Adapting the COSP Radar Simulator to Compare GCM Output and GPM Precipitation Observations

EMILY M. RILEY DELLARIPA, a AARON FUNK, b COURTNEY SCHUMACHER, b HEDANQIU BAI, b and THOMAS SPANGEHL c

a Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado
b Department of Atmospheric Sciences, Texas A&M University, College Station, Texas
c German Weather Service, Offenbach, Germany

ABSTRACT: Comparisons of precipitation between general circulation models (GCMs) and observations are often confounded by a mismatch between model output and instrument measurements, including variable type and temporal and spatial resolutions. To mitigate these differences, the radar-simulator Quickbeam within the Cloud Feedback Model Intercomparison Project (CFMIP) Observation Simulator Package (COSP) simulates reflectivity from model variables at the subgrid scale. This work adapts Quickbeam to the dual-frequency precipitation radar (DPR) on board the Global Precipitation Measurement (GPM) satellite. The longer wavelength of the DPR is used to evaluate moderate to heavy precipitation in GCMs, which is missed when Quickbeam is used as a cloud radar simulator. Latitudinal and land–ocean comparisons are made between COSP output from the Community Atmosphere Model version 5 (CAM5) and DPR data. Additionally, this work improves the COSP subgrid algorithm by applying a more realistic, nondeterministic approach to assigning GCM gridbox convective cloud cover when convective cloud is not provided as a model output. Instead of assuming a static 5% convective cloud coverage, DPR convective precipitation coverage is used as a proxy for convective cloud coverage. For example, DPR observations show that convective rain typically only covers about 1% of a 2° grid box, but that the median convective rain area increases to over 10% in heavy rain cases. In our CAM5 tests, the updated subgrid algorithm improved the comparison between reflectivity distributions when the convective cloud cover is provided versus the default 5% convective cloud-cover assumption.

KEYWORDS: Precipitation; Algorithms; Radars/Radar observations; Diagnostics; General circulation models; Model evaluation/performance

1. Introduction

Comparisons between general circulation models (GCMs) and observations are critical to assess the fidelity of a GCM and pinpoint GCM deficiencies, such as precipitation biases (Dai 2006; Sun et al. 2006; Stephens et al. 2010; Tan et al. 2018). Many satellite–GCM comparisons are done using surface rainfall, which is a highly derived quantity from both satellite measurements and GCM equations. In addition, focusing only on surface rain severely limits model evaluation since it is the processes above the surface and throughout the depth of the troposphere that are important for understanding the microphysics and dynamics of precipitating cloud systems and their impact on the larger-scale circulation (Hartmann et al. 1984; Schumacher et al. 2004). A potentially more useful exercise is to examine the vertical structure of model precipitation compared to observations to assess whether the dynamics and microphysics of precipitation production are being reasonably captured by the physics of the GCM.

Comparison of GCM output and satellite radar observations is full of potential owing to the three-dimensionality (3D) of radar data, but also challenging because of the incompatibility of observed radar signals and model output. Radars measure the returned power of a transmitted radio pulse scattered by a volume of hydrometeors (i.e., radar reflectivity in units of dBZ hereafter referred to as reflectivity) resulting in a measurement with high vertical and horizontal spatial resolution, whereas GCMs simulate hydrometeor characteristics such as number concentration, size distribution, and mixing ratio over an entire grid at fixed model heights. These GCM versus radar observation differences for precipitation can be reconciled by either 1) converting the observed reflectivity to model hydrometeor variables such as mixing ratios or rain rate using an inverse model or 2) deriving reflectivity from the model’s hydrometeor information using a forward operator (e.g., Masunaga et al. 2010; Bodas-Salcedo et al. 2011; Di Michele et al. 2012). Converting model output to reflectivity is preferred, since the derivation of precipitation from reflectivity relies on numerous assumptions to constrain the inverse model due to uncertainties in the radar retrieval such as particle size distribution (PSD). The assumptions in the inverse model can lead to large errors and multiple solutions of rain or snow rate for a given reflectivity value (e.g., Stout and Mueller 1968; Battan 1973). Additionally, transferring reflectivity to microphysical measurements such as ice or liquid water content often relies on a priori information or the use of multiple satellite retrievals (e.g., Waliser et al. 2009; Delanoë and Hogan 2010). Conversely, deriving reflectivity from model hydrometeor information is only limited by the accuracy of the forward operator; known biases in the observation process are accounted for as best as possible and assumptions such as PSD can be chosen to be

DOI: 10.1175/JTECH-D-20-0089.1

© 2021 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).
consistent with the GCM’s PSD assumption (Masunaga et al. 2010; Bodas-Salcedo et al. 2011; Carlin et al. 2016).

The Cloud Feedback Model Intercomparison Project (CFMIP) Observation Simulator Package (COSP) contains several forward operators that simulate both passive and active satellites, including the radar simulator Quickbeam (Haynes et al. 2007; Bodas-Salcedo et al. 2011). To date, Quickbeam has only been run as a cloud radar simulator despite its ability to be configured to any radar frequency. Quickbeam, by default, simulates the space-based CloudSat cloud profiling radar (94 GHz; Stephens et al. 2002; Haynes et al. 2007; Bodas-Salcedo et al. 2011), but has been adapted to simulate the Atmospheric Radiation Measurement (ARM) program’s ground-based cloud radars (35 GHz; Zhao et al. 2017; Zhang et al. 2018). Utilizing Quickbeam as a cloud radar simulator in conjunction with other COSP simulators provides a fairly complete 3D evaluation of a model’s ability to simulate clouds (e.g., Zhang et al. 2010; Cesana and Chepfer 2012; Kay et al. 2012; Kodama et al. 2012; Franklin et al. 2013a,b; Nam et al. 2014; Jin et al. 2017a,b; Webb et al. 2017). However, attenuation of reflectivity is significant in heavy rainfall at cloud radar frequencies (e.g., Stephens et al. 2002; Matrosov 2007; Haynes et al. 2009), which means CloudSat observations can only be used to evaluate light to moderate precipitation in models (Bodas-Salcedo et al. 2008; Nam and Quaas 2012).

Therefore, the first goal of this work is to adapt the Quickbeam radar simulator inside COSP to the spaceborne Global Precipitation Measurement (GPM) mission dual-frequency precipitation radar (DPR; Hou et al. 2014) to be able to evaluate the full range of simulated rain rates in GCMs. The DPR Ku band is much less prone to attenuation than CloudSat and its sensitivity starts near the upper limit of a cloud radar’s sensitivity (i.e., 12–14 dBZ; Toyoshima et al. 2015) and extends up to about 60 dBZ. Therefore, simulating the GPM precipitation radar allows for a more complete evaluation of model precipitation processes than cloud radar simulators. After optimizing Quickbeam and COSP for this application, we applied the simulator to output from the Community Atmospheric Model version 5 (CAM5; Neale et al. 2012) and evaluated how well CAM5 captured latitudinal variations and land–ocean differences in the vertical distribution of precipitation relative to DPR observations, although many other applications are possible for model assessment.

In addition to the differences in radar observations and GCM output variables that make GCM–observation comparisons difficult, radar observations and GCM comparisons are further confounded by a mismatch in their horizontal resolutions. Spaceborne precipitation radars have a horizontal resolution (i.e., footprint) on the order of 5 km, while typical GCM grid spacing is on the order of 50–200 km. COSP mitigates these differences by distributing GCM grid cloud cover and precipitation across the GCM subgrid scale before calculating reflectivity in Quickbeam (Haynes et al. 2007; Zhang et al. 2010; Bodas-Salcedo et al. 2011). Kay et al. (2018) discusses the importance of “downscaling” the GCM gridbox mean cloud cover and precipitation to subcolumns of the GCM grid box in order to make fair model–observation comparisons. For each GCM grid box, the COSP subgrid algorithm equally and separately distributes convective and stratiform precipitation at each model level among a user-defined number of subcolumns to be consistent with the GCM’s cloud-cover amount and overlap assumption (Zhang et al. 2010). For instances where model precipitation occurs in a grid box but no model cloud-cover information is provided, large-scale precipitation is distributed equally over 100% of the subcolumns, whereas convective precipitation is distributed equally over 5% of the subcolumns (Zhang et al. 2010). Further details on the COSP subgrid algorithm are given in section 3a.

