

Characterizing and understanding radiation budget biases in CMIP3/CMIP5 GCMs, contemporary GCM, and reanalysis

J.-L. F. Li,¹ D. E. Waliser,¹ G. Stephens,¹ Seungwon Lee,¹ T. L'Ecuyer,² Seiji Kato,³ Norman Loeb,³ and Hsi-Yen Ma⁴

Received 30 July 2012; revised 26 March 2013; accepted 27 March 2013; published 2 August 2013.

[1] We evaluate the annual mean radiative shortwave flux downward at the surface (RSDS) and reflected shortwave (RSUT) and radiative longwave flux upward at top of atmosphere (RLUT) from the twentieth century Coupled Model Intercomparison Project Phase 5 (CMIP5) and Phase 3 (CMIP3) simulations as well as from the NASA GEOS5 model and Modern-Era Retrospective Analysis for Research and Applications analysis. The results show that a majority of the models have significant regional biases in the annual means of RSDS, RLUT, and RSUT, with biases from -30 to 30 W m^{-2} . While the global average CMIP5 ensemble mean biases of RSDS, RLUT, and RSUT are reduced compared to CMIP3 by about 32% (e.g., -6.9 to 2.5 W m^{-2}), 43%, and 56%, respectively. This reduction arises from a more complete cancellation of the pervasive negative biases over ocean and newly larger positive biases over land. In fact, based on these biases in the annual mean, Taylor diagram metrics, and RMSE, there is virtually no progress in the simulation fidelity of RSDS, RLUT, and RSUT fluxes from CMIP3 to CMIP5. A persistent systematic bias in CMIP3 and CMIP5 is the underestimation of RSUT and overestimation of RSDS and RLUT in the convectively active regions of the tropics. The amount of total ice and liquid atmospheric water content in these areas is also underestimated. We hypothesize that at least a part of these persistent biases stem from the common global climate model practice of ignoring the effects of precipitating and/or convective core ice and liquid in their radiation calculations.

Citation: Li, J.-L. F., D. E. Waliser, G. Stephens, S. Lee, T. L'Ecuyer, S. Kato, N. Loeb, and H.-Y. Ma (2013), Characterizing and understanding radiation budget biases in CMIP3/CMIP5 GCMs, contemporary GCM, and reanalysis, *J. Geophys. Res. Atmos.*, 118, 8166–8184, doi:10.1002/jgrd.50378.

1. Introduction

[2] Atmospheric radiative structures, such as fluxes and the vertical/horizontal distributions of heating, are one of the most important factors determining global weather and climate. Clouds can exert a strong influence on regional radiative balance by reflecting shortwave (SW) radiation back to space and trapping longwave (LW) radiation and radiating it back to the surface, providing one of the strongest feedback models in the climate system. The balance of these fluxes is essential for understanding Earth's climate system and constraining the energy balance for climate models

[Stephens, 2005]. However, representing clouds and cloud radiation feedback in global climate models (GCMs) to reduce and quantify uncertainties associated with climate change projections remains a big challenge. Global constraints and information for developing and evaluating clouds and radiation in GCM simulations were typically derived from radiation budget observations from the Earth Radiation Budget Experiment/Clouds and the Earth's Radiant Energy System (ERBE/CERES) [Wielicki *et al.*, 1996] and from cloud cover observations from the International Satellite Cloud Climatology Project (ISCCP) and related products [e.g., Han *et al.*, 1999; Rossow and Zhang, 1995; Rossow and Schiffer, 1999]. In the last decade, the first satellite simulator was available for the International Satellite Cloud Climatology Project (ISCCP) [Rossow and Schiffer, 1999] to serve for evaluation and intercomparison of climate model clouds [e.g., Norris and Weaver, 2001; Lin and Zhang, 2004; Zhang *et al.*, 2005; Schmidt *et al.*, 2006; Cole *et al.*, 2011; Kay *et al.*, 2012]. More recently, the Cloud Feedback Model Intercomparison Project (CFMIP) [e.g., Bony *et al.*, 2011] has been coordinating development of the CFMIP Observation Simulator Package (COSIP) and includes a number of new satellite observations from Multiangle Imaging Spectroradiometer, Moderate Resolution Imaging Spectroradiometer (MODIS), CloudSat,

Additional supporting information may be found in the online version of this article.

¹Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA.

²University of Wisconsin-Madison, Madison, Wisconsin, USA.

³NASA/Langley Research Center, Hampton, Virginia, USA.

⁴Lawrence Livermore National Laboratory, Livermore, California, USA.

Corresponding author: J.-L. F. Li, JPL/CalTech/NASA, Jet Propulsion Laboratory, MS 233-306K, 4800 ak Grove Drive, Pasadena, CA, 91109, United States. (Juilin.F.Li@jpl.nasa.gov)

©2013. American Geophysical Union. All Rights Reserved.
2169-897X/13/10.1002/jgrd.50378

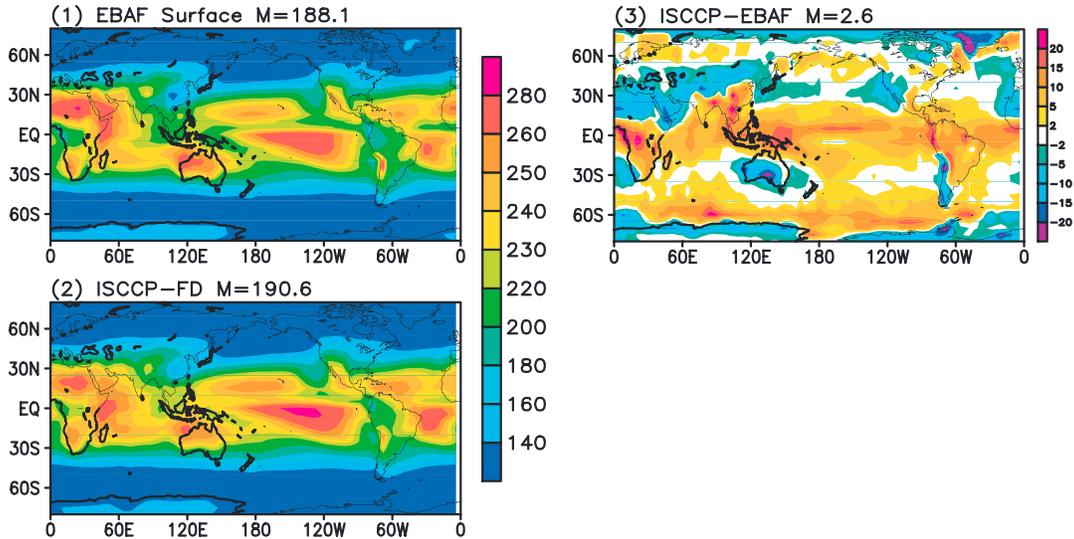


Figure 1. Annual mean maps of surface shortwave downward flux (RSDS; W m^{-2}) estimated from satellite estimates of (1) EBAF-Surface, (2) ISCCP-FD (refer to as ISCCP), and (3) the differences between EBAF-Surface and ISCCP. The ISCCP estimates are used for one reference boundary for the model evaluations for the Taylor plot presented in this paper.

Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO), and Polarization and Anisotropy of Reflectances for Atmospheric Sciences Coupled With Observations From a Lidar (PARASOL). COSP has been used widely to understand and quantify climate model cloud biases [e.g., *Chepfer et al.*, 2008; *Bodas-Salcedo et al.*, 2008, 2011; *Zhang et al.*, 2010; *Kay et al.*, 2012; *Kodoma et al.*, 2012, etc; *Nam and Quaas*, 2012]. For example, with the ISCCP simulator, *Zhang et al.* [2005] identified the “too few, too bright problem” in the early stage of GCMs, while recently, *Klein et al.* [2012] have shown that the representation of clouds, in particular the “too few, too bright” problem, has improved in CMIP5 over CMIP3.

[3] A key element of obtaining an accurate top of atmosphere (TOA) and surface radiation budget is the representation of clouds, which for GCMs and Earth radiation budget considerations can be roughly broken down into cloud

cover, cloud water mass, and cloud particle sizes. While ISCCP and other products have provided some very useful information for the constraints on cloud cover, the latter two quantities have been largely unconstrained due to the lack of observations for cloud water mass and particle size. This is especially the case for the vertical structure information leaving too many degrees of freedom unconstrained. The ramifications of this issue for cloud mass are clearly evident in the wide disparity in the cloud ice and liquid water content (CIWC and CLWC) values exhibited in present-day models, including those contributing to Phases 3 and 5 of the Coupled Model Intercomparison Project (CMIP3 and CMIP5) [e.g., *Waliser et al.*, 2009; *Li et al.*, 2012].

[4] The recent availability of the first tropospheric vertically resolved cloud radar reflectivity and derived ice/liquid profiles from CloudSat [*Austin et al.*, 2009] and combined with CALIPSO [*Deng et al.*, 2010, 2013; *Delanoë and Hogan*, 2008, 2010] provides new means for global cloud mass evaluation [e.g., *Chepfer et al.*, 2008; *Li et al.*, 2012];

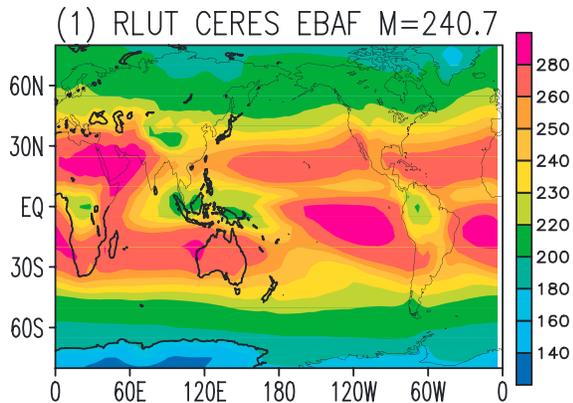


Figure 2. Annual mean maps of top-of-atmosphere (TOA) radiative longwave upward flux at TOA (RLUT: W m^{-2}) from satellite measurements CERES EBAF which are used for the model evaluations presented in this paper. M represents the global area mean value.

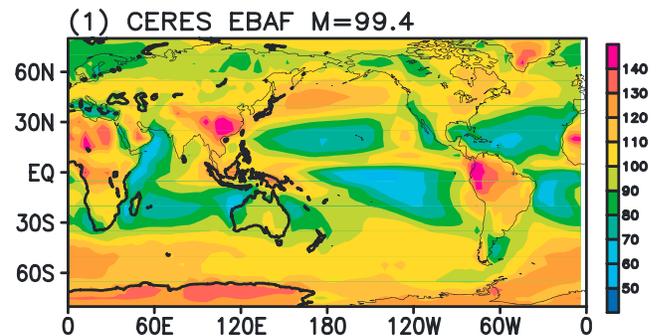


Figure 3. Same as Figure 2 but for annual mean maps of radiative shortwave upward at TOA (RSUT: W m^{-2}) from satellite measurement of CERES EBAF which are used for the model evaluations presented in this paper. M represents the global area mean value.

