Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE
10.1029/2017JD027680

Sensitivities of Simulated Satellite Views of Clouds to Subgrid-Scale Overlap and Condensate Heterogeneity

B. R. Hillman[1], **R. T. Marchand**[2], **and T. P. Ackerman**[2,3]

[1]Sandia National Laboratories, Albuquerque, NM, USA, [2]Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA, [3]Joint Institute for the Study of the Atmosphere and Ocean, Seattle, WA, USA

Key Points:
- Simulated satellite retrievals from global model outputs are sensitive to subgrid assumptions
- Using maximum-random overlap and horizontally homogeneous condensate leads to errors in simulated retrievals
- Errors in simulated satellite retrievals are reduced using more realistic overlap and horizontal variability

Correspondence to:
B. R. Hillman,
bhillma@sandia.gov

Citation:
Hillman, B. R., Marchand, R. T., & Ackerman, T. P. (2018). Sensitivities of simulated satellite views of clouds to subgrid-scale overlap and condensate heterogeneity. *Journal of Geophysical Research: Atmospheres*, *123*, 7506–7529. https://doi.org/10.1029/2017JD027680

Received 30 AUG 2017
Accepted 5 MAY 2018
Accepted article online 8 JUN 2018
Published online 21 JUL 2018

Abstract Satellite simulators are often used to account for limitations in satellite retrievals of cloud properties in comparisons between models and satellite observations. The purpose of this framework is to enable more robust evaluation of model cloud properties, so that differences between models and observations can more confidently be attributed to model errors. A critical step in this process is accounting for the difference between the spatial scales at which cloud properties are retrieved with those at which clouds are simulated in global models. In this study, we create a series of sensitivity tests using 4-km global model output from the multiscale modeling framework to evaluate the sensitivity of simulated satellite retrievals to common assumptions about cloud and precipitation overlap and condensate variability used in climate models whose grid spacing is many tens to hundreds of kilometers. We find the simulated retrievals are sensitive to these assumptions. Using maximum-random overlap with homogeneous cloud and precipitation condensate leads to errors in Multiangle Imaging Spectroradiometer and International Satellite Cloud Climatology Project-simulated cloud cover and in CloudSat-simulated radar reflectivity that are significant compared to typical differences between the model simulations and observations. A more realistic treatment of unresolved clouds and precipitation is shown to substantially reduce these errors. The sensitivity to these assumptions underscores the need for the adoption of more realistic subcolumn treatments in models and the need for consistency among subcolumn assumptions between models and simulators to ensure that simulator-diagnosed errors are consistent with the model formulation.

1. Introduction

Large-scale (global) climate models are often evaluated by comparing simulations with observations of present-day climate (e.g., Gleckler et al., 2008). Satellite remote sensing retrievals provide an attractive baseline for evaluation of cloud properties due to their good spatial coverage and increasingly long time records, but comparisons between remote sensing retrievals and models are often challenged by limitations and uncertainties in retrieving cloud properties from space (e.g., Marchand et al., 2010). Satellite instrument simulators (Klein & Jakob, 1999) have emerged as means of accounting for some of the limitations in order to put remote sensing retrievals and model simulations on a common ground for comparison. The goal of the satellite simulator approach is to simulate (or mimic) what a specific satellite instrument would retrieve given a model-simulated description of the cloudy atmosphere. The result is a *pseudoretrieval* that is nominally more directly comparable to the corresponding satellite retrieval product than the raw model fields. The advantage of this approach is that it turns the problem of comparing satellite observations to model results into that of a forward modeling problem (that is, reproducing the observations given the model state) rather than the inverse problem (determining from the observations a specific model property). With respect to cloud occurrence in particular, using the simulator framework has the benefit of providing a definition of cloud that is consistent not only between a given model and some specific observational data set but also between different models as well, which has been especially useful for multimodel intercomparison studies.

The simulator approach has been used in many comparisons between models and observations (Bodas-Salcedo et al., 2011; Kay et al., 2012; Klein & Jakob, 1999; Klein et al., 2013; Lin & Zhang, 2004; Marchand & Ackerman, 2010; Pincus et al., 2012; Webb et al., 2001; Wyant et al., 2006; Zhang et al., 2005). Simulators have been developed for a variety of satellite platforms, many of which have been collected and packaged together

with a common interface into the Cloud Feedback Model Intercomparison Project (Webb et al., 2017) Observation Simulator Package (COSP; Bodas-Salcedo et al., 2011). COSP includes simulators for the International Satellite Cloud Climatology Project (Klein & Jakob, 1999; Webb et al., 2001), the Multiangle Imaging Spectroradiometer (Marchand & Ackerman, 2010), the Moderate Resolution Imaging Spectroradiometer (Pincus et al., 2012), the CloudSat cloud radar (Haynes et al., 2007), and the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) lidar (Chepfer et al., 2008). The goal of COSP and of these individual simulators is to account for limitations and ambiguities in the retrievals, but not all ambiguities in these comparisons can be removed by the simulator approach (Pincus et al., 2012).

Simulating satellite retrievals in large-scale models (e.g., global climate models or GCMs) is essentially a three-step process, involving (1) inferring observed- or pixel-scale properties from the large-scale description provided by models, (2) simulating (or emulating) the pixel-scale satellite retrievals from the modeled pixel-scale properties, and finally, (3) aggregating the simulated pixel-scale retrievals into statistical summaries consistent with gridded, global summary products distributed by the respective satellite teams (often referred to as *level 3* products). In general, there may be uncertainties associated with each of these steps, but we focus on the first of these in this study.

Inferring pixel-scale properties, including cloud properties, is necessary because the horizontal resolution of current GCMs is much coarser (on the order of hundreds of kilometers) than the scales at which satellite retrievals are performed (on the order of kilometers or finer). This is somewhat problematic in that most climate model parameterizations have not (at least historically) been designed to provide a pixel- or kilometer-scale description of the atmosphere, even in a statistical sense. In particular, the need to represent fine-scale cloud properties is not unique to the problem of simulating satellite retrievals, but it is also important for calculating radiative fluxes and heating rates within models. Typically, the impact of overlapping partly cloudy layers on the radiative fluxes in GCMs is accounted for by applying an *overlap assumption* that describes how two partly cloudy layers should overlap in a statistical sense. These overlap assumptions have traditionally been rather simply defined and have included (1) random overlap, in which the cloudy portions of any two layers are uncorrelated, (2) maximum overlap, in which the cloudy portions of different layers are perfectly correlated, and (3) maximum-random overlap, in which clouds in vertically continuous layers are maximally overlapped and cloud layers separated by one or more clear layers are randomly overlapped (Geleyn & Hollingsworth, 1979; Tian & Curry, 1989). Different choices of overlap assumptions can have a substantial effect on calculated radiative fluxes (e.g., Barker et al., 1999; Morcrette & Fouquart, 1986; Stubenrauch et al., 1997), and these simple assumptions (including the maximum-random assumption) have been shown to be insufficient in capturing the complexity of cloud overlap seen in observations (Barker, 2008; Hogan & Illingworth, 2000; Mace & Benson-Troth, 2002). Sensitivity tests using high-resolution model simulations have shown that unrealistic overlap assumptions can lead to instantaneous errors in calculated fluxes in excess of $50 W/m^2$ (Barker et al., 1999; Wu & Liang, 2005).

Subgrid-scale horizontal variability in cloud condensate has traditionally been neglected in GCMs, despite the fact that clouds can exhibit large horizontal variability on scales much smaller than GCM gridboxes (e.g., Stephens & Platt, 1987). This is problematic because radiative fluxes and heating rates calculated from model radiative transfer parameterizations are sensitive to subgrid-scale variations in cloud condensate (e.g., Barker et al., 1999; Oreopoulos et al., 2012; Wu & Liang, 2005). Barker et al. (1999) demonstrate instantaneous flux errors due to unresolved horizontal cloud variability in excess of $100 W/m^2$, and Oreopoulos et al. (2012) demonstrate global mean cloud radiative effect errors on the order of $5 W/m^2$, with much larger regional errors. While some models have begun to account for the impact of subgrid-scale condensate variability on radiative transfer calculations (e.g., Canadian Fourth Generation Atmospheric Global Climate Model; von Salzen et al., 2013), others have not (e.g., Community Atmosphere Model version 5 [CAM5]; Neale, Gettelman, et al., 2010), making this problem still relevant today. The sensitivity of radiative fluxes to both cloud overlap and condensate horizontal variability suggests difficulty for instrument simulator calculations as well. A recent study by Cesana and Waliser (2016) examining CALIPSO-simulated cloud amount in Coupled Model Intercomparison Project Phase 5 GCMs supports the notion that more realistic overlap schemes can reduce biases between observations and climate models.

In COSP the overlap and variability assumptions are applied using a *subcolumn generator* that creates an ensemble of stochastic *subcolumns* for each model gridbox to sample from the distribution of possible subgrid profiles that are consistent with the gridbox mean profiles and with the selected overlap assumption

(Klein & Jakob, 1999; Webb et al., 2001). The overlap assumption is ideally selected to be consistent with the overlap assumption used in the radiative transfer calculations in the model, and the current implementation in COSP includes options for random, maximum, or maximum-random overlap (all with horizontally homogeneous condensate amounts). The purpose of this study is to quantify sensitivities in COSP to these assumptions and to investigate an improved treatment of overlap and variability that reduces errors that arise due to these assumptions. It is shown here that the simulated satellite retrievals from COSP are sensitive to both the maximum-random overlap and the homogeneous condensate assumptions. This underscores the need to select overlap assumptions within COSP that are consistent with those used in the specific model being evaluated if the simulator output is to be used to make an unbiased and fair assessment of the model. As with previous studies that quantify radiative flux errors, this result also suggests that models should seek to develop parameterization that realistically represents the fine-scale variability, including horizontal heterogeneity of cloud condensate. Subcolumn precipitation structure is also shown to be important because it affects the simulated radar reflectivity, and the current scheme in COSP is shown to overestimate precipitation occurrence, leading to large errors in simulated radar reflectivity. These errors are reduced when an improved subcolumn generator (described below) is used.

2. Method

Generating stochastic subcolumns of cloud and precipitation properties in COSP is itself a multistep process. First, stochastic subcolumns of binary cloud occurrence are generated using the Subcolumn Cloud Overlap Profile Sampler (SCOPS; Klein & Jakob, 1999; Webb et al., 2001). Subcolumn binary precipitation occurrence profiles are then generated following the algorithm described by Zhang et al. (2010) and implemented in the PREC_SCOPS routine. Condensate amounts (mixing ratios) are then assigned to the cloudy and precipitating subcolumn elements. The current implementation in COSP assumes a constant in-cloud (and in-precip) condensate mixing ratio at each level within each gridbox, so that each subcolumn at a given level is assigned the gridbox-mean in-cloud (or in-precip) condensate mixing ratio when the binary mask indicates cloud (or precip) is present.

The PREC_SCOPS precipitation treatment implemented in COSP associates precipitation with cloud but fails to account for any estimate of precipitation fraction (the fraction of the gridbox that contains precipitation at any level) that may be diagnosed by the model. Rather, the precipitation treatment essentially assumes that once precipitation is diagnosed at a particular level in a subcolumn, it falls all the way down to the surface unless a precipitation-free layer (that is, a layer in which gridbox mean condensate is equal to zero) is encountered. This can lead to a gross overestimation of the number of precipitating subcolumns and, consequently, a gross overestimation of the occurrence of large values of simulated radar reflectivity factor (the increase in radar reflectivity is mitigated somewhat by spreading the in-precipitation mean over a larger area, but this effect is secondary to the overestimate in precipitation occurrence).

An adjustment to the subcolumn precipitation occurrence is introduced here, following the work of Di Michele et al. (2012), in which subcolumn precipitation is either added or removed at each level until the fraction of subcolumns with precipitation at a given level matches an input precipitation fraction. Precipitation is added preferentially to columns with more (vertically integrated) cloudy levels and removed preferentially from columns with less cloudy levels. This is similar to the "PEVAP" adjustment described by Di Michele et al. (2012) (so-called because it represents an evaporation of precipitation), and the improvement to simulated radar reflectivity in response to this adjustment will be evaluated in section 6 along with the results from the improved subcolumn generator scheme described below. The precipitation adjustment algorithm works as follows.

1. For each gridbox, PREC_SCOPS is first used to determine the binary precipitation mask $p_{i,k}$, where

$$p_{i,k} = \begin{cases} 0 & \text{if subcolumn is not precipitating} \\ 1 & \text{if subcolumn is precipitating} \end{cases}$$

2. The number of cloudy levels c_i^{num} at each subcolumn i is calculated by summing the binary cloud occurrence $c_{i,k}$ over all levels, such that

$$c_i^{\text{num}} = \sum_{k=1}^{n_{\text{levels}}} c_{i,k}.$$

3. The number of cloudy levels c_i^{num} is then sorted from least cloudy to most cloudy to obtain the sorted array $\hat{c}_i^{\text{num}}$, retaining the indices to the original column order I_i such that $\hat{c}_{I_i}^{\text{num}} = c_i^{\text{num}}$.
4. For each level, if the fraction of precipitating subcolumns in the binary mask exceeds the input precipitation fraction P_k at that level, then precipitation is removed (binary mask elements are zeroed out) from subcolumns in order of least cloudy columns to most cloudy until the fractions are consistent. If the fraction of precipitating subcolumns in the binary mask is less than the input precipitation fraction at that level, then precipitation is added to subcolumns in order of most cloudy columns to least cloudy.

In addition to evaluating the default and precipitation-adjusted COSP schemes, we also evaluate the performance of a different subcolumn scheme that is an extension of that presented by Räisänen et al., 2004 (2004, hereafter R04). The key features of the R04 scheme are a more flexible overlap treatment and the ability to generate horizontally heterogeneous subcolumn condensate distributions. Overlap is assumed to be a linear combination of maximum and random, with the weighting between maximum and random overlap determined by a parameter that depends only on the separation between cloudy layers (which needs to be parameterized). This type of overlap is referred to as *generalized overlap* (or *exponential overlap* due to the fact that the overlap parameter is often parameterized as a decaying exponential function of separation distance). Condensate variability is handled by sampling condensate amounts from an assumed probability distribution. This sampling is done in such a way as to be consistent also with an assumption about the condensate rank correlation (the degree to which the distributions in adjacent layers are correlated). We extend the R04 scheme to handle precipitation by using PREC_SCOPS to determine the subcolumn precipitation occurrence and then using the R04 algorithm for sampling the subcolumn condensate amounts. More details on the scheme and the implementation are provided in Appendix C, and our implementation of the scheme (including the precipitation adjustment scheme described above) is publicly available on GitHub (https://github.com/brhillman/genvar_subcol).