Some GCMs do not provide convective cloud cover despite providing convective precipitation (e.g., ECHAM; Rast et al. 2012) such that convective precipitation is always divided among 5% of the COSP subcolumns. Assuming a static 5% convective precipitation coverage (or any static percent coverage) for all convective rain rates is incompatible with observations that show convective rain area varies with rain rate (e.g., Yuter and Houze 1998; section 5a and Fig. 9 of this study). Thus, the second goal of this work is to improve the static 5% convective coverage assumption in the COSP subgrid algorithm with a more realistic nondeterministic algorithm based on DPR observations. Our new algorithm was tested using output from the CAM5.

The following section briefly describes the satellite radar observations and the GCM used in this study. Section 3 reviews how Quickbeam operates within COSP and discusses using it as a precipitation radar simulator. Section 4 discusses comparisons between CAM5 reflectivity generated by the precipitation radar simulator and DPR observations. Section 5 details the observationally based improvements for convective cloud-cover assumptions implemented in the COSP subgrid algorithm for this study. Section 6 summarizes this work and offers suggestions to further improve COSP for precipitation radar/GCM comparisons.

2. Data and model description

a. GPM DPR

The GPM satellite (Hou et al. 2014; Skofronick-Jackson et al. 2017) was launched February 2014 as a successor to and extension of the Tropical Rainfall Measuring Mission (TRMM) satellite (Kummerow et al. 1998) to observe precipitation and snow across the tropics and midlatitudes. The GPM satellite flies at an altitude of 407 km and makes observations from 65°S to 65°N. The DPR on board the GPM satellite consists of two phased-array radars operating at 13.6 GHz (Ku band) and 35.5 GHz (Ka band). The TRMM Precipitation Radar (PR) operated only at Ku band with an orbit covering 35°S–35°N from late 1997 to mid-2015.

Previous work configured Quickbeam in COSP for cloud radar frequencies [i.e., 94 GHz (e.g., Haynes et al. 2007; Bodas-Salcedo et al. 2011) and 35 GHz (Zhao et al. 2017; Zhang et al. 2018)]; however, we implemented the Ku-band wavelength into COSP because it is the main precipitation frequency on the TRMM and GPM radars and it is much less affected by attenuation than Ka band. Similar to the TRMM PR, the DPR Ku band has a 5-km footprint at nadir with 250-m vertical
resolution and a 245-km cross-track width (i.e., swath width). However, the GPM DPR is more sensitive than the TRMM PR owing to a higher transmitted peak power (Hou et al. 2014) and makes observations into higher latitudes. Therefore, we focus on the use of the DPR observations in the rest of this work.

The analysis in this study was done with June 2017 of the GPM level-2A DPR (2ADPR) version 6 Ku-band orbital data (Iguchi and Meneghini 2017). June 2017 was chosen to represent generally neutral El Niño–Southern Oscillation (ENSO) conditions. The addition of more months and years did not impact the conclusions we draw from the results in any noticeable way, although a longer period of GPM observations would likely be necessary if model comparisons were done over smaller regions because of the impact of the relatively narrow DPR swath width on sampling. DPR observations used in this study were limited to the inner swath (i.e., 25 inner rays) where dual-frequency correction algorithms are applied to coincident Ku- and Ka-band rays. The dual-frequency algorithms can provide more accurate estimates of instantaneous rain rates for a radar volume, which is important for this work since near-surface DPR rain-rate estimates are used to develop the improved subgrid algorithm.

Daily 2ADPR near-surface rain-rate estimates from the inner swath were used to generate mean rain rates as well as estimates of the precipitation frequency of occurrence (PFO) for convective and stratiform rain types on fixed grid sizes analogous to the resolution of GCM output (i.e., from 0.25° to 2°). The height of the near-surface rain rates typically ranges between 1 and 2 km. The PFO is the ratio of raining samples to the total (i.e., raining and nonraining) samples over a grid. Most grid boxes will contain reflectivity profiles from only a single GPM overpass per day since the average revisit time of the DPR’s inner swath is 1.79 days. Temporal averaging occurs at the highest latitudes when more than one overpass can exist in a grid box for a given day. Therefore, most of the DPR reflectivity values used to construct the reflectivity distributions are only spatially averaged and not temporally averaged. We only used DPR overpasses that had at least 171 total samples in a 2.0° grid. This threshold is equivalent to the area covered by a 70-km grid box, which ensures sufficient sampling at GCM grid scales. Using a much smaller or no threshold allows unphysical results in the analysis, while a much larger threshold has little impact on the results but lessens the amount of data available.

b. CAM5

CAM5 reflectivity generated by Quickbeam within COSP is compared to GPM observations. At the time our CAM5 simulations were run, CAM5 was the latest version of the global atmospheric model from the National Center for Atmospheric Research (NCAR). It uses the Zhang and McFarlane (1995) convective parameterization for deep convection with a modified dilute plume calculation following Raymond and Blyth (1986, 1992) and convective momentum transport by Richter and Rasch (2008). There is a separate shallow convection scheme based on Park and Bretherton (2009). Large-scale rain and snow are diagnosed in CAM5 using a two-moment bulk microphysics scheme detailed in Morrison and Gettelman (2008).

The primary analysis in this work is based on a simulation of CAM5 run at a grid spacing of 1.9° × 2.5° for June 2017. However, an additional simulation using 0.9° × 1.25° grid spacing was done to evaluate the sensitivity of our results to model resolution. These two simulations will be referred to as the 2°- and 1°-resolution runs throughout the text. The model time step was 30 min with output saved hourly. The default version of COSP 1.3 was run inline with the simulation in order to output COSP-appropriate variables that were used as input to COSP 2.0 (COSP2 hereafter), which is the latest version of COSP (Swales et al. 2018). COSP2 was run offline because it had not yet been integrated into CAM5 when we ran our 2017 simulation. COSP-appropriate variables include large-scale and convective precipitation as fluxes, where convective precipitation is the combination of precipitation from the deep and shallow convective schemes. Saving hourly output preserved consistency between the temporal resolution of variables input into our offline version of COSP2 and the frequency that COSP 1.3 was called when run inline with CAM5, which was hourly at every radiation time step. Hourly data are also a compromise between computational costs and a desire to have time-averaged model output comparable to the instantaneous GPM radar observations. We also note that since the model time step is 30 min, hourly averaged output is not very different than instantaneous (i.e., every time step) output.

3. GPM precipitation radar simulator in COSP

a. Background on how COSP with Quickbeam works

In the next three sections, we review the workflow of COSP with the radar simulator Quickbeam (Fig. 1) and outline necessary modifications to run COSP with CAM5 and the wavelength of the GPM DPR. More details on how COSP works can be found in Bodas-Salcedo et al. (2011), Zhang et al. (2010), and the COSP user’s manual (Bodas-Salcedo 2010). First, gridbox mean vertical profiles of GCM variables are input into COSP (first step in Fig. 1). The input variables required to run COSP with the GPM precipitation radar simulator are the same as the original COSP setup with the CloudSat cloud radar simulator, namely, profiles of temperature, pressure, relative humidity, total cloud cover, mixing ratios for convective and large-scale cloud liquid and ice, and convective and large-scale precipitation fluxes. Instead of precipitation fluxes, convective and large-scale precipitation mixing ratios can be used as input to COSP, but are not used in this study. Optional inputs are hydrometeor effective radii and number concentrations, both of which are used in this study to mitigate microphysical assumptions (discussed in section 3b). Convective cloud cover can also be input when available from the GCM. Convective cloud cover is available from CAM5 and was used as input with the default version of COSP2, but was excluded in the sensitivity runs discussed in section 4.