Table 1. Model Label, Resolution, Institution, and Model Name for the CMIP5 GCMs Examined in This Study (See Section 3 for More Details)

Model Label	Model Resolution	Institution/Model Name
BCC-CSM1	64 × 128 × 26	Beijing Climate Center, China/BCC-CSM1-1
BCC	64 × 128 × 26	Beijing Climate Center, China/BCC-CSM1-1_esm
CanESM2	64 × 128 × 35	Canadian Centre for Climate Modelling and Analysis, Canada/CanESM2
CCSM4	288 × 192 × 26	National Center for Atmospheric Research, USA/CCSM4
CNRM-CM5	128 × 256 × 17	Centre National de Recherches Meteorologiques, France/CNRM-CM5
CSIRO	96 × 192 × 18	Australian Commonwealth Scientific and Industrial Research Organization, Australia/CSIRO-Mk3-6-0
GISS-E2-H	90 × 144 × 29	NASA/Goddard Institute for Space Studies, USA/GISS-E2-H
GISS-E2-R	90 × 144 × 29	NASA/Goddard Institute for Space Studies, USA/GISS-E2-R
HadGEM2-ES	145 × 192 × 38	Hadley Centre for Climate Prediction and Research/Met Office, UK/HadGEM2-ES
Inmcm4	120 × 180 × 21	Institute for Numerical Mathematics, Russian/Inmcm4
Inmcm4ESM	120 × 180 × 21	Institute for Numerical Mathematics, Russian/Inmcm4_esm
IPSL	96 × 96 × 39	Institute Pierre Simon Laplace, France/IPSL-CM5A-LR
MIROC	64 × 128 × 80	University of Tokyo, NIES, and JAMSTEC, Japan/MIROC-ESM-CHEM
MIROC4h	320 × 640 × 56	University of Tokyo, NIES, and JAMSTEC, Japan/MIROC4h
MIROC5	128 × 256 × 40	University of Tokyo, NIES, and JAMSTEC, Japan/MIROC5
MRI-CGCM3	160 × 320 × 35	Meteorological Research Institute, Japan/MRI-CGCM3
NorESM	96 × 144 × 26	Norwegian Climate Centre, Norway/NorESM1-M
MPI	96 × 192 × 25	Max Planck Institute for Meteorology/MPI-ESM-LR
GFDL-CM3	90 × 144 × 40	NOAA GFDL/GFDL-CM3
CESM1	192 × 288 × 30	NSF-DOE-NCAR/CESM1-CAM5

Waliser et al., 2009; Chen et al., 2011; Jiang et al., 2012; Gettelman et al., 2010; Klein and Jakob, 1999; Webb et al., 2001; Delanoë and Hogan, 2008, 2010; Delanoë et al., 2011; Bodas-Salcedo et al., 2008, 2011; Zhang et al., 2010; Kay et al., 2012; Kodoma et al., 2012]. Among them, Li et al. [2011, 2012] and Waliser et al., [2009, 2011] strove to point out that considerable care and caution are required in order to make judicious comparisons/interpretations regarding atmospheric liquid/ice and its associated interactions with radiation. This is because most GCMs typically only represent the “suspended” hydrometeors associated with some/most clouds, while satellite observations include both clouds and falling hydrometeors (e.g., rain or snow) as well as convective core cloud mass. For example, the observations from sensors such as the CloudSat Radar and the Clouds and the Earth’s Radiant Energy System (CERES) instruments are sensitive to a broader range of particles for ice/liquid water mass, including clouds, falling snow/rain, as well as convective core water mass. In contrast, most GCMs, including all CMIP3 models and most of the models in CMIP5, only model the radiation impacts from the cloud-related hydrometeors and, in some cases, not even all the clouds (e.g., deep convection). Given that most models from CMIP3 and CMIP5, for example, significantly underestimate (or do not explicitly model all) the total water mass, this may result in possible biases in the radiation fields. Evidence for systematic biases in GCM radiation fields, for example, in CMIP3, was given in the analysis by Trenberth and Fasullo [2009] that showed too much absorbed SW and outgoing LW in conjunction with heavy precipitation regions, e.g., the Intertropical Convergence Zone (ITCZ). An observation-based modeling study by Waliser et al. [2011] led to the hypothesis that the typical practice of ignoring the impacts of precipitating hydrometeors would account for at least a portion of this systematic bias. As the persistence of this practice continues, it is worth examining if the same systematic bias might be evident in CMIP5.

[5] In this study, we examine systematic radiation budget biases in CMIP3 and CMIP5, in particular by evaluating the RSDS at the surface and RSUT and RLUT fluxes at the TOA. These three quantities were chosen because they are

the radiative fluxes that are most directly influenced by convections and clouds, and we are interested in examining their biases in light of the liquid and ice water biases mentioned above. Note that other fluxes, such as the downward longwave flux at the surface, are as much influenced by clouds; they are also influenced by other species such as water vapor, which are not covered in this study. The model simulations considered in this study are from the twentieth century CMIP3 and CMIP5 simulations as well as the NASA GEOS5 AGCM with prescribed SSTs and Modern-Era Retrospective Analysis for Research and Applications (MERRA) reanalysis data. Observation-based reference data for the TOA fluxes are derived from contemporary satellite radiation measurements, while surface fluxes are derived from satellite-constrained model calculations using a radiative transfer model. In section 2, we describe these observational data sets, including the way the different retrievals and other methodologies are combined to form a robust observational estimate with some quantitative information on uncertainty. In section 3, we briefly describe the models and Modern-Era Retrospective Analysis for Research and Applications (MERRA) reanalysis data sets utilized in this evaluation study. In section 4, we illustrate and discuss the results of our model evaluation. Section 5 summarizes and draws conclusions.

2. Radiation Data Sources

[6] To evaluate model radiative fluxes and to help account for observational uncertainty, we use reflected shortwave and outgoing longwave radiative fluxes at the TOA from satellite measurements and computed downward shortwave flux at the surface from two different algorithms. They are the following.

[7] 1. *Surface downward radiative flux data.* The sources of surface downward radiation fluxes, referred to here as RSDS, are from EBAF-Surface and ISCCP-derived products.

[8] 1.1 *EBAF-Surface.* This surface flux radiation product is constrained by TOA CERES-derived flux with Energy Balanced and Filled (EBAF) adjustments [Kato et al., 2011, 2012a, 2012b] and is based on two CERES data products.

Table 2. Outline of Shortwave and Longwave Radiation Representations Used in the Selected CMIP5 Models, as Well as GEOS5 AGCM and MERRA

Models	Prognostic Cloud Variables	Shortwave Radiation Scheme	Longwave Radiation Scheme	Note
GEOS5/MERRA	Single mixing ratio of cloud condensate	Absorption due to water vapor, O ₃ , O ₂ , CO ₂ , clouds, and aerosols. Interactions among the absorption and scattering by clouds, aerosols, molecules (Rayleigh scattering), and the surface are fully taken into account. Fluxes are integrated virtually over the entire spectrum, from 0.175 to 10 μm [Chou and Suarez, 1999].	This includes the absorption due to major gaseous absorption (water vapor, CO ₂ , O ₃) and most of the minor trace gases (N ₂ O, CH ₄ , and CFCs), as well as clouds and aerosols with optical properties specified as input parameters.	Chou and Suarez [1999]; Rienecker et al. [2008]
BCC-CSM1	Single mixing ratio of total water	Morcrette and Fouquart [1986]	Morcrette et al. [1986]	Collins et al., [2004]; Rashe and Rasch and Kristjánsson, [1998]; Sundqvist, [1988]
BCC	Single mixing ratio of total water	Morcrette and Fouquart, [1986]	Morcrette and Fouquart, [1986]	Collins et al., [2004]; Rashe and Rasch and Kristjánsson, [1998]; Sundqvist, [1988]
CanESM2	Mixing ratio of cloud liquid and ice	A two-stream solver, along with a delta-Eddington approximation, is used for calculations of radiative transfer in the atmosphere, which leads to a numerically efficient approach that has linear dependence on the number of vertical levels.	A two-stream solver is also used along with a methodology to efficiently account for the scattering by cloud and aerosol particles [Li, 2002].	
CCSM4	Single mixing ratio of total water	Delta-Eddington with exponential-sum fit representation of near-IR, truncated ICA (TICA) for all-sky	RRTMG absorptivity-emissivity method for clear skies, plus truncated independent column approximation (TICA) for all-sky	Gent et al. [2011]
	Diagnostic cloud liquid water, PDF based	From Fouquart and Bonnel [1980], integrates the fluxes over the whole shortwave spectrum between 0.2 and 4 μm. The scheme includes Rayleigh scattering, absorption by water vapor and ozone, both varying in space and time, and by CO ₂ , N ₂ O, CO, CH ₄ , and O ₂ , which are treated as uniformly mixed gases.	Rapid Radiation Transfer Model (RRTM) [Mlawer et al., 1997] included in the IFS ECMWF model. The radiativetransfer equation is solved by a two-stream method. The RRTM scheme computes fluxes in the spectral range encompassing the 10–3000 cm ⁻¹ band. The computation is organized in 16 spectral bands and includes line absorption by H ₂ O, CO ₂ , O ₃ , CH ₄ , N ₂ O, CFC-11, CFC-12, and aerosols.	
CSIRO-Mk3.6.0	Mixing ratio of cloud liquid and ice, diagnostic falling snow	The shortwave radiation scheme is a two-stream code with 12 bands and 24k terms [Grant and Grossman, 1998; Grant et al., 1999].	The longwave scheme has 10 bands—a combination of <i>k</i> distribution and precomputed transmittances [Grant et al., 1999].	
GISS-E2-H	Single mixing ratio of total water, diagnostic falling snow			
GISS-E2-R	Single mixing ratio of total water, diagnostic falling snow			
HadGEM2-ES	Cloud ice water content (both cloud ice and snow);			
Inmcm4	Mixing ratio of cloud liquid and ice			

Table 2. (continued)

Models	Prognostic Cloud Variables	Shortwave Radiation Scheme	Longwave Radiation Scheme	Note
Inmcm4ESM	Mixing ratio of cloud liquid and ice			
IPSL-CM5A-LR	Single mixing ratio of total water			
MIROC (MIROC-ESM; MIROC-ESM-CHEM)	Mixing ratio of cloud liquid and ice			
MIROC4h	Mixing ratio of cloud liquid and ice			
MIROC5	Mixing ratio of cloud liquid and ice			
MRI-CGCM3	Mixing ratio of cloud liquid and ice			
NorESM1	Single mixing ratio of total water			
MPI	Mixing ratio of cloud liquid and ice			
GFDL-CM3	Mixing ratio of cloud liquid and ice			
CESM1-CAM5	Number concentrations (N_c , N_i) and mixing ratios (q_c , q_i) of cloud droplets (subscript c) and cloud ice (subscript i). Diagnostic two-moment treatment of rain and snow			

Edition 3-lite SYN1deg-Month provides computed irradiances to be adjusted, and EBAF Ed2.6r [Loeb *et al.*, 2009, 2012] provides the constraint. In addition, temperature and humidity profiles used in the computations are from the Goddard Earth Observing System (EOS) Data Assimilation System reanalysis (GEOS-4 and GEOS-5). MODIS-derived cloud properties [Minnis *et al.*, 2011] are combined with geostationary satellite (GEO)-derived cloud properties [Minnis *et al.*, 1994] to resolve the diurnal cycle used in the SYN1deg-Month flux computations. Note that unlike TOA irradiances, the global estimate of irradiance at the surface is only possible by using a radiative transfer model. The errors in cloud and atmospheric properties used as inputs therefore directly affect the accuracy and stability of modeled surface irradiances [Kato *et al.*, 2012a]. In addition, model-computed TOA irradiances do not necessarily agree with CERES-derived observed TOA irradiances. Therefore, to mitigate these problems, computed fluxes in the EBAF-Surface product are constrained to be consistent with CERES-derived observed TOA fluxes to within their uncertainties. In the constraining process, CloudSat radar-derived and CALIPSO lidar-derived cloud vertical profiles [Kato *et al.*, 2010] as well as Atmospheric Infrared Sounder (AIRS)-derived temperature and humidity profiles are used to determine the uncertainty in cloud and atmospheric properties. The fluxes from EBAF-Surface have been extensively compared to surface observations taken at a number of sites. Kato *et al.* [2012b] showed that the bias (root-mean-square difference) between computed and observed monthly mean irradiances calculated with 10 years of data is 4.7 (13.3) W m^{-2} for downward shortwave and -2.5 (7.1) W m^{-2} for downward longwave over ocean and -1.7 (7.8) W m^{-2} for downward shortwave and -1.0 (7.6) W m^{-2} for downward longwave irradiances over land. Both the downward shortwave and longwave flux differences are within the monthly modeled gridded flux uncertainty of 10 W m^{-2} for shortwave and 14 W m^{-2} for longwave estimated by Kato *et al.* [2012a]. In addition, these differences are also within the

uncertainty of surface measurements at a given site; measurement uncertainty at a buoy reported by Colbo and Weller [2009] is $5\text{--}6 \text{ W m}^{-2}$ for daily or annual mean downward shortwave and 4 W m^{-2} for daily and annual mean downward longwave flux [cf. Waliser *et al.*, 1999; Medovaya *et al.*, 2002]. The data used in this study are the monthly mean collected from January 2000 to December 2010.