Because the treatment of subcolumn precipitation is critical to obtaining reasonable simulations of radar reflectivity factor from large-scale model output (see section 6), we extend the R04 subcolumn generator to also generate stochastic subcolumns of precipitation condensate with horizontally heterogeneous condensate amount in order to also improve the treatment of unresolved precipitation for use with the simulators. As an initial approach to extending this subcolumn scheme to handle precipitation, the subcolumn cloud occurrence is first generated using the subcolumn generator described above. The subcolumn precipitation occurrence is then generated using the PREC_SCOPS routine from COSP, with the precipitation adjustment described above to constrain the number of precipitating subcolumns by the precipitation fraction from the model. The subcolumn precipitation condensate amount is then prescribed in a similar manner to the subcolumn cloud condensate amount but with a separate rank correlation for precipitation and in general a separate assumed probability distribution. More elaborate approaches could certainly be designed to handle the precipitation, but the results in section 6 suggest that this simple approach performs quite well.

In order to use the R04 scheme, we must provide (1) the overlap parameter that determines the weighting between maximum and random overlap between each pair of neighboring layers, (2) the condensate rank correlation between adjacent layers for both cloud and precipitation, and (3) a probability distribution for the horizontal variability of condensate. Here two different approaches are used: one in which these quantities are parameterized based on high-resolution model output, and one in which these quantities are calculated exactly from high-resolution model output. The combination enables evaluating both the ability of the subcolumn generator to reproduce realistic cloud properties assuming *perfect parameterization* (that is, testing the generalized overlap and assumed distributional form for the variability, where the values needed are perfectly known and hence testing the limits of the subgrid representation itself), as well as the sensitivity of the outputs to the parameterization of these quantities in a manner that might be implemented in a GCM. Parameterization of these values are described in Appendix D. As we will discuss further in section 7, we are not suggesting that the parameterization used here be implemented in models. Rather, we argue that it is critical that the subgrid generator used with the simulator be consistent with assumptions made in the host model. Our intent here is to demonstrate (and quantify) how assumptions in the subgrid generator impact the simulated fields.

3. Evaluation Framework

The modular nature of COSP enables bypassing the subcolumn generation step if model fields with sufficiently high resolution are available, such as when using COSP with a cloud-resolving model (CRM; Marchand & Ackerman, 2010; Marchand et al., 2009). This enables testing of various subgrid assumptions by creating from high-resolution (resolved) cloud condensate fields a set of modified fields that mimic various subgrid assumptions, running the modified fields through the individual simulators, and then comparing the COSP-simulated outputs with a baseline simulation obtained using the original unmodified high-resolution fields. A similar approach has been used to evaluate sensitivities of radiative fluxes and heating rates using output from a limited-area CRM to provide the resolved cloud condensate fields (Barker et al., 1999; Wu & Liang, 2005).

In this study, output from a multiscale modeling framework (MMF; Randall et al., 2003) simulation is used as the source for resolved cloud condensate fields (rather than output from a limited-area CRM) to provide more comprehensive spatial context for sensitivities. The MMF replaces the convection and cloud parameterizations in a traditional GCM with a two-dimensional CRM running within each gridbox. This framework has been implemented using the National Center for Atmospheric Research (NCAR) CAM as the host GCM and the System for Atmospheric Modeling as the embedded CRM (Super-Parameterized Community Atmosphere Model or *SP-CAM*; Khairoutdinov & Randall, 2001), but the MMF has also been implemented with a completely different GCM and CRM (Tao et al., 2009) and with a variety of different schemes for handling turbulence, clouds, and aerosols (Cheng & Xu, 2011, 2013). MMF model output provides sufficiently high-resolution (typically 4 km or finer) cloud and precipitation fields to run the simulators directly (bypassing the subcolumn generator in COSP) and also provides the global coverage necessary to evaluate the impact of modifying the inputs on both the global and regional diagnostics that are typically used to evaluate the performance of large-scale models (e.g., Gleckler et al., 2008). The SP-CAM simulation used in this study was run with 2° horizontal resolution with 26 vertical levels for the host GCM, and with an embedded east-west oriented two-dimensional CRM with 64 columns, 4 km horizontal resolution, 24 vertical levels (sharing the bottom 24 levels with the host GCM), and single-moment bulk cloud microphysics. Further details of the model configuration and simulation are given by Khairoutdinov et al. (2005) and Marchand et al. (2009). Model outputs for the single month of July 2000 are used for this study.

In order to separately evaluate the sensitivity of COSP diagnostics to cloud overlap and condensate heterogeneity assumptions, modified cloud and precipitation condensate fields are created from the original CRM fields from the SP-CAM simulation. These modifications are described below, and total cloud and precipitation condensate amount from each case are shown for an example gridbox (00 UTC 1 July 2000, 10°N, 180°E) in Figure 1. The first row shows resolved condensate from the CRM, the second two rows demonstrate homogenizing and regenerating subcolumns using the default scheme in COSP, and the last three demonstrate regenerating subcolumns following modifications to the subcolumn generator.

CRM. The original, unmodified cloud and precipitation condensate fields from the CRM embedded in SP-CAM. This is used as the baseline against which sensitivities are evaluated.

CRM-HOM. Cloud condensate amounts from the CRM are homogenized by replacing the condensate amount in each cloudy CRM column in each gridbox with the gridbox in-cloud average (for each level). This is repeated for precipitation and is done separately for each hydrometeor type (cloud liquid, cloud ice, precipitating liquid, and precipitating ice). No change is made to the binary cloud or precipitation occurrence (which columns and levels are cloudy or clear), so this modification retains the exact cloud and precipitation occurrence overlap with the original CRM fields.

MRO-HOM. CRM-scale cloud and precipitation fields are *regenerated* from the original CRM fields by first calculating the gridbox mean cloud and precipitation condensate fields and the gridbox-mean cloud profiles (by averaging over all of the CRM columns within a gridbox) and then using the subcolumn generator in COSP (SCOPS and PREC_SCOPS, Appendix A) to sample stochastic subcolumns consistent with the gridbox means, maximum-random cloud overlap, and horizontally homogeneous cloud, and precipitation condensate amounts. This example illustrates the particularly problematic (although not uncommon) situation in which cloud fraction (and thus precipitation occurrence in the default scheme) is large at some level. Because PREC_SCOPS assumes that precipitation falls through all lower levels in a given subcolumn, precipitation occurrence is overestimated throughout the lower levels.

MRO-HOM-PADJ. CRM-scale cloud and precipitation fields are regenerated from the original CRM fields as described above for the MRO-HOM case, but precipitation occurrence is then adjusted to match the known

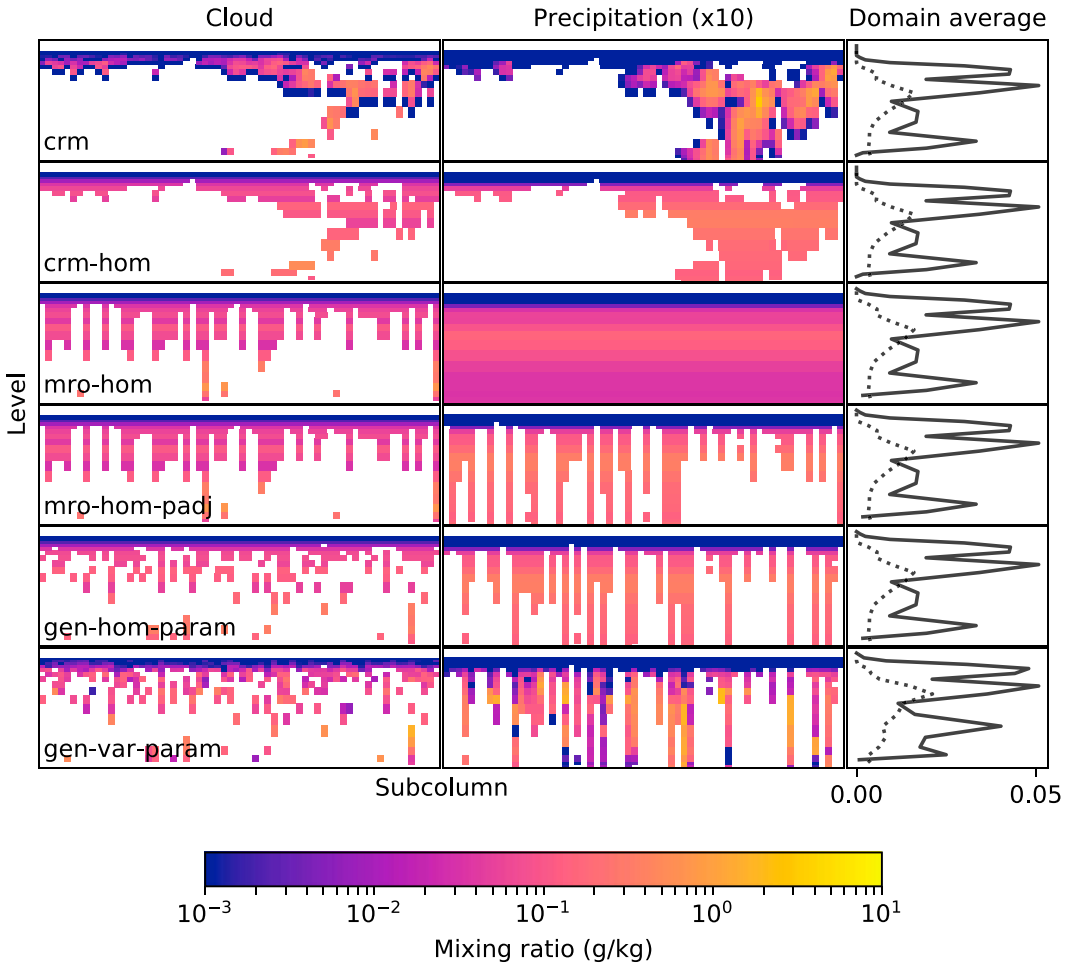

Figure 1. Total cloud (left) and precipitation (right) mixing ratios for each case described in the text, generated from a single gridbox (00 UTC 1 July 2000, 10°N, 180°E). Vertical profiles of the domain averages are shown in the far right panels, with solid lines representing the domain-averaged cloud water mixing ratio and dotted lines the precipitation mixing ratio.

CRM precipitation occurrence using the technique described in Appendix B. This *precipitation adjustment* can be used with models that predict profiles of precipitation occurrence.

GEN-VAR-PARAM and GEN-VAR-EXACT. CRM-scale cloud and precipitation fields are regenerated from the original CRM fields similar to the MRO-HOM case, but using the R04 generator. Two different configurations are used, one with the cloud overlap, condensate rank correlation, and variance parameterized as described in Appendix D (which we refer to as GEN-VAR-PARAM), and one in which the overlap, condensate rank correlation, and variance are calculated directly from the CRM-scale cloud and precipitation fields at each time step (which we refer to as GEN-VAR-EXACT). Comparing these two different configurations allows separation of sensitivities that arise due to errors in the parameterization from those that arise due to fundamental shortcomings of the subcolumn generator and underlying assumptions (i.e., overlap that can be described as a linear combination of maximum and random, and gamma-distributed condensate).

GEN-HOM-PARAM and GEN-HOM-EXACT. The same as GEN-VAR-PARAM and GEN-VAR-EXACT but using homogeneous condensate amounts.

CRM-scale fields from these cases are run through an offline driver we have written to run COSP on archived output from the MMF in order to obtain COSP outputs from each of these cases (this driver code is publicly available on Github (https://github.com/brhillman/cosp_mmf). The COSP outputs from these cases are compared to evaluate the sensitivity to the various subgrid treatments as follows:

$$
\begin{aligned}
E_{\text{hom}} &= X_{\text{crm-hom}} - X_{\text{crm}} \\
E_{\text{mro}} &= X_{\text{mro-hom-padj}} - X_{\text{crm-hom}} \\
E_{\text{precip}} &= X_{\text{mro-hom}} - X_{\text{mro-hom-padj}} \\
E_{\text{total}} &= E_{\text{hom}} + E_{\text{mro}} + E_{\text{precip}},
\end{aligned}
$$

where X represents a simulator output (such as Multiangle Imaging Spectroradiometer [MISR]-simulated cloud cover) and E represents the error due to homogeneity assumption (E_{hom}), the maximum-random overlap assumption (E_{mro}), the PREC_SCOPS precipitation treatment (E_{precip}), or due to the combination of all of these components (total). Similarly, errors arising due to each component of the new subcolumn generator are calculated by taking the following differences:

$$
\begin{aligned}
E_{\text{gen-var}} &= X_{\text{gen-var}} - X_{\text{crm}} \\
E_{\text{gen}} &= X_{\text{gen-hom}} - X_{\text{crm-hom}} \\
E_{\text{var}} &= E_{\text{gen-var}} - E_{\text{gen}}.
\end{aligned}
$$

4. MISR-Simulated Cloud Cover

The MISR simulator estimates the cloud top heights that would be retrieved by the MISR instrument from the model input. The cloud top heights are aggregated together with the column cloud optical depth (which is input to the simulator) into joint histograms consistent with those produced by the MISR instrument team and available as a global, gridded data product. Cloud cover for specific cloud types can be calculated from these joint histograms by summing appropriate bins in the joint histograms. For example, the midtopped cloud cover can be calculated by summing the histogram components representing clouds with cloud tops between 3 and 7 km.

Figure 2 shows errors in MISR-simulated monthly-mean cloud cover for cloud tops at all altitudes (optical depth $\tau > 0.3$), high-topped cloud cover (cloud top height $z_c > 7$ km, $\tau > 0.3$), midtopped cloud cover ($3 < z_c < 7$ km, $\tau > 0.3$), and low-topped cloud cover ($z_c < 3$ km, $\tau > 0.3$) arising from the maximum-random overlap and homogeneous condensate assumptions. From here onward, the expression *all* will be used to refer to clouds with tops at all altitudes (i.e., the sum of high-, middle-, and low-topped cloud cover). The left column shows the total error ($E_{\text{total}} = E_{\text{mro-hom}}$) in regenerating condensate from gridbox means, the middle column shows the error due to homogenizing the cloud condensate within each gridbox (E_{hom}), and the right column shows the errors resulting from the maximum-random overlap assumption (E_{mro}). We note that precipitation has no impact on the MISR simulator and so E_{precip} is zero.

