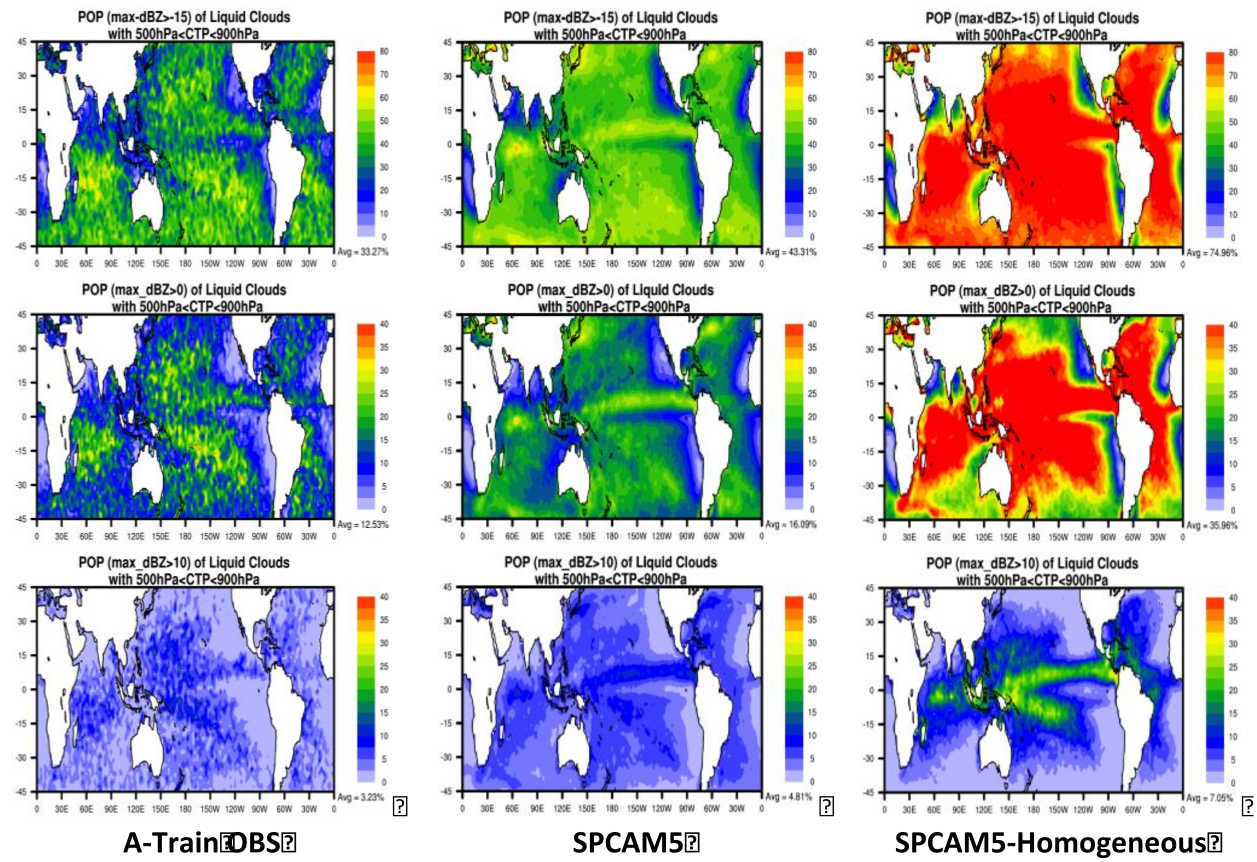


Figure 5. Probability of precipitation (POP) of liquid clouds between 500hPa and 900hPa levels in the satellite observations (left panel), the SPCAM5 COSP simulation (middle panel) and the SPCAM5-Homogeneous COSP simulation (right panel). Three categories of precipitation: drizzle plus rain (column Z_{max} > -15 dBZ, top panels), rain (column Z_{max} > 0 dBZ, middle panels), and strong rain only (column Z_{max} > 10 dBZ , bottom panels). Unit of POP is %.