[9] 1.2 ISCCP-FD. These products are calculated from an advanced radiation scheme using ISCCP-D1 input data that include global observations of the key variables, including better treatment of ice clouds, revision of aerosol climatology, diurnal variance of surface skin and air temperatures, revision of the water vapor profiles, and refinement of the land surface albedos and emissivities [Zhang *et al.*, 2004]. An 18 year flux record has been created at 3 h time steps at 280 km intervals globally based on comparisons of monthly and regional mean values to the ERBE and the CERES TOA fluxes and to BSRN surface fluxes. Zhang *et al.* [2004] reported that the overall uncertainties are $10\text{--}20 \text{ W m}^{-2}$ for shortwave fluxes. Comparisons to BSRN also suggest that biases in monthly shortwave fluxes are less than 5 W m^{-2} . The time period of the data used in this study is from January 1984 to December 2004.

[10] The A-Train fluxes derived from EBAF-Surface are especially noteworthy because unlike the other estimates such as ISCCP-FD, the flux values quoted are based on actual cloud profile observations, notably including the critical new information about cloud base derived from CloudSat and CALIPSO [Mace *et al.*, 2009]. In this study, the RSDS flux product used as a reference for direct model evaluation is EBAF-Surface, which has been adjusted and constrained with CloudSat and CALIPSO cloud-derived vertical profiles, AIRS-derived temperature/humidity profiles, and CERES-derived TOA fluxes to within their uncertainties. The ISCCP-FD values are used for an additional measure of observational uncertainty for the model-observation comparisons.

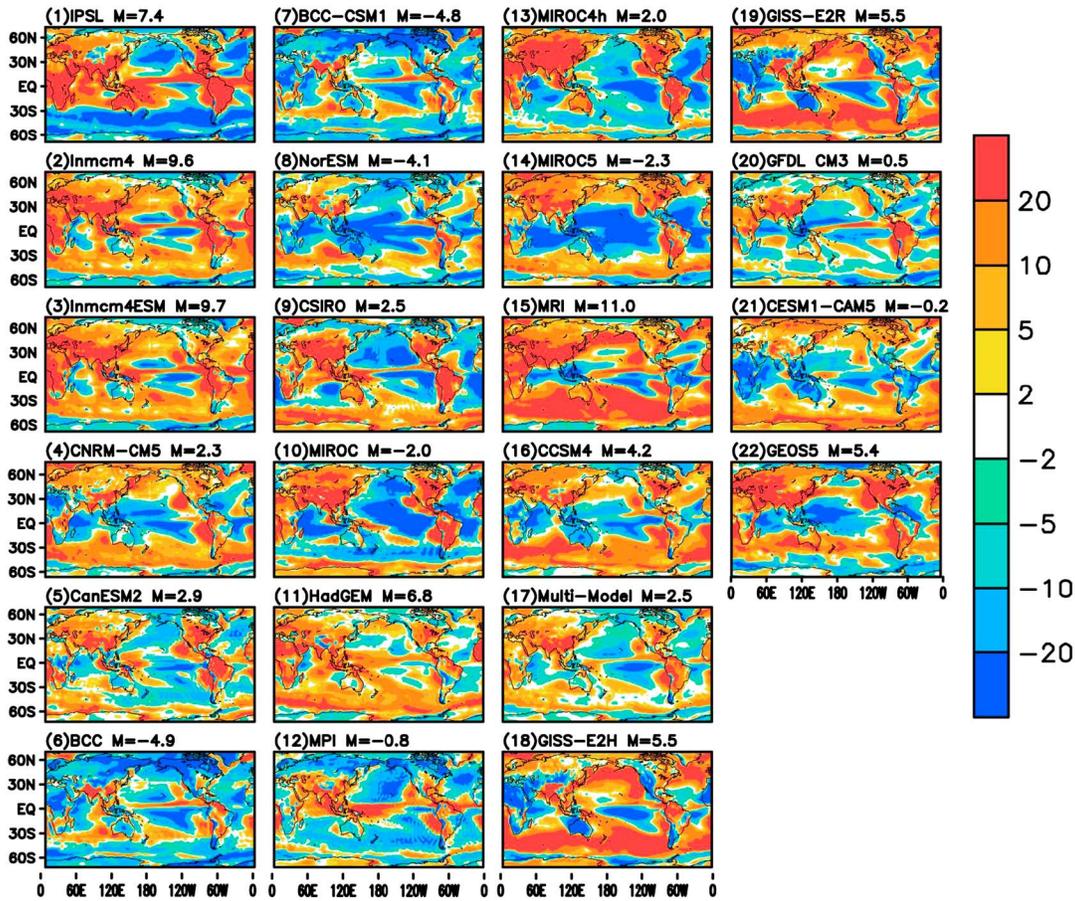


Figure 4. (1)–(16) Bias of the annual mean radiative shortwave downward at the surface (RSDS; W m^{-2}) from the twentieth century GCM simulations contributed to CMIP5, with their multimodel mean bias shown in (17); (18) and (19), which are the two GISS models; (20) GFDL CM3; (21) NCAR CESM1-CAM5; and (22) NASA GEOS5 AGCM against the EBAF-Surface estimate shown in Figure 1, model 1 (see section 2).

[11] 2. Radiative longwave upward at TOA (RLUT) and radiative shortwave wave upward at TOA (RSUT) fluxes. Emitted longwave and reflected solar radiation at the TOA, referred to in this study as RLUT and RSUT, respectively, are from the CERES EBAF product (CERES EBAF-TOA_Ed2.6r) [Loeb *et al.*, 2012, 2009]. The CERES EBAF product includes the latest instrument calibration improvements, algorithm enhancements, and other updates. CERES TOA SW and LW fluxes in the EBAF product are adjusted within their range of uncertainty to remove the inconsistency between average global net TOA flux and heat storage in the Earth-atmosphere system, as determined primarily from ocean heat content anomaly (OHCA) data (see Supplementary Information in Loeb *et al.* [2012] for more details). The data used in this study is the monthly mean collected from January 2000 to December 2010.

[12] In this study, the RLUT and RSUT fluxes used for direct model evaluation are CERES EBAF, which is directly measured and adjusted/balanced with global energy.

[13] Figure 1 shows the annual mean maps of RSDS estimated from EBAF-Surface (1) and ISCCP (2) and of the difference between ISCCP and EBAF-Surface (3). ISCCP RSDS is highly biased relative to EBAF-Surface over most of the globe (global average 2.6 W m^{-2} ; Figure 1, model

3), except for few subtropical regions, where stratocumulus and trade wind shallow cumulus clouds dominate, and it is lowly biased. Figure 2 shows the annual mean maps of RLUT (W m^{-2}) from CERES EBAF. Figure 3 shows the annual mean map of RSUT from CERES EBAF. For more details on the observed radiation reference data sets and their uncertainties, readers are referred to Kato *et al.* [2012a, 2012b] for RSDS as well as to Loeb *et al.* [2012] for RSUT and RLUT.

3. Modeled Values of Radiative Fluxes

[14] Using the observations described in the previous section, we evaluate RSDS, RLUT, and RSUT in coupled atmosphere-ocean GCMs (CGCMs) from CMIP3, from CMIP5, and one additional GCM, the NASA GEOS5 AGCM. The CMIP5 simulations are listed in Table 1. Table 2 is an outline of the SW/LW radiation parameterizations used in the selected CMIP5 models and the GEOS5 model. While the GEOS5 model has land interactions and complex interactions within the atmosphere (clouds, boundary layer, dynamics, etc.), the CMIP5 models, in addition, have ocean interaction. Thus, the performance of their radiation simulations is not solely determined by the

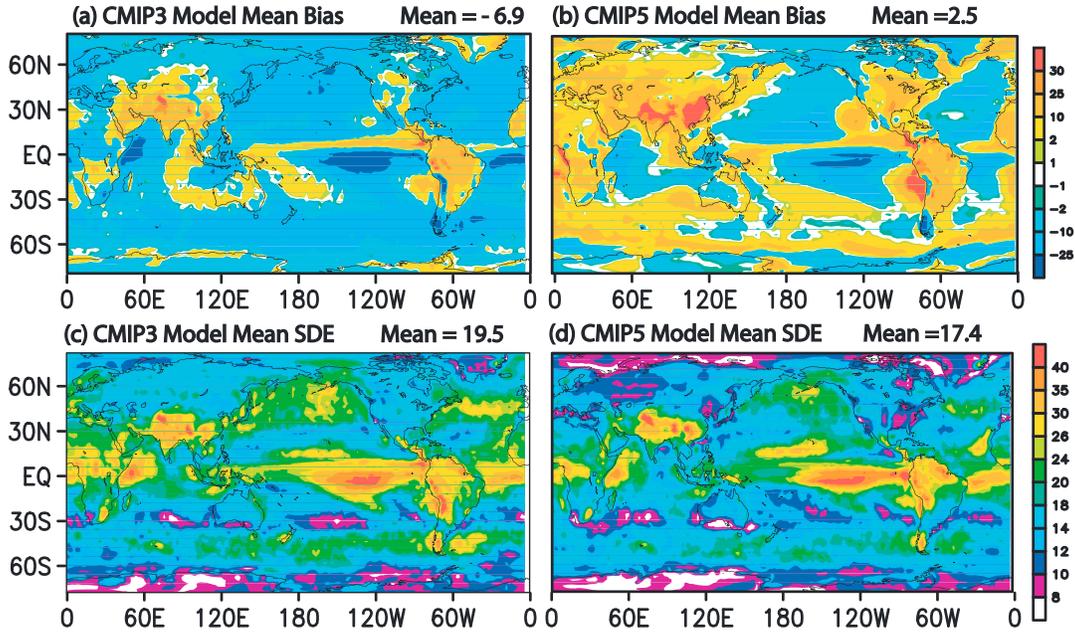


Figure 5. (a) Multimodel mean bias of RSDS (W m^{-2}) for the CMIP3 GCMs shown in Figure A1. (b) Multimodel mean bias for the CMIP5 GCMs in Figure 4. (c) Same as Figure 5a except for the standard deviation of the error (SDE). (d) Same as Figure 5b except for the SDE for CMIP5. In all cases, the observed reference is shown in Figure 1, model 1. Note that the mean bias has been removed in the calculation of the SDE.

radiation schemes used but rather through complex interactions arising from a fully coupled system. Therefore, the behavior is not likely to be simply explained by any single component or scheme, but rather by details of the model's specific schemes and the coupling among schemes as well as the interactions with sea surface temperatures (SSTs) [e.g., Donner *et al.*, 2011; Ma *et al.*, 2012; Gettleman *et al.*, 2010; Li *et al.*, 2012; Randall *et al.*, 2007; Webb *et al.*, 2001]. The specific experimental scenario in CMIP5 is the historical twentieth century simulation, which used observed twentieth century greenhouse gas, ozone, aerosol, and solar forcing. The time period used for the long-term mean is 1970–2005, and if a model provided an ensemble of simulations, only one of them was chosen for this evaluation. For both the GCM and observational data sets, all fields have been regridded and mapped onto common $2^\circ \times 2^\circ$ latitude-by-longitude grids.