To mitigate the spatial-scale differences between the GCM grid box and radar observations, COSP distributes GCM gridbox mean convective and total cloud cover across a user-defined number of gridbox subcolumns (second step in Fig. 1) using the COSP Subgrid Cloud Overlap Profile Sampler (SCOPS)
algorithm (Klein and Jakob 1999; Webb et al. 2001). The cloud distribution is consistent with the GCM’s radiative cloud overlap assumption (i.e., maximum, random, or maximum–random overlap). Using the subcolumn distribution of total and convective cloud cover from SCOPS, the PREC_SCOPS algorithm developed by Zhang et al. (2010) then flags cloudy subcolumns with large-scale and/or convective precipitation (third step in Fig. 1). Sometimes precipitation occurs in grid boxes with no cloud cover. In such instances, large-scale precipitation is assigned to 100% of the subcolumns, while convective precipitation is assigned to 5% of the subcolumns. This unphysical behavior occurs on the order of 15% of the time for each rain type, but the rain rates are generally very weak so do not contribute strongly to the resulting reflectivity.

Once the subcolumns have been flagged for large-scale and/or convective precipitation, the precipitation flux is divided equally among the precipitation-flagged subcolumns to give the GCM gridbox mean precipitation value. For example, if the GCM gridbox mean convective rain rate is 1 mm h$^{-1}$ and 2 out of 10 subcolumns contain convective precipitation, then each of the two precipitating subcolumns would be assigned a convective rain rate of 5 mm h$^{-1}$.

b. Modifications to COSP and Quickbeam for consistency with CAM5 microphysics

The precipitation fluxes are then converted to subgrid profiles of mixing ratio (fourth step in Fig. 1) following Khairoutdinov and Randall (2003), who use an exponential (i.e., Marshall and Palmer 1948) PSD for precipitating hydrometeors, which is the same PSD CAM5 uses for large-scale rain and snow (Neale et al. 2012; Morrison and Gettelman 2008). In Khairoutdinov and Randall (2003) a static intercept parameter is used in the exponential PSD. However, in CAM5, the intercept parameter varies as a function of number concentration and slope parameter, where the latter can be calculated using hydrometeor effective radii (Morrison and Gettelman 2008). Therefore, to be consistent with CAM5, we varied the intercept parameter by inputting rain and snow effective radius and number concentrations into COSP. Since CAM5 does not provide convective rain and snow effective radius and number concentration, we used the large-scale values as the convective values. This follows the in-line version of COSP1.3 in CAM5, which sets the convective rain and snow effective radius values to the large-scale values. Leaving the convective values as zero and using a static intercept parameter of 8 $\times$ 10$^6$ m$^{-4}$ for rain and 3 $\times$ 10$^6$ m$^{-4}$ for snow, which are the values from Khairoutdinov and Randall (2003), only slightly altered the reflectivity distribution, but did not impact the main conclusions of the paper (discussed further below).

The subgrid profiles of each hydrometer’s mixing ratios are then used as inputs into Quickbeam (fifth step in Fig. 1) along with the GCM gridbox mean vertical profiles of temperature, pressure, and relative humidity. Within Quickbeam, the phase, mass–diameter relationship [i.e., $m(D) = \alpha D^b$], and PSD of each hydrometer type is specified to match the parent GCM (Haynes et al. 2007). Both large-scale and convective rain and cloud liquid were assigned the liquid phase, while large-scale and convective snow and cloud ice were assigned the ice phase. In the mass–diameter relationship, a constant density of 1000 g kg$^{-1}$ was used for rain and cloud liquid, 500 g kg$^{-1}$ for cloud ice, and 100 g kg$^{-1}$ for snow, which are the same values used in CAM5. The density values were applied to both the large-scale and convective hydrometers. Using the constant density, $\alpha$ was solved with $\alpha = (\pi b/6)\rho$ and $\beta$ was equal to 3, since the hydrometers are assumed to be spherical, which is the same assumption that is used in CAM5. When solving for reflectivity in Quickbeam, the ice-phase hydrometers are treated as pure ice, so the discrepancy between the density of a pure ice sphere and the density of the prescribed ice hydrometeors (i.e., 500 and 100 g kg$^{-1}$ for CAM5 ice and snow) is accounted for by solving the equivalent-volume sphere problem. This means the diameters of the pure ice spheres are solved for such that the mass of the pure ice spheres is equal to the mass of the prescribed hydrometers. Using the mass–diameter relationship above this works out to $D_{eq} = D(\rho/\rho_{ice})^{1/b}$, where $D_{eq}$ is the equivalent diameter of the pure ice sphere, $D$ is the diameter of the spheres with the prescribed density, $\rho$ is the prescribed density, and $\rho_{ice}$ is the density of pure ice (i.e., 917 g kg$^{-1}$).

The PSDs followed CAM5 and were exponential for large-scale rain, snow, and ice and gamma for large-scale liquid. The convective hydrometers were assigned the same PSDs as their large-scale hydrometer counterparts. To follow CAM5’s double-moment microphysics, the intercept, slope, and shape parameters in the PSDs were solved for using effective radius and number concentration (Neale et al. 2012; Morrison and Gettelman 2008). CAM5 provides the effective radii for large-scale rain, snow, liquid, and ice and convective liquid and ice.

As stated above, since convective rain and snow effective radius are not provided, they are set to the respective large-scale values following the in-line version of COSP1.3 in CAM5. Similarly, CAM5 only provides number concentration for the large-scale hydrometers. We therefore set the convective hydrometer number concentration values to the large-scale values as was done for the convective rain and snow effective radius in the in-line version of COSP1.3 in CAM5. We tested leaving the convective rain and snow effective radius as zero...
and all the convective number concentrations as zero and using a static intercept parameter in the exponential distribution for convective rain and snow and a static shape parameter in the gamma distribution for convective liquid. We varied the static shape parameter value between 2 and 15, which is the range of possibilities within CAM5 (Neale et al. 2012), for the gamma distribution for convective liquid. Using the static intercept and shape parameters did not alter the distribution of reflectivity significantly nor impact the main conclusions of our paper (not shown).

Quickbeam then calculates reflectivity following Mie theory using either full Mie calculations or precalculated lookup tables (Haynes et al. 2007). In this work, the full Mie calculations were done. Currently, Quickbeam does not account for melting. The reflectivity for all of the subcolumns in a GCM grid box is then aggregated into a joint histogram of reflectivity versus height (last step in Fig. 1), referred to as a contoured frequency by altitude diagram (CFAD; Yuter and Houze 1995). The user may define the number of height bins for the CFAD, which can be different from the number of GCM levels, and designate a minimum reflectivity value to include in the CFAD. For our CFAD output, the height bins were 250 m and the minimum reflectivity was 12 dB to be consistent with the vertical resolution and detection limits, respectively, of GPM. In addition to CFADs, another GCM-precipitation diagnostic available from the COSP package is maps of near-surface precipitation frequency stratified by precipitation intensity (Kay et al. 2018). These precipitation frequency maps were originally stratified based on CloudSat thresholds and while they could be adapted for this work, we prefer to use the CFADs in order to fully evaluate the vertical structure of precipitation and not just the near-surface characteristics.

c. Modifications to COSP and Quickbeam for the GPM radar

Several modifications were made within COSP to adapt Quickbeam from a CloudSat simulator to a GPM radar simulator. First, the radar frequency was changed to 13.6 GHz to match the frequency of the DPR Ku band. Second, the number of subcolumns over which GCM cloud and precipitation are distributed was tested. The default number is 20 when running COSP offline, while the default number is 50 when running COSP inline. Using 50 subcolumns enhances the precision from 5% when 20 subcolumns are used to 2% (Fig. 2a). A 2% precision means that for every 2% change in cloud cover, one additional subcolumn in COSP contains cloud and precipitation. Higher precision is required when dealing with convective cloud and precipitation because of its highly variable spatial occurrence. For example, Fig. 2b shows the histogram for CAM5 convective cloud cover at 821 mb (roughly the lowest level that the DPR regularly observes without clutter) between 65°S and 65°N for June 2017 for both the 1° and 2° runs. Based on CAM5, convection often covers only a few percent of a grid, has a secondary peak between 15% and 20%, and rarely covers more than 30% of a grid. These characteristics do not change when going from a 1° (dash lines) to 2° (solid lines) resolution in CAM5. The use of a 2% bin size (black lines) better differentiates the distribution of convective cloud cover compared to the 5% bin size (red lines), especially at values less than 10%. Thus, we recommend using 50 subcolumns to more accurately represent the precipitation flux at the subgrid scale.

