Evaluation of a general circulation model by the CERES Flux-by-cloud type simulator

Zachary A. Eitzen[1], Wenying Su[2], Kuan-Man Xu[2], Norman Loeb[2], Moguo Sun[1], David Doelling[2], Fred Rose[1], and Alejandro Bodas-Salcedo[3]

[1]Science Systems and Applications, Inc.

[2]NASA Langley Research Center

[3]Met Office Hadley Centre for Climate Change

Corresponding author: Zachary Eitzen (Zachary.A.Eitzen@nasa.gov)

Key Points:

- The CERES FluxByCloudTyp data product assigns TOA fluxes to cloud types that are defined by cloud optical depth and cloud top pressure.
- The CERES Flux-by-cloud type simulator is used to assign subgrid-scale fluxes to GCM grid cells.

Abstract

In this work, we use the Clouds and the Earth's Radiant Energy System (CERES) FluxByCloudTyp data product, which calculates TOA shortwave and longwave fluxes for cloud categories defined by cloud optical depth (τ) and cloud top pressure (p_c), to evaluate the HadGEM2-A model with a simulator. The CERES Flux-by-cloud type simulator is comprised of a cloud generator that produces subcolumns with profiles of binary cloud fraction, a cloud property simulator that determines the (τ, p_c) cloud type for each subcolumn, and a radiative transfer model that calculates TOA fluxes. The identification of duplicate atmospheric profiles reduces the number of radiative transfer calculations required by approximately 97.6%. In the Southern Great Plains region in JFD (January, February, and December) 2008, the simulator shows that simulated cloud tops are higher in altitude than observed, but also have higher values of OLR than observed, leading to a compensating error that results in an average value of OLR that is close to observed. When the simulator is applied to the Southeast Pacific stratocumulus region in JJA 2008, the simulated cloud tops are primarily low in altitude; however, the clouds tend to be less numerous, and have higher optical depths than are observed. In addition to the increase in albedo that comes from having too many clouds with higher optical depth, the HadGEM2-A albedo is higher than observed for those cloud types that occur most frequently. The simulator is also applied to the entire 60º N – 60º S region, and it is found that there are fewer clouds than observed for most cloud types, but there are also higher albedos for most cloud types, which represents a compensating error in terms of the shortwave radiative budget.

1 Introduction

Traditionally, general circulation models (GCMs) have been evaluated using gridded, monthly-averaged quantities such as cloud cover, top-of-atmosphere (TOA) outgoing longwave radiation (OLR), and shortwave albedo. While these evaluations have led to many model improvements, there can be compensating errors (particularly with radiative quantities) that combine to produce a result that is close to observed. One example of this is that in stratocumulus regions, some GCMs simulate clouds which have too little areal coverage but are also too bright, combining to produce a relatively small bias in the shortwave energy budget.

Recently, instrument simulators have been developed to help evaluate GCMs. These simulators are meant to emulate what a remote sensing instrument would measure and/or retrieve as it travels over a model atmosphere. Examples of these simulators are included in the CFMIP (Cloud Feedback Model Intercomparison Project) Observation Simulator Package (COSP; *Bodas-Salcedo et al.* [2011]). Within COSP, there are simulators of the International Satellite Cloud Climatology Project (ISCCP; *Klein and Jakob* [1999]) product, CloudSat radar reflectivities [*Haynes et al.*, 2007], the Cloud-Aerosol Lidar with Orthogonal Polarization [*Chepfer et al.*, 2008], and the Moderate Resolution Imaging Spectroradiometer (MODIS; *Pincus et al.* [2012]).

Although there are now many ways to evaluate GCMs, the CERES Flux-by-Cloud Type Simulator that will be described in this study has the potential to offer additional insight. First, the cloud frequencies and fluxes are matched within 1.5 hours to the closest CERES overpass (assuming 3-hourly model output is available). This is important because there are large diurnal cycles in cloud fraction, cloud top pressure (p_c) and cloud optical depth (τ), in many locations (e.g., [*Burleyson and Yuter*, 2015]; [*Wood et al.*, 2002]). Second, calculating the fluxes by cloud type can help isolate physical parameterizations that are problematic (e.g., convective clouds, boundary-layer parameterizations, or processes involving surface albedo), and also provide a test for updated parameterizations. Third, having the radiative properties for each (τ, p_c) cloud type provides more information than simply knowing the cloud frequencies alone, since the albedo and OLR can vary significantly within a cloud type (see [*Hartmann et al.*, 2001]; [*Zelinka et al.*, 2012]). Finally, model evaluations that use the CERES Flux-by-Cloud Type Simulator (hereafter abbreviated as FBCTSim) and the CERES FluxByCloudTyp data product (hereafter abbreviated as FBCTObs), when combined with cloud frequency of occurrence, can help determine whether an unrealistically small or large occurrence of a given cloud type results in a significant radiative impact for a given region.

The FBCTSim shares some broad similarities with the work of *Cole et al.* [2011]. They used a cloud generator and the Monte Carlo independent column approximation (McICA; *Pincus et al., 2003*; *Räisänen and Barker*, 2004) to calculate TOA shortwave and longwave fluxes along the *Terra* satellite path, and compare them to CERES SSF (Single Scanner Footprint) observations. While the FBCTSim also uses a cloud generator, the radiative transfer model it uses is designed to provide accurate flux calculations for individual atmospheric profiles, while McICA produces substantial random errors for individual profiles (but very small biases when many profiles are used) with its flux calculations [*Pincus et al.*, 2003].

Cloud radiative kernels have been used by *Zelinka et al.* [2012] to calculate how shortwave and longwave cloud feedbacks change with the cloud fraction of each ISCCP cloud type. In the course of this analysis, they compute the TOA fluxes for each cloud type based on an average of the fluxes calculated at the four (τ, p_c) corners of each bin. The impact of this assumption on estimated cloud feedbacks is quantified in *Zelinka et al.* [2012]. In this work, the fluxes within each bin correspond to the distributions of (τ, p_c) within each bin for both FBCTObs and FBCTSim.

This paper introduces both the CERES FluxByCloudTyp data product and the CERES Flux-by-Cloud Type simulator. A simplified view of the inputs (represented by ellipses) and processes (represented by rectangles) involved in both the data product and simulator is shown in Figure 1. For the FBCTObs, we begin with MODIS imager radiances, which are used to derive CERES-MODIS cloud property retrievals [*Minnis et al.*, 2011]. Two of these properties are cloud top pressure and cloud optical depth, which can be used to form a histogram of cloud frequency, similar to those seen using the ISCCP data set [*Rossow and Schiffer*, 1999]. Then the TOA fluxes and cloud properties from the CERES SSF data product are combined to produce TOA fluxes by cloud type (see Section 2.1 for details).

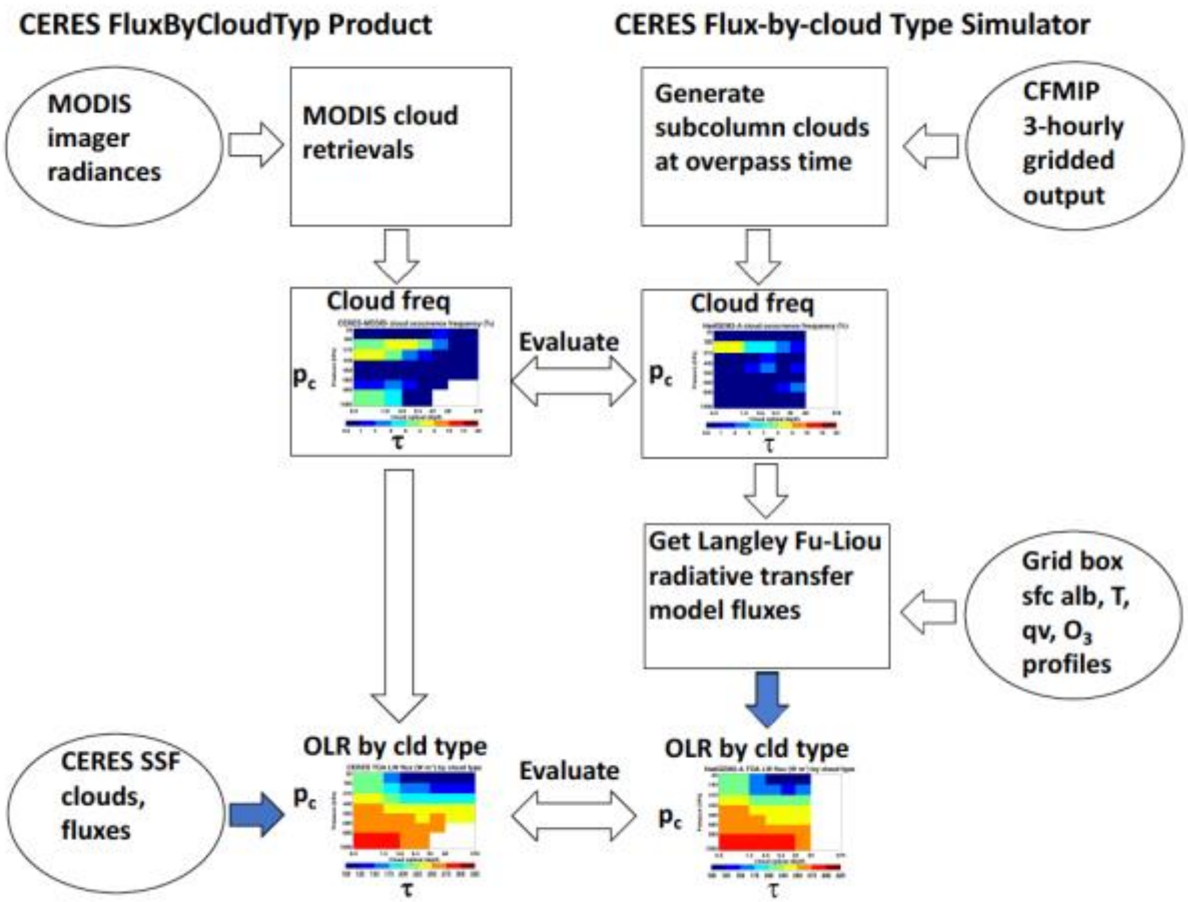

Figure 1. Schematic diagram of processes involved in producing the CERES FluxByCloudTyp data product (left side of diagram) and the CERES Flux-by-cloud type simulator (right side of diagram).