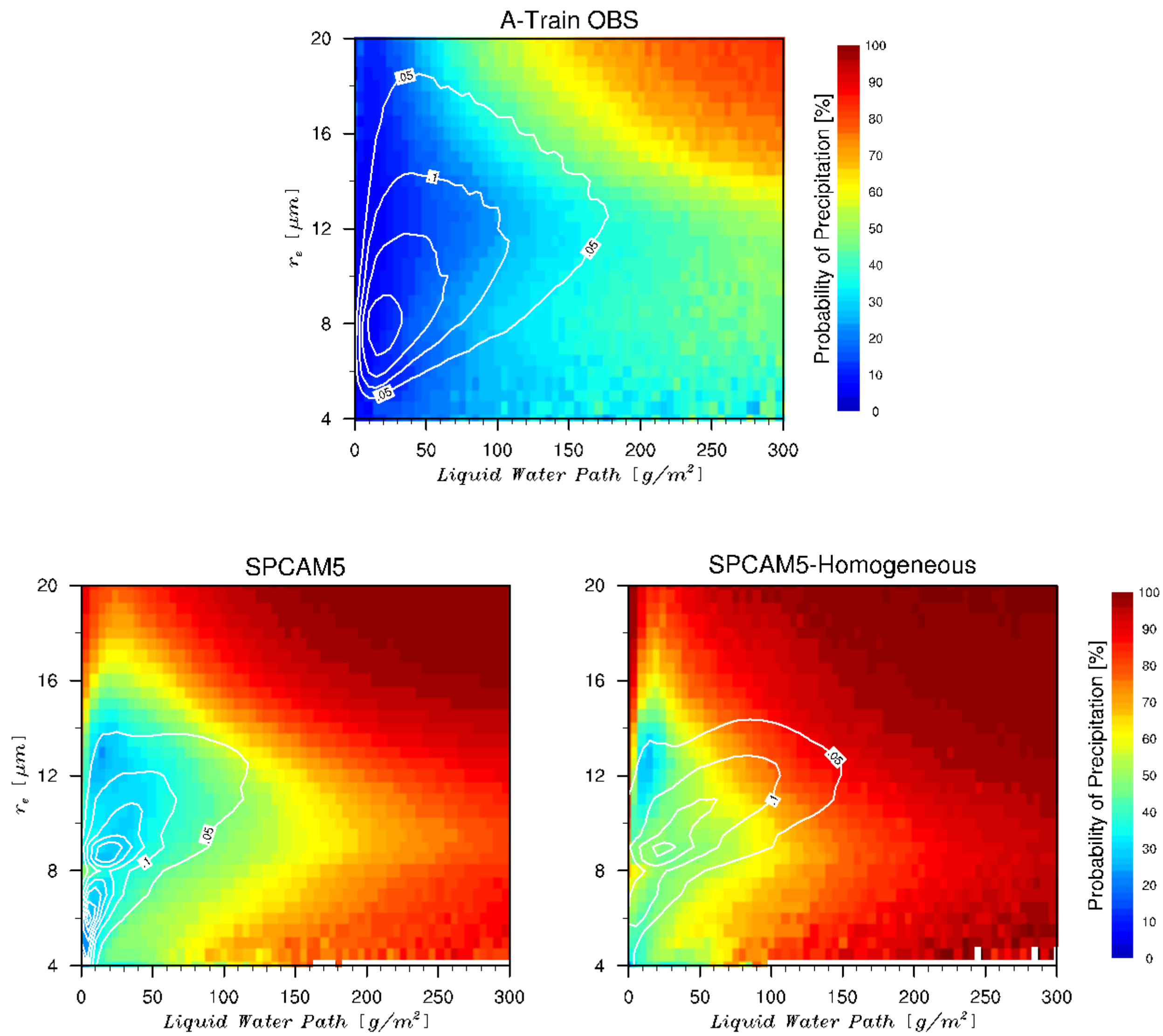

A-Train OBS
SPCAM5
SPCAM5-Homogeneous
r_e [μm]
Liquid Water Path [g/m^2]
Probability of Precipitation [%]
.05
.1
0
50
100
150
200
250
300
4
8
12
16
20
10
20
30
40
60
70
80
90

Figure 6. POP (drizzle or rain) of liquid clouds at each LWP and liquid cloud effective radius in the satellite observations (top), the SPCAM5 COSP simulation (bottom left) and the SPCAM5-Homogeneous COSP simulation (bottom right). The white solid contours are joint PDF of LWP and liquid cloud effective radius. Units of POP and PDF are %.