One might ask, why not run COSP with more than 50 subcolumns? As a test, COSP was run using 1600 subcolumns so that the size of each subcolumn was roughly equivalent to the GPM DPR footprint (i.e., 5 km). COSP ran significantly slower with 1600 subcolumns and the differences in subcolumn reflectivity were negligible in the CAM5 CFADs evaluated in

---

1 Running a single hourly CAM5 file through COSP using 20 subcolumns took approximately 10 s on the UCAR Cheyenne login node, whereas using 1600 subcolumns took 5 min with one core on Cheyenne.
this study, which included reflectivity over various latitude bands. We determined that negligible differences occurred in the reflectivity distributions because over a large-enough area, rounding errors in the number of subcolumns that contain cloud and precipitation will be inconsequential since over- versus underestimation of the reflectivity will average out and result in a similar CFAD regardless of a reasonable subcolumn choice.

The differences in reflectivity per subcolumn, though, may be manifest in CFADs that are produced for a small-enough area and short enough time where the over- versus underestimation of reflectivity due to rounding errors in the number of subcolumns that contain cloud do not cancel out. Ultimately, users should select the number of subcolumns to meet their needs knowing that individual subcolumn reflectivity values can change by altering the number of subcolumns since the COSP subgrid algorithm equally divides precipitation flux among the number of cloudy subcolumns. Consideration of the number of subcolumns for radar simulators was also discussed by Di Michele et al. (2012). Future work will test non-homogeneously distributing precipitation among cloudy subcolumns as previous works have already explored (Posselt et al. 2012; Song et al. 2018; Hillman et al. 2018). This study is an initial step at implementing the GPM simulator into COSP within its default framework and comparing it to modifying one assumption within the COSP subgrid algorithm, namely, the convective precipitation coverage at the subgrid scale when convective cloud cover is not available (detailed in section 5). Recall that sometimes convective cloud cover is not provided by a GCM (e.g., ECHAM; Rast et al. 2012) or there are some instances where a GCM produces precipitation at a given grid box and level but no cloud cover (like in CAM5). Since CAM5 does provide convective cloud cover in most cases, we remove it as input to the simulator to test our modifications to the COSP subgrid algorithm, which is discussed in section 4.

The third modification within COSP to adapt Quickbeam from a CloudSat simulator to a GPM radar simulator was a reduction in the reflectivity bin size for the CFADs from 5 to 2 dB. While a 5-dB bin size may be appropriate for cloud radars like CloudSat that have a dynamic range starting at very low reflectivity values for cloud, extending up into moderate reflectivity for light to moderate rain rates, it is not appropriate for heavy rainfall or delineating finer structure details of the CFAD. The logarithmic nature of reflectivity means that a 5-dB bin size could give rain rates ranging from approximately 60% to 140% of the central value (Houze et al. 2004), allowing an overly large range of rain rates within each bin and smearing out reflectivity distributions in general.

Figure 3 shows the June 2017 CAM5 2° CFADs for the tropics and subtropics (i.e., 35°S–35°N) for reflectivity bin sizes of 5 and 2 dB. The 1° CFADs were very similar and are not shown. CFAD data output by COSP at each model time step were aggregated over relevant latitudes and all longitudes to generate a single mean CFAD. The mean CFAD was then normalized by the maximum value at or above 2 km in height for each CFAD since spaceborne radar observations are not always reliable below 2 km due to surface clutter. Thus, we only plot model data starting at this level as well. Figure 3 shows that details of the vertical structure of reflectivity are lost
using 5-dB bin sizes. For example, the mode is much broader throughout the depth of the troposphere using a 5-dB bin size (Fig. 3a) compared to a 2-dB bin size (Fig. 3b). Overall, the 2-dB bin size provides finer details of the reflectivity distribution associated with precipitation.

The final modification to adapt Quickbeam to the GPM radar simulator is the elimination of the simulation of gaseous and hydrometeor attenuation. At cloud radar frequencies, and to a lesser extent DPR frequencies, gaseous and hydrometeor attenuation can be significant (e.g., Matrosov 2007; Iguchi and Meneghini 1994; Takahashi et al. 2006; Haynes et al. 2009). Quickbeam, by default, adjusts radar reflectivity profiles to include estimates of gaseous and hydrometeor attenuation. However, the GPM 2ADPR product has been corrected for gaseous and hydrometeor attenuation. Therefore, in our modified version of COSP for the GPM radar simulator, gaseous and hydrometeor attenuation are no longer simulated. If raw (i.e., uncorrected) reflectivity observations are compared to Quickbeam reflectivity, then the simulation of gaseous and hydrometeor attenuation should be retained.

The effects of attenuation for the GPM precipitation radar observations and CAM5 reflectivity from Quickbeam run at the DPR and CloudSat frequencies and 2° resolution are shown in Fig. 4. Note that section 4 will describe the DPR averaging methodology in more detail. The top row is the DPR Ku-band reflectivity CFADs between 35°S and 35°N corrected for attenuation (Fig. 4a) and not (Fig. 4b) and the percent difference between the two (Fig. 4c). The second and third rows are the equivalent set of plots for COSP output for CAM5 run for the different radars. The DPR reflectivity that has been corrected for the effects of attenuation in the level 2 product 2ADPR is analogous to the CAM5 reflectivity from COSP that did not simulate the effects of attenuation (Figs. 4a,d,g). Likewise, the DPR-measured reflectivity contains the effects of attenuation and is analogous to the CAM5 reflectivity from COSP that simulated the effects of attenuation (Figs. 4b,e,h).

Figures 4c and 4f show that for both the DPR and COSP, correcting for or eliminating the effects of attenuation shifts the reflectivity distribution to higher values at all levels. The impact is especially true at lower levels (z < 5 km) where the GPM CFAD has a strong shift from reflectivities below 32 dBZ to reflectivities above that value. This pattern is not surprising given that the effects of hydrometeor attenuation increase as hydrometeor size increases (i.e., by $D^3$) and more intense reflectivity values typically indicate larger hydrometeor size from heavier rainfall. COSP has a similar but weaker shift, which indicates that the forward model deals with attenuation differently than the DPR algorithm. Because the DPR attenuation correction has been well vetted by the satellite community and is being continually assessed through ground validation efforts, we recommend that users do not apply attenuation when they run COSP and use the attenuation-corrected DPR data.

Figure 4 also illustrates how the CAM5 CFADs look when COSP is run with the shorter wavelength of the CloudSat cloud radar. As expected, the simulated reflectivity distribution at CloudSat’s W band (Fig. 4g) is shifted to lower values compared to the simulated reflectivity at the DPR’s Ku band (Fig. 4d). This comparison shows that Quickbeam is providing reasonable results. In addition, COSP accounts for the significant attenuation at W band (Fig. 4h) and indicates the complete attenuation at reflectivity > 20 dBZ (Fig. 4i). The lack of ability of W band to capture moderate to large reflectivities further motivates the need for COSP to be run at longer wavelengths (like Ku band) when studying precipitation processes in models.