4. Results

[15] There are 20 CMIP5 models used in this study, which are listed in Tables 1 and 2. Among them, 16 are conventional models, i.e., no radiative convective core cloud mass and falling hydrometeors included [Li *et al.*, 2012]. The GFDL-CM3, CESM1-CAM5, and two GISS models, which include radiative falling cloud hydrometeors, are not considered in the multimodel mean (Figure 4, model 17). Figure 4 shows the long-term annual mean spatial distributions of simulated bias values of RSDS with EBAF-Surface as a reference from 16 conventional CMIP5 CGCMs (see Table 1, models 1–16), with the multimodel ensemble mean bias from these 16 CMIP5 models (17), three CGCMs which include precipitation-radiation interactions (18–21), and the

GEOS5 atmosphere-only model (22). The magnitudes of the RSDS bias can be quite different regionally and across models, ranging from -25 to more than 30 W m^{-2} . Except for CNRM (4), NorESM (8), MIROC (10), MIROC5 (14), and CCSM4 (16), most of the models and the CMIP5 multimodel mean bias values show overestimated RSDS fluxes in strongly convective regions of the tropics (e.g., ITCZ and South Pacific Convergence Zone (SPCZ), Indian Monsoon, and warm pool), in the storm tracks and the Southern Ocean, and over most of landmasses, especially the convectively active continental areas such as central Africa, South America, and the Antarctic. A number of models, i.e., IPSL (1), Inmcm4 (2), Inmcm4ESM (3), HadGEM (11), MRI (15), and GISS-E2H (18), and GISS-E2R (19), significantly overestimate RSDS globally (up to 30 W m^{-2}), except the subtropical regions over the oceans. On the other hand, the two BCC models (6, 7) and NorESM (8) underestimate RSDS globally, except the convective regions over the ITCZ and tropical landmasses. The GEOS5 AGCM (22) significantly underestimates RSDS (by more than -25 W m^{-2}) in the warm pool and the Indian Ocean. While the above model-observation differences in RSDS bias are substantial in many regards, it is worth noting that the CMIP5 multiensemble RSDS bias values (17) appear to exhibit improvement over the CMIP3 ensemble models (Figure A1, model 13); this will be discussed and quantified further below. The two GISS models (19) and (20) and the GFDL-CM3 model (21), as well as NCAR CESM1-CAM5 (21), include the impacts from either diagnostic precipitating grid mean mass or convective core mass on radiation. The GFDL models further include shallow cumulus, deep cumulus cells, and convective mesoscale clouds, weighted by their respective area fractions for radiation calculation. However, in the GFDL-

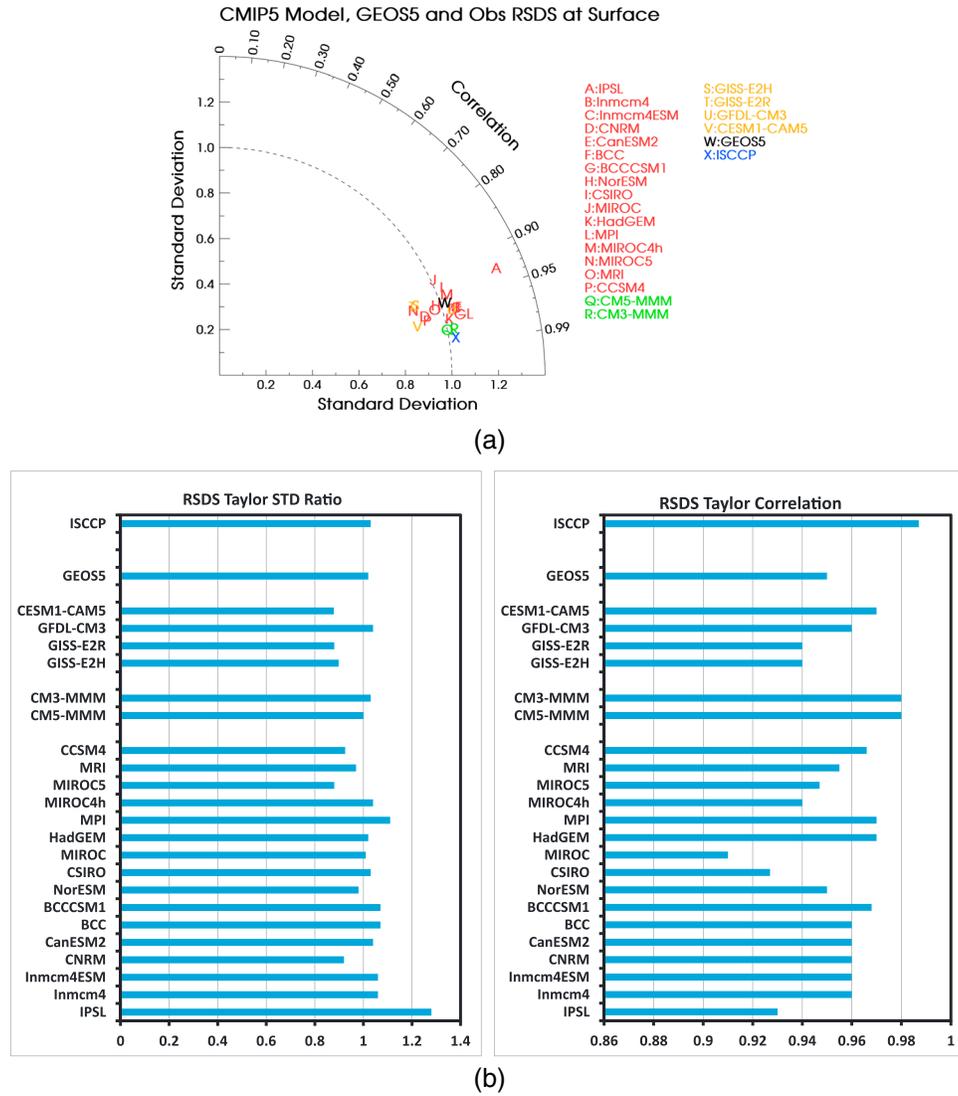


Figure 6. (a) Taylor diagram for the GCM annual mean (80°N – 80°S) radiative shortwave downward flux at the surface (RSDS; W m^{-2}), where the reference value is the observed estimate from the EBAF-Surface shown in Figure 1, model 1. ISCCP-FD is also included. The ratio of the standard deviation (x and y axes) exhibits the relative amplitude of the simulated and the “observation” variations, whereas the correlation (angle) indicates the degree of similarity of variation between the two. The “centered” root-mean-square error between the reanalysis/simulations and “reference,” i.e., “observation”, is proportional to the distance between the points for the given data set and the reference data set. (b) The CMIP5 RSDS correlation and standard deviation ratios against Obs Est. EBAF-Surface (Figure 1, model 1) plotted in the form of a bar chart for GCM annual mean (80°N – 80°S) RSDS (W m^{-2}), where the reference value is the observed estimate EBAF-Surface shown in Figure 1, model 1, while ISCCP-FD is used as another reference bound. Modeled values include the CMIP5 models, multimodel mean CMIP3 GCMs (CM3-MMM), multimodel mean CMIP5 GCMs (CM5-MMM), as well as GEOS5 GCM. The MERRA data are not included. The vertical lines represent the values of these quantities associated with observational uncertainty, which provide quantitative levels for indicating good model-observation agreement in variability.

CM3 model, precipitating ice that has fallen out of large-scale stratiform clouds and into clear areas is not included [Donner et al., 2011]. The GFDL-CM3 model exhibits relatively good RSDS in the tropical ITCZ and convectively active continental regions but still underestimates RSDS in the warm pool and continental regions. The NCAR CESM1-CAM5 model (21), on the other hand, includes diagnostic radiative snow, shows similar patterns to that in

CCSM4 (16) but with less bias over the Southern Ocean, and still underestimates RSDS in the warm pool and continental regions.

[16] To summarize the multimodel performance of CMIP3 and CMIP5 models in representing the time-mean pattern of RSDS, Figure 5 illustrates the multimodel mean biases (top) and the multimodel mean standard deviation of the error, which is defined as a root-mean-square error