For the FBCTSim, the initial input is CFMIP (Cloud Feedback Model Intercomparison Project) 3-hourly gridded output. If a given GCM grid cell has a satellite overpass within 1.5 hours of the output time, subcolumns are produced using SCOPS (Subgrid Cloud Overlap Profile Sampler; [*Klein and Jakob*, 1999]). SCOPS uses the model's overlap assumption and grid-mean vertical profile of cloud fraction and optical depth to generate subcolumn profiles of binary (0.0 or 1.0) cloud fraction, with the model's cloud-mean optical depth assigned to each level with a cloud. In some models, the cloud fraction is split between stratiform and convective clouds, and SCOPS returns a trinary (0.0, 1.0, or 2.0) result, and the relevant stratiform cloud-mean or convective cloud-mean optical depth is assigned at each level with a 1.0 or 2.0, respectively. These subcolumn properties are run through the MODIS simulator (see section 3 and [*Pincus et al.*, 2012]), providing a histogram with frequencies of occurrence for each (τ, p_c) cloud category. This histogram from the MODIS simulator can be compared to the cloud frequency histogram from the FBCTObs product. Additional grid-scale output (e.g., surface albedo; profiles of temperature, water vapor, cloud phase, cloud particle size and ozone) are combined with the cloud subcolumns as inputs to the Langley Fu-Liou radiative transfer model (see section 3), which calculates TOA LW and SW fluxes. The average fluxes are calculated for each cloud type, and they can then be compared to the FBCTObs observations.

2 Data Sources

2.1. CERES FluxByCloudTyp Product

The CERES FluxByCloudTyp single satellite daily file product is a gridded (1°x1°), instantaneous product that uses the CERES-MODIS cloud retrievals and CERES TOA fluxes to derive fluxes for each (τ, p_c) cloud type along either the *Terra* or *Aqua* orbit. As outlined in *Minnis et al.* [2011], the CERES SSF product includes information about properties for up to two cloud levels and the clear portion (if any) within each CERES footprint. In cases where there is a single cloud level or the footprint is entirely clear, the FBCTObs fluxes assigned to the footprint are the same as those in the SSF data product. For footprints with multiple cloud types, the average narrowband MODIS radiance is converted to a broadband radiance using a narrowband-to-broadband regression for each cloud type using a method similar to [*Loeb et al.*, 2009]. The broadband radiance is then converted to an estimated TOA flux for a given cloud type ($F_{FBCTO}(\tau, p_c)$) using CERES angular distribution models [*Loeb et al.*, 2005]. The sub-footprint fluxes from the different cloud types are then normalized so that their average equals the CERES SSF TOA flux, as shown below.

$$F^{n}_{FBCTO}(\tau, p_c) = \frac{\overline{F_{SSF}}}{\overline{F_{FBCTO}}} F_{FBCTO}(\tau, p_c)$$

Here, $\overline{F_{SSF}}$ is the footprint-level CERES SSF TOA flux, and $\overline{F_{FBCTO}}$ is the FluxByCloudTyp (FBCT) footprint-mean flux averaged over the cloud types within the footprint. The flux for a particular (τ, p_c) cloud type is denoted by $F_{FBCTO}(\tau, p_c)$ prior to normalization and $F^{n}_{FBCTO}(\tau, p_c)$ after normalization.

Note that the results shown in this work from the FBCTObs product are produced with a preliminary version of the product that uses Edition 3 of the CERES SSF product. A publicly-available version of the FBCTObs product based on Edition 4 of the CERES SSF product is expected in late 2017.

2.2. HadGEM2-A Model

The HadGEM2 family of models is described in *Martin et al.* [2011]; also see *Martin et al.* [2006, 2010]. The HadGEM2-A model is an “atmosphere-only” configuration with prescribed sea surface temperatures (SSTs). The HadGEM2-A output that is evaluated here is a year of Atmospheric Model Intercomparison Project (AMIP)-style output, with many fields available at 3-hourly intervals. This output was obtained from the CMIP5/CFMIP-2 (Coupled Model Intercomparison Project Phase 5/Cloud Feedback Model Intercomparison Project 2) archive, which contained relatively few models that contained the cloud and atmosphere data necessary to run the simulator. The HadGEM2-A OLR and incoming solar radiation fields that were in the archive are not consistent with the instantaneous output from the radiation scheme, and were replaced with appropriate values provided by A. Bodas-Salcedo.

There are 38 vertical levels used in the model, with a coordinate system that is height-based in the free atmosphere, and terrain-following near the lower boundary. The vertical coordinate has higher resolution near the surface and a model top near 40 km [*Martin et al.*, 2011]. The horizontal grid resolution is 1.875º in the zonal direction and 1.25º in the meridional direction.

We will be looking at 3-month seasonal aggregates of data to compare the HadGEM2-A output with the CERES FBCT data product. Three of the seasons (MAM, JJA, and SON) are self-explanatory, but the winter season is denoted by JFD to indicate that the months used are January, February and December of 2008. The three-hourly cloud output necessary for this study was only available for calendar year 2008. Note that monthly-mean aerosol optical depths were only available through November 2008, so December 2007 aerosol optical depths were used in conjunction with the other December 2008 fields. When the December 2008 validation data was examined in isolation, the shortwave and longwave flux biases and RMS errors were similar to those from January and February of 2008 (or the three-month JFD average shown in Section 4), which indicates that using the December 2007 aerosol optical depths did not have a substantial impact on the results.

3. Description of CERES Flux-by-Cloud Type Simulator

The first element of the FBCTSim is the cloud generator, SCOPS, which takes a grid-mean profile of cloud fraction and generates subcolumns with profiles of trinary (0.0, 1.0, or 2.0) SCOPS flag, consistent with the maximum-random overlap assumption used in the HadGEM2-A model. In this study, the cloud generator produces 1000 subcolumns per grid cell. As noted in Section 1, only those grid cells with a daytime *Aqua* satellite overpass within 1.5 hours of the output time are used. When the SCOPS flag is 1.0 (stratiform) or 2.0 (convective), it is assigned the grid-mean stratiform or convective optical depth at that vertical level.

Another component of the FBCTSim represents MODIS cloud retrievals, similar to the MODIS simulator described in *Pincus et al.* [2012]. In this simulator, the vertically integrated optical depth is simply the sum of the optical depths for each subcolumn. If the total optical depth is less than 0.3, the column is considered clear (although these undetected clouds are retained for the radiative transfer flux calculation). The cloud top pressure is determined by calculating the mean extinction-weighted pressure of the cloudy portion of the atmosphere, integrating downward from TOA to τ=1 (or the lowest cloud base, if the total optical depth of the subcolumn is less than 1). A difference between the MODIS section of the simulator and that of *Pincus et al.* [2012] is that they used the ISCCP simulator to determine the cloud top pressure of low clouds, while the simulator described here uses the procedure described above for all clouds.

When calculating fluxes, FBCTSim uses the Langley Fu-Liou radiative transfer code [*Fu and Liou*, 1993; *Kato et al.*, 2005; *Rose et al.*, 2013]. For the purpose of FBCTSim, this code is operated with direct cloud inputs, which specify the phase, cloud particle diameter or radius, and optical depth for each model layer. For layers with both water and ice cloud, the phase with the higher optical depth is used, and the combined (water plus ice) optical depth is used. The relationship between optical depth and liquid/ice water content for a given cloud particle diameter is the same that is used in the CERES-MODIS cloud retrievals.

Radiative transfer calculations are computationally expensive, and the cost of performing 1000 of them per grid cell would be prohibitive. Fortunately, because the maximum-random overlap assumption is used, the actual number of distinct profiles per grid cell is approximately 24, when averaging over all HadGEM2-A grid cells in 2008. Note that there can be more than one distinct profile with the same (τ, p_c) cloud type. These distinct profiles are identified (as well as the number of subcolumns that have the same profile) and one radiative transfer calculation is performed per distinct profile, causing a 97.6% decrease in the number of calculations.