4. GPM DPR and CAM5 reflectivity comparisons

a. Spatial averaging and sensitivity considerations

While section 3c discussed our recommended changes to COSP for the simulation of reflectivity at precipitation radar wavelengths using GCM output, this section describes issues and best practices to ensure the fairest comparison between the spaceborne radar observations and output from COSP, including spatial resolution and instrument sensitivity. The native resolution of the GPM DPR is its 5-km footprint, which allows for a wide range of reflectivity values in a typical GCM grid box. However, COSP has little subgrid variation in reflectivity since precipitation fluxes are distributed equally among cloudy subcolumns of the same type (i.e., convective and large scale; section 3a). Therefore, COSP generates at most three unique reflectivity values among the subcolumns of a given GCM grid box and level—a reflectivity value for subcolumns that only contain convective precipitation, for subcolumns that only contain large-scale precipitation, and for subcolumns that contain both precipitation types. We thus averaged the instantaneous DPR reflectivity over grid boxes equivalent to the grid spacing of the CAM5 simulations before calculating the CFADs to eliminate subgrid-scale variability in reflectivity. Similar averaging approaches were done by Zhang et al. (2010) and Di Michele et al. (2012).

Figure 5 shows the difference in the CFAD structure of reflectivity from the GPM DPR using the native 5-km resolution versus using reflectivity averaged over 0.25°, 1°, and 2° grid boxes. Also shown is the CFAD of TRMM PR reflectivity averaged over 2° grid boxes for June 2008, since TRMM was no longer flying in 2017. Comparisons between June 2008 and June 2017 are fair as they are both neutral ENSO months.

The GPM CFADs from the gridbox averaged radar data (Figs. 5b–d) have an increasingly narrow distribution of reflectivity aloft (i.e., above 5 km) compared to the native 5-km scale (Fig. 5a). The narrower distribution is expected given the reduction in reflectivity variability due to the averaging over the various gridbox sizes. There is also an increasingly higher value mode at low levels because reflectivity is on a logarithmic scale such that a few large reflectivity values will overwhelm many small reflectivity values when converted to a linear scale and averaged. Thus, if a 2° box contains many weak and shallow convective cells and only a few deeper and stronger convective cells (as is common over the tropical oceans except over cooler waters where shallow, isolated convection dominates; Funk et al. 2013), the deep convection will dictate the 2° averaged reflectivity. However, in most rainy regions, stratiform rain is actually the predominant rain type in terms of
counts so the mean CFADs take on more of a stratiform-like structure (e.g., Schumacher and Houze 2006).

It is also important to acknowledge radar sensitivity when constructing and analyzing CFADs. Even though the TRMM and GPM Ku-band radars have a similar footprint size and frequency (13.8 and 13.6 GHz, respectively), they have different sensitivities to weaker reflectivity echo. TRMM PR minimum reflectivity at the footprint scale is limited to around 18 dBZ, while GPM DPR minimum reflectivity is approximately 13 dBZ. The different sensitivities slightly alter the shape and mode of the TRMM and GPM CFADs averaged over 2° grid boxes. The GPM CFAD (Fig. 5d) is deeper and wider than the TRMM CFAD (Fig. 5e) and has a larger concentration of reflectivity below 22 dBZ over all heights. The GPM CFAD also has a mode slightly lower than TRMM’s (i.e., centered at 31 vs 33 dBZ, respectively). The variability between the GPM and TRMM radar CFADs should be considered when comparing spaceborne radar data to model output that does not suffer

**FIG. 4.** (a) GPM CFAD for the tropics (35°S–35°N) for June 2017 using daily 2°-averaged reflectivity that has been corrected for the effects of gaseous and hydrometeor attenuation, (b) As in (a), but for the measured GPM reflectivity (i.e., the effects of attenuation are in the reflectivity measurement). (c) The difference between the GPM corrected and measured reflectivity CFADs. Differences are only shown where both the GPM corrected and measured data are nonzero and the GPM corrected and/or measured data are within a 10% contour. (d) CAM5 2° CFAD for the tropics for June 2017 using the COSP-Quickbeam setup altered for the GPM Ku-band radar that does not include the effects of gaseous and hydrometeor attenuation in the calculation of reflectivity. (e) As in (d), except the effects of attenuation are included in the calculation of CAM5 reflectivity. (f) The difference between CAM5 reflectivity without attenuation and with attenuation. Differences are only shown where both the COSP CFADs with and without attenuation have data and the CFAD with attenuation has data within a 10% contour. (g)–(i) CAM5 CFAD as in (d)–(f), but at the CloudSat W-band frequency. Crosshatched area in (i) represents bins that are missing in (h) due to attenuation. For all the CFADs, the reflectivity and height bin sizes are 2 dB and 250 m, respectively. Color shading and contour lines in (a), (b), (d), (e), (g), and (h) are spaced at 10% and the values represent the percentage of the maximum count value occurring at 2 km or higher.
from sensitivity issues. For example, because the contours in the 1°- and 2°-averaged GPM CFADs cease below 16 dBZ, the remaining CAM5 CFADs in this section (i.e., Figs. 6–8) are plotted starting at 16 dBZ. An even more stringent threshold should be used when comparing model data to TRMM reflectivity CFADs, especially at coarser resolutions.

b. GPM DPR and CAM5 reflectivity CFADs by latitude

As an example of the type of comparison possible with the GPM radar simulator, Fig. 6 shows the reflectivity CFADs for each latitude band using DPR and CAM5 data for June 2017. Since the GPM CFADs are very similar between 1° and 2° resolution (Fig. 5), only 2° GPM CFADs are shown in Fig. 6, while CFADs for both resolutions are shown for CAM5. Figure 6 shows a distinct latitudinal variation in the 2° CFADs from both the GPM DPR (top row) and CAM5 COSP (middle row), with variations in low-level reflectivity and upper-level echo extent consistent with the region and season. For example, the strength and depth of the low-level mode and the height of the echo tops is largest in the tropics (35°S–35°N; Figs. 6b,e), where it is warmest, and weakest in the SH mid-latitudes (35°–65°S; Figs. 6c,f), where it is auroral winter and coldest. The NH midlatitudes (35°–65°N; Figs. 6a,d), which are experiencing boreal summer, represent a reflectivity distribution in between these other regions. Similar latitudinal variations are seen in the CAM5 1° CFADs (Figs. 6g–i).

While there are general consistencies between the DPR and CAM5 reflectivity distributions, we now focus on their differences by latitude. We note that the microphysics between CAM5 and COSP are as consistent as possible; however, assumptions remain within COSP including setting the convective number concentration values to the large-scale values and a forward simulator that does not account for melting. These assumptions could impact the shape of the CFAD, but comparisons like Fig. 6 still offer insight into how well CAM5 is capturing variations in the vertical precipitation structure relative to observations. With the addition of convective number concentration in newer models, future work will examine the sensitivity of the reflectivity distribution to the assumption of equating convective number concentration to the large-scale values. Regardless of assumptions necessary in the COSP forward model, we would not expect the GPM and CAM5 CFADs to match exactly since GCM rain-rate distributions are notorious for having too much light rain rate compared to observations (i.e., the “drizzle” problem; Stephens et al. 2010). This issue could explain why CAM5 simulates significantly more weak reflectivity at low levels than the DPR at all latitudes.

In the NH midlatitudes, however, the CAM5 2° run also produces higher maximum reflectivity (and thus stronger rain rates) at low levels than the DPR (Figs. 6a,d). The NH midlatitudes are convectively active during boreal summer, so it could be argued that CAM5’s convective parameterization is being overactive in the midlatitudes during this season. However, the CAM5 2° run does not simulate deep enough echo tops over most of the reflectivity values compared to the DPR suggesting that the convective parameterization or the large-scale ice microphysics are not being active enough. When CAM5 is run at 1° resolution (Fig. 6g), there is a shift of high reflectivity (>30 dBZ) occurrence to lower reflectivity values at low levels and echo tops extend much higher to align more closely to the GPM echo tops. Thus, higher resolution appears to create more realistic precipitation structures for the warm season midlatitudes in CAM5 although there remains a disconnect between warm precipitation processes (e.g., the overly high reflectivities at low levels) and cold precipitation processes [i.e., the distribution of reflectivity at higher levels that has been enhanced (arguably too much so) but is still shifted left] in CAM5 compared to the DPR. Separating the CAM5 and GPM CFADs by precipitation type could shed more light on these discrepancies and is the topic of current research.