Errors in MISR-simulated total cloud cover due to homogenizing the cloud condensate (top row, middle panel) are everywhere positive. By homogenizing the cloud condensate, the total number of CRM columns that contain cloud condensate has not actually been changed nor have those columns been rearranged in any way. Rather, the increase in the simulated total cloud cover is explained in terms of how *cloud* is defined using the MISR simulator outputs. In order to make more reasonable comparisons with satellite observations, which have finite detection capabilities, columns are considered cloudy only if the total column optical depth exceeds some threshold value, nominally $\tau > 0.3$. Homogenizing the condensate changes the distribution of optical depth. This happens because CRM columns with low condensate amounts (and thus lower resulting optical depths) often occur alongside columns with larger condensate amounts within the same gridbox, such that taking the average results in a squeezing of the distribution of condensate (less occurrence in the tails of the distribution and more near the mode), so a greater number of columns exceed the optical depth threshold. The increase in total cloud cover due to this effect is modest, and only results in an increase of 2% in the global mean but with larger regional errors on the order of 4 to 6%. These errors (and all errors discussed in this text) are absolute errors rather than relative errors; for example, in this case the homogenization results in an increase in the global-average cloud cover from about 53% to 55%, for an absolute error of 2%. Errors due to this effect are larger for the diagnosis of high-topped cloud cover and can exceed 8–10% in the deep tropics, especially over the tropical warm pool region, over the Maritime Continent, and over the Indian Ocean. These regions are dominated by deep convective cloud systems with associated cirrus that often have very low optical depths.

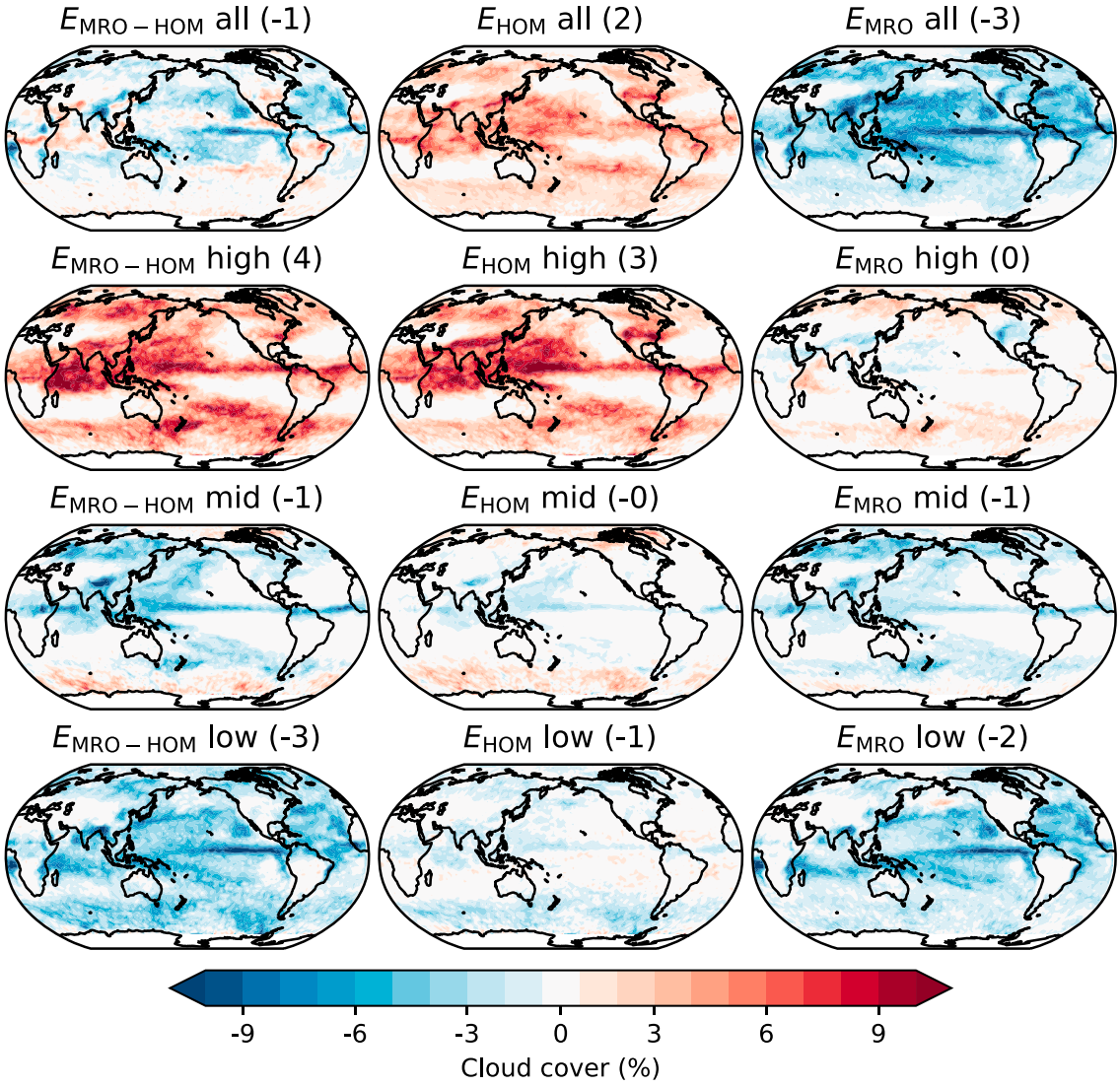

Figure 2. Sensitivities in MISR-simulated all-, high-, middle-, and low-topped cloud cover (simulated July 2000 means) due to regenerating subcolumns with maximum-random overlap and homogeneous condensate. Shown from left to right are the total error in using the subcolumn generator relative to the MISR-simulated outputs from the baseline CRM fields (MRO-HOM minus CRM), the component of the error due to using horizontally uniform condensate amounts (CRM-HOM minus CRM), and the component of the error due to using the maximum-random overlap treatment (MRO-HOM minus CRM-HOM). Numbers in parenthesis indicate the area-weighted global means. MISR = Multiangle Imaging Spectroradiometer; CRM = cloud-resolving model.

While errors due to homogenizing cloud condensate are primarily positive, the errors in cloud cover (clouds with tops at all altitudes) due to using the maximum-random overlap assumption are negative nearly everywhere, showing that implementing maximum-random overlap tends to decrease the total vertically projected cloud cover. The decrease in cloud cover is a result of the maximum-random overlap assumption tending to overestimate the vertical correlation in adjacent cloudy layers (e.g., Barker, 2008; Hogan & Illingworth, 2000; Mace & Benson-Troth, 2002). The decrease is only 3% in the global mean but can reach values exceeding 10% regionally, especially in the tropics. The decrease is largest for the low-topped clouds. This is due to increased *shielding* of low-topped (and mid-topped) clouds by high-topped clouds due to the increased vertical correlation introduced using maximum-random overlap, such that low-topped clouds tend to exist too frequently beneath high-topped clouds. This has a minimal affect on the high-topped clouds, and in fact high-topped cloud cover actually increases slightly throughout some regions in middle to high latitudes. This is because increasing vertical correlation of cloudy layers tends to increase the cloud water path (and hence the cloud optical depth), which increases the number of columns where the high-level cloud optical depth exceeds the $\tau > 1$ threshold.

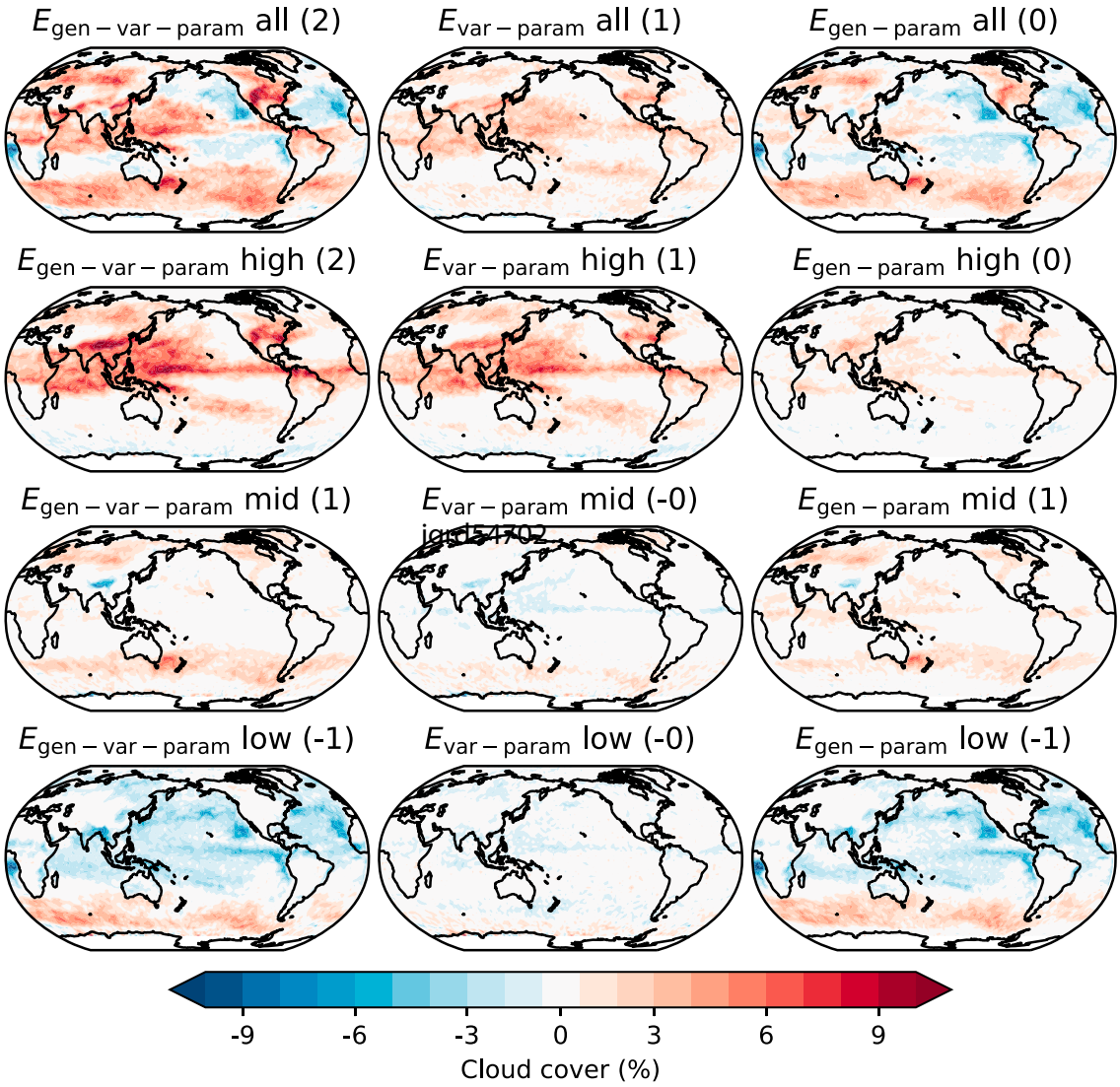

Figure 3. Sensitivities in MISR-simulated cloud cover (simulated July 2000 means) due to using the R04 subcolumn generator (GEN-VAR-PARAM) with generalized overlap and horizontally variable condensate, with overlap, condensate rank correlation, and condensate variance parameterized based on output from the MMF. Shown from left to right are the total error in using the subcolumn generator relative to the outputs from the baseline CRM fields (GEN-VAR-PARAM minus CRM), the component of the error due to the treatment of horizontal heterogeneity in condensate ($E_{\text{GEN-VAR-PARAM}}$ minus $E_{\text{GEN-HOM-PARAM}}$), and the component of the error due to the generalized overlap treatment (GEN-HOM-PARAM minus CRM-HOM). Numbers in parenthesis indicate the area-weighted global means. MISR = Multiangle Imaging Spectroradiometer; MMF = multiscale modeling framework; CRM = cloud-resolving model.

The errors in MISR-simulated cloud cover due separately to homogenizing cloud condensate and using MRO are mostly compensatory in regards to cloud cover for clouds with tops at all altitudes (top row of Figure 2) but produce noteworthy errors in high-, middle-, and low-topped cloud cover (lower three rows, respectively). The effect on simulated high-topped clouds due to the two components of the error are both positive in sign, so that these components of the error combine to produce much larger errors in simulated high-topped cloud, with almost 5% increase in the global mean and an increase greater than 10% throughout much of the deep tropics. Low-topped cloud, on the other hand, is decreased by 4% in the global average. Combined with the 2% decrease in midtopped clouds, this almost completely compensates the increase in high-topped cloud, such that errors in the total cloud cover (discussed above) appear small.

Figure 3 shows the errors in MISR-simulated cloud cover by cloud top height that arise due to using the R04 subcolumn generator with parameterized values of overlap, condensate rank correlation, and condensate variance. The errors due to the treatment of variability are reduced relative to the errors that arise from using homogeneous condensate (middle column) but are of the same sign. The persistence of the overestimation is

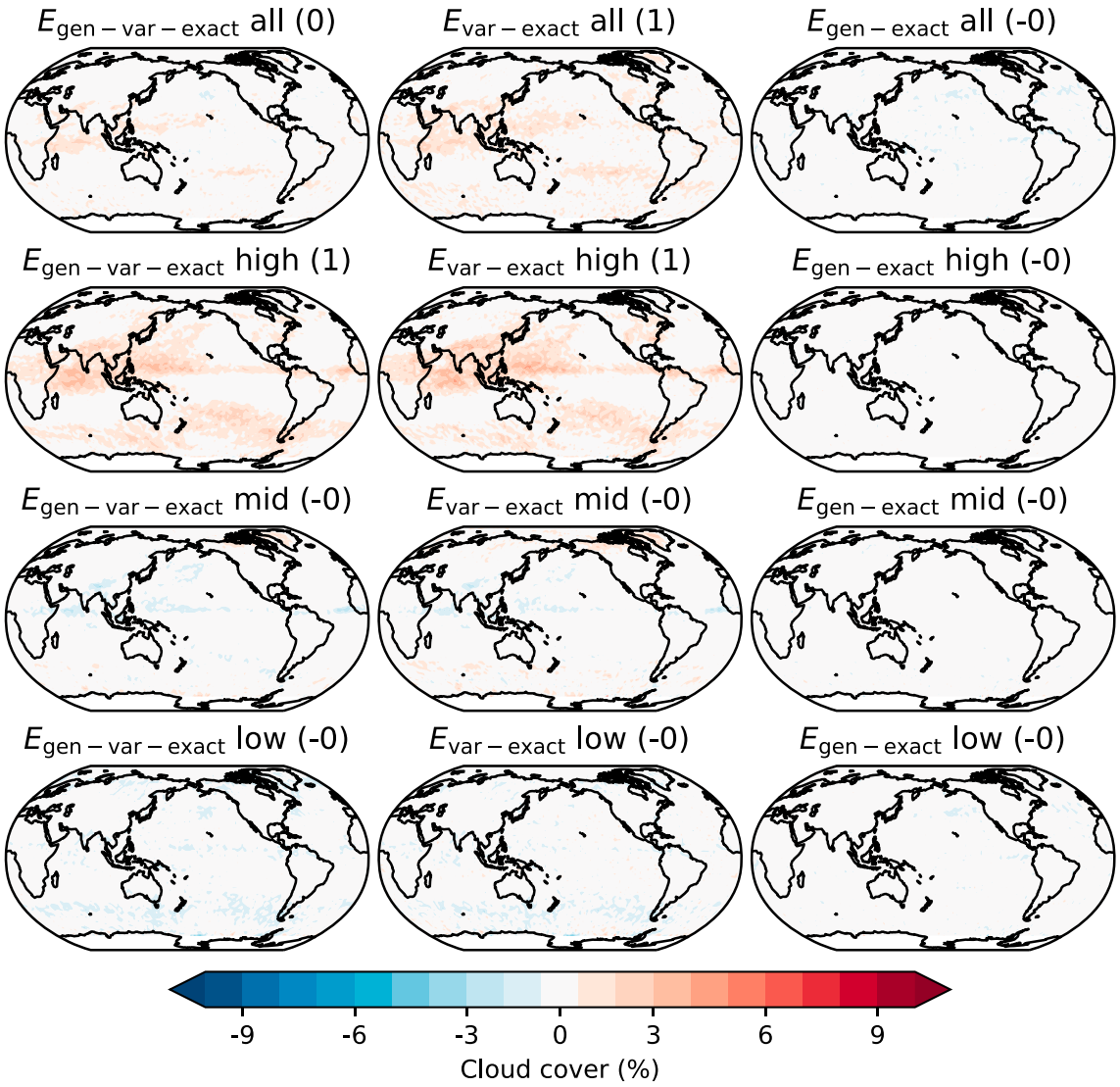

Figure 4. As in Figure 3 but using exact overlap, rank correlation, and variance calculated directly from MMF fields. MMF = multiscale modeling framework.

a consequence of the parameterization of the variance failing to capture the full spread of condensate values present in the original CRM fields and is discussed in Appendix D. Much of the error shown in Figure 3 is due to the parameterization rather than the new subgrid framework, and Figure 4 shows the errors using the R04 subcolumn generator with these parameters calculated directly from the MMF fields at each output time. Figure 4 shows that the errors are quite small compared with Figure 3.