The FBCTSim is currently run offline on GCM output rather than run simultaneously within the GCM. There is a possibility of reconfiguring the code so that it runs inline, in a manner similar to those in the COSP group of simulators. With the additional computational expense of using an outside radiative transfer model, it may be prohibitively expensive to run the FBCTSim inline for long periods of time. Another option is for a model to use its own radiative transfer model on subcolumns, and in this case, the FBCTSim would be primarily used to aggregate fluxes by cloud type.

4. Validation

We wish to evaluate the ability of FBCTSim to produce TOA radiative fluxes that are similar to those produced within HadGEM2-A. First, we sum up the subcolumn fluxes calculated by FBCTSim within a HadGEM2-A grid cell. The arithmetic mean of these fluxes can then be compared to the TOA fluxes calculated by the HadGEM2-A model itself. SW and LW flux biases and RMS differences are shown in Table 1. Here, the biases are calculated by subtracting the HadGEM2-A fluxes from the Langley Fu-Liou grid-mean fluxes for each grid cell between 60° N and 60° S.

Table 1. Biases and RMS flux errors (W m^{-2}) associated between HadGEM2-A grid-cell fluxes and grid-cell mean fluxes from the simulator for three-month periods in 2008.

	TOA Reflected Shortwave			TOA OLR		
	Mean	Bias	RMS	Mean	LW Bias	LW RMS
JFD 2008	242.6	-1.5	14.4	246.2	-1.6	3.4
MAM 2008	238.4	-0.9	15.0	249.6	-1.4	3.8
JJA 2008	223.9	-0.9	14.4	254.2	-1.4	3.9
SON 2008	233.9	-1.6	14.3	252.4	-1.6	3.5

For each season, both the TOA shortwave and longwave biases are negative, with magnitudes that are less than 2 W m^{-2}. There are a number of possible reasons for differences in the fluxes, including the fact that HadGEM2-A uses a different radiative transfer scheme [*Edwards and Slingo*, 1996; *Cusack et al.*, 1999]. The shortwave RMS errors (14-15 W m^{-2}) are much larger than the longwave RMS errors (3-4 W m^{-2}), which makes sense because the dynamic range of TOA reflected shortwave flux is much larger than that of OLR.

5. Results

In order to compare the HadGEM2-A FBCT to those observed, we first normalize each cloud type's fluxes by the HadGEM2-A output fluxes:

$$F^n_{FBCTS}(\tau, p_c) = \frac{\overline{F_H}}{\overline{F_{FBCTS}}} F_{FBCTS}(\tau, p_c).$$

Here $\overline{F_H}$ is the grid-mean flux from HadGEM2-A, $\overline{F_{FBCTS}}$ is the grid-mean flux from the Langley Fu-Liou model, $F_{FBCTS}(\tau, p_c)$ is the average flux for a given (τ, p_c) cloud type from the Langley Fu-Liou model prior to normalization, and $F^n_{FBCTS}(\tau, p_c)$ is the flux after normalization. This is similar to the normalization used for the FBCTObs product, as shown in Section 2a. This normalization allows us to calculate flux differences by cloud type while preserving the grid-scale difference between the HadGEM2-A output and the CERES FBCT product. When comparing albedos and fluxes by cloud type between observations and model output it is useful to weight the results by the frequency of occurrence of each cloud type in order to identify cloud types with albedo or longwave fluxes that have important differences from those observed. The weighting that is used is

$$\Delta F_{cf}(\tau, p_c) = 0.5(C_{FBCTS} + C_{FBCTO})(F^n_{FBCTS} - F^n_{FBCTO})$$

where $\Delta F_{cf}(\tau, p_c)$ is the cloud fraction weighted flux difference for a given cloud type, C_{FBCTS} is the FBCTSim cloud fraction of that type, and C_{FBCTO} is the FBCTObs cloud fraction of that type. Although there are many ways that a cloud fraction weighted flux difference could be defined, this was chosen in order to preserve the sign of the unweighted flux difference. In addition, the weighting quantity $0.5(C_{FBCTS} + C_{FBCTO})$ will be large if C_{FBCTS} and/or C_{FBCTO} are large, ensuring that large unweighted flux differences will also appear large after weighting.

5.1. Southern Great Plains Region

In the Southern Great Plains (SGP) Region (defined here as 29.375º-40.625º N, 89.0625º-100.3125º W), there is a primary maximum in cloud occurrence at both low altitude (p_c> 800 hPa), with a secondary maximum at medium-high altitude (310 hPa < p_c< 440 hPa), as shown in Figure 2a. The cloud frequency histogram simulated by HadGEM2-A in the SGP region also has maxima at low and high altitudes, but the high-altitude maximum is stronger and at a higher altitude than observed (Figures 2b, 2c).

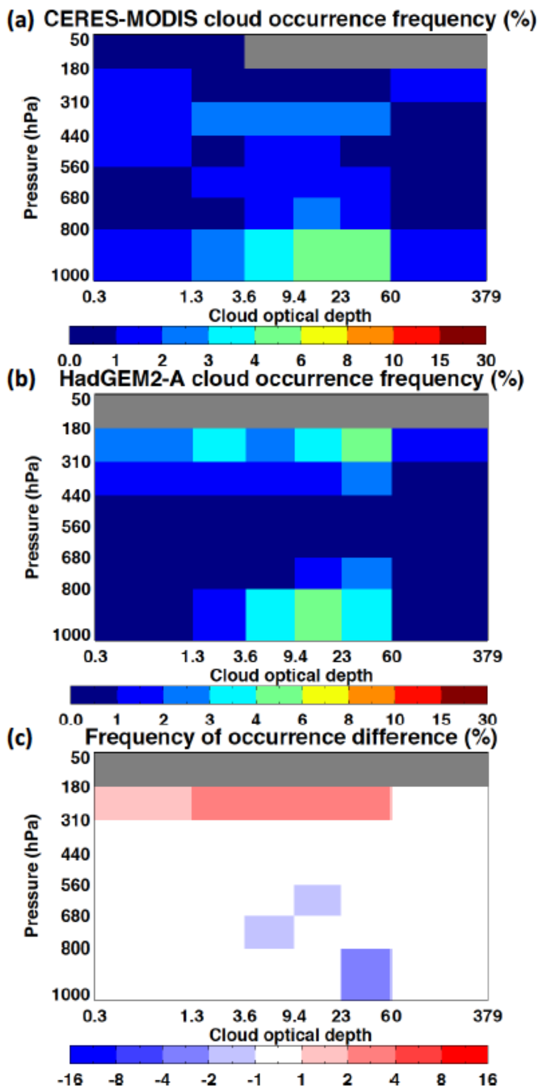

Figure 2. Average JFD 2008 cloud frequency of occurrence by cloud type over Southern Great Plains region for (a) CERES FluxByCloudTyp data, (b) HadGEM2-A model, (c) average difference (HadGEM2-A minus CERES FluxByCloudTyp). Missing types are denoted by gray shading.

As one might expect, the TOA albedo increases for cloud types with higher optical depths while remaining relatively unchanged with p_c, as shown in Figure 3a. This is also the case for the HadGEM2-A model (Figure 3b). The TOA albedo by cloud type simulated by the HadGEM2-A model over the SGP region tends to be lower than observed for most cloud types, except for clouds with p_c> 680 hPa and optical depths less than 23 (Figure 3c). Part of the reason for this difference may be that the HadGEM2-A clear-sky albedo over the SGP region (0.168) is lower than that observed (0.192; Table 2). After weighting for cloud fraction, the patterns are similar (Figure 3d), but the lower HadGEM2-A albedos for high, thin cloud types are more prominent than in the unweighted difference plot.