In the SH midlatitudes (Figs. 6c,f), the shape of the CAM5 2° CFAD is similar to the GPM CFAD, but shifted left by many dB. Thus, CAM5 run at 2° does reasonably well in capturing the vertical extent of precipitation but misses the higher end of the reflectivity observed by the DPR. In this instance, it is possible that CAM5’s convective parameterization is having trouble with the shallow convection that occurs at higher latitudes in the auroral winter hemisphere (e.g., Kulie et al. 2016). The 1° CAM5 CFAD (Fig. 6i) shifts the reflectivity distribution...
to the right, thus improving the overall model precipitation structure compared to the DPR, although echo tops become overestimated as a result.

Finally, we compare the DPR and CAM5 reflectivity CFADs for the tropics. The GPM CFAD (Fig. 6b) is strongly influenced by stratiform rain that accounts for a majority of the rain area in the tropics (Schumacher and Houze 2003). Stratiform rain occurs from depositional processes that form snow and ice aloft; these hydrometeors slowly fall and melt at the 0°C level, falling more rapidly as smaller water droplets below cloud base
(e.g., Leary and Houze 1979). These processes are reflected in the GPM CFAD in Fig. 6b, which shows a narrow distribution of weak reflectivity aloft (deposition); a region of enhanced reflectivity (melting) near 5 km, which is the climatological 0°C level in the tropics; and a broad distribution of reflectivity at low levels (rain). Convective rain broadens the reflectivity distribution and intensity at all heights via enhanced vertical motions and stronger collision and collection processes (Houze 1997; Schumacher and Houze 2006; Funk et al. 2013).

The CAM5 CFAD in Fig. 6e appears to have trouble simulating the depositional growth aloft and broadening via accretion and riming and it does not capture the bright band near 5 km (although we do not expect it to since Quickbeam does not account for melting). The CAM5 low-level mode also occurs at a much lower reflectivity value, although the range of the overall reflectivity distribution is generally similar to the DPR. The comparison of the CAM5 and GPM CFADs suggest that CAM5 does not produce strong-enough tropical convection in the mean and that there are significant issues with the ice microphysics aloft that may relate to the ability of GCMs to simulate stratiform rain associated with deep convection. The ability to represent correct precipitation processes aloft and their associated heating profiles have strong implications for large-scale circulations and climate (e.g., Schumacher et al. 2004).

c. GPM DPR and CAM5 reflectivity CFADs matched by rain rate

Rain-rate distributions from GCMs can vary widely compared to satellite observations, especially in the tropics (e.g., Fiedler et al. 2020). Since variations by surface rain rate would be expected to extend into the vertical, we isolated the DPR and CAM5 grids that agreed with each other in terms of near-surface rain rate to better assess the fidelity of the COSP forward operator when the CFAD inputs are constrained to be more consistent (Fig. 7). During June 2017, there were 26,168 hourly grid points in which the DPR and CAM5 both had nonzero rain rates from 35°S to 35°N. The CAM5 rain rates were biased toward lower values (i.e., 94% of the CAM5 nonzero precipitation flux occurred between 0 and 1 mm h⁻¹, whereas only 64% of the DPR precipitation flux did), an indication of the drizzle problem (Stephens et al. 2010). Only 6% of the grids had DPR and CAM5 rain rates matching each other within ±25% of CAM5, and 12% of the grids had rain rates matching each other within ±50% of CAM5, so we chose ±50% for a larger sample size to make rain-matched CFADs, although the ±25% rain-matched CFADs were generally similar. Note, CFADs using rain rates matched within ±50% of the DPR gave similar comparisons (not shown).
Figure 7a shows that the rain-matched GPM CFAD is consistent with the fully sampled GPM CFAD in Fig. 6b; however, there is a notable shift in the mode toward lower reflectivity at all heights (e.g., the low-level mode shifts left to 27 dBZ from 31 dBZ and the maximum occurrence at 6 km shifts left to 19 dBZ from 23 dBZ), presumably because high-rain-rate situations not simulated by CAM5 are no longer included. For the CAM5 CFADs, the low-level mode shifts in the opposite direction between the rain-matched (Fig. 7b) and fully sampled (Fig. 6e) CFADs (i.e., to 29 dBZ from 25 dBZ) and there is much less occurrence at low levels in the lower reflectivity bins. Both of these changes bring the CAM5 CFAD into better agreement with the GPM CFAD at low levels, since the model drizzle problem is mitigated and large rain rates in the DPR dataset not simulated by CAM5 are no longer included.

For the CAM5 CFADs, the low-level mode shifts in the opposite direction between the rain-matched (Fig. 7b) and fully sampled (Fig. 6e) CFADs (i.e., to 29 dBZ from 25 dBZ) and there is much less occurrence at low levels in the lower reflectivity bins. Both of these changes bring the CAM5 CFAD into better agreement with the GPM CFAD at low levels, since the model drizzle problem is mitigated and large rain rates in the DPR dataset not simulated by CAM5 are no longer included.

Regardless, Fig. 7 shows that the version of the Quickbeam simulator presented in this work for GCM output is capable of producing realistic reflectivity structures at precipitation radar wavelengths when compared to GPM satellite radar observations.

d. GPM DPR and CAM5 reflectivity CFADs for land versus ocean

Figure 8 provides yet another type of comparison possible with the COSP precipitation radar simulator, in this case differences between the DPR and CAM5 reflectivity distributions over land and ocean. Fiedler et al. (2020) showed that the ratio of daily mean tropical precipitation over land versus ocean ranges from 0.78 to 0.82 for CMIP3 through CMIP6, whereas observational satellite estimates range from 0.86 to 0.99. Thus, CMIP models simulate land precipitation amounts that are too low compared to over ocean with only slight improvement over the past 15 years. We argue that the COSP precipitation radar simulator can be used to look at potential land–ocean model biases beyond just comparing daily surface rain averages.

Figure 8 shows the 2° GPM and CAM5 CFADs over land and ocean in the tropics for June 2017 and the difference between the land and ocean CFADs at both 2° and 1° resolution. The DPR (top row) indicates that there is significantly stronger convection over land than ocean in intensity as evidenced by...
the large shift in reflectivity to higher values at all heights. The largest differences are at midlevels between 3.5 and 5.5 km (near the climatological 0°C level in the tropics) with values exceeding 20% up to 7 km. The mixed phase region exists between 5 and 7 in deep convection and the higher reflectivity values in this region suggest more robust graupel production. Averaging the observations over different grid sizes does not alter these results.

CAM5 (bottom row, Fig. 8) shows much smaller increases in reflectivity over land and little to no preference for enhancement at midlevels. In addition, an increase in model resolution from 2° to 1° does not appear to ameliorate the lack of land/ocean reflectivity differences. While this may be in part because Quickbeam does not account for melting, CAM5 only differentiates between ice and snow so is not able to fully represent precipitation growth processes in the mixed phase region. Thus, a potential root cause of the land/ocean precipitation bias in CAM5 (and other GCMs) may reside in how the convective parameterization simulates the physics associated with graupel production.

5. Improvement to the COSP subgrid algorithm

a. New subroutine

The COSP subgrid algorithm assigns convective precipitation to 5% of the subcolumns for instances when the GCM grid box has convective precipitation, but no convective cloud-cover information. However, a static 5% convective cloud cover for all convective precipitation fluxes is incompatible with GPM radar observations. The GPM DPR convective rain frequency of occurrence (or PFO) was calculated over all GPM latitudes (i.e., 65°S–65°N) for 2° grid boxes from April 2014 to March 2020 as the ratio of near-surface convective rain pixels to total (i.e., rain plus no-rain) pixels. Similarly, convective rain rates were calculated as the mean of the near-surface convective rain rate over all GPM latitudes in 2° grid boxes. Recall that the inner swath of the GPM DPR cannot cover an entire 2° grid box but the 171-pixel threshold requires that at least a reasonable portion of the grid contains radar observations with which to represent sampling over GCM grid scales. Near-surface values are used because the GPM DPR cannot see all the way to the surface. Recall, the height of the near-surface rain rates typically ranges between 1 and 2 km.