Errors due to using generalized overlap with parameterized decorrelation length (right column) show clear spatial patterns, with overestimation of cloud cover especially in the Southern Ocean and also somewhat in the tropical western Pacific and over the continents, and an underestimation of cloud cover elsewhere. In the Southern Ocean, these errors appear to manifest primarily in the low-topped cloud. The clear spatial structure in these errors is a consequence of using a globally constant decorrelation length for cloud occurrence overlap (as was done here), which is not entirely sufficient to characterize the overlap of clouds simulated by SP-CAM (and likely real clouds in the physical atmosphere). Nonetheless, even this simplistic assumption notably reduces the errors relative to using maximum-random overlap.

Globally averaged and including clouds at all altitudes, the total error in using the default COSP scheme ($E_{\text{mro-hom}}$, upper left column of Figure 2) is actually lower than the total error in using the R04 scheme ($E_{gen-var}$, upper left column of Figure 3). This is due to a decreased compensation of the components of the error that arise due to the overlap and variability assumptions when using the R04 scheme: in the MRO-HOM case, the

errors due to the homogeneous and overlap assumptions are opposite in sign, but the errors due to the variability and overlap assumptions in the GEN-VAR-PARAM case are both positive. Again, using the exact overlap calculated directly from the MMF fields at each output time (GEN-VAR-EXACT), the framework itself is capable of reproducing the overlap structure of the MMF fields, but our simple parameterization using a globally constant decorrelation length scale falls short.

5. ISCCP-Simulated Cloud Cover

The International Satellite Cloud Climatology Project (ISCCP)-simulated output (and ISCCP retrievals) are similar to MISR histograms, but based on somewhat different retrievals (Marchand et al., 2010), and are saved as cloud top pressure (p_c) and cloud optical depth joint histograms. Low-topped clouds are defined as those with $p_c > 680$ hPa, mid-topped clouds are those with $680 < p_c < 440$ hPa, and high-topped clouds are those with $p_c < 440$ hPa. Figure 5 shows the errors in ISCCP-simulated cloud area by cloud top pressure using the COSP subgrid scheme (SCOPS with maximum-random overlap and homogeneous condensate). Figure 6 shows the errors using the R04 scheme with parameterized overlap and condensate rank correlation and variance, and Figure 7 shows the errors using exact parameters calculated directly from the MMF fields. The errors in ISCCP-simulated all-cloud area (clouds at all altitudes, top panels) are nearly identical to the MISR-simulated all-cloud errors, but there are subtle differences in the errors for clouds partitioned by cloud top height. The largest differences are in the high-topped cloud, where the overlap and homogeneous errors (second row, middle and right columns) combine to produce a globally averaged error of 4% cloud area in MISR-simulated high-topped cloud but only 2% cloud area in ISCCP-simulated high-topped cloud. This is due primarily to a larger compensation of errors with the ISCCP simulator. Specifically, the ISCCP simulator has a larger underestimate in the occurrence of high-topped cloud due to cloud overlap (second row, right column) than MISR. This underestimate partially compensates the overestimate resulting from the homogeneous assumption (which affects both simulators equally). These differences likely arise because ISCCP high-topped cloud area is based on IR detections and is not as sensitive as MISR to the increased cloud water path that results from the MRO approximation. However, it is also possible that some of these differences are due to ISCCP using cloud top pressure bins as opposed to cloud top height bins, such that definitions of low-, middle-, and high-topped clouds are not quite the same.

6. CloudSat-Simulated Reflectivity and Hydrometeor Occurrence

The 94-GHz radar reflectivity (Z_e) retrieved by the CloudSat Cloud Profiling Radar (Tanelli et al., 2008) is simulated in COSP using the Quickbeam radar simulator (Haynes et al., 2007). Quickbeam accounts for attenuation due to both hydrometeors and gases. Because the CloudSat cloud radar has difficulty detecting hydrometeors with reflectivity below −27.5 dBZ, this threshold is often used when comparing simulated reflectivity from models to CloudSat observations (Marchand et al., 2009). The fraction of profiles with radar reflectivity above this threshold can be taken as a measure of the *hydrometeor occurrence*, which is the fraction of radar volumes containing either cloud or precipitation, or both.

Figure 8 shows the errors in CloudSat-simulated hydrometeor occurrence ($Z_e > -27.5$ dBZ; July 2000 mean) using the MRO-HOM, GEN-VAR-PARAM, and GEN-VAR-EXACT fields as well as the components of the errors due separately to the treatment of overlap, precipitation occurrence, and condensate variability. Homogenizing the cloud and precipitation condensate amounts (top left column) results in an increase in simulated hydrometeor occurrence at all altitudes. These errors are especially large in the deep tropics and in both Northern Hemisphere and Southern Hemisphere midlatitudes. The causes of these errors are discussed in the following paragraphs, but we note here that these errors are substantially reduced using the new scheme, even with the simple parameterization of overlap, condensate rank correlation, and variance ($E_{\text{GEN-VAR-PARAM}}$). The errors in simulated hydrometeor occurrence due to overlap (third row) are small relative to that of variability (second row). Errors due to using the unconstrained precipitation occurrence in PREC_SCOPS are noteworthy, especially in the tropics. This is consistent with the results published by Di Michele et al. (2012) and shows that it is helpful to constrain the precipitation occurrence in each layer. However, errors due to the treatment of horizontal variability remain problematically large.

The causes of the errors in hydrometeor occurrence are understood more fully by examining the CloudSat-simulated reflectivity with height histograms. Figure 9 shows the simulated radar reflectivity with height histograms for the Northern Hemisphere tropics (0°N to 10°N latitude) using the full CRM fields (top

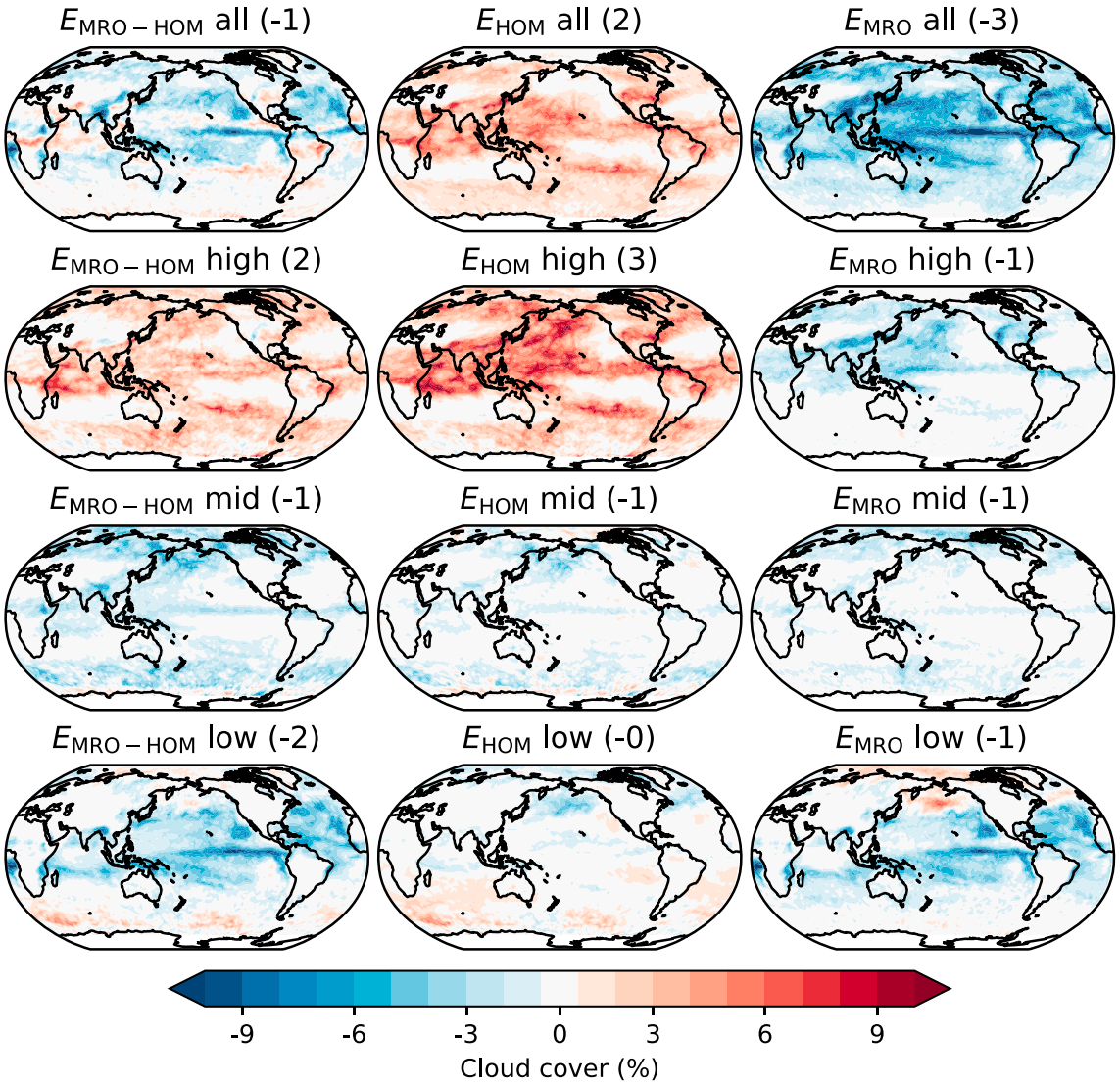

Figure 5. As in Figure 2 but for International Satellite Cloud Climatology Project-simulated clouds.

row, left panel), using the subcolumn generator with the default COSP scheme (maximum-random overlap, homogeneous condensate, and no precipitation adjustment; MRO-HOM, top row, second panel), and using the R04 subcolumn generator with generalized overlap and heterogeneous condensate (GEN-VAR-PARAM and GEN-VAR-EXACT, top row, third and fourth panels). The histograms all have similar patterns of high frequency along a characteristic curve (see discussion in ; Marchand et al., 2009). While the default COSP scheme has a similar curve, there is much larger occurrence along the characteristic curve than the CRM case and lower occurrence off the characteristic curve. This is clearer in the bottom row of Figure 9, which shows the errors (difference) between the two (bottom left panel). The other panels in the bottom row of Figure 9 show the errors with the precipitation adjustment (bottom row, second panel) and using generalized overlap with heterogeneous condensate, with overlap, rank correlation, and variance both parameterized and calculated directly from MMF fields (bottom row, third and fourth panels). As in Figure 8, overlap errors are comparatively small and are not shown. Similar to the errors in MISR-simulated cloud cover, the source of these errors is driven by the squeezing of the distribution of condensate (toward its modal value) that results from replacing the subgrid distributions of condensate with the gridbox averages, which effectively reduces the tails of the distribution by removing the within-gridbox variability. The within-gridbox variability is restored, if imperfectly, by the (gamma distributed) heterogeneous condensate in the GEN-VAR-PARAM and GEN-VAR-EXACT cases, and thus, these errors are consequently reduced. The small differences between the GEN-VAR-PARAM