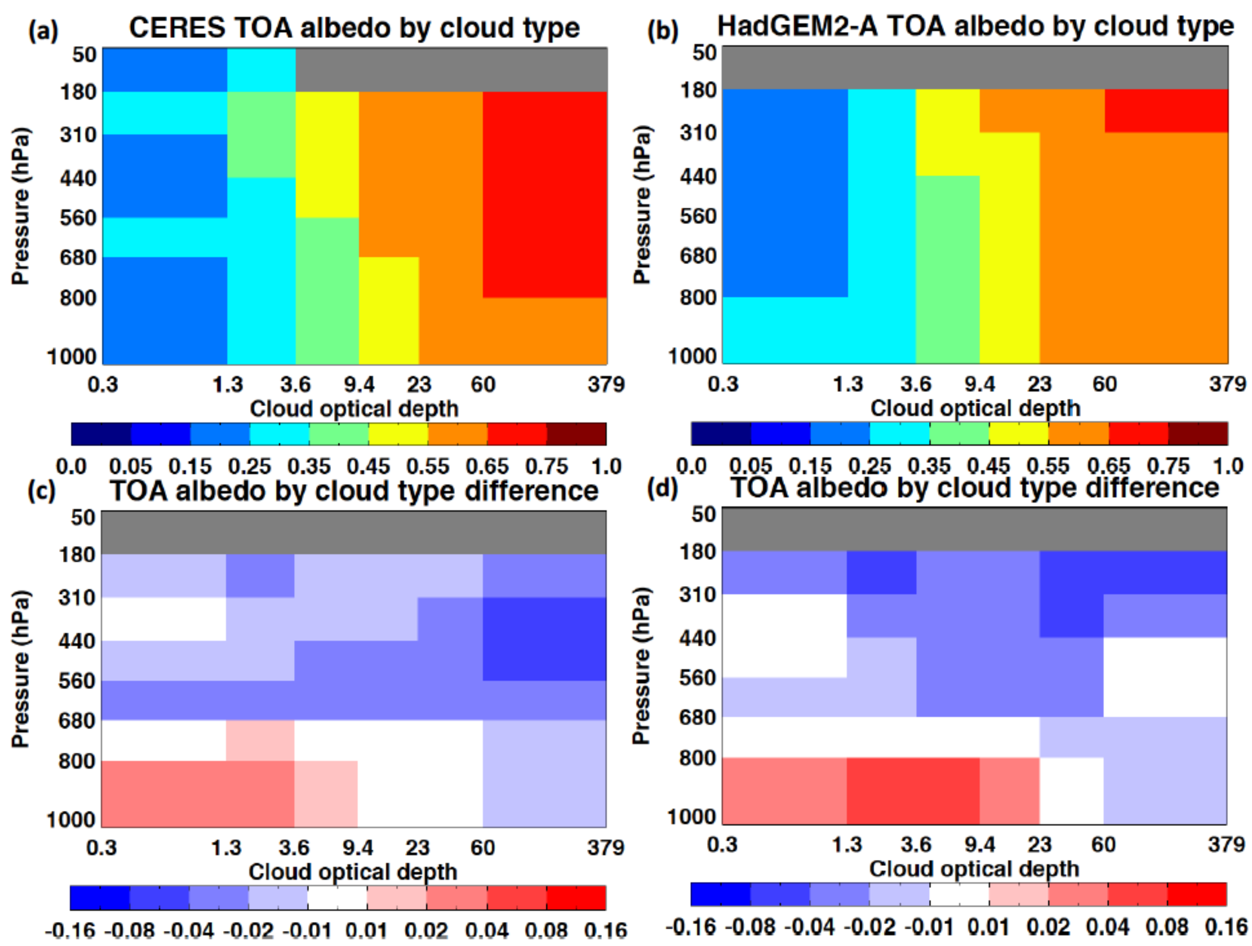

Figure 3. Average JFD 2008 TOA shortwave albedo by cloud type difference over the Southern Great Plains region (a) average difference (HadGEM2-A minus CERES FluxByCloudTyp), (b) cloud fraction-weighted difference.

Because cloud top temperature increases with p_c and emissivity increases with τ, the relationship between cloud type and OLR is less straightforward than that between cloud type and albedo. The CERES FlxByCloudTyp TOA outgoing longwave radiation over the SGP region generally decreases with optical depth at a given value of p_c and decreases with altitude for a given value of τ (Figure 4a). This is also the case for the HadGEM2-A model, except that for clouds with p_c> 800 hPa, the lowest values of OLR are with the lowest optical depths (Figure 4b). Looking at the difference plots, the HadGEM2-A model produces OLR that is significantly higher than observed for most cloud types with $\tau > 1.3$ (Figures 4c, d). This likely indicates a simulated atmosphere over the SGP region that is warmer than observed, which is consistent with the HadGEM2-A clear-sky OLR (269.3 W m^{-2}) being higher than that of FBCTObs (262.7 W m^{-2}). To test this hypothesis, the Atmospheric Radiation Measurement Best Estimate (ARMBE) temperature profile [*Xie et al.*, 2010] for JFD 2008 at the ARM SGP site is compared to the average HadGEM2-A temperature profile at the model grid cell that covers the ARM SGP site in Figure 5. Here, we see that the simulated temperatures are indeed warmer than observed at altitudes below the 200 hPa level, which where most of the simulated and observed cloud tops are.

Despite these large differences in OLR by cloud type, the HadGEM2-A average OLR for this region in JFD 2008 is 236.7 W m^{-2}, which is close to the corresponding observed regional average of 233.4 W m^{-2} (Table 2). The regionally averaged HadGEM2-A cloud fraction (0.520) is also close to observed (0.558). It appears that the HadGEM2-A bias towards high clouds (Figure 2c) compensates the higher OLRs that occur for most cloud types.

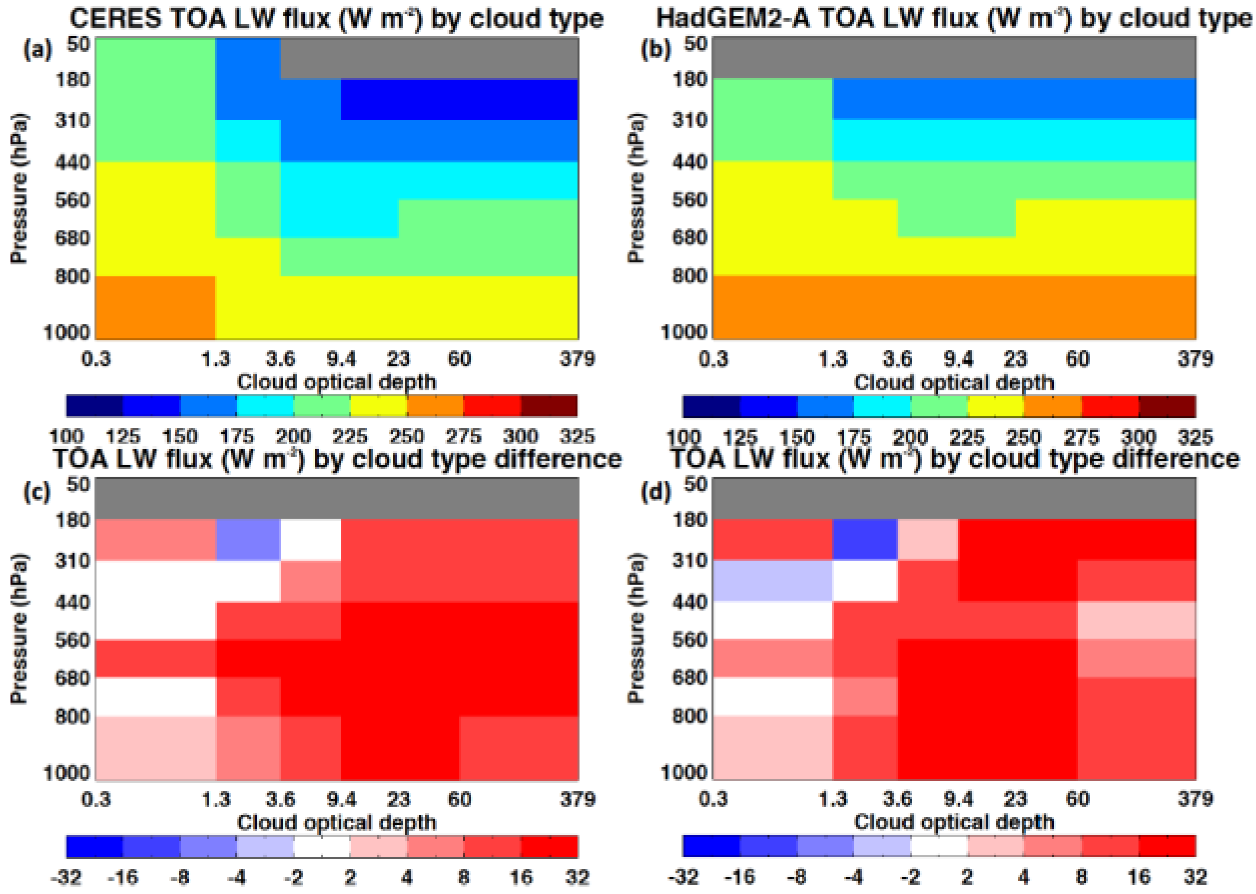

Figure 4. Average JFD 2008 TOA outgoing longwave flux by cloud type over the Equatorial Pacific region for (a) CERES FluxByCloudTyp data, (b) HadGEM2-A model, (c) average difference (HadGEM2-A minus CERES FluxByCloudTyp), (d) cloud fraction-weighted difference.