While the relationship between GPM near-surface convective rainfall and PFO can vary widely (Fig. 9), the distribution is somewhat bimodal. One arm is limited to low convective rain rates (<10 mm day⁻¹) and extends vertically, reaching convective PFO values > 60% in the extreme. The counts in Fig. 9 are on a log scale, though, such that the median convective PFO for convective rain rates < 10 mm day⁻¹ is only 0.8% (much lower than the COSP 5% assumption), while the 25th- and 75th-percentile convective PFOs are only 0.3% and 2.2%, respectively. Note also that 97% of the GPM convective samples occur at convective rain rates < 10 mm day⁻¹. However, the less common but larger convective rain rates can have an oversized impact on humans (e.g., via flooding and other extreme weather) and even the large-scale circulation. Thus, we separately analyze the other arm of the distribution that extends horizontally to high rain rates. While there is a wide range of possible convective PFOs at convective rain rates of 10 mm day⁻¹, the distribution narrows at higher rain rates as the convective PFO mode increases. The median convective PFO for convective rain rates > 10 mm day⁻¹ is 11% (much higher than the COSP 5% assumption), while the 25th- and 75th-percentile convective PFOs range from 8% to 15%.

We incorporate the distribution of GPM radar convective rain rates with respect to convective PFOs into the COSP subgrid algorithm to improve the algorithm and make it more compatible with observations. The updated COSP subgrid algorithm treats GPM convective PFO as a proxy for GCM convective cloud cover. Specifically, the GCM convective cloud cover is determined by pairing GCM convective precipitation to GPM convective PFOs. A simple linear relationship between convective PFO and rainfall (solid line, Fig. 9) is insufficient at capturing their relationship because GPM convective PFOs vary widely within each 1 mm day⁻¹ rainfall bin. Instead, to adequately account for the spread in convective PFO with convective rain rates, probability distribution functions (PDFs) and cumulative distribution functions (CDFs) of convective PFOs were made for each 1 mm day⁻¹ rainfall bin. Figure 10 shows examples of the PDFs and CDFs of convective PFO from three of the rainfall bins in Fig. 9. From the CDFs, cloud cover for a given GCM rain rate is found nondeterministically by locating the GPM convective PFO that pairs with a randomly selected number between 0 and 1 along the CDF.
To succinctly include the convective PFO PDFs and CDFs for each rain-rate bin into COSP, six-term Gaussian fits were applied to each PDF. The six-term Gaussian fit is defined as

$$f(x) = A_0 \exp\left(-\frac{z^2}{2}\right) + A_1 + A_2 x + A_3 x^2,$$

where $z = (x - A_1)/A_2$ and $x$ for our cases is the PFO bins from 0% to 100%. The six-term Gaussian fits apply well to the original PDFs and CDFs (Fig. 10). The Gaussian coefficients for the PDFs and their corresponding CDFs of convective PFO for each rainfall bin were used to write a new subroutine within the COSP subgrid algorithm. Though the Gaussian fits used in

---

**Fig. 10.** (left) Probability distribution functions (PDFs) of GPM convective-precipitation frequency of occurrence (PFO) over 65°S–65°N for the (a) 1.5, (c) 10.5, and (e) 80.5 mm day$^{-1}$ rain-rate bins. (right) As in the left column, but for cumulative distribution functions (CDFs) for the (b) 1.5, (d) 10.5, and (f) 80.5 mm day$^{-1}$ rain-rate bins. The rain-rate bin values indicate the center of the 1 mm day$^{-1}$ rain-rate bins used in Fig. 9. In each panel, the six-term Gaussian fits to the 65°S–65°N PDF or CDF are shown (i.e., red dashed lines) along with the six-term Gaussian fits to the tropical (35°S–35°N) PDFs and CDFs (i.e., blue dash–dotted lines), where the true tropical PDFs and CDFs are not shown.
the new subroutine were based on the PDFs of GPM convective PFO over all GPM latitudes, the all-latitude fits can be used for individual latitude bands since there is little difference in the all-latitude Gaussian fits and the Gaussian fits for the NH and SH midlatitudes (not shown) and tropics (blue dashed line, Fig. 10). The NH and SH midlatitude Gaussian fits are not shown because they fell almost exactly on top of the Gaussian fits for the tropics.

The steps of the new subroutine called Gaussian_Fit are shown in Fig. 11. The subroutine inputs vertical profiles of GCM convective precipitation flux (in units of kg m\(^{-2}\) s\(^{-1}\)) and converts the precipitation flux to rain rate (in units of mm day\(^{-1}\)) to be compatible with the rain-rate units from the GPM observations. Next, the subroutine finds the GCM convective rainfall that is closest to the surface to use for determining cloud cover. The subroutine then selects the 1 mm day\(^{-1}\) rainfall bin (i.e., the x axis in Fig. 9) that the GCM rainfall occurs in and reconstructs the PDFs and CDFs of the convective PFO based on the 6-Gaussian coefficients associated with that rainfall bin. Finally, a point along the CDF is selected based on a random number between 0 and 1 and paired to a convective PFO value. The convective PFO value is used as a proxy for the convective-GCM cloud cover that is input into the PREC_SCOPS algorithm of COSP. PREC_SCOPS then assigns precipitation to the gridbox subcolumns based on cloud cover (discussed in section 3a). These steps are done iteratively for each model grid point. Grid points that do not contain convective precipitation are skipped in the PREC_SCOPS subroutine.

As an example of how the new Gaussian_Fit subroutine works with the PREC_SCOPS algorithm, a model convective precipitation flux of 1.18 \times 10^{-3} \text{kg m}^{-2} \text{s}^{-1} corresponds to a rain rate of 10.2 mm day\(^{-1}\), which falls into the rain-rate bin centered on 10.5 mm day\(^{-1}\). The six-Gaussian coefficients corresponding to the 10.5 mm day\(^{-1}\) rain-rate bin are used to reconstruct the convective PFO PDF and CDF of that rain-rate bin (i.e., Figs. 10c,d). A random number between 0 and 1 is then used to select a convective PFO value. If the random number is 0.5, the PFO would be 10%. The 10% PFO would then be used as a proxy for convective cloud cover in the PREC_SCOPS algorithm, which means that 5 of the 50 subcolumns would be filled with convective rain. Each of the 5 cloudy subcolumns would then be assigned a convective rain rate of 102 mm day\(^{-1}\) (1.18 \times 10^{-3} \text{kg m}^{-2} \text{s}^{-1}) in order to return the 10.2 mm day\(^{-1}\) gridbox mean rain rate.

b. Example of updated subgrid algorithm

COSP was run twice for the CAM5 2° run using the original subgrid algorithm. The first iteration of COSP included CAM5 convective cloud cover as input and is referred to as the default-cloud run (Fig. 12a), while the second iteration excluded CAM5 convective cloud cover and is referred to as the 5%-assumption run (Fig. 12b) because it relies on the 5% convective cloud-cover assumption inside the unmodified subgrid algorithm when convective precipitation is present with no convective cloud-cover information. COSP was then run for the CAM5 2° run excluding CAM5 convective cloud cover as input, but with the updated subgrid algorithm that contains the Gaussian_fit subroutine (Fig. 12c). Ten iterations were run to ensure the robustness of our results to the random number component in the Gaussian_fit subroutine. Recall, a random point along the GPM convective PFO CDF is used to determine GCM convective cloud cover for a given GCM convective rain rate. There is little variation in the CFADs among the 10 runs with a standard deviation less than 7 counts for each bin of the CFAD (Fig. 12d). Therefore, only one iteration is shown and is referred to as the Gaussian-fit run.