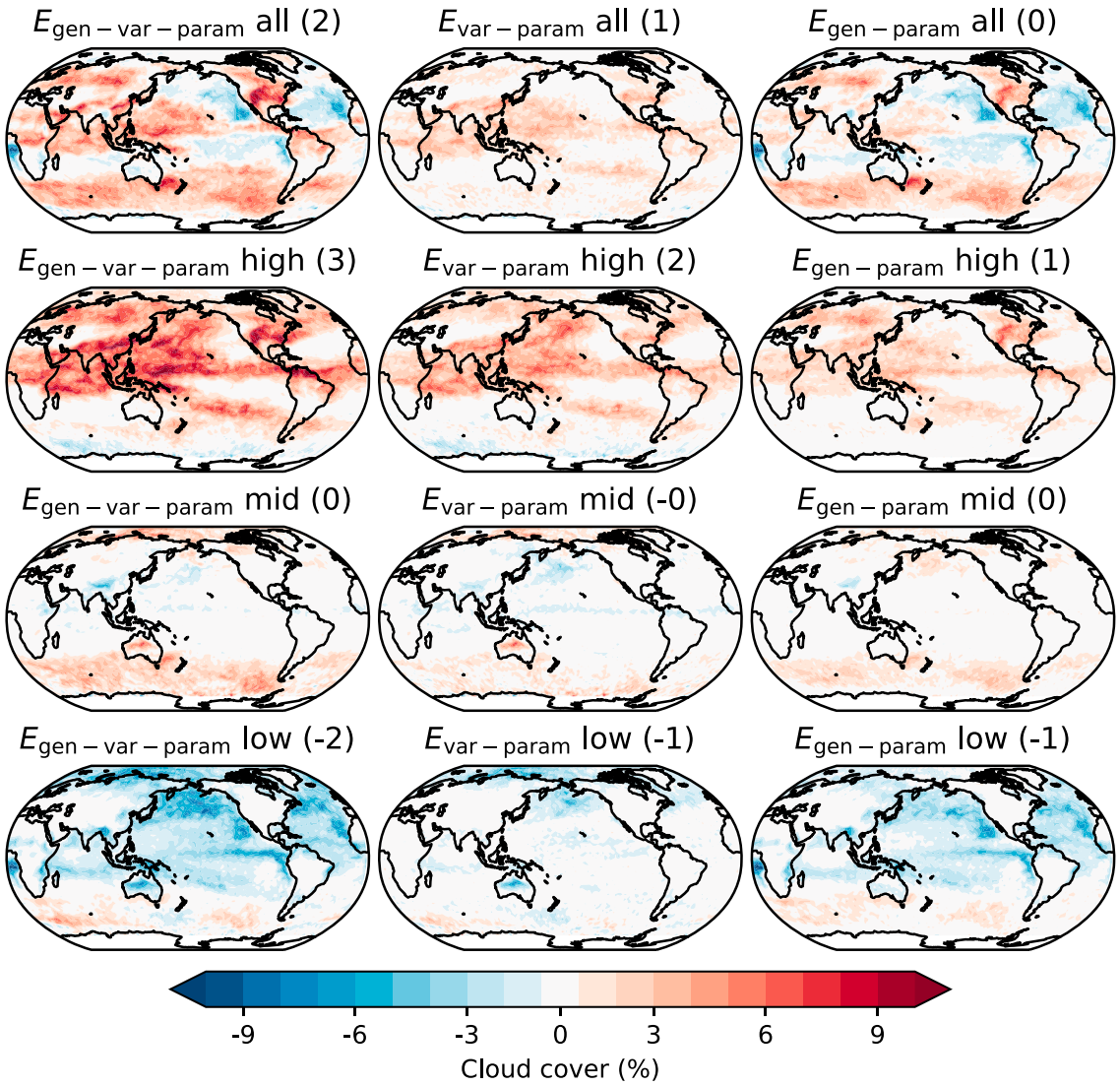

Figure 6. As in Figure 3 but for International Satellite Cloud Climatology Project-simulated clouds.

and GEN-VAR-EXACT errors suggest that most of the improvement in using the R04 subcolumn generator is realized just by reintroducing horizontally variable condensate. The remaining errors are likely due to the underlying assumptions about the condensate distributions (i.e., assuming gamma-distributed condensate).

Constraining precipitation fraction is also important, and failing to faithfully reproduce the CRM precipitation fraction leads to large increases in the occurrence of columns with large reflectivity values. The effect is more pronounced at low to middle levels (heights $z < 5$ km). This error is not surprising given the situation illustrated in Figure 1, which shows that the PREC_SCOPS subcolumn precipitation generator can dramatically overestimate the number of precipitating subcolumns.

7. Summary and Discussion

Current global models do not resolve individual cloud elements but rather represent most cloud-scale variability by way of parameterization. A common simplification made in large-scale models is that cloud (and precipitation) condensate is horizontally homogeneous on the gridbox scale, and that cloud occurrence follows maximum-random overlap. Previous authors have shown that these assumptions lead to biases in vertically integrated cloud cover (e.g., Cesana & Waliser, 2016) and in calculated fluxes and heating rates (e.g., Barker et al., 1999; Oreopoulos et al., 2012; Wu & Liang, 2005), and it is shown here that these assumptions also affect satellite-simulated cloud property retrievals.

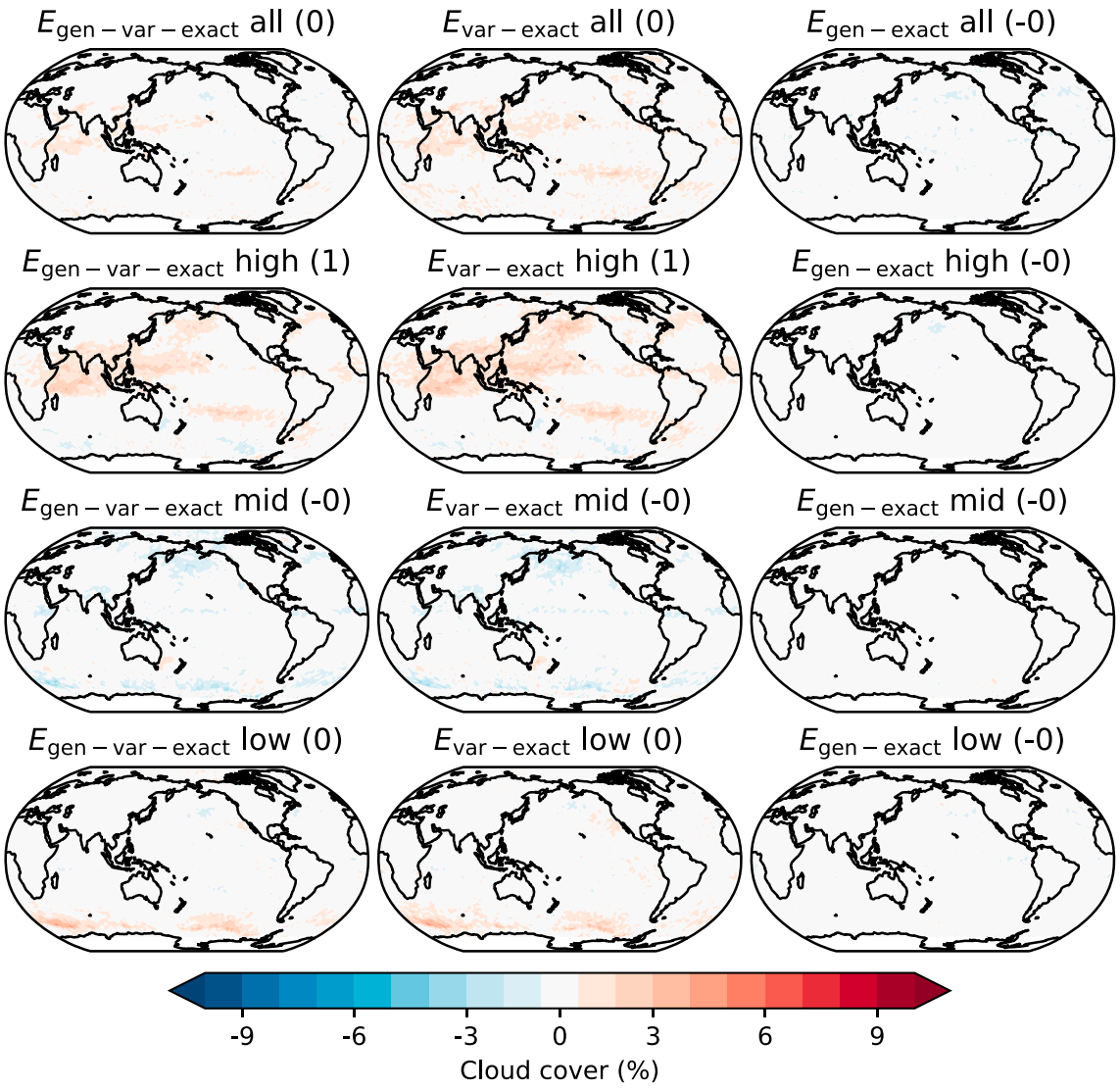

Figure 7. As in Figure 4 but for International Satellite Cloud Climatology Project-simulated clouds.

The assumption of homogeneous cloud properties tends to inflate MISR and ISCCP-simulated cloud cover (when counting all clouds with an optical depth $\tau > 0.3$) because columns with small optical depths in the tail of the distribution are sometimes shifted to values above the cutoff threshold by averaging with columns with larger optical depths. These errors occur primarily in high-topped clouds, and high-topped cloud occurrence can be overestimated by as much as 10% in regions with a lot of high-topped optically thin cloud (most notably throughout the deep tropics).

The maximum-random overlap assumption tends to decrease the cloud cover because it overestimates the overlap of vertically continuous clouds (Barker, 2008; Hogan & Illingworth, 2000; Mace & Benson-Troth, 2002). The global mean underestimate in total cloud cover due to this effect is small (only 3%), but regional errors are much larger (up to 10%). We note that the underestimation of cloud cover (especially of vertically integrated high-level cloud cover) when using maximum-random overlap is consistent with systematic biases in multimodel mean CALIPSO-simulated cloud cover (against CALIPSO retrievals) shown by Cesana and Waliser (2016), who also provide evidence that these biases are at least partly attributable to the maximum-random overlap assumption.

The errors in MISR and ISCCP-simulated cloud cover due to these two effects are generally opposite in sign and result in a partial cancelation. The error in all cloud cover (meaning clouds at all altitudes) is less than 2%, with high-topped cloud overestimated by 5% and low-topped cloud underestimated by 4% in the global mean.

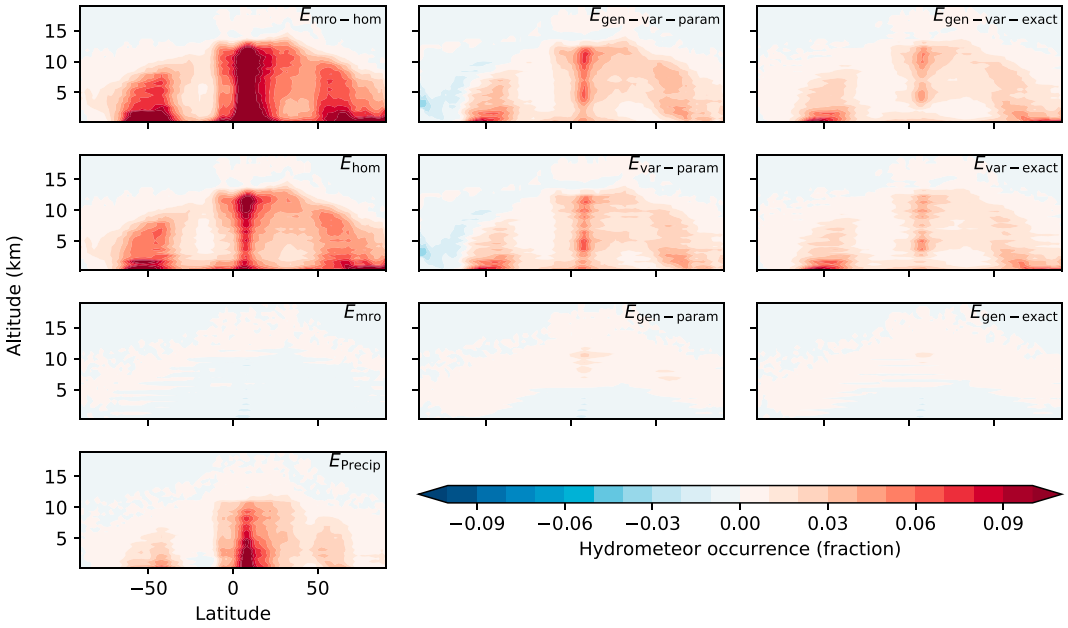

Figure 8. Sensitivities in zonally averaged CloudSat-simulated hydrometeor occurrence fraction with height (simulated July 2000 means). Hydrometeor occurrence fraction is defined as the fraction of columns at a particular height with radar reflectivity $Z_e > -27.5$ dBZ. In the top row are the total errors in the current (default) COSP subcolumn generator (MRO-HOM) and in the R04 subcolumn generator with parameterized (GEN-VAR-PARAM) and exact (GEN-VAR-EXACT) values relative to the baseline CRM case. The lower rows show how the total error in the top panel of each column is divided between errors due to horizontal variability (second row) and overlap schemes (third row). Neglecting precipitation fraction also increases the error in the default COSP scheme (bottom left panel). COSP = CFMIP Observation Simulator Package; CRM = cloud-resolving model.

The sensitivity in MISR-simulated total cloud cover identified here is generally less than the errors identified in GCMs (Bodas-Salcedo et al., 2011; Kay et al., 2012; Klein et al., 2013) and roughly equal to the spread in estimates of total cloud cover from satellite remote sensing retrievals (Marchand et al., 2010; Pincus et al., 2012). However, the regional errors in MISR-simulated cloud cover by cloud top height identified here are large and exceed the uncertainty of the MISR-retrieved high-topped cloud cover, which is estimated to be about 5% (Hillman et al., 2017). Thus, the sensitivity of MISR and ISCCP-simulated cloud cover to homogeneous cloud condensate and maximum-random overlap should not be ignored, especially as representations of clouds in GCMs improve.

Simulated CloudSat radar reflectivity is found to be sensitive to the treatment of condensate variability and precipitation occurrence but generally insensitive to the treatment of cloud occurrence overlap. Homogenizing the cloud and precipitation condensate leads to a narrowing of the distribution of simulated radar reflectivity, increasing the frequency of occurrence along the characteristic curve in reflectivity-height space. Employing a reflectivity cut-off to determine hydrometeor occurrence then results in an apparent increase in the hydrometeor occurrence when homogenizing the cloud and precipitation properties and an apparent increase in precipitation occurrence. The increase in simulated hydrometeor occurrence fraction reaches a value of 10% in high altitudes in the tropics and in low altitudes in middle to high latitudes.