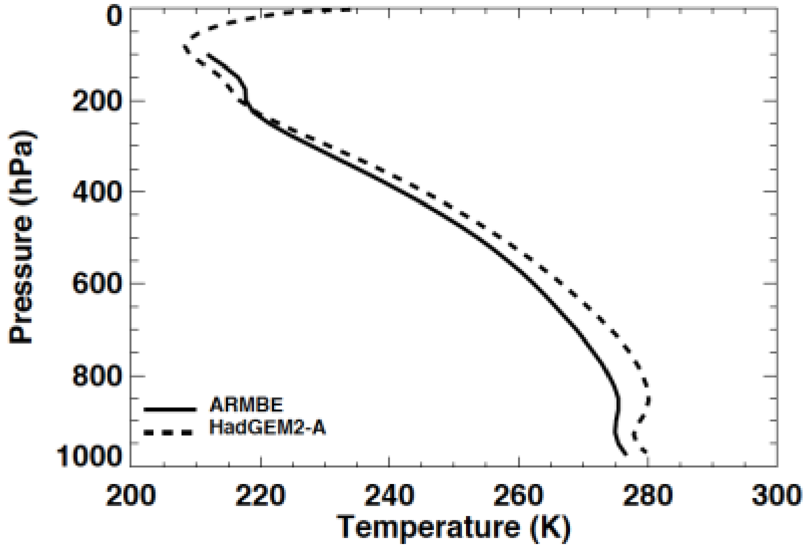

Figure 5. Average JFD 2008 temperature as a function of pressure for the ARMBE product at the ARM SGP site and HadGEM2-A grid cell.

Table 2. Average FBCTObs and HadGEM2-A cloud fractions and radiative fluxes for Southern Great Plains, Southeast Pacific, and Equatorial Pacific regions.

	SGP (JFD 2008)	SE Pacific (JJA 2008)	Eq. Pacific (JJA 2008)
Cloud Fraction (FBCTObs)	0.558	0.801	0.633
Cloud Fraction (HadGEM2-A)	0.520	0.658	0.403
All-sky TOA OLR, W m^{-2} (FBCTObs)	233.4	279.9	233.8
All-sky TOA OLR, W m^{-2} (HadGEM2-A)	236.7	289.0	252.6
Clear-sky TOA OLR, W m^{-2} (FBCTObs)	262.7	287.1	280.0
Clear-sky TOA OLR, W m^{-2} (HadGEM2-A)	269.3	298.7	285.4
All-sky TOA SW albedo (FBCTObs)	0.350	0.274	0.210
All-sky TOA SW albedo (HadGEM2-A)	0.337	0.324	0.186
TOA SW albedo (FBCTObs)	0.192	0.093	0.074
TOA SW albedo (HadGEM2-A)	0.168	0.095	0.078

5.2. Southeast Pacific Region

The Southeast Pacific region (defined here as 9.375º-20.625º S, 79.6875º-90.9375º W; similar to the “Peruvian region” in *Klein and Hartmann* [1993]) is dominated by stratocumulus clouds. This can be seen in Fig. 6a, which shows the CERES-MODIS JJA 2008 cloud occurrence frequency. The observed clouds tend to have $p_c > 800$ hPa, and low to moderate optical thicknesses with τ between 1.3 and 23. In its simulation of the same region, the HadGEM2-A model also mostly produces low clouds (Fig. 6b), but these clouds tend to have higher optical depths than observed, as shown in the difference plot (Fig 6c). As is shown in Table 2, the total

HadGEM2-A cloud fraction over the Southeast Pacific (0.658) is somewhat lower than that from CERES-MODIS (0.801).

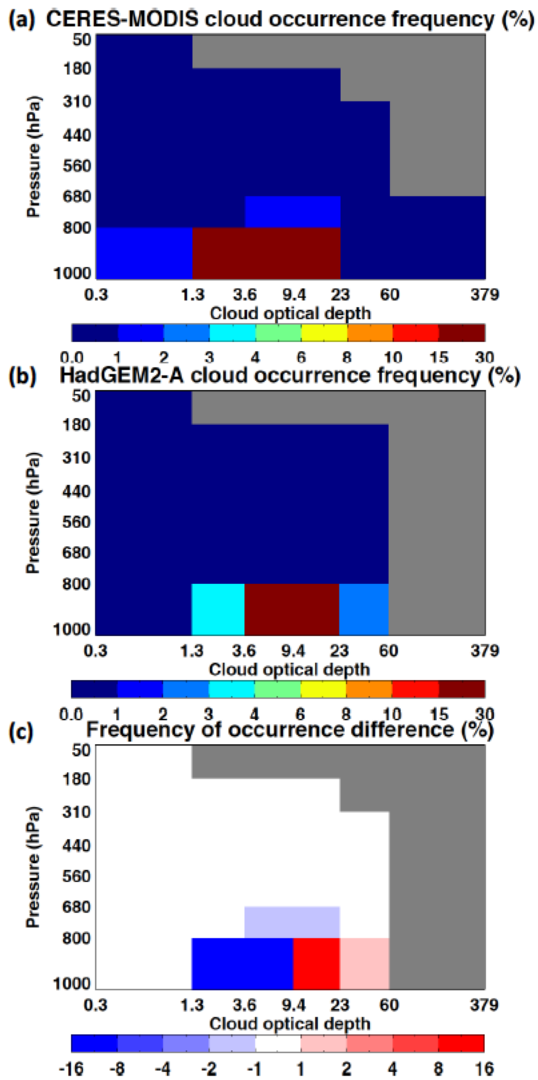

Figure 6. Average JJA 2008 cloud frequency of occurrence by cloud type over Southeast Pacific region for (a) CERES FluxByCloudTyp data, (b) HadGEM2-A model, (c) average difference (HadGEM2-A minus CERES FluxByCloudTyp).

As was the case in the SGP region, the observed and simulated TOA albedo increases for cloud types with higher optical depths while remaining relatively unchanged with p_c, as shown in Figures 7a and 7b. However, when the CERES albedo by cloud type fields is subtracted from that of the HadGEM2-A model, we see that the HadGEM2-A albedos are higher than those observed for most cloud types, except for the highest and optically thinnest clouds (Figure 7c). One possible explanation for this difference is that the optical depths within each category may be higher than those observed. When the albedo differences are weighted by cloud fraction, we see that the HadGEM2-A albedos are higher for the low clouds that dominate this region (Figure

7d). For clear scenes in the Southeast Pacific region, the albedo is similar for the HadGEM2-A model (0.095) and for the FBCTObs (0.092), as shown in Table 2.

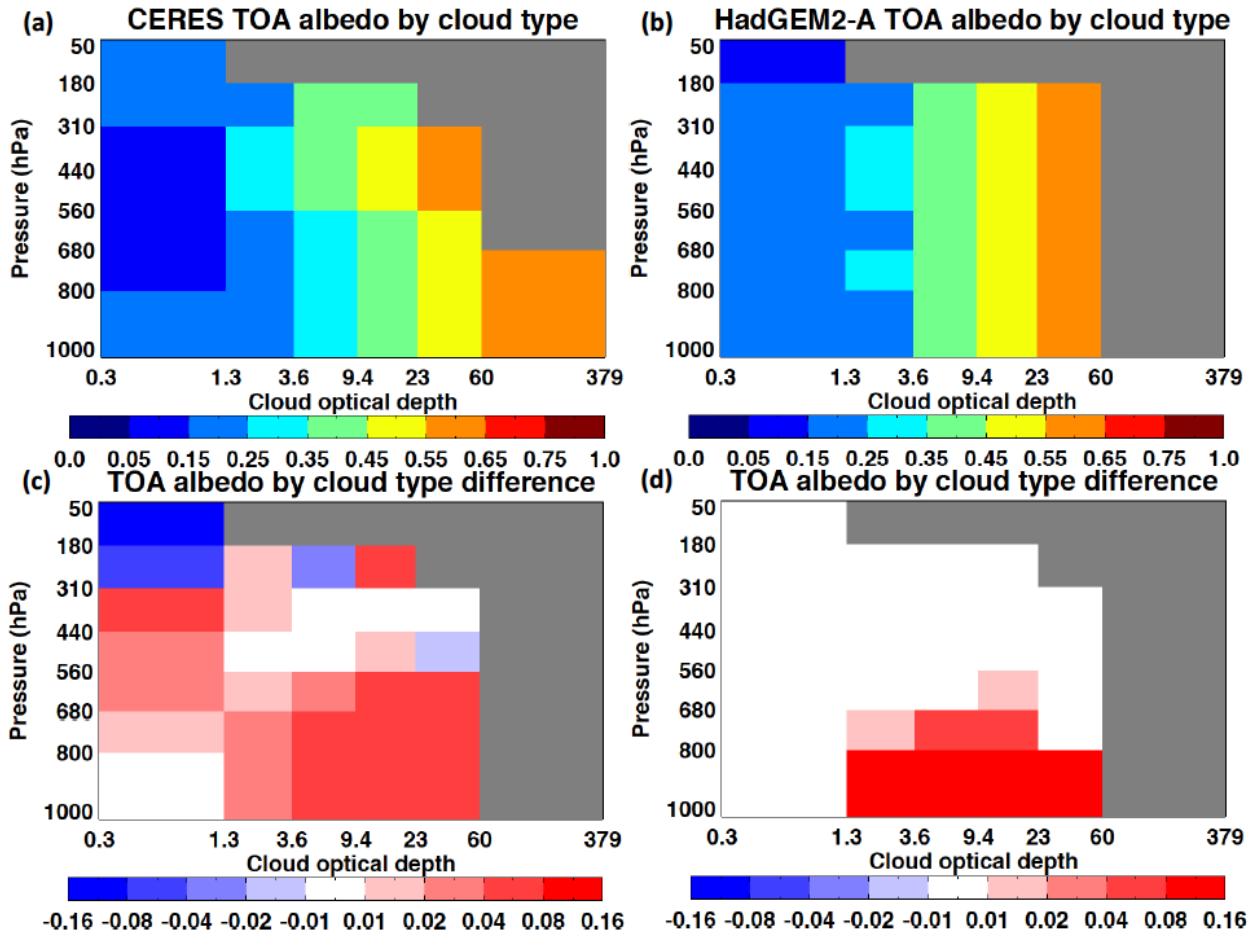

Figure 7. Average JJA 2008 TOA shortwave albedo by cloud type over the Southeast Pacific region for (a) CERES FluxByCloudTyp data, (b) HadGEM2-A model, (c) average difference (HadGEM2-A minus CERES FluxByCloudTyp), (d) cloud fraction-weighted difference.