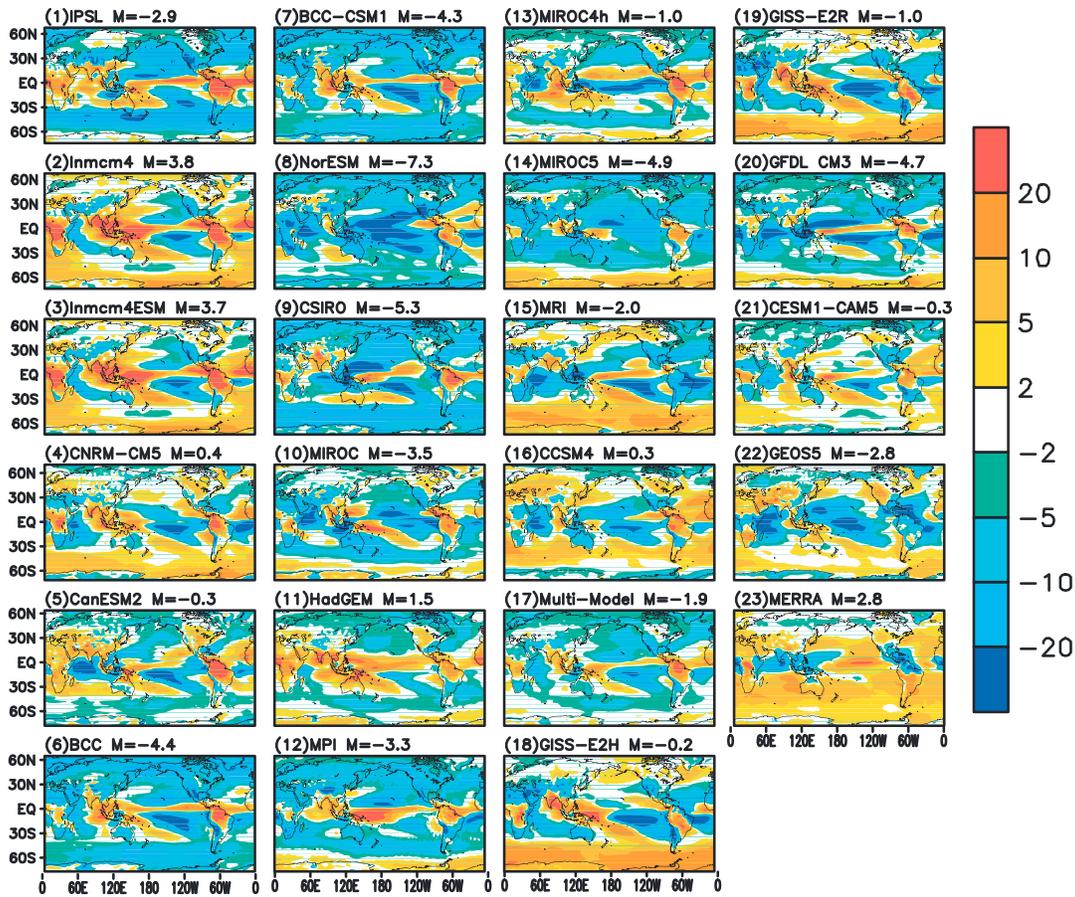


Figure 7. (1)–(16) Bias of the annual mean maps of radiative longwave upward flux at TOA (RLUT; W m^{-2}) from the twentieth century GCM simulations contributed to CMIP5, with their multimodel mean shown in (17); (18) and (19), which are the two GISS models; (20) GFDL CM3; (21) NCAR CESM1-CAM5; (22) NASA GEOS5 2.5 AMIP; and (23) MERRA against the CERES EBAF observed estimates shown in Figure 2, model 1 (see section 2).

with the mean bias removed (hereafter SDE) (bottom) against the observed estimate, which is calculated across the models for each of the ensembles. Overall, CMIP3 RSDS (Figure 5a) shows low biases globally and more uniformly, except the ITCZ, off the coast of the Peru/California regions and over Indian Monsoon regions. While the CMIP5 global area average bias (Figure 5b) is reduced ($\sim 30\%$) to the value of 2.5 W m^{-2} from the value of -6.9 W m^{-2} in CMIP3, it exhibits more distinct spatial gradients with greater local extreme biases and a higher bias over most land regions. While the bias figure emphasizes the sign and magnitude of systematic biases across the two model archives, the SDE figure emphasizes errors irrespective of the sign. The similar pattern of the SDE between the two model ensembles indicates that CMIP3 and CMIP5 share many of the same systematic errors in simulating radiation fluxes. The fact that the magnitude of the SDE is about the same between CMIP3 and CMIP5 indicates that little improvement from CMIP3 has been afforded. The high SDE values in the equatorial regions of the Pacific and Atlantic (Figures 5c and 5d) combined with the low bias in the same regions (Figures 5a and 5b) indicate that significant disparity in the manner models represent surface radiation in these regions. This is similarly true over the mountainous regions

of Asia and South America and, to a lesser extent, over the storm tracks.

[17] To further quantify and synthesize the comparative information discussed above, we use a Taylor diagram [Taylor, 2001] with the CERES EBAF-Surface in Figure 1, model 1, used as the reference data set and the ISCCP-derived RSDS estimates used to help quantify observational uncertainty. The Taylor diagram used in this study relates two statistical measures of model fidelity: the spatial correlation and the spatial standard deviations [Taylor, 2001]. These statistics are calculated for the long-term time mean and over the global domain (area weighted). The reference data set is plotted along the x axis at the value of 1.0. The radial distance from the origin is proportional to the ratio of the standard deviations of the given data set relative to the reference data set. The azimuthal angle represents the spatial correlation between the given data set and the reference data set. The ratio of the standard deviation exhibits the relative amplitude of the simulated and the “reference” variations, whereas the correlation indicates the degree of similarity of variation between the two.

[18] Figure 6a shows the Taylor diagrams for CMIP3 and CMIP5 annual mean RSDS with the same information shown as bar charts in Figure 6b. Not surprisingly, the

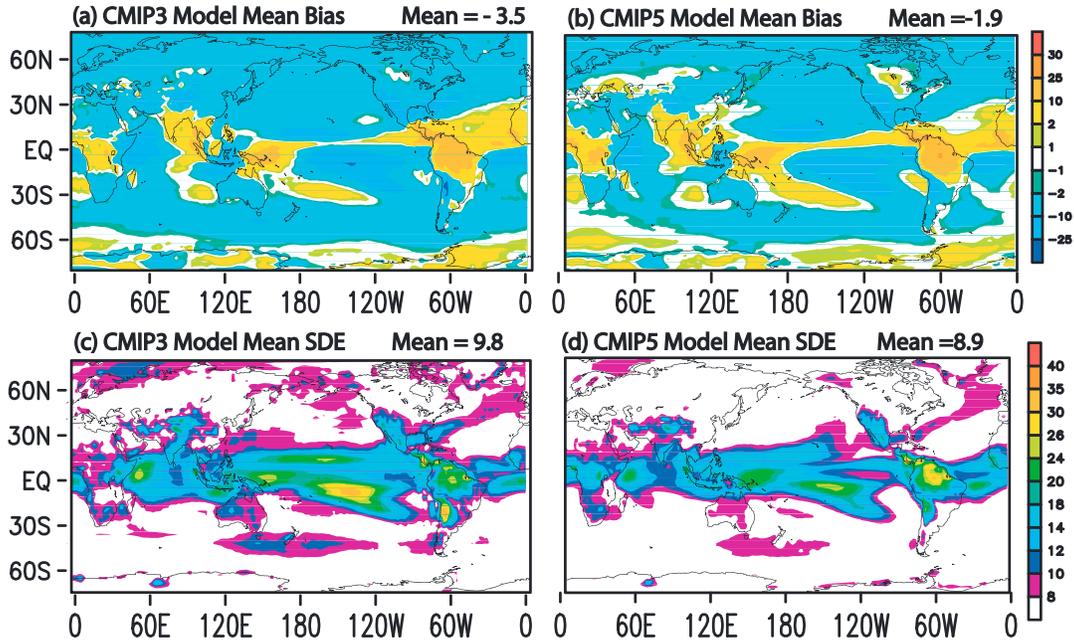


Figure 8. (a) Multimodel mean bias of RLUT (W m^{-2}) for the CMIP3 GCMs shown in Figure A2. (b) Multimodel mean bias for the CMIP5 GCMs in Figure 7. (c) Same as Figure 8a except for the standard deviation of the error (SDE). (d) Same as Figure 8b except for the SDE. In all cases, the observed reference is shown in Figure 2, model 1. Note that the mean bias has been removed in the calculation of the SDE.

ISCCP (23) observed estimate shows the best agreement with the EBAF-Surface values, having a correlation of about 0.98 or better and standard deviation ratios close to 1.0. For the CMIP values, most of them have very good correlations between about 0.91 and 0.97 with standard deviation ratios between 0.9 (MIROC5) to 1.3 (IPSL). Note that none of the individual CMIP5 models fall within the range of the observational uncertainty, represented here by the ISCCP (23) value, although the CMIP3 and CMIP5 multimodel means are very close. Also notable and consistent with the discussion above, there is very little difference in the overall performance of the CMIP3 and CMIP5 multimodel mean values of RSDS. In regard to specific models, the best performing CMIP5 model by this metric is HadGEM (11), although most of the models fare relatively well according to Figure 6. One significant outlier is IPSL (1) with a standard deviation ratio of about 1.3 and a somewhat poor correlation value, relative to the others, of 0.93.

[19] Next, we examine RLUT. Figure 7 shows the long-term annual mean spatial distributions of bias values of RLUT against CERES EBAF (Figure 2, model 1) from 16 CMIP5 CGCMs (see Table 1, models 1–16), with the multimodel ensemble mean bias from the 16 CMIP5 models (17), and GEOS5 (22). The relative magnitudes of bias can be quite different across models and regionally, ranging from -25 to about 30 W m^{-2} . Similar to RSDS, most of the models and the multimodel mean bias CMIP5 (17) values show overestimated RLUT in many of the strongly convective region of the tropics such as the warm pool, ITCZ/SPCZ, and convectively active continental regions. Inmcm4 (2), Inmcm4ESM (3), CNRM (4), MRI (15), and CCSM4 (16) overestimate RLUT over southern middle/high latitudes and Southern and Antarctic Oceans. The GEOS5

AGCM model (22) underestimates (more than 25 W m^{-2}) RLUT in the subsiding regions of the tropical Pacific and Indian Oceans, while it overestimates RLUT in much of the Southern and Antarctic Oceans. MERRA (22) exhibits an overall high RLUT bias, except in the warm pool region. The ensemble mean (17) of the CMIP5 RLUT bias values appears to exhibit some degree of improvement compared to the ensemble mean of CMIP3 models (Figure A2); this will be discussed and quantified further below. The GFDL CGCM model (20) again does relatively well over most of the globe compared to the other models. It exhibits a relatively good spatial distribution of RLUT in the tropical ITCZ and convectively active continental regions but moderately underestimates RLUT in most of the oceans compared to the observed RLUT. The CESM1-CAM5 model (21) exhibits the best performance in representing RLUT over most of the globe compared to the other models with bias within -5 to 5 W m^{-2} . It exhibits a relatively good spatial distribution of RLUT in the Southern Ocean and convectively active continental regions but moderately overestimates RLUT in the warm pool to the observed RLUT. Note that most of the CMIP5 models simulate cloud top heights that are highly biased compared to the cloud top height values estimate from CloudSat/CALIPSO IWC, suggesting that the bias to the RLUT is not from the underestimation of cloud top height.