The sensitivities in satellite-simulated cloud and precipitation properties that arise due to overlap and variability assumptions provide further motivation for the need to better represent subgrid-scale cloud and precipitation overlap and horizontal condensate variability in models and show that consistently treating overlap and horizontal variability between model physics and COSP is crucial to properly diagnosing inherent model biases from COSP outputs. Sensitivities in COSP outputs to alternative treatments of overlap and variability are shown by generating subcolumns using the algorithm of Räisänen et al. (2004) (extended and applied to both cloud and precipitation condensate). The Räisänen et al. (2004) generator allows for a more realistic representation of cloud overlap by representing overlap as a linear combination of maximum and random overlap, as well as heterogeneous cloud and precipitation condensate amount sampled from assumed probability distributions. Implementing the new subcolumn generator with overlap and variability

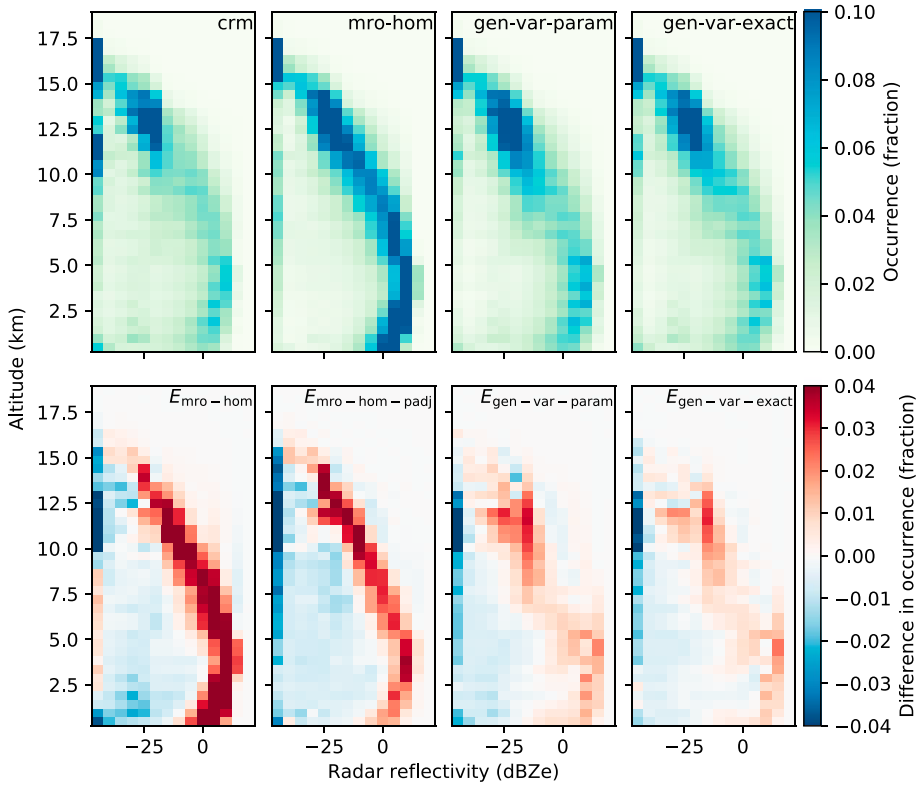

Figure 9. CloudSat-simulated reflectivity-height histograms (top row) and errors in reflectivity-height histograms due to overlap and variability assumptions (bottom row) averaged over the North Hemisphere tropics (0°N to 10°N; simulated July 2000 means).

parameterized (using very simple globally averaged values obtained from the MMF) substantially reduces the errors in both MISR-simulated cloud cover and in CloudSat-simulated radar reflectivity factor when compared against the baseline COSP output from the MMF (in which no subcolumn assumptions are needed). Nonetheless, some systematic errors remain with our application of the Räisänen et al. (2004) subcolumn generator. In particular, there is room for improvement in the parameterization of cloud overlap, which, for example, might be parameterized based on environmental conditions such as large-scale vertical velocity. While the results presented using the Räisänen et al. (2004) are in much closer agreement with baseline COSP outputs derived from resolved condensate fields, we stress that improving subcolumn assumptions in COSP should only be done consistent with improvements in these assumptions throughout the physical parameterizations of a given model.

Given the errors in the radar simulator, one might reasonably ask how can one use the radar simulator in COSP with the current homogeneous condensate assumption. As demonstrated in Figure 9, the homogeneous condensate assumption causes an overestimate of hydrometeor occurrence along the characteristic curve, but it does not change the shape (or position) of the curve. The shape of the curve is controlled much more by the hydrometeor microphysics. A good use of the radar simulator is evaluating to what degree a model is able to reproduce the correct shape, rather than asking if the intensity (or amount) is too large. In previous work we found that the MMF model (as compared to CloudSat observations) has a characteristic shape with reflectivity values that are too large between 5 and 10 km in the tropics (Marchand et al., 2009). In effect the curves shown in Figure 9 are too *upright* and do not decrease with altitude as quickly as they should in the MMF. This indicates that this version of the MMF tends to have too much condensate or particles that are too large much of the time (and likely both).

The sensitivities presented here to unresolved variability are not unique to simulation of satellite-observable cloud diagnostics. It has been recognized that subgrid-scale variability affects many important processes in large-scale models, and researchers are working to develop explicit subgrid treatments for GCMs. This includes the so-called *statistical* or *assumed probability distribution* schemes, which predict the evolution of not only

the mean but also the probability distribution of total water (and to a degree the cloud and precipitation condensate) within each gridbox (e.g., Golaz et al., 2002; Guo et al., 2014; Thayer-Calder et al., 2015; Tompkins, 2002). Because these schemes are formulated using a probability distribution for the subgrid variability of condensate, they are a natural fit to the stochastic treatment of subgrid clouds and precipitation used in COSP. Thus, we are optimistic that in the future a separate parameterization of condensate variability (such as developed here based on MMF output) would not be needed in the subcolumn generator, but rather subgrid information on condensate variability would be obtained from a given model subgrid parameterization and applied in a consistent fashion to COSP and across all model processes. Indeed, the most recent release of COSP (version 2.0) has been designed to allow (and even encourages) modelers to input to COSP subcolumns generated independently of COSP (bypassing the SCOPS scheme), with the expectation that subcolumns input to COSP would be consistent with those used in the radiative transfer parameterization even as those schemes improve. Certainly, this would be the ideal solution, and we do not recommend that a new scheme be implemented only with COSP, independent of the subcolumn treatment in a model radiative transfer or microphysics parameterization. Rather, the results presented here demonstrate the need for consistency among these schemes; the sensitivities of COSP outputs to subgrid assumptions suggest that if these assumptions are inconsistent, then the simulators may misrepresent the model physics and tell a story that is inconsistent with how the model behaves internally.

Appendix A: Generating Subcolumns in COSP

SCOPS can generate subcolumns obeying random, maximum, or maximum-random overlap and can separately treat convective and stratiform cloud if such a distinction is made in the model. If the model distinguishes between convective and stratiform cloud, convective cloud is maximally overlapped and the remaining stratiform cloud may follow a separate overlap assumption (one of random, maximum, or maximum random). SCOPS takes as input the gridbox-mean total cloud fraction profile $\overline{c}_k$ (the fraction of the gridbox at each level k containing either stratiform or convective cloud) and the gridbox-mean convective cloud fraction profile $\overline{c}_k^{\text{conv}}$ and then outputs an ensemble of n_{col} binary subcolumn cloud occurrence profiles $c_{i,k}$, where for each subcolumn i and at each level k,

$$c_{i,k} = \begin{cases} 0 & \text{if subcolumn is clear} \\ 1 & \text{if subcolumn is stratiform cloud} \\ 2 & \text{if subcolumn is convective cloud} \end{cases}$$

Homogeneous cloud condensate amounts $q_{i,k}$ are then assigned to each cloudy subcolumn element $c_{i,k}$ by setting $q_{i,k} = \overline{q_k}$ for all $c_{i,k} > 0$, where $\overline{q_k}$ is the in-cloud gridbox-mean cloud condensate passed as input from the large-scale model. If the model distinguishes between stratiform and convective cloud, then this criteria is applied separately for stratiform and convective cloud (that is, for $c_{i,k} = 1$ and $c_{i,k} = 2$, with $\overline{q_k}$ given by the appropriate stratiform and convective cloud condensate amounts).

PREC_SCOPS takes as input the subcolumn cloud occurrence profiles $c_{i,k}$ as determined by SCOPS and either the gridbox-mean precipitation condensate amount (mixing ratio) or the gridbox-mean precipitation fluxes and outputs a similar ensemble of binary subcolumn precipitation occurrence profiles $p_{i,k}$. A predicted gridbox-mean precipitation fraction is *not* used in the default PREC_SCOPS scheme. Rather, the binary precipitation occurrence in each subcolumn element i, k is determined by considering each level k in turn from the top model level down to the bottom. If the gridbox mean precipitation (flux or mixing ratio) is nonzero at level k, then precipitation is assumed to occupy each subcolumn element i at level k that contains cloud ($c_{i,k} > 0$) *or* has precipitation in the level *above* ($p_{i,k-1} > 0$). That is, $p_{i,k} = \max(c_{i,k}, p_{i,k-1})$. Thus, once precipitation is established in a given subcolumn i, it will fall through successive layers until a completely precipitation free layer is encountered (one that has a zero gridbox-mean flux or mixing ratio). The algorithm includes additional provisions for cases in which precipitation is not assigned by this simple rule and is described by Zhang et al. (2010). Appendix B describes an extension of this scheme that constrains the binary precipitation occurrence by an input gridbox-mean precipitation fraction, if available. As in SCOPS, homogeneous precipitation condensate amounts are assigned to each precipitating subcolumn element by assigning $q_{i,k} = \overline{q_k}$ for all $p_{i,k} > 0$, again separately for stratiform and convective clouds if the model makes that distinction.

Appendix B: Constraining Precipitation Fraction

The PREC_SCOPS precipitation treatment implemented in COSP associates precipitation with cloud but fails to account for any estimate of precipitation fraction (the fraction of the gridbox that contains precipitation at any level) that may be diagnosed by the model. Furthermore, the precipitation treatment essentially assumes that once precipitation is diagnosed at a particular level in a subcolumn, it falls all the way down to the surface unless a precipitation-free layer (that is, a layer in which gridbox mean condensate is equal to zero) is encountered. This can lead to a gross overestimation of the number of precipitating subcolumns and, consequently, has a large impact on the radar simulator.

An adjustment to this default precipitation scheme is introduced here, following the work of Di Michele et al. (2012), in which subcolumn precipitation fraction is assumed to be available from the model (although not all models provide this information) and used to constrain the precipitation occurrence by either adding or removing the occurrence (or existence) of precipitation at each level until the fraction of subcolumns with precipitation at a given level matches the input precipitation fraction. Precipitation is added preferentially to columns with more (vertically integrated) cloudy levels and removed preferentially from columns with less cloudy levels. This is similar to the *PEVAP* adjustment described by Di Michele et al. (2012) (so-called because it represents an evaporation of precipitation), and the improvement to simulated radar reflectivity in response to this adjustment will be evaluated in section 6 along with the results from the improved subcolumn generator scheme described below.

Appendix C: An Improved Subcolumn Generator for COSP

Räisänen et al. (2004) (hereafter R04) introduce a stochastic subcolumn cloud generator that can handle both horizontally variable cloud condensate and *generalized* cloud overlap. In the generalized overlap assumption, the vertically projected cloud cover from the combination of two layers j and k is assumed to be a linear combination of that which would arise from maximum and random overlap, that is

$$\bar{c}_{j,k}^{\text{gen}} = \alpha_{j,k}\bar{c}_{j,k}^{\text{max}} + (1-\alpha_{j,k})\bar{c}_{j,k}^{\text{ran}} \tag{C1}$$

where $\bar{c}_{j,k}^{\text{gen}}$ is the combined (vertically projected) cloud cover (fraction) that would result from generalized overlap, $\bar{c}_{j,k}^{\text{max}}$ is the cloud cover that would result if the layers were maximally overlapped, $\bar{c}_{j,k}^{\text{ran}}$ is the cloud fraction that would result if the layers were randomly overlapped, and $\alpha_{j,k}$ is the *overlap parameter* that specifies the weighting between maximum and random overlap. The theoretical combined cloud fractions $\bar{c}_{j,k}^{\text{max}}$ and $\bar{c}_{j,k}^{\text{ran}}$ are defined as

$$\bar{c}_{j,k}^{\text{max}} = \max(\bar{c}_j, \bar{c}_k)$$
$$\bar{c}_{j,k}^{\text{ran}} = \bar{c}_j + \bar{c}_k - \bar{c}_j\bar{c}_k$$

where $\bar{c}_j$ and $\bar{c}_k$ are the partial cloud fractions of layers j and k, respectively (i.e., the fraction of the gridbox at levels j and k that are cloudy).

In general, equation (C1) is assumed to apply to any two pairs of layers, but for the practical implementation of the subcolumn generator R04 consider only adjacent layer pairs. Given $\alpha_{k,k-1}$ and the gridbox-mean cloud fraction $\bar{c}_k$ at each layer k, R04 describe a straightforward algorithm to stochastically generate a binary subcolumn clear/cloudy flag with n_{col} subcolumns that obeys the above overlap relationship by stepping down from the top of the atmospheric column and considering only adjacent layer pairs.

Once the cloud occurrence subcolumns are created, cloud condensate is assigned to the cloudy elements by drawing from a specified probability distribution for condensate amount. The choice of distribution should be consistent with the given model assumptions about subgrid variability of condensate, if such exists. Traditionally, models have assumed homogeneous subcolumn condensate (e.g., CAM4; Neale, Richter, et al., 2010), but some models have started to incorporate subgrid-scale variability into their cloud microphysics schemes. For example, the most recent version of the NCAR CAM assumes condensed cloud water follows a gamma distribution, but cloud ice and precipitation are still assumed to be homogeneous (CAM5; Neale, Gettelman, et al., 2010).

Condensate values are drawn such that the subcolumn distributions of condensate obey a specified rank correlation $\rho_{j,k}$ for condensate amount between layers j and k, where $\rho_{j,k}$ is the Pearson Product-Moment

Correlation coefficient of the ranks r_j and r_k of condensate at levels j and k, defined by

$$\rho_{j,k} = \frac{\mathrm{cov}(r_j, r_k)}{\sigma_{r_j}\sigma_{r_k}} = \frac{\sum_{i=1}^{n_{\mathrm{col}}}(r_{i,j} - \overline{r_j})(r_{i,k} - \overline{r_k})}{\sqrt{\sum_{i=1}^{n_{\mathrm{col}}}(r_{i,j} - \overline{r_j})^2}\sqrt{\sum_{i=1}^{n_{\mathrm{col}}}(r_{i,k} - \overline{r_k})^2}} \tag{C2}$$

where the overbars denote horizontal averages over all n_{col} subcolumns. Again, only rank correlations between adjacent layers are considered in the R04 approach, so that all that is needed to describe the rank correlation is $\rho_{k,k-1}$ for each level k.

The problem of generating stochastic subcolumns of cloud condensate with generalized occurrence overlap and heterogeneous condensate distributions then reduces to specifying the parameters $\alpha_{k,k-1}$ and $\rho_{k,k-1}$ for each pair of adjacent layers within a gridbox and specifying an appropriate probability distribution from which to sample condensate amount at each gridbox and time step.

Previous studies (based largely on cloud radar) have shown that the cloud occurrence overlap can be fit to an inverse exponential function of the separation between layers, such that

$$\alpha_{j,k} = \exp\left(-\frac{|z_j - z_k|}{z_0}\right) \tag{C3}$$

where z_j and z_k are the heights of layers j and k, and z_0 is the *decorrelation length* for cloud overlap that specifies how quickly the vertical correlation in cloud occurrence decays from maximal to random (Barker, 2008; Hogan & Illingworth, 2000; Mace & Benson-Troth, 2002; Pincus et al., 2005; Räisänen et al., 2004; Tompkins & Di Giuseppe, 2015). Räisänen et al. (2004) and Pincus et al. (2005) further suggest that the same exponential relationship can describe the rank correlation of condensate but in general using a separate decorrelation length. These studies have suggested decorrelation lengths for cloud occurrence overlap between 1.5 and 2.5 km and somewhat smaller decorrelation lengths for condensate rank correlation. Overlap and decorrelation lengths will be parameterized in the following section for use with the SP-CAM output used in this study.