In stratocumulus regions, a number of GCMs have the error of too little cloud fraction with a compensating error of the clouds that do form there being too bright (the "too few, too bright" problem described in *Nam et al.* [2012]). This combination of errors can bring the total albedo close to that observed, while the albedo associated with an amount of cloud cover is higher than observed. For the HadGEM2-A model, there are too few clouds in the Southeast Pacific region, and those that are present are generally have higher optical depths than observed, and within each (τ, p_c) cloud type, the albedos are too high. This leads to a HadGEM2-A all-sky albedo (0.324) that is higher than that for FBCTObs (0.274; Table 2). This final assessment is only possible with the FBCTObs product and the FBCTSim. A similar "too few, too bright" error was found in other seasons over the Southeast Pacific and also over the Southeast Atlantic (not shown).

5.3. Equatorial Pacific Region

The Equatorial Pacific region (defined here as 10.625º N-10.625º S, 154.6875º-175.3125º E, near the island of Nauru) has a wide variety of clouds, including deep convection. In Figure 8a, we see that there is a maximum in the frequency of cloud tops at low ($p_c > 800$ hPa) and high (180 hPa < p_c< 440 hPa) altitudes for JJA 2008. The simulated HadGEM2-A $p_c - \tau$ frequency diagram for the Equatorial Pacific region includes a maximum at high altitude, but it is weaker than observed, and there are far fewer cloud tops at low altitudes than observed (Figures 8b, 8c). Overall, the HadGEM2-A model simulates fewer clouds (0.403) than observed (0.633) in this region (Table 2).

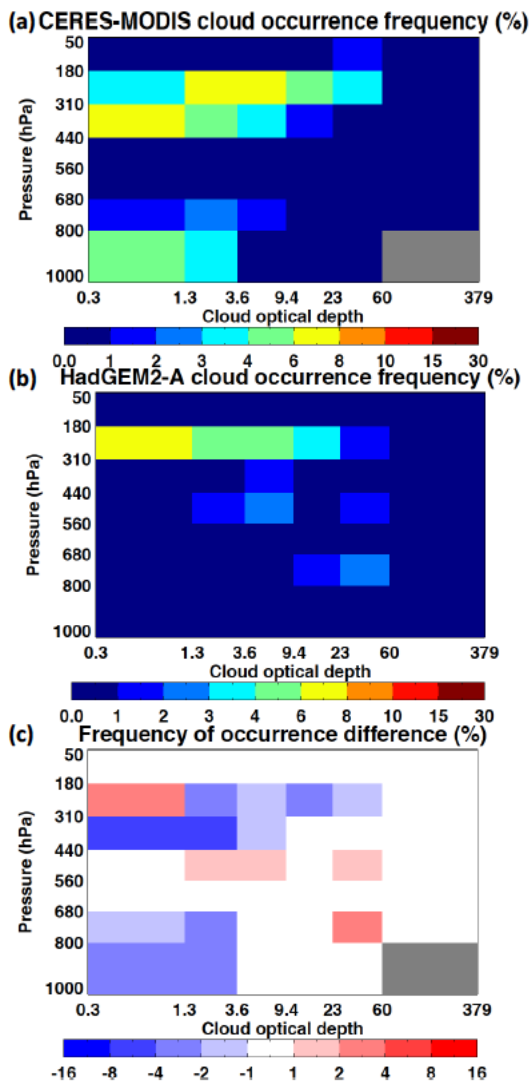

Figure 8. Average JJA 2008 cloud frequency of occurrence by cloud type over Equatorial Pacific region for (a) CERES FluxByCloudTyp data, (b) HadGEM2-A model, (c) average difference (HadGEM2-A minus CERES FluxByCloudTyp).

The overall pattern of the TOA longwave flux by cloud type histogram is for the OLR to decrease with both optical depth and cloud top height, for both observed and simulated clouds in the Equatorial Pacific (Figures 9a, 9b). The simulated HadGEM2-A longwave fluxes tend to be higher for clouds with low tops, but the HadGEM2-A fluxes are lower for cloud types with $p_c <$ 310 hPa and τ between 1.3 and 23 (Figures 9c, 9d). These high clouds with moderate optical depths are among the most common in nature and in the GCM (Figures 7a, 7b), causing the cloud fraction-weighted flux difference to be strongly negative for these types (Figure 9d). It is interesting to note that despite having lower fluxes for these cloud types, the regionally averaged JJA 2008 OLR is 252.6 W m^{-2} for the HadGEM2-A model, compared to 233.8 W m^{-2} observed (Table 2). This is likely due to the much smaller cloud fraction in this region, and also because the HadGEM2-A clear-sky OLR (285.4 W m^{-2}) is higher than that of FBCTObs (280.0 W m^{-2}).

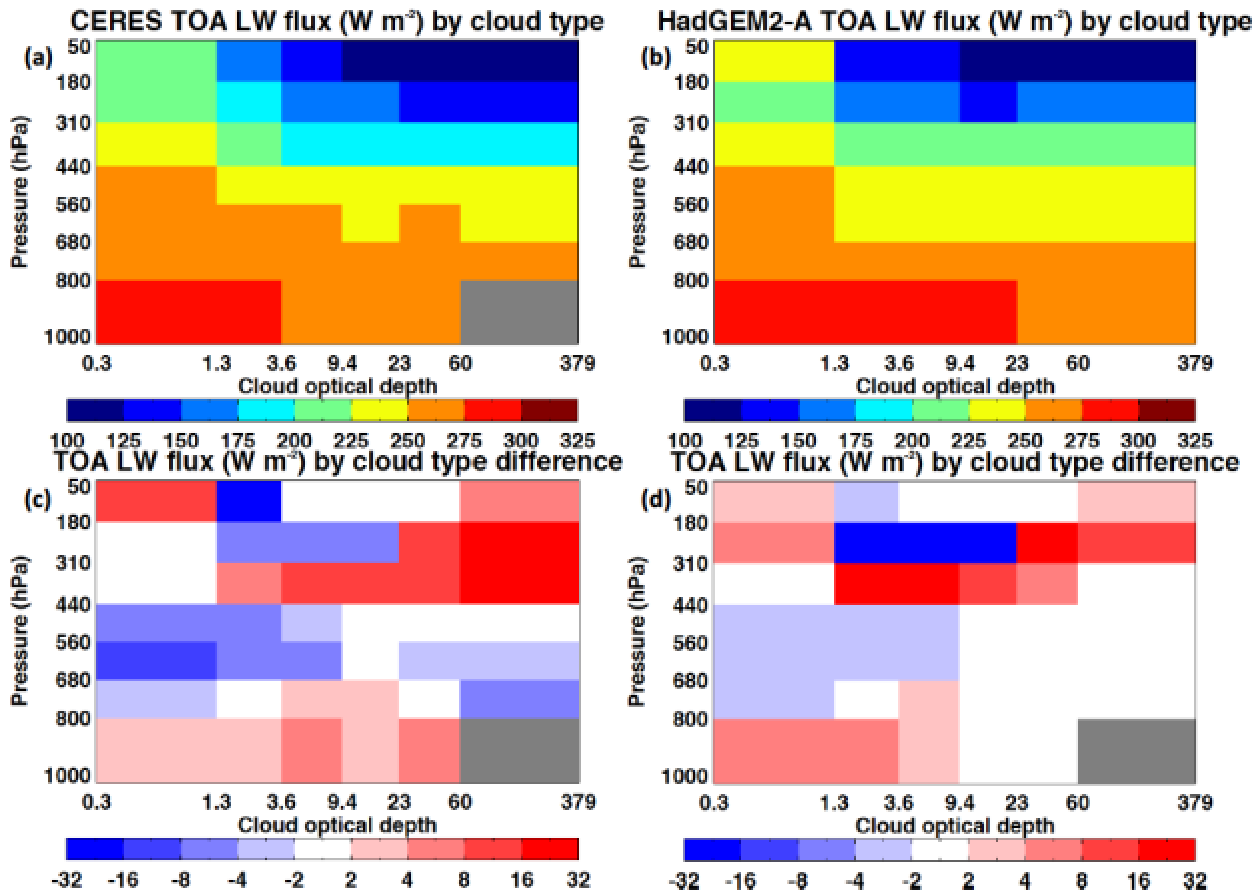

Figure 9. Average JJA 2008 TOA outgoing longwave flux by cloud type over the Equatorial Pacific region for (a) CERES FluxByCloudTyp data, (b) HadGEM2-A model, (c) average difference (HadGEM2-A minus CERES FluxByCloudTyp), (d) cloud fraction-weighted difference.