[20] To summarize the multimodel performance of CMIP3 and CMIP5 in representing the time-mean pattern of RLUT, Figure 8 illustrates the multimodel mean bias (top) and the multimodel mean SDE (bottom) against the observed estimate. Interestingly, both CMIP3 (Figure 8a) and CMIP5 (Figure 8b) exhibit very similar RLUT bias patterns and magnitudes. Notable is the consistency of having a high

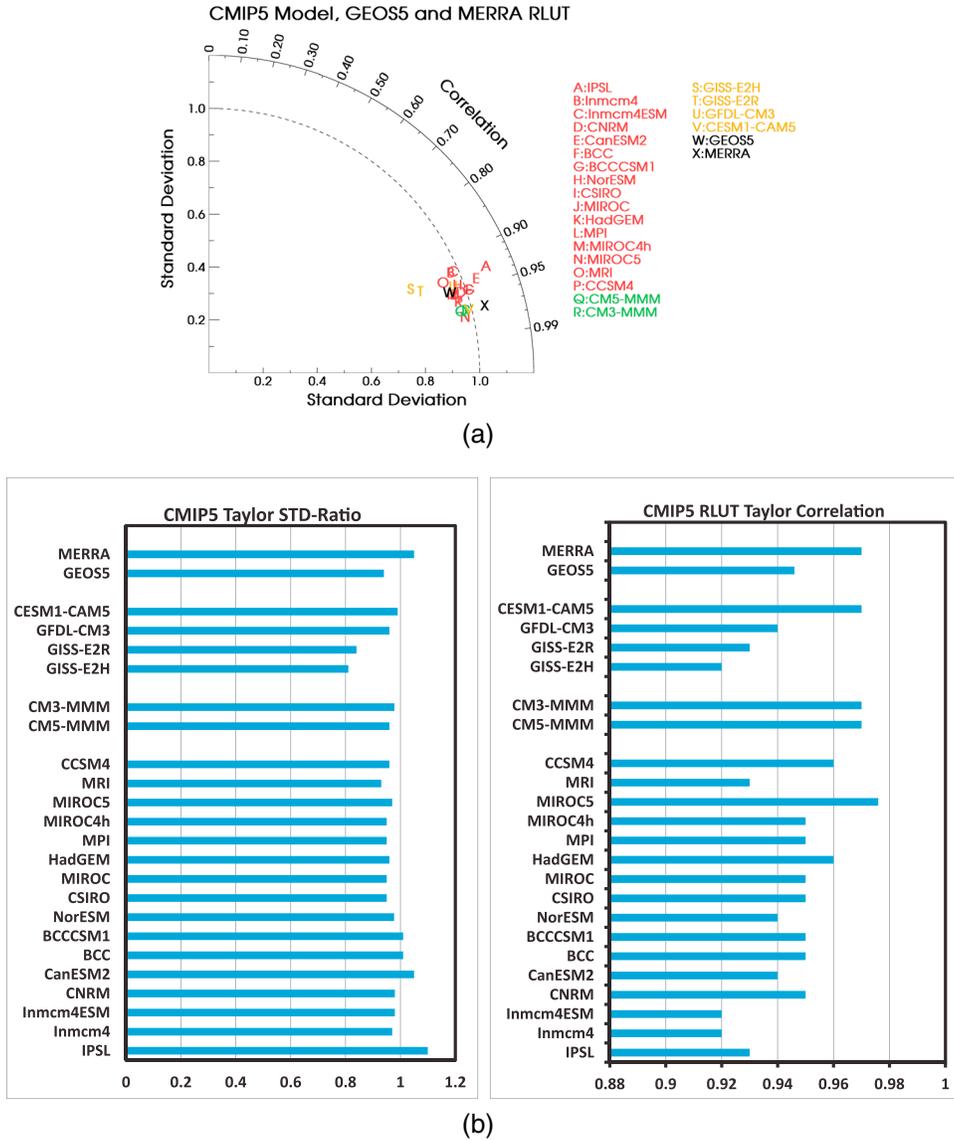


Figure 9. (a) Taylor diagram for the GCM annual mean (80°N–80°S) radiative longwave upward flux at TOA (RLUT; $W m^{-2}$), where the reference value is the observed estimate from CERES EBAF shown in Figure 2, model 1. (b) The CMIP5 RLUT correlation and standard deviation ratios against CERES EBAF plotted in the form of a bar chart for GCM annual mean (80°N–80°S) RLUT ($W m^{-2}$), where the reference value is the observed estimate shown in Figure 3, model 1. Modeled values include the CMIP5 models, multimodel mean CMIP3 GCMs (CM3-MMM), multimodel mean CMIP5 GCMs (CM5-MMM), as well as GEOS5 GCM and MERRA.

bias over the ITCZ, the SPCZ, a part of the Southern Ocean and tropical continents, and the Indian Monsoon regions. The CMIP5 global area average RLUT bias value of $-1.9 W m^{-2}$ is about a factor of 2 smaller compared to the CMIP3 value of $-3.9 W m^{-2}$. The SDE figure indicates a similar pattern of systematic errors in the tropics, with no substantial change in the global mean SDE, CMIP3 ($9.8 W m^{-2}$), and CMIP5 ($8.9 W m^{-2}$). The bias and SDE together indicate that the models are performing relatively well in the middle and high latitudes but exhibit significant shortcomings in the tropics. The presence of bias and SDE errors suggests that some aspects of the errors are arising from the same cause (i.e., bias) and some aspects result from different causes (i.e., SDE). Given that all but one of these

models are coupled, it is possible that some part of the SDE errors could result from spatial variations in the location of the ITCZ and the associated cloud structure which have a substantive impact on the RLUT field.

[21] Figure 9a shows a Taylor diagram for CMIP3 and CMIP5 annual mean RLUT, while the same information is shown as bar charts in Figure 9b. The reference value used for this Taylor diagram is the observation estimate from CERES EBAF shown in Figure 2, model 1. The CMIP5 models generally have correlations of 0.9 or better and standard deviation ratios ranging from 0.8 to 1.15. The two GISS models and IPSL represent the greatest outliers within the group. The CESM1-CAM5 model (21) exhibits the best correlations (0.97) and the best standard deviation ratio

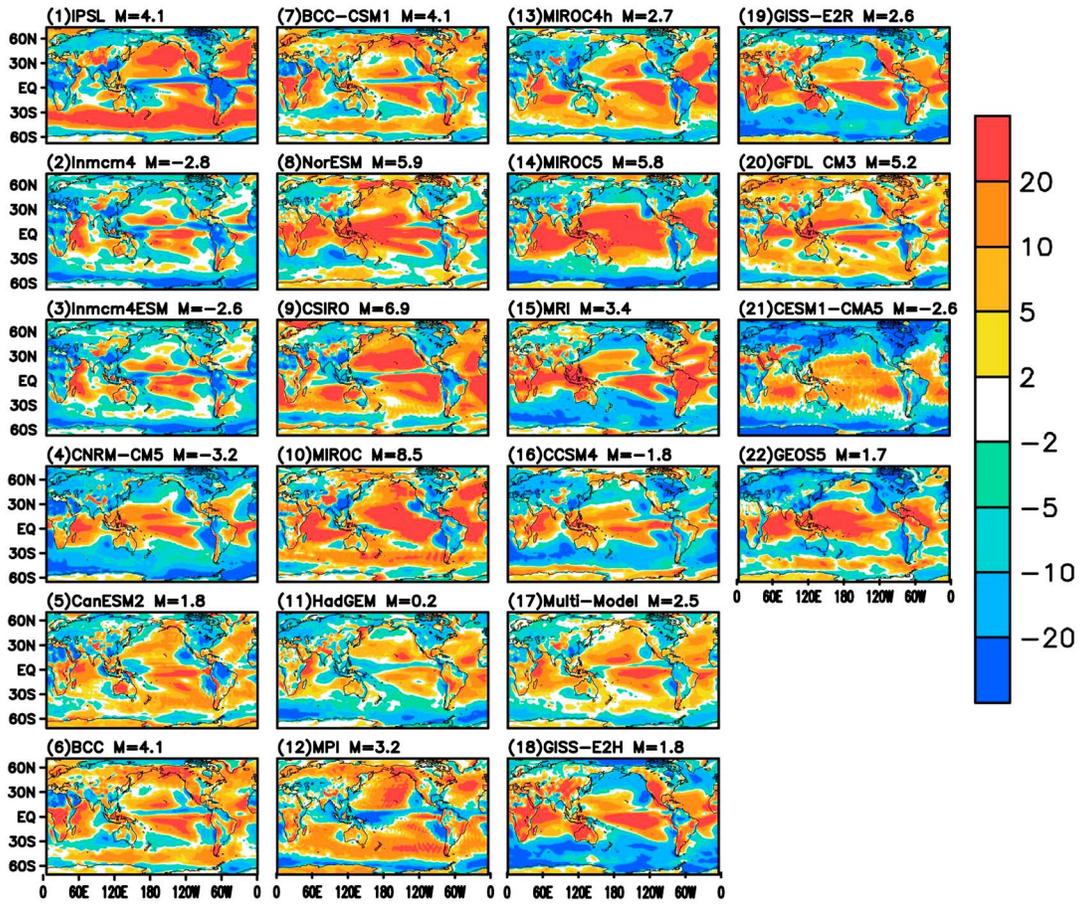


Figure 10. (1)–(16) Bias of the annual mean maps of radiative shortwave upward flux at TOA (RSUT; W m^{-2}) from the twentieth century GCM simulations contributed to CMIP5, with their multimodel mean shown in (17); (18) and (19), which are the two GISS models; (20) GFDL CM3; (21) NCAR CESM1-CAM5; and (22) NASA GEOS5 2.5 AMIP against the CERES EBAF observed estimates shown in Figure 3, model 1 (see section 2).

(0.99) among the models. Similar to RSDS, the CMIP5 (18) multimodel mean does not exhibit any significant improvement over the CMIP3 mean (17) in representing RLUT.

[22] Finally, we examine RSUT. Figure 10 shows the long-term annual mean spatial distributions of simulated bias values of RSUT from 16 CMIP5 CGCMs (see Table 1, models 1–16), with the multimodel ensemble mean bias from the 16 CMIP5 models (17), three CGCMs which include precipitation-radiation interactions (18–21), and the GEOS5 atmosphere-only model (22). Apparent is that most of the CMIP5 models tend to reflect too much shortwave radiation to space in the tropics and subtropics and too little in the middle to high latitudes. This indicates either clouds that occur too frequently or clouds that are overly reflective in terms of cloud cover and/or microphysical characteristics. Notable, however, is that even within this overall high bias in the tropics, there is a systematic underestimation of RSUT in the strongly convective region of the tropics such as the ITCZ/SPCZ, the Indian Monsoon, and over the land surfaces. For example, the relative minimum in the ITCZ region is evident in the multimodel mean and in most individual models, except, for example, for CanESM and MIROC5. The CMIP5 models also tend to be lowly biased over many of the subtropical stratocumulus regimes (e.g., northeastern and southeastern Pacific against North America and South

America). Both of these latter features imply too few clouds or insufficiently reflecting clouds in these regions. Similar to RSDS and RLUT, the ensemble mean performance of the CMIP5 RSUT (16) bias patterns appears to be very similar compared to the ensemble mean CMIP3 models (Figure A3, model 13). Note that despite GEOS5 having the advantage of specified SSTs and that the GFDL model (20) and the two GISS models (17, 18) consider total LWC/IWC in the radiation calculations, none of them exhibit much of an improvement relative to the other CGCMs examined here, while the CESM1-CAM5 model (21) shows overestimation of RSUT in subtropical and tropical oceans with underestimation of RSUT elsewhere.

[23] Figure 11 illustrates the multimodel mean biases (top) and the SDE (bottom) against the observed estimate. Similar to RSDS and RLUT, the RSUT SDE and bias in both CMIP3 and CMIP5 exhibit very similar patterns, and there is a clear systematic underestimation over the ITCZ, the SPCZ, and a part of the Southern Ocean and tropical continents, and the Indian Monsoon regions. While the CMIP5 global area average RSUT bias is about a factor of 2 smaller (2.5 W m^{-2}) compared to the CMIP3 value of 4.5 W m^{-2} , they share very similar pattern distributions in bias and SDE. The SDE figure indicates a slight improvement of CMIP5 (14.1 W m^{-2}) over CMIP3 (14.7 W m^{-2}).