Appendix D: Parameterizing Overlap and Variability Statistics

Because the goal of this study is to evaluate the sensitivity of the simulated satellite diagnostics to different assumptions of subgrid-scale cloud and precipitation structure and variability, overlap parameters and condensate variability are parameterized using a month of output from the SP-CAM. This approach allows for direct comparison with diagnostics calculated from unmodified SP-CAM model outputs, thus allowing quantification of the sensitivities to these parameters. However, it is important to recognize that the SP-CAM fields from which the parameterizations are developed here are model fields, not observations. As such, the parameterization developed here is subject to limitations of the model used and may not be entirely consistent with observations of clouds in the physical atmosphere. Nonetheless, the analysis that follows offers a new perspective on overlap and variability, providing a global description of both overlap and condensate variability while previous efforts to quantify overlap have been limited in duration (Räisänen et al., 2004), limited to smaller domains for specific cases using CRMs (Pincus et al., 2005) or ground-based cloud radar (Hogan & Illingworth, 2000; Mace & Benson-Troth, 2002), or used satellite-based retrievals that come with their own limitations (Barker, 2008; Oreopoulos et al., 2012).

With the high-resolution model output provided by the SP-CAM, the occurrence overlap can be directly calculated for each gridbox from the subcolumn cloud condensate amount by solving equation (C1) for the overlap parameter $\alpha_{j,k}$ and assuming that the *true* combined cloud fraction between layers j and k can be described by generalized overlap, so that $\overline{c}_{j,k}^{\mathrm{true}} = \overline{c}_{j,k}^{\mathrm{gen}}$. This yields for the overlap parameter

$$\alpha_{j,k} = \frac{\overline{c}_{j,k}^{\mathrm{true}} - \overline{c}_{j,k}^{\mathrm{ran}}}{\overline{c}_{j,k}^{\mathrm{max}} - \overline{c}_{j,k}^{\mathrm{ran}}} \tag{D1}$$

For each gridbox and for each pair of layers j and k the overlap parameter $\alpha_{j,k}$ can be calculated by first calculating the true combined cloud fraction between the two layers $\overline{c}_{j,k}^{\mathrm{true}}$ and the theoretical maximally and randomly overlapped cloud fractions $\overline{c}_{j,k}^{\mathrm{max}}$ and $\overline{c}_{j,k}^{\mathrm{ran}}$, and then using these in equation (D1). Using this, overlap

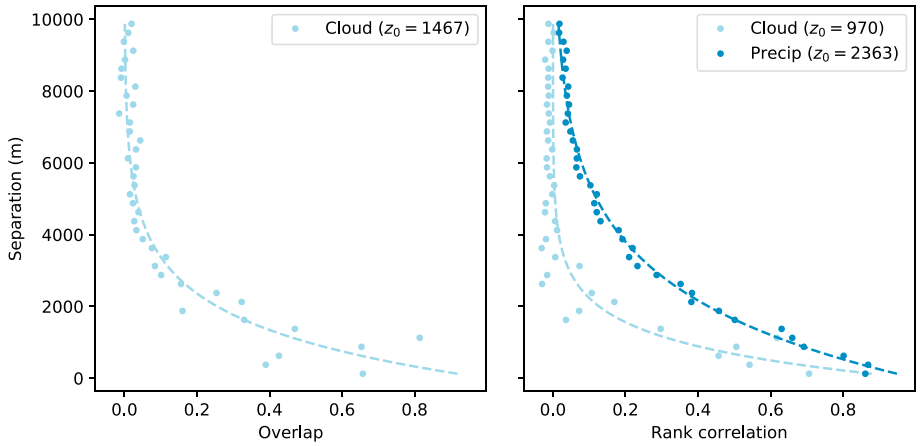

Figure D1. Global average cloud occurrence overlap parameter (left) and condensate rank correlation (right) as a function of separation distance between layers from a month of SP-CAM output. Also shown are fits to equation (C3). The decorrelation lengths from these fits are shown in the legends in each panel.

is calculated for pairs of layers in each model gridbox and at each archived 3-hourly snapshot of the SP-CAM output. The results are then binned by separation, and the monthly-averaged overlap as a function of separation (at each gridbox) is then calculated by summing the binned overlap and dividing by the number of valid counts in each bin. Rank correlation of total cloud and total precipitation condensate is similarly calculated at each gridbox and level for each 3-hourly snapshot and binned using the same separation bins used to calculate the average overlap.

Figure D1 shows the globally averaged overlap and condensate rank correlation for total (liquid plus ice) cloud and precipitation condensate as a function of separation distance (the area-weighted average of the overlap and rank correlation calculated using equation D1 at each latitude-longitude gridbox for each pair of layers). Overlap and rank correlation are fit to equation (C3) using nonlinear least squares, and the fits are plotted in Figure D1 along with the scatterplots of overlap and rank correlation. The overlap and rank correlation statistics shown in Figure D1 demonstrate the general tendency for both overlap and rank correlation to decrease as the separation between layers increases, and especially for distant layers the inverse exponential dependence on separation distance following equation C3 fits the data well. There is, however, generally larger spread in cloud overlap and rank correlation for small layer separations. This spread in overlap and rank correlation for small separations is not seen in previous analyses (e.g., Pincus et al., 2005), but those analyses have been primarily limited to much smaller domains, and consequently may consist of a smaller subset of cloud regimes than the global, month-long simulation considered here. R04 show overlap decorrelation lengths for a single day of SP-CAM output as a function of latitude and vertical (pressure) level, and it is evident from that analysis that overlap statistics vary substantially with both location and height, with decorrelation lengths varying from less than 0.5 km near the surface up to 10 km in the upper troposphere (see Figure 3 in R04). Some amount of spread in Figure D1 can be expected then, due to grouping all profiles at all levels together. Nonetheless, for simplicity we use the decorrelation lengths from the fits shown in Figure D1 for the sensitivity tests in this study.

The statistical distribution of condensate in the physical atmosphere has been studied using a variety of data sources, including aircraft observations (Larson et al., 2001; Wood & Field, 2000), tethered balloon observations (Price, 2001), satellite retrievals using passive sensors (Barker et al., 1996), and more recently using CloudSat retrievals (Lee et al., 2010). High-resolution model simulations from CRMs and large eddy simulations have also been used to study the statistical distribution of cloud condensate (Lewellen & Yoh, 1993; Xu & Randall, 1996a, 1996b). Cloud condensate distributions have been fit with a variety of different statistical distributions based on these studies. Lee et al. (2010) in particular fit retrievals of cloud liquid water content from CloudSat to a variety of distributions, including gamma, lognormal, exponential, Gaussian, Weibull, beta, and uniform distributions. They find that the CloudSat retrievals most closely follow either a lognormal or gamma distribution, depending on a number of conditions including geolocation, altitude, temperature, and the presence of precipitation. They also note that the data are reasonably well fit by the beta distribution, and

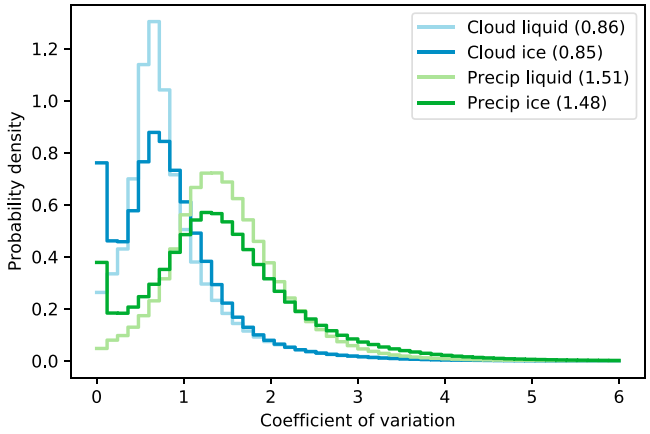

Figure D2. Histograms of the coefficient of variation of condensate mixing ratio from a month (simulated July 2000) of SP-CAM output, normalized to approximate the probability density. The mean of the coefficient of variation for each hydrometeor type is indicated in the legend.

Oreopoulos et al. (2012) subsequently use the beta distribution in their parameterization of subgrid variability for use in a GCM radiative transfer code.

These studies support the selection of the gamma distribution to represent subgrid-scale condensate distributions, and in fact, the cloud microphysics scheme in the most recent version of the NCAR Community Atmosphere Model (CAM5) assumes that subgrid-scale condensed cloud liquid follows a gamma distribution, although cloud ice and both precipitating liquid and ice are still treated as homogeneous (Neale, Gettelman, et al., 2010).

The gamma distribution has probability density

$$p_{k,\theta}(q) = \frac{1}{\Gamma(k)\theta^k} q^{k-1} e^{-q/\theta} \tag{D2}$$

where k and θ are the shape and scale parameters of the distribution and Γ is the gamma function. The distribution has mean $\mu = k\theta$ and variance $\sigma^2 = k\theta^2$. Using the method of moments (e.g., Wilks, 2011), the population mean and variance are equated with the sample mean $\overline{q}$ and variance σ_q^2, and this system of two equations is solved to estimate the shape and scale parameters $k = \mu^2/\sigma_q^2$ and $\theta = \sigma_q^2/\mu$. Taking q to be the condensate amount (mixing ratio) for a given hydrometeor type the distribution of condensate is completely specified in terms of the gridbox mean $\overline{q}$ and variance σ_q^2 using this formulation.

Cloud physics parameterizations in large-scale (global) models diagnose gridbox-mean cloud and precipitation condensate amounts but have not traditionally diagnosed (or even implicitly assumed) a nonzero gridbox variance. Previous authors have quantified condensate heterogeneity using the coefficient of variation $h_q = \sigma_q/\overline{q}$ (also referred to as the "fractional standard deviation" or "heterogeneity parameter"), where σ_q and $\overline{q}$ are the standard deviation and mean of condensate (Boutle et al., 2014; Hill et al., 2012; Hogan & Illingworth, 2003; Shonk et al., 2010). Shonk et al. (2010) conclude that a fixed coefficient of variation is sufficient to specify the heterogeneity, with a value of 0.75 ± 0.18. This would imply that gridbox standard deviation in cloud condensate can be calculated by simply scaling the mean, but other studies have suggested that the heterogeneity might vary with cloud fraction (Boutle et al., 2014; Hill et al., 2012; Oreopoulos et al., 2012), region (Lebsock et al., 2013), or regime (Ahlgrimm & Forbes, 2016; Hill et al., 2015). These studies suggest that the heterogeneity in the physical atmosphere may be better parameterized in terms of additional quantities, but the results of Shonk et al. (2010) suggest that a reasonable first attempt might be to parameterize the gridbox standard deviation with a linear dependence on the mean by assuming a constant coefficient of variation, and we argue that this is sufficient for the sensitivity tests presented here.

Figure D2 shows histograms (normalized to represent probability density) of the coefficient of variation h_q for each hydrometeor type from a month of SP-CAM output. The distributions of h_q for cloud condensate have mean values of 0.86 and 0.85 for cloud liquid and ice, respectively. These values are consistent with the results

obtained by Shonk et al. (2010) using a variety of observations, suggesting that the heterogeneity of clouds in SP-CAM is consistent with those in the physical atmosphere. Heterogeneity for precipitation condensate is somewhat larger and less sharply peaked than for cloud condensate, with mean values of 1.51 and 1.48 for precipitating liquid and ice, respectively. The spread in these distributions suggests again that an improved parameterization may be realized by representing heterogeneity by additional parameters, but for simplicity these constant values of heterogeneity are used in this study.

Acknowledgments
Funding for this research was provided by the MISR project at the NASA Jet Propulsion Laboratory (under contract 1318945) and in part by funding from the Center for Multiscale Modeling of Atmospheric Processes (a NSF Science and Technology Center) at Colorado State University under subaward ATM-0425247 to the University of Washington. The authors would like to thank David Diner, Principal Investigator of the MISR instrument, for his continual interest in and support of this research. The subcolumn generator and the offline simulator driver codes used in this manuscript are publicly available on GitHub (https://github.com/brhillman), and the MMF output used to derive cases for the sensitivity tests is available on Zenodo (https://zenodo.org/record/1246857).

References

Ahlgrimm, M., & Forbes, R. M. (2016). Regime dependence of cloud condensate variability observed at the atmospheric radiation measurement sites. *Quarterly Journal of the Royal Meteorological Society*, *142*(697), 1605–1617. https://doi.org/10.1002/qj.2783

Barker, H. W. (2008). Overlap of fractional cloud for radiation calculations in GCMs: A global analysis using CloudSat and CALIPSO data. *Journal of Geophysical Research*, *113*, D00A01. https://doi.org/10.1029/2007JD009677

Barker, H. W., Stephens, G. L., & Fu, Q. (1999). The sensitivity of domain-averaged solar fluxes to assumptions about cloud geometry. *Quarterly Journal of the Royal Meteorological Society*, *125*(558), 2127–2152. https://doi.org/10.1256/smsqj.55809

Barker, H. W., Wielicki, B. A., & Parker, L. (1996). A parameterization for computing grid-averaged solar fluxes for inhomogeneous marine boundary layer clouds. part II: Validation using satellite data. *Journal of the Atmospheric Sciences*, *53*(16), 2304–2316. https://doi.org/10.1175/1520-0469(1996)053<2304: APFCGA>2.0.CO;2

Bodas-Salcedo, A., Webb, M. J., Bony, S., Chepfer, H., Dufresne, J.-L., Klein, S. A., et al. (2011). COSP: satellite simulation software for model assessment. *Bulletin of the American Meteorological Society*, *92*(8), 1023–1043. https://doi.org/10.1175/2011BAMS2856.1

Boutle, I., Abel, S., Hill, P., & Morcrette, C. (2014). Spatial variability of liquid cloud and rain: Observations and microphysical effects. *Quarterly Journal of the Royal Meteorological Society*, *140*(679), 583–594.

Cesana, G., & Waliser, D. E. (2016). Characterizing and understanding systematic biases in the vertical structure of clouds in CMIP5/CFMIP2 models. *Geophysical Research Letters*, *43*, 10,538–10,546. https://doi.org/10.1002/ 2016GL070515

Cheng, A., & Xu, K.-M. (2011). Improved low-cloud simulation from a multiscale modeling framework with a third-order turbulence closure in its cloud-resolving model component. *Journal of Geophysical Research*, *116*, D14101. https://doi.org/10.1029/2010JD015362

Cheng, A., & Xu, K.-M. (2013). Evaluating low-cloud simulation from an upgraded multiscale modeling framework model. Part III: Tropical and subtropical cloud transitions over the northern Pacific. *Journal of Climate*, *26*(16), 5761–5781. https://doi.org/10.1175/JCLI-D-12-00650.1

Chepfer, H., Bony, S., Winker, D., Chiriaco, M., Dufresne, J.-L., & Sèze, G. (2008). Use of CALIPSO lidar observations to evaluate the cloudiness simulated by a climate model. *Geophysical Research Letters*, *35*, L15704. https://doi.org/10.1029/2008GL034207

Di Michele, S, Ahlgrimm, M., Forbes, R., Kulie, M., Janisková, R. B. M., & Bauer, P. (2012). Interpreting an evaluation of the ECMWF global model with CloudSat observations: Ambiguities due to radar reflectivity forward operator uncertainties. *Quarterly Journal of the Royal Meteorological Society*, *138*, 2047–2065.