5.4. 60º N – 60º S Results

In addition to evaluating HadGEM2-A on regional scales, it is also of interest to examine whether the model has similar behavior on a global scale. To accomplish this, (τ, p_c) histograms of the differences between HadGEM2-A and the FBCTObs product were calculated for cloud frequency of occurrence, TOA shortwave albedo, and TOA OLR at each HadGEM2-A grid cell between 60º N and 60º S. These histograms were then combined, weighting by each grid cell's surface area. Here, we use MAM 2008, because the other three seasons produced similar difference histograms. This was repeated for land (grid cells with land fraction greater than 50%) and ocean (grid cells with land fraction less than 50%) grid cells. Since most of the Earth's surface is ocean, we expect the ocean histograms to be similar to those produced for all surfaces, but the land histograms can be quite different.

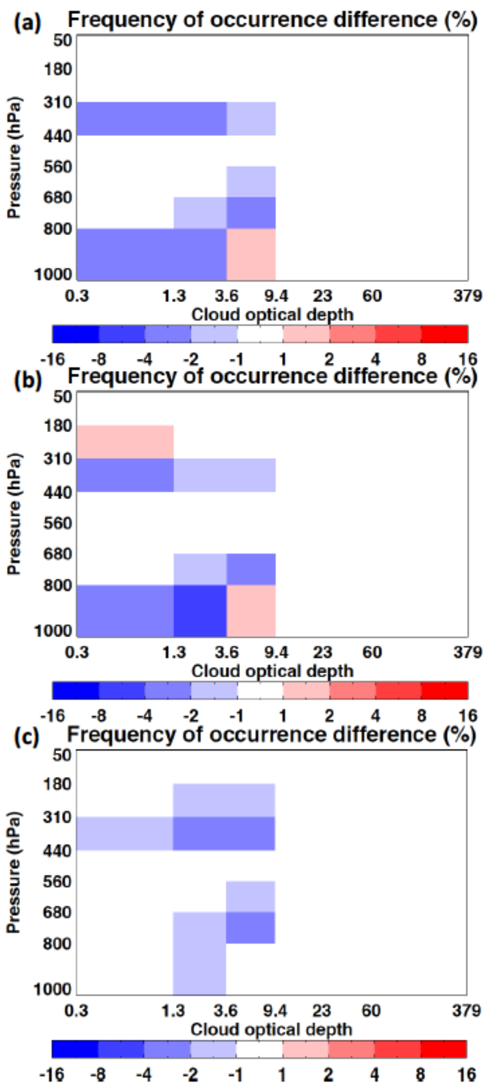

Figure 10. Average MAM 2008 cloud frequency of occurrence by cloud type differences (HadGEM2-A minus CERES FluxByCloudTyp) over 60° N – 60° S for (a) all surfaces, (b) ocean surfaces, (c) land surfaces.

The mean cloud frequency of occurrence differences between the HadGEM2-A and FBCTObs product for MAM 2008 are shown in Figure 10. For many cloud types, the difference between the model and observations is relatively small; however, over both combined land and ocean and ocean-only surfaces (Figures 11a, 11b), HadGEM2-A simulated fewer optically thin clouds with $p_c > 800$ hPa and with 310 hPa $< p_c <$ 440 hPa. The model simulates more low clouds with τ between 3.6 and 9.4 over ocean and combined surfaces. The net low-cloud behavior over ocean and combined surfaces could be characterized as “too few, too bright”, as was seen over the Southeast Pacific (Fig. 2c). Over land, the HadGEM2-A model produces too few clouds with τ between 1.3 and 9.4 at both medium and high altitudes (Fig. 11c).

When we examine the differences in albedo by cloud type for MAM 2008, we see that over combined, land-only, and ocean-only surfaces, most simulated clouds are brighter than observed, except for those with τ greater than 60, which are less reflective than observed (Figures 11a, 11b, 11c). Since there are relatively few clouds with such high optical depths, the net effect is for the cloud albedo to be higher than observed. This helps to offset the effects of having lower cloud cover than observed.

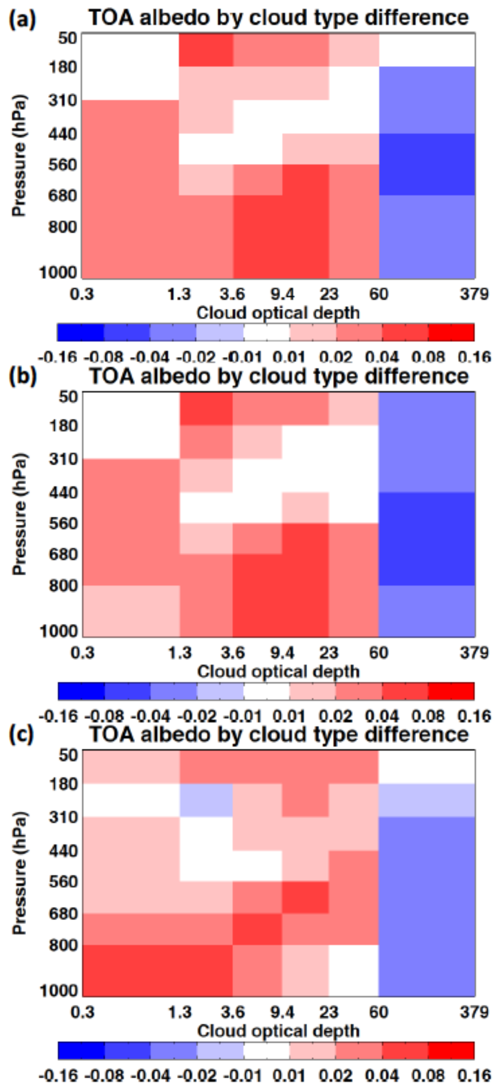

Figure 11. Average MAM 2008 TOA shortwave albedo by cloud type differences (HadGEM2-A minus CERES FluxByCloudTyp) over 60º N – 60º S for (a) all surfaces, (b) ocean surfaces, (c) land surfaces.

The HadGEM2-A values of OLR by cloud type for MAM 2008 are generally lower than those observed for $\tau < 3.6$, while the simulated OLR tends to be higher than observed for cloud types with $\tau > 23$ (Figure 12a). The OLR differences are similar for ocean surfaces (Figure 12b), but are much stronger (with the same sign for most cloud types) over land (Figure 12c).

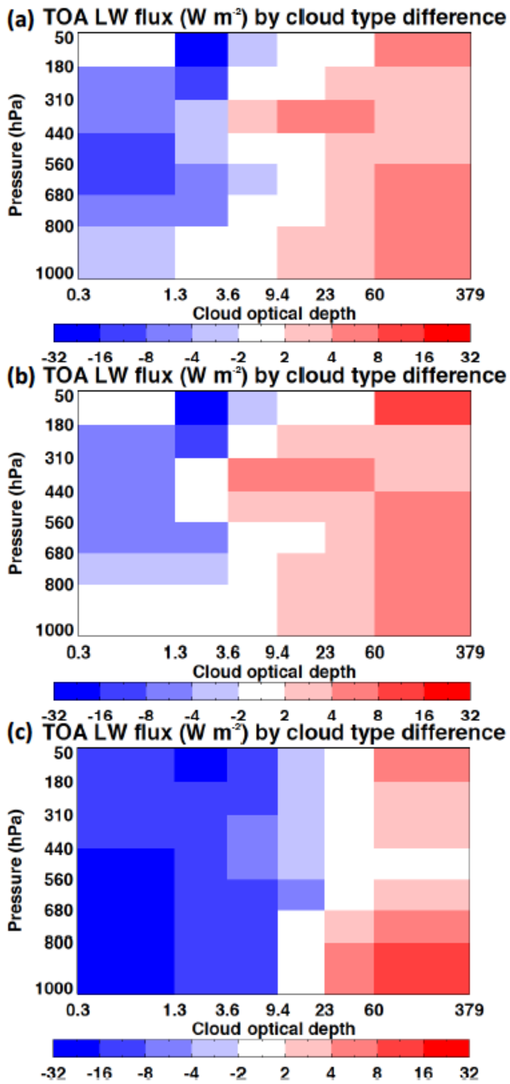

Figure 12. Average MAM 2008 TOA outgoing longwave flux by cloud type differences (HadGEM2-A minus CERES FluxByCloudTyp) over 60º N – 60º S for (a) all surfaces, (b) ocean surfaces, (c) land surfaces.

6. Conclusions

This paper has introduced the CERES FluxByCloudTyp data product. This product provides instantaneous gridded (τ, p_c) histograms of daytime cloud fraction and TOA outgoing shortwave and longwave fluxes for both the *Terra* and *Aqua* CERES instruments along their respective orbits. This data product can be used to characterize the frequency of occurrence and fluxes associated with each cloud type within 1ºx1º between 60º N and 60º S. The FBCTObs product can be used to evaluate GCMs with the additional step of applying the FBCTSim on high-frequency output.