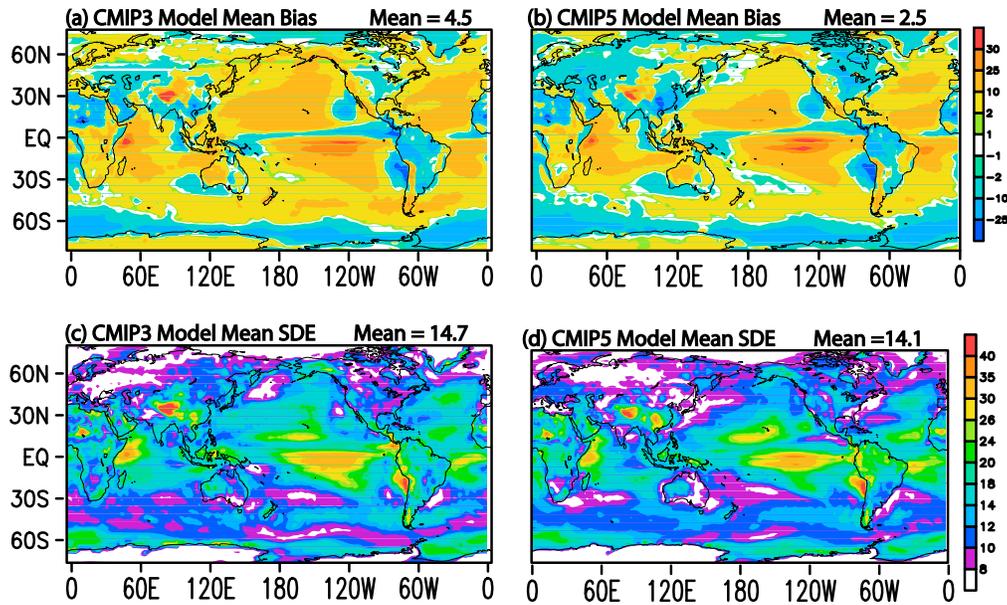


Figure 11. (a) Multimodel mean bias of RSUT (W m^{-2}) for the MIP3 GCMs shown in Figure A3. (b) Multimodel mean bias for the CMIP5 GCMs in Figure 10. (c) Same as Figure 11a, except for the SDE. (d) Same as Figure 11b, except for the SDE. In all cases, the observed reference is shown in Figure 3, model 1.

[24] A Taylor diagram for CMIP3 and CMIP5 annual mean RSUT is shown in Figure 12a, while the same information is shown as bar charts in Figure 12b. Of the three quantities shown, i.e., RSDS, RLUT, and RSUT, it is clear that the overall performance for RSUT is considerably worse in regard to the Taylor diagram metrics (i.e., compare to Figures 6 and 9). For the CMIP5 values, most of them have poor correlations between about 0.4 and 0.8 with standard deviation ratios ranging from 0.8 to 1.1. The CMIP3 (16) and CMIP5 (17) multimodel means do not show much difference in representing RSUT. In regard to specific models, the best performing CMIP5 model by this metric is HadGEM (11) and a group of models, i.e., CanESM2 (5), BCC (6), BCCesm (7), and NorESM (8), as the next best performers with standard deviation ratios ranging from 0.9 to 1.0. The GEOS5 model (22) performs poorly relative to the others in this group with the correlation of 0.65 and the ratio of 1.1.

5. Summary and Discussion

[25] The objective of this study is to evaluate the representation of surface and TOA fluxes of atmospheric radiation in GCMs, namely, CMIP GCMs, with the focus on the fluxes most strongly influenced by clouds (i.e., RSDS, RSUT, and RLUT). Apart from a general assessment of the fidelity of the models' representation of radiation, we seek to relate the impacts of ignoring the radiation interaction with precipitating and convective clouds, a common practice in most of the CGCMs contributing to CMIP3 and CMIP5 (see cloud liquid and ice evaluations in *Li et al.* [2012]). Observational reference values and their uncertainties for RSDS are addressed by using two different estimates, each based on satellite plus model retrieval approaches. These include the EBAF-Surface and ISCCP products [see section

2]. For the TOA radiative fluxes of RLUT and RSUT, the CERES EBAF-based estimates are used. In this case, even though both CERES RSDS and ISCCP RSDS are estimates derived from the model, only the CERES data are used; the ISCCP data are used only to give a sense of observational uncertainty but not directly for model-data comparison (e.g., to produce difference maps). The models evaluated (see section 3 and Tables 1 and 2) include 20 simulations of present-day climate available to date in CMIP5 and one GCM of interest (GEOS-5 AGCM). The evaluation of RLUT also includes one modern reanalysis (MERRA). In addition, we include comparative evaluations of 12 of the GCMs from CMIP3, allowing for an assessment of whether the GCM fidelity, in regard to RSDS, RLUT, and RSUT, has improved between CMIP3 and CMIP5.

[26] Overall, there is a fairly wide disparity in the fidelity of RSDS, RLUT, and RSUT representations in the models examined. Even for the annual mean bias maps considered, there are local biases easily as high (low) as 30 (-30) W m^{-2} . Further, as a basis of comparison between the three quantities, it is worth noting the significantly poorer and more varied fidelity measures for the RSUT compared to RSDS and RLUT in the Taylor diagrams (Figures 6a, 9a, and 12a). Considering individual models, some models examined here perform rather well in regard to the Taylor diagram metrics for RSDS, RLUT, and RSUT; these include, for example, HadGEM (RSDS, see Figures 4 and 6; RSUT, see Figures 10 and 12) and MIROC5 (RLUT, see Figures 7 and 9). HadGEM RSDS/RSUT has the best correlations (0.97/0.82) and the second best standard deviation ratio of RSDS (1.02) and RSUT with a ratio above 0.9. Four models, on the other hand, perform particularly poorly in terms of having either a very large standard deviation ratio and/or a small correlation, showing their poor representations of the spatial variability relative to the observation

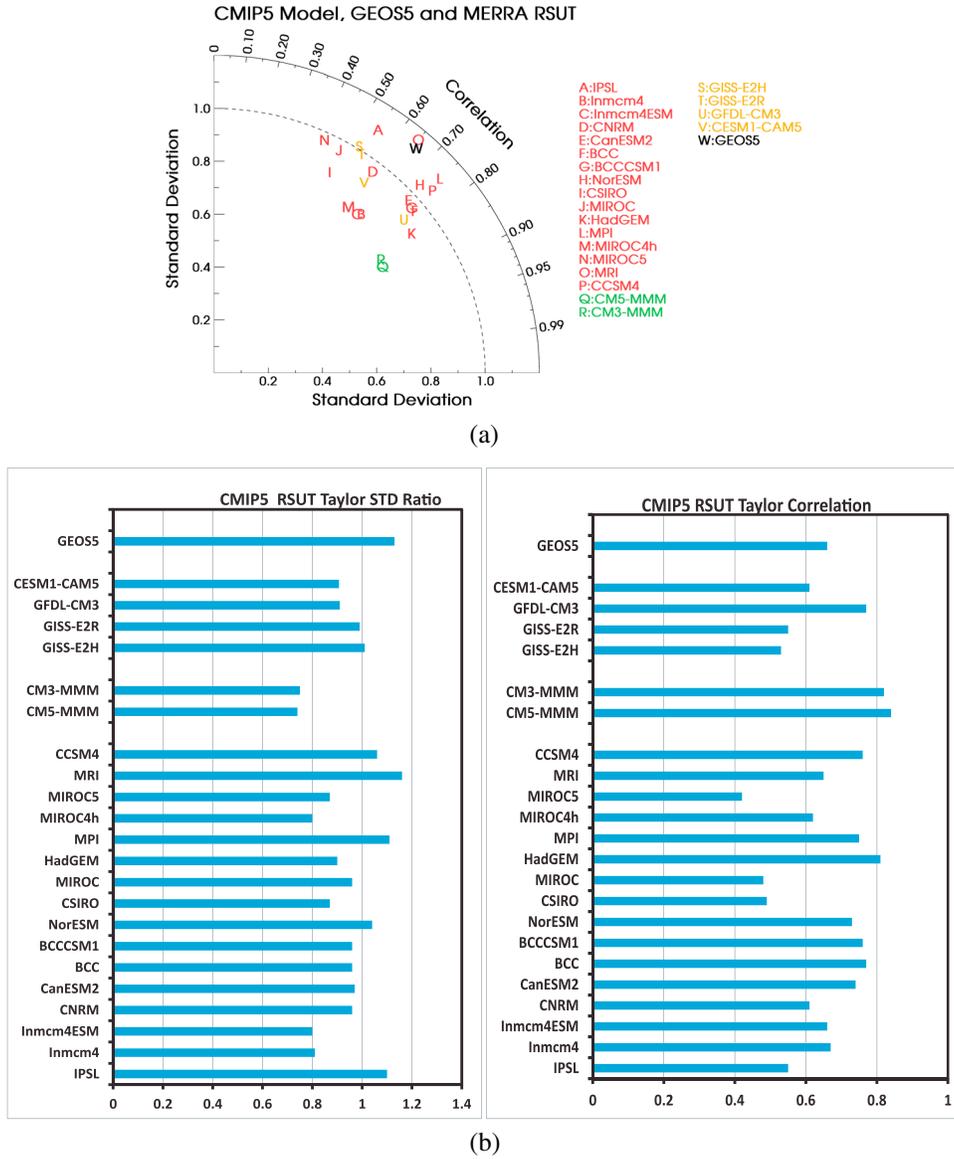


Figure 12. (a) Taylor diagram for the GCM annual mean (80°N–80°S) radiative shortwave upward flux at TOA (RSUT; $W m^{-2}$), where the reference value is the observed estimate shown in Figure 3, model 1. (b) The CMIP5 RSUT correlation and standard deviation ratios against CERES EBAF plotted in the form of a bar chart for GCM annual mean (80°N–80°S) RSUT ($W m^{-2}$), where the reference value is the observed estimate shown in Figure 3, model 1. Modeled values include the CMIP5 models, multimodel mean CMIP3 GCMs (CM3-MMM), multimodel mean CMIP5 GCMs (CM5-MMM), as well as GEOS5 GCM.

reference data (i.e., MIROC, IPSL, INMCM4, and INMCM4-ESM). In comparison, the GEOS5 GCM model was found to perform relatively poorly in comparison to most of the CMIP5 models, for RSDS (Figure 6a), RLUT (Figures 9a), and RSUT (Figure 12a). This is despite the fact that GEOS5 simulation utilized observed SSTs, while the CMIP5 models were coupled ocean-atmosphere models.

[27] One topic of interest in this study is to compare the overall performance of CMIP3 to that of CMIP5. Based on a number of diagnostics, there has been a small degree of improvement in the representation of RSDS, RLUT, and RSDS from CMIP3 to CMIP5. This is demonstrated, in terms of global mean, by the reduction of RSDS (by about 30%), RLUT (about 50%), and RSUT (about 40%)

in the multimodel mean bias. In particular, the multimodel mean bias of RSDS has been reduced from CMIP3 ($-6.9 W m^{-2}$) to CMIP5 ($\sim 2.5 W m^{-2}$). However, this is mostly due to the positive biases over land that became larger, while the negative biases over ocean remained about the same. In addition, an indication of overall improvement in the representation of the quantities studied from CMIP3 to CMIP5 is not evident when considering the SDE computed across the models (i.e., Figures 5, 8, and 11) or in regard to the Taylor diagram metrics (Figures 6, 9, and 12).