Geleyn, J.-F., & Hollingsworth, A. (1979). An economical analytical method for the computation of the interaction between scattering and line absorption of radiation. *Beitraege zur Physik der Atmosphaere*, *52*, 1–16.

Gleckler, P. J., Taylor, K. E., & Doutriaux, C. (2008). Performance metrics for climate models. *Journal of Geophysical Research*, *113*, D06104. https://doi.org/10.1029/2007JD008972

Golaz, J.-C., Larson, V. E., & Cotton, W. R. (2002). A PDF-based model for boundary layer clouds. part I: Method and model description. *Journal of the Atmospheric Sciences*, *59*(24), 3540–3551. https://doi.org/10.1175/ 1520-0469(2002)059<3540:APBMFB>2.0.CO;2

Guo, Z., Wang, M., Qian, Y., Larson, V. E., Ghan, S., Ovchinnikov, M., et al. (2014). A sensitivity analysis of cloud properties to CLUBB parameters in the Single-column Community Atmosphere Model (SCAM5). *Journal of Advances in Modeling Earth Systems*, *6*(3), 829–858. https://doi.org/10.1002/2014MS000315

Haynes, J. M., Marchand, R. T., Luo, Z., Bodas-Salcedo, A., & Stephens, G. L. (2007). A multipurpose radar simulation package: Quickbeam. *Bulletin of the American Meteorological Society*, *88*(11), 1723–1727. https://doi.org/10.1175/BAMS-88-11-1723

Hill, P. G., Hogan, R. J., Manners, J., & Petch, J. C. (2012). Parametrizing the horizontal inhomogeneity of ice water content using CloudSat data products. *Quarterly Journal of the Royal Meteorological Society*, *138*(668), 1784–1793. https://doi.org/10.1002/qj.1893

Hill, P., Morcrette, C., & Boutle, I. (2015). A regime-dependent parametrization of subgrid-scale cloud water content variability. *Quarterly Journal of the Royal Meteorological Society*, *141*(691), 1975–1986.

Hillman, B. R., Marchand, R. T., Ackerman, T. P., Mace, G. G., & Benson, S. (2017). Assessing the accuracy of MISR and MISR-simulated cloud top heights using CloudSat- and CALIPSO-retrieved hydrometeor profiles. *Journal of Geophysical Research: Atmospheres*, *122*, 2878–2897. https://doi.org/10.1002/2016JD025510

Hogan, R. J., & Illingworth, A. J. (2000). Deriving cloud overlap statistics from radar. *Quarterly Journal of the Royal Meteorological Society*, *126*, 2903–2909. https://doi.org/10.1256/smsqj.56913

Hogan, R. J., & Illingworth, A. J. (2003). Parameterizing ice cloud inhomogeneity and the overlap of inhomogeneities using cloud radar data. *Journal of the Atmospheric Sciences*, *60*(5), 756–767.

Kay, J. E., Hillman, B. R., Klein, S. A., Zhang, Y., Medeiros, B., Pincus, R., 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*, 5190–5207. https://doi.org/10.1175/ JCLI-D-11-00469.1

Khairoutdinov, M. F., & Randall, D. A. (2001). A cloud-resolving model as a cloud parameterization in the NCAR community climate system model: Preliminary results. *Geophysical Research Letters*, *28*(18), 3617–3620.

Khairoutdinov, M., Randall, D., & DeMott, C. (2005). Simulations of the atmospheric general circulation using a cloud-resolving model as a superparameterization of physical processes. *Journal of the Atmospheric Sciences*, *62*(7), 2136–2154. https://doi.org/10.1175/JAS3453.1

Klein, S. A., & Jakob, C. (1999). Validation and sensitivities of frontal clouds simulated by the ECMWF model. *Monthly Weather Review*, *127*(10), 2514–2531. https://doi.org/10.1175/1520-0493(1999)127<2514: VASOFC>2.0.CO;2

Klein, S. A., Zhang, Y., Zelinka, M. D., Pincus, R., Boyle, J., & Gleckler, P. J. (2013). Are climate model simulations of clouds improving? An evaluation using the ISCCP simulator. *Journal of Geophysical Research: Atmospheres*, *118*, 1329–1342. https://doi.org/10.1002/jgrd.50141

Larson, V. E., Wood, R., Field, P. R., Golaz, J.-C., Haar, T. H. V., & Cotton, W. R. (2001). Small-scale and mesoscale variability of scalars in cloudy boundary layers: One-dimensional probability density functions. *Journal of the Atmospheric Sciences*, *58*(14), 1978–1994.

Lebsock, M., Morrison, H., & Gettelman, A. (2013). Microphysical implications of cloud-precipitation covariance derived from satellite remote sensing. *Journal of Geophysical Research: Atmospheres*, *118*, 6521–6533. https://doi.org/10.1002/jgrd.50347

Lee, S., Kahn, B. H., & Teixeira, J. A. (2010). Characterization of cloud liquid water content distributions from CloudSat. *Journal of Geophysical Research*, *115*, D20203. https://doi.org/10.1029/2009JD013272

Lewellen, W. S., & Yoh, S. (1993). Binormal model of ensemble partial cloudiness. *Journal of the Atmospheric Sciences*, *50*(9), 1228–1237.

Lin, W. Y., & Zhang, M. H. (2004). Evaluation of clouds and their radiative effects simulated by the NCAR Community Atmospheric Model against satellite observations. *Journal of Climate*, *17*(17), 3302–3318. https://doi.org/10.1175/1520-0442(2004)017<3302:EOCATR>2.0.CO;2

Mace, G. G., & Benson-Troth, S. (2002). Cloud-layer overlap characteristics derived from long-term cloud radar data. *Journal of Climate*, *15*, 2505–2515. https://doi.org/10.1175/1520-0442(2002)015<2505:CLOCDF>2. 0.CO;2

Marchand, R., & Ackerman, T. (2010). An analysis of cloud cover in multiscale modeling framework global climate model simulations using 4 and 1 km horizontal grids. *Journal of Geophysical Research*, *115*, D16207. https://doi.org/10.1029/2009JD013423

Marchand, R., Ackerman, T., Smyth, M., & Rossow, W. B. (2010). A review of cloud top height and optical depth histograms from MISR, ISCCP and MODIS. *Journal of Geophysical Research*, *115*, D16206. https://doi.org/10.1029/2009JD013422

Marchand, R., Haynes, J., Mace, G. G., & Ackerman, T. (2009). A comparison of simulated radar output from the multiscale modeling framework global climate model with CloudSat cloud radar observations. *Journal of Geophysical Research*, *114*, D00A20. https://doi.org/10.1029/2008JD009790

Morcrette, J.-J., & Fouquart, Y. (1986). The overlapping of cloud layers in shortwave radiation parameterizations. *Journal of the Atmospheric Sciences*, *43*(4), 321–328. https://doi.org/10.1175/1520-0469(1986) 043<0321:TOOCLI>2.0.CO;2

Neale, R. B., Gettelman, A., Park, S., Chen, C.-C., Lauritzen, P. H., Williamson, D. L., et al. (2010). Description of the NCAR Community Atmosphere Model (CAM 5.0) (NCAR Technical Note TN-486+STR). Boulder, CO: NCAR.

Neale, R. B., Richter, J. H., Conley, A. J., Park, S., Lauritzen, P. H., Gettelman, A., et al. (2010). Description of the NCAR Community Atmosphere Model (CAM 4.0) (*NCAR Technical Note TN-485+STR*). Boulder, Colorado, USA: NCAR.

Oreopoulos, L., Lee, D., Sud, Y. C., & Suarez, M. J. (2012). Radiative impacts of cloud heterogeneity and overlap in an atmospheric general circulation model. *Atmospheric Chemistry and Physics*, *12*, 9097–9111. https://doi.org/10.5194/acp-12-9097-2012

Pincus, R., Hannay, C., Klein, S. A., Xu, K.-M., & Hemler, R. (2005). Overlap assumptions for assumed probability distribution function cloud schemes in large-scale models. *Journal of Geophysical Research*, *110*, D15S09. https://doi.org/10.1029/2004JD005100

Pincus, R., Platnick, S., Ackerman, S. A., Hemler, R. S., & Hofmann, R. J. P. (2012). Reconciling simulated and observed views of clouds: MODIS, ISCCP, and and the limits of instrument simulators. *Journal of Climate*, *25*, 4699–4720. https://doi.org/10.1175/JCLI-D-11-00267.1

Price, J. D. (2001). A study of probability distributions of boundary-layer humidity and associated errors in parametrized cloud-fraction. *Quarterly Journal of the Royal Meteorological Society*, *127*(573), 739–758.

Räisänen, P., Barker, H. W., Khairoutdinov, M. F., Li, J., & Randall, D. A. (2004). Stochastic generation of subgrid-scale cloudy columns for large-scale models. *Quarterly Journal of the Royal Meteorological Society*, *130*, 2047–2067. https://doi.org/10.1256/qj.03.99

Randall, D., Khairoutdinov, M., Arakawa, A., & Grabowski, W. (2003). Breaking the cloud parameterization deadlock. *Bulletin of the American Meteorological Society*, *84*(11), 1547–1564. https://doi.org/10.1175/ BAMS-84-11-1547

Shonk, J. K. P., Hogan, R. J., Edwards, J. M., & Mace, G. G. (2010). Effect of improving representation of horizontal and vertical cloud structure on the Earth's global radiation budget. Part I: Review and parametrization. *Quarterly Journal of the Royal Meteorological Society*, *136*(650), 1191–1204.

Stephens, G. L., & Platt, C. M. R. (1987). Aircraft observations of the radiative and microphysical properties of stratocumulus and cumulus cloud fields. *Journal of Climate and Applied Meteorology*, *26*(9), 1243–1269. https://doi.org/10.1175/1520-0450(1987)026<1243:AOOTRA>2.0.CO;2

Stubenrauch, C. J., Genio, A. D. D., & Rossow, W. B. (1997). Implementation of subgrid cloud vertical structure inside a GCM and its effect on the radiation budget. *Journal of Climate*, *10*, 273–287.

Tanelli, S., Durden, S. L., Im, E., Pak, K. S., Reinke, D. G., Partain, P., et al. (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. https://doi.org/10.1109/TGRS.2008.2002030

Tao, W.-K., Chern, J.-D., Atlas, R., Randall, D., Khairoutdinov, M., Li, J.-L., et al. (2009). A multiscale modeling system: Developments, applications, and critical issues. *Bulletin of the American Meteorological Society*, *90*(4), 515.

Thayer-Calder, K., Gettelman, A., Craig, C., Goldhaber, S., Bogenschutz, P., Chen, C.-C., et al. (2015). A unified parameterization of clouds and turbulence using CLUBB and subcolumns in the community atmosphere model. *Geoscientific Model Development*, *8*(12), 3801.

Tian, L., & Curry, J. A. (1989). Cloud overlap statistics. *Journal of Geophysical Research*, *94*(D7), 9925–9935.

Tompkins, A. M. (2002). A prognostic parameterization for the subgrid-scale variability of water vapor and clouds in large-scale models and its use to diagnose cloud cover. *Journal of the Atmospheric Sciences*, *59*, 1917–1942. https://doi.org/10.1175/1520-0469(2002)059<1917:APPFTS>2.0.CO;2

Tompkins, A. M., & Di Giuseppe, F. (2015). An interpretation of cloud overlap statistics. *Journal of the Atmospheric Sciences*, *72*(8), 2877–2889.

von Salzen, K., Scinocca, J. F., McFarlane, N. A., Li, J., Cole, J. N. S., Plummer, D., et al. (2013). The Canadian Fourth Generation Atmospheric Global Climate Model (CanAM4): Part I: Representation of physical processes. *Atmosphere–Ocean*, *51*, 104–123. https://doi.org/10.1080/07055900.2012.755610

Webb, M. J., Andrews, T., Bodas-Salcedo, A., Bony, S., Bretherton, C. S., Chadwick, R., et al. (2017). The Cloud Feedback Model Intercomparison Project (CFMIP) contribution to CMIP6. *Geoscientific Model Development*, *10*(1), 359–384. https://doi.org/10.5194/gmd-10-359-2017

Webb, M., Senior, C., Bony, S., & Morcrette, J.-J. (2001). Combining ERBE and ISCCP data to assess clouds in the Hadley Centre, ECMWF and LMD atmospheric climate models. *Climate Dynamics*, *17*(12), 905–922. https://doi.org/10.1007/s003820100157

Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences* (Vol. 100). Academic press.

Wood, R., & Field, P. R. (2000). Relationships between total water, condensed water, and cloud fraction in stratiform clouds examined using aircraft data. *Journal of the Atmospheric Sciences*, *57*(12), 1888–1905.

Wu, X., & Liang, X.-Z. (2005). Radiative effects of cloud horizontal inhomogeneity and vertical overlap identified from a monthlong cloud-resolving model simulation. *Journal of the Atmospheric Sciences*, *62*, 4105–4112.

Wyant, M. C., Bretherton, C. S., Bacmeister, J. T., Kiehl, J. T., Held, I. M., Zhao, M., et al. (2006). A comparison of low-latitude cloud properties and their response to climate change in three AGCMs sorted into regimes using mid-tropospheric vertical velocity. *Climate Dynamics*, *27*(2-3), 261–279.

Xu, K.-M., & Randall, D. A. (1996a). A semiempirical cloudiness parameterization for use in climate models. *Journal of the Atmospheric Sciences*, *53*(21), 3084–3102.

Xu, K.-M., & Randall, D. A. (1996b). Explicit simulation of cumulus ensembles with the gate phase III data: Comparison with observations. *Journal of the Atmospheric Sciences*, *53*(24), 3710–3736.

Zhang, Y., Klein, S. A., Boyle, J., & Mace, G. G. (2010). Evaluation of tropical cloud and precipitation statistics of Community Atmosphere Model version 3 using CloudSat and CALIPSO data. *Journal of Geophysical Research*, *115*, D12205. https://doi.org/10.1029/2009JD012006

Zhang, M., Lin, W. Y., Klein, S. A., Bacmeister, J. T., Bony, S., Cederwall, R. T., et al. (2005). Comparing clouds and their seasonal variations in 10 atmospheric general circulation models with satellite measurements. *Journal of Geophysical Research*, *110*, D15S02. https://doi.org/10.1029/2004JD005021