The CERES Flux-by-cloud type simulator is comprised of a cloud generator that produces subcolumns with profiles of binary cloud fraction, a cloud property simulator that determines the (τ, p_c) cloud type for each subcolumn, and a radiative transfer model that calculates TOA fluxes. Because the maximum-random cloud overlap scheme is used in the cloud generator (consistent with the GCM), the simulator is only required to perform an average of 24 calculations per grid cell. The simulator produces shortwave and longwave fluxes that have a small (less than 2.0 W m^{-2} in magnitude) negative bias relative to the HadGEM2-A grid-mean TOA fluxes, and RMS errors of less than 15.0 W m^{-2} in the shortwave and less than 4.0 W m^{-2} in the longwave.

Over the Southern Great Plains in JFD 2008, the HadGEM2-A model produces a similar amount of cloud cover to that observed, but more clouds with high tops than are observed. Normally, one would expect the simulated OLR to be lower with the presence of more high clouds, but the flux-by-cloud type analysis shows that the HadGEM2-A model produced higher values OLR than observed for most cloud types. The compensating errors of too many high clouds, and too much OLR by cloud type leads to a realistic OLR in the Southern Great Plains region (236.7 W m^{-2}, which is only slightly higher than the 233.4 W m^{-2} observed).

When the simulator is applied to the Southeast Pacific stratocumulus region for JJA 2008, the simulated cloud tops are primarily low in altitude, which is similar to those observed. However, the clouds tend to be less numerous, and have higher optical depths than are observed, which is consistent with the “too few, too bright” problem with tropical low clouds noted by *Nam et al.* [2012]. In addition to the increase in albedo that comes from having too many clouds with higher optical depth, the HadGEM2-A albedo is higher than observed for those cloud types that occur most frequently. This diagnosis on standard GCM gridded output is only possible with an approach similar to the one used here.

Over the Equatorial Pacific for JJA 2008, HadGEM2-A produces some high clouds, but not as many as are observed, and much fewer low clouds than are observed. The overall cloud cover is much lower than observed (0.403 versus 0.633). However, the lack of high cloud cover is associated with the OLR higher than observed (252.6 versus 233.8 W m^{-2}) despite many cloud types having lower simulated values of OLR than observed.

When the flux-by-cloud type simulator is applied to the entire 60º N – 60º S region, it is shown that the simulated albedo is higher than observed for most cloud types with optical depths below 60. Since most clouds are optically thinner than this value, it points to an overall bright bias in simulated clouds. In the longwave, the HadGEM2-A model appears to have lower OLR than observed for optically thin cloud types, and higher OLR than observed for optically thick cloud types. These trends are much stronger over land than ocean, possibly indicating that land-surface processes are a factor in this bias.

We plan to publish a more comprehensive paper focused on the CERES FluxByCloudTyp data product when Edition 4 of the product is completed. We would also like to use the CERES flux-by-cloud type simulator to evaluate additional climate models in the future.

Acknowledgments

This work has been supported by the NASA Clouds and the Earth's Radiant Energy System (CERES) project.

References

Bodas-Salcedo, A., and Coauthors, 2011: COSP satellite simulation software for model assessment. *Bull. Amer. Meteor. Soc.*, **92**, 1023-1043.

Burleyson, C. D., and S. E. Yuter, 2015: Patterns of diurnal marine stratocumulus cloud fraction variability. *J. Appl. Meteor. Climatol.*, **54,** 847-866.

Chepfer, H., S. Bony, D. Winker, M. Chiriaco, J.-L. Dufresne, and G. Sèze, 2008: Use of CALIPSO lidar observations to evaluate the cloudiness simulated by a climate model. *Geophys. Res. Lett.*, **35**, L15704, doi:10.1029/2008GL034207.

Cole, J., H. W. Barker, N. G. Loeb, and K. von Salzen, 2011: Assessing simulated clouds and radiative fluxes using properties of clouds whose tops are exposed to space. *J. Climate*, **24**, 2715-2727.

Edwards, J. M., and A. Slingo, 1996: Studies with a flexible new radiation code. I: Choosing a configuration for a large-scale model. *Quart. J. Roy. Meteor. Soc*., **122**, 689-719.

Fu, Q., and K. N. Liou, 1993: Parameterization of the radiative properties of cirrus clouds. *J. Atmos. Sci.*, **50**, 2008-2025.

Hartmann, D. L., L. A. Moy, and Q. Fu, 2001: Tropical convection and the energy balance at the top of the atmosphere. *J. Climate*, **14**, 4495-4511.

Haynes, J. M., R. T. Marchand, Z. Luo, A. Bodas-Salcedo, and G. L. Stephens, 2007: A multipurpose radar simulation package: Quickbeam. *Bull. Amer. Meteor. Soc.*, **88**, 1723-1727.

Kato, S., F. G. Rose, and T. P. Charlock, 2005: Computation of domain-averaged irradiance using satellite-derived cloud properties. *J. Atmos. Ocean. Tech.*, **22**, 146-164.

Klein, S. A., and D. L. Hartmann, 1993: The seasonal cycle of low stratiform clouds. *J. Climate*, **6**, 1587-1606.

Klein, S. A., and C. Jakob, 1999: Validation and sensitivities of frontal clouds simulated by the ECMWF model. *Mon. Wea. Rev.*, **127**, 2514-2531.

Loeb, N. G., Kato, S., Loukachine, K., and Manalo-Smith, N., 2005: Angular distribution models for top-of-atmosphere radiative flux estimation from the Clouds and the Earth's Radiant Energy System instrument on the *Terra* satellite, Part I: Methodology, *J. Atmos. Ocean. Tech.*, **22**, 338–351, 2005.

Loeb, N. G., B. A. Wielicki, D. R. Doelling, G. L. Smith, D. F. Keyes, S. Kato, N. Manalo-Smith, and T. Wong, 2009: Toward optimal closure of the Earth's top-of-atmosphere radiation budget. *J. Climate*, **22**, 748-766.

Martin, G. M., M. A. Ringer, V. D. Pope, A. Jones, C. Dearden, and T. J. Hinton, 2006: The physical properties of the atmosphere in the new Hadley Centre Global Environmental Model (HadGEM1). Part I: Model description and global climatology. *J. Climate*, **19**, 1274-1301.

Martin, G. M., S. F. Milton, C. A. Senior, M. E. Brooks, S. Ineson, T. Reichler, and J. Kim, 2010: Analysis and reduction of systematic errors through a seamless approach to modeling weather and climate. *J. Climate*, **23**, 5933-5957.

Martin, G. M. and Coauthors, 2011: The HadGEM2 family of Met Office Unified Model climate configurations. *Geosci. Model Dev.*, **4**, 723-757.

Minnis, P., and Coauthors, 2011: CERES Edition-2 cloud property retrievals using TRMM VIRS and TERRA and AQUA MODIS data, Part I: Algorithms. *IEEE Trans. Geosci. Remote Sens.*, **49**, 4374–4400, doi:10.1109/TGRS.2011.2144601.

Nam, C. S. Bony, J.-L. Dufresne, and H. Chepfer, 2012: The ‘too few, too bright tropical low-cloud problem in CMIP5 models. *Geophys. Res. Lett.*, **39**, L21801, doi:10.1029/2012GL053421.

Pincus, R., H. W. Barker, and J.-J. Morcrette, 2003: A fast, flexible, approximate technique for computing radiative transfer in inhomogeneous cloud fields. *J. Geophys. Res.*, **108**, 4376, doi:10.1209/2002JD003322

Pincus, R., S. Platnick, S. A. Ackerman, R. S. Hemler, and R. J. P. Hofmann, 2012: Reconciling simulated and observed views of clouds: MODIS, ISCCP, and the limits of instrument simulators. *J. Climate*, **25**, 4699-4720.

Räisänen, P., and H. W. Barker, 2004: Evaluation and optimization of sampling errors for the Monte Carlo Independent Column Approximation. *Quart. J. Roy. Meteor. Soc.*, **130**, 2069–2085.

Rose, F. G., D. A. Rutan, T. Charlock, G. L. Smith, and S. Kato, 2013: An algorithm for the constraining of radiative transfer calculations to CERES-observed broadband top-of-atmosphere irradiance. *J. Atmos. Ocean. Tech.*, **30**, 1091-1106.

Rossow, W. B., and R. A. Schiffer, 1999: Advances in understanding clouds from ISCCP. *Bull. Amer. Meteor. Soc.*, **80**, 2261-2287.

Wood, R., C. S. Bretherton, and D. L. Hartmann, 2002: Diurnal cycle of liquid water path over the subtropical and tropical oceans. *Geophys. Res. Lett.*, **29**, 2092, doi:10.1029/2002GL015371.

Zelinka, M. D., S. A. Klein, and D. L. Hartmann, 2012: Computing and partitioning cloud feedbacks using cloud property histograms. Part I: Cloud radiative kernels. *J. Climate*, **25**, 715-735.