The goal of Fig. 12 is to evaluate how well the 5% assumption and Gaussian fit subroutine replicate the inputs to the CFAD compared to when convective cloud cover is available from the model. Therefore, Fig. 12 compares the counts that go into each CFAD for the tropics (35°S–35°N). The midlatitude (i.e., 35°–65°N, 35°–65°S) and all-latitude (i.e., 65°S–65°N) figures are not shown, though they had consistent but subtler differences between the three COSP runs when compared to the differences among the three runs for the tropical-latitude CFADs. Figure 12 also includes data starting at 12 dBZ and all the way to the surface instead of just 2 km. We felt this was necessary to show the full impact of exchanging the 5% assumption with the Gaussian-fit subroutine and we also wanted readers to see what inputs to the CAM5 CFAD look like when including data below 2 km. Because there are more data below 2 km, the CFAD structure for an otherwise identical run would change if data were plotted all the way to the surface because CFADs are normalized to show the percent of the maximum count.

Figure 12 shows that the 5%-assumption counts (Fig. 12b) are significantly different from the default-cloud counts (Fig. 12a). While counts decrease substantially in the 5% run in most bins, there is a significant increase in counts in the higher reflectivity bins at each height. Thus, even though both runs are based on the same input precipitation flux, the 5%-assumption effectively produces a lower mean cloud cover and higher mean rain rate.

While the Gaussian-fit counts (Fig. 12c) do not perfectly reproduce the default-cloud counts, they appear to better capture important features of the default-cloud COSP run compared to the 5%-assumption COSP run by producing gridbox cloud cover and rain rates that are more consistent with the default. This perhaps becomes more intuitive if one
considers the convective cloud cover simulated by CAM5 at low levels (Fig. 2b) compared to the GPM near-surface convective PFO (Fig. 9). While not exactly matched at values > 10% (CAM5 shows a secondary peak in occurrence near 15% while the DPR shows an exponential decrease), both show a very high occurrence at values less than 2%, differentiating them from the static 5% assumption. Overall, the Gaussian fit in COSP better represents the distribution of CAM5 reflectivity compared to the 5% assumption for instances when the model has convective precipitation but no convective cloud-cover information.

6. Summary and conclusions

COSP and the radar simulator Quickbeam within COSP were adapted to simulate the Ku-band precipitation radar aboard the GPM satellite. The default CloudSat frequency of 94 GHZ was changed to match the Ku-band frequency of 13.6 GHz. Several other modifications were implemented. First, in the offline version of COSP, the number of subcolumns was set to 50 to mitigate rounding errors in the subgrid distribution of cloud and precipitation. Though the subcolumn choice did not affect our CFADs since the CFADs encompassed a large-enough area such that rounding errors in the number of subcolumns that contained cloud and precipitation were offset, subcolumn choice should be carefully considered by users to meet their needs. Additionally, the CFAD reflectivity bin size was changed from 5 to 2 dB to reduce the range of rain rates allowed per bin and allow for a more detailed representation of the reflectivity distribution. Finally, the simulation of gaseous and hydrometeor attenuation inside Quickbeam was eliminated since the CAM5 reflectivity was compared to DPR reflectivity that had already been corrected for the effects of attenuation.

To make effective comparisons of the CAM5 versus DPR reflectivity distributions, the spatial and temporal scales of the reflectivity were matched as best as possible between the model and observations because differences in the scales could skew interpretation of the model–observation comparison. The DPR reflectivity was averaged to grid boxes consistent with the CAM5 grid spacing. Spatially averaging the observations was necessary due to the homogeneous distribution of CAM5 reflectivity at the subgrid scale. Spatial averaging would...
not be necessary if the model convective and large-scale reflectivity were allowed to vary at the subgrid scale. Future work will modify the COSP precipitation subgrid algorithm to allow for inhomogeneous distribution of precipitation fluxes so that reflectivity will vary at the subgrid scale and align more closely to observations.

Comparisons between the CAM5 and GPM reflectivity CFADs were shown for the NH midlatitudes (35°–65°N), tropics (35°S–35°N), and SH midlatitudes (35°–65°S) and for land versus ocean in the tropics during June 2017. While there were general consistencies between the model and observations (e.g., weaker and less deep echo over the cold season midlatitudes compared to the warm season midlatitudes and stronger and deeper echo over tropical land compared to tropical ocean), a number of differences in the reflectivity distributions were described. These differences could be due to issues with the CAM5 physics and resolution, remaining assumptions in the COSP forward operator, or a combination of the two. However, each comparison highlighted precipitation processes to further analyze in CAM5 including the relationship between warm and cold rain processes in the warm season midlatitudes, the ability to produce shallow convection in the cold season midlatitudes, the need to capture stratiform rain and ice processes in the tropics, and the implications of having insufficient graupel representation over land. More detailed diagnosis of CAM5 and DPR reflectivity distributions will be shown in forthcoming work that separates the COSP CFADs into convective and large-scale/stratiform rain types, but the applications of this technique remain wide ranging.

Part of the COSP precipitation subgrid algorithm, which assigns precipitation to gridbox subcolumns based on cloud cover, was updated to align with observations. The original subgrid algorithm assumes 5% convective cloud cover for instances when a model produces convective precipitation, but no convective cloud information. However, the static 5% assumption is inconsistent with CAM5 output and DPR observations that show a wide distribution of convective precipitation coverage. Therefore, we wrote a new subroutine (i.e., Gaussian_Fit) within COSP that uses the DPR convective precipitation coverage as a proxy for GCM convective cloud cover. In the new subroutine, the DPR convective coverage PDF is produced based on the DPR rain-rate bin that the GCM precipitation falls into. The convective coverage that matches a random number between 0 and 1 along the CDF is used as the GCM convective cloud-cover value for a given grid point and level. In the absence of CAM5 convective cloud-cover information, CFADs produced using the new subroutine were better matched to CFADs that had CAM5 convective cloud information (and therefore did not need to make assumptions about cloud cover) than the CFADs that used the 5% cloud-cover assumption.

By adapting COSP and its radar simulator Quickbeam to the GPM precipitation radar, the vertical structure of precipitation processes can now be more fully evaluated within GCMs by simulating the heavier rainfall in the GCM that a cloud radar simulator misses. While future work is warranted to improve COSP’s performance, the modifications to COSP and Quickbeam outlined herein, as well as the guidelines for effectively comparing reflectivity in models versus observations, offer guidance for others to successfully use COSP and Quickbeam as a precipitation radar simulator.

Acknowledgments. We are grateful to Dustin Swales who helped us understand and use COSP2. Discussions with Hugh Morrison helped us understand the microphysics inside CAM5, while John Haynes and Roj Marchand clarified questions we had regarding Quickbeam. We thank the anonymous reviewers for providing constructive feedback that greatly improved the manuscript. This research was supported by the NASA Data for Operation and Assessment (NDOA) program Grant NNX17AH45G. Emily Riley Dellaripa was also supported by NASA CYGNSS Grant NNX17AH77G. We would like to acknowledge high-performance computing support from Cheyenne (doi:10.5065/D6RX99HX). We would like to acknowledge high-performance computing support from Cheyenne (doi:10.5065/D6RX99HX) provided by NCAR’s Computational and Information Systems Laboratory, sponsored by the National Science Foundation, and from the Texas A&M Supercomputing Center.

Data availability statement. The GPM 2ADPR version 6 Ku-band orbital data are included in Iguchi and Meneghini (2017). The original COSP version 2 is available on GitHub at https://github.com/CMIP/COSPv2.0, while our version of COSP2 with all the modifications mentioned in this paper is available on GitHub at https://github.com/afunktamu/COSPv2.0. GPM. CAM5 can be downloaded from http://www.cesm.ucar.edu/models/csm1.2/cam/. The output from our CAM5 simulations that are analyzed in this paper can be obtained from https://doi.org/10.25675/10217/232461.

References