[28] Persistent and systematic spatial pattern biases across most of the models with the multimodel ensemble mean values are underestimated in RSUT and overestimated in RSDS and RLUT in the convectively active regions of the

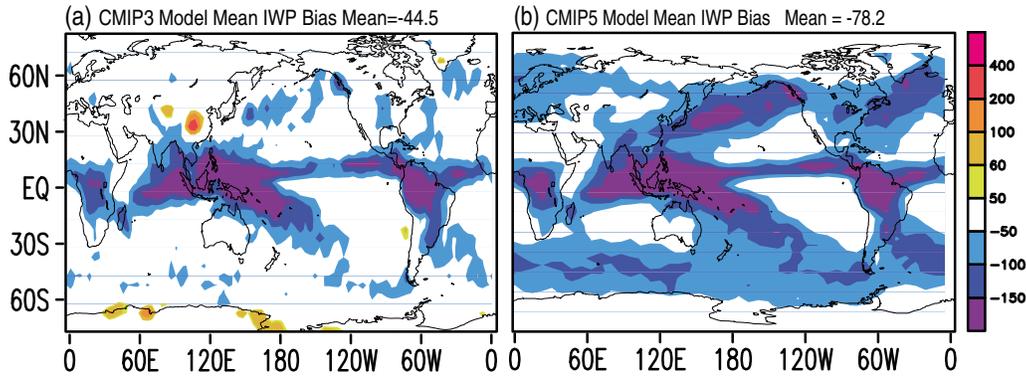


Figure 13. (a) Multimodel mean bias (g m^{-2}) of cloud ice water path (CIWP=cloud IWP) from the CMIP3 GCMs against observed estimates of ice water path under all conditions (TIWP=cloud + convective core + precipitating IWP). (b) Same as Figure 13a but for the multimodel mean bias of CIWP in the CMIP5 GCMs. The observed estimate is from the ensemble mean of the total ice water path (TIWP) using the standard CloudSat product CWC, a version of DARDAR, and a version of 2C-ICE satellite products. For details, refer to *Li et al.* [2012].

tropics (i.e., ITCZ/SPCZ, warm pool, Indian Monsoon, as well as South America and central Africa) (i.e., Figures 5, 8, and 11). Given that a number of these RSDS, RLUT, and RSUT biases occur in conjunction with heavy precipitation and with biases in cloud liquid and ice water biases [*Li et al.*, 2012], we hypothesize that at least a part of these persistent radiation biases stem from GCMs ignoring the effects of precipitating and/or convective core ice and liquid in their radiation calculations [e.g., *Waliser et al.*, 2011].

[29] Given that viable observed estimates of TOA radiation fields and observation-driven modeled values at the surface have been available for many years, and yet the biases are still sizable, suggests challenges to utilizing the observations by the modeling groups or that there are still too many degrees of freedom unconstrained (e.g., cloud cover, cloud mass, particle size, vertical structure, and particle shape). There is certainly evidence of this in regard to cloud liquid and ice content [e.g., *Waliser et al.*, 2009; *Li et al.*, 2012]. In addition, GCMs have typically tuned their radiation and cloud fields to the observations, which naturally are sensitive to all/most hydrometeors in the atmosphere despite the fact that most of the models typically

only represent the suspended hydrometeors associated with clouds, and usually, this does not include ice and liquid in convective cores. Thus, contributions from falling/precipitating hydrometeors are unaccounted for and/or erroneously accounted for by other processes such as interaction with radiation calculation and hydrological cycle.

[30] In order to illustrate the possible shortcomings of GCMs ignoring the cloud mass associated with precipitating hydrometeors (e.g., snow and rain) and convective core ice and liquid mass, Figures 13a and 13b present CMIP3 and CMIP5 multimodel mean biases, respectively, of total ice water path (TIWP = cloud + convective core + precipitating), where the observed cloud ice estimate is from the ensemble mean of three total ice water path observed estimates from the standard CloudSat, a version of DARDAR [*Delanoë and Hogan*, 2008, 2010], and a version of 2C-ICE [*Deng et al.*, 2013] satellite products. For details, see *Li et al.* [2012]. Given that most models do not represent the convective core or precipitating ice, the strong negative bias shown in Figure 13 in the tropics is not surprising, yet it is important to be highlighted, given that this missing water mass will have some interaction with radiation.

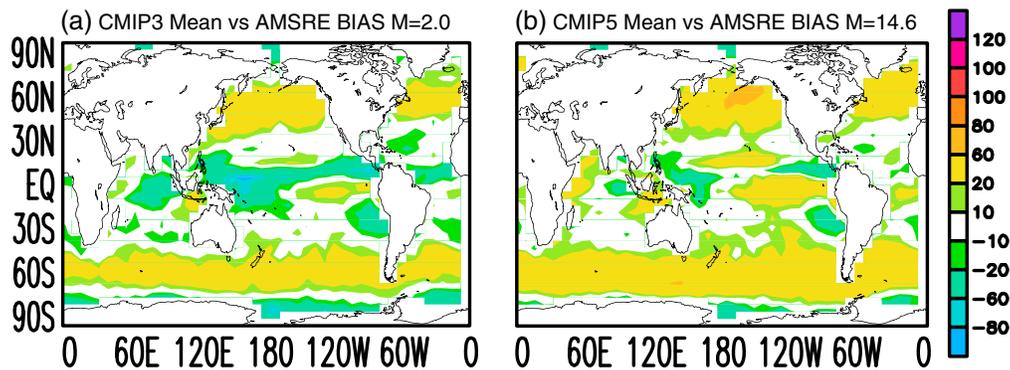


Figure 14. (a) Multimodel mean cloud LWP (CLWP) bias (g m^{-2}) against the Advanced Microwave Scanning Radiometer (AMSR-E) total liquid water path (TLWP cloud + convective + precipitating) for the CMIP3 GCMs (see *Li et al.*, 2012). (b) Multimodel mean bias for CMIP5 against AMSR-E.

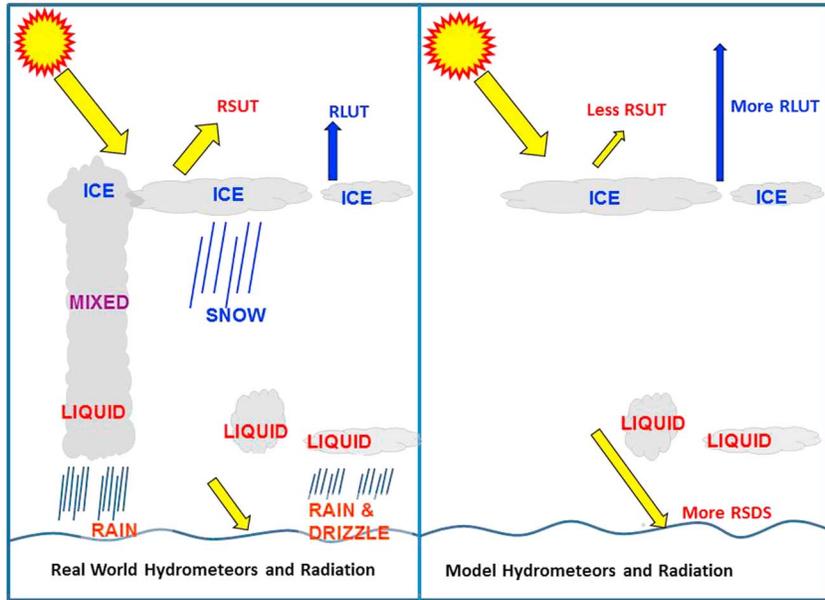


Figure 15. Conceptual illustration of cloud-precipitation-radiation interactions (left) for the real-world GCMs and (right) for conventional GCMs. RSUT: outgoing shortwave radiation, RLUT: outgoing longwave radiation, and RSDS: downward shortwave radiation. The light grey clouds represent floating ice clouds and floating liquid clouds in the real world and the model. The darker grey cloud represents the convective cloud core in the real world. The blue (grey) lines represent falling snow (rain) in the real world.

[31] The modeled ice water path (IWP) values include only the contributions of suspended clouds, as that is all they typically represent, with no contribution from convective cores or precipitation. The observed IWP values are based on an ensemble of CloudSat+CALIPSO estimates [Li et al., 2012], which do include contributions from precipitation and all clouds, including convective cores. It shows that both CMIP3 and CMIP5 models significantly underestimate ice mass over the ITCZ, the SPCZ, a part of the Southern Ocean and tropical continents, and the Indian Monsoon regions. Thus, the models are trying to achieve radiative balance without representing all the ice mass in the atmosphere [e.g., Waliser et al., 2011]. Figure 14 is for the multimodel mean cloud liquid water path (CLWP) bias against the Advanced Microwave Scanning Radiometer (AMSR-E) liquid water path with all conditions (TLWP = no precipitating + convective + precipitating) for the CMIP3 GCMs (Figure 14a) and the CMIP5 multimodel mean bias (Figure 14b). Figure 14 shows that similar underestimates of liquid water occur in the same regions [see Li et al., 2012] and exacerbates the challenge of obtaining an accurate representation of the radiation fields.

[32] The combined results of Figures 13 and 14 suggest that the total (liquid and ice) cloud mass are underestimated over the ITCZ, the SPCZ, a part of the Southern Ocean and tropical continents, and the Indian Monsoon regions, which also tend to be all high-precipitation regions. In addition, Figures 13 and 14 suggest the IWP and LWP biases against the observed total IWP/LWP are both significantly worse in CMIP5 models compared to CMIP3 models (in the multimodel means), yet the radiative fields are slightly improved overall in terms of global area average shown in Figure 5. It exhibits more distinct spatial gradients with

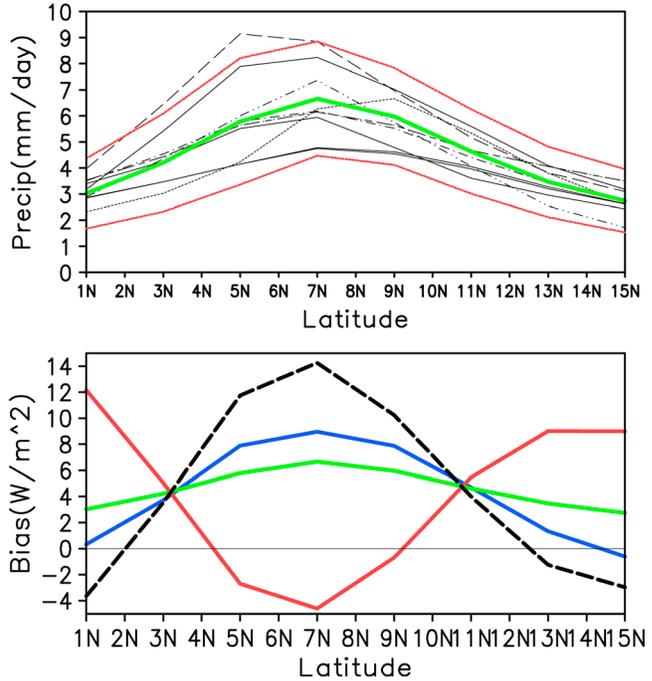


Figure 16. (top) Models and multimodel mean (thick green) of precipitation (mm d^{-2}) for the CMIP5 GCMs. The red lines represent the one standard deviation from model population. (bottom) Multimodel mean bias of (blue) RLUT, (dashed black) RSDS, and (red) RSUT for CMIP5. In all cases, the observed reference is from CERES EBAF for RLUT and RSUT, while RSDS is from SRB EBAF. Zonal mean is taken from the longitude section from 180° to 360° and latitude from 0°N to 15°N for the ITCZ section.

greater local extreme biases and a higher bias over most land regions in the CMIP5 RSDS bias (Figure 5b) than in the CMIP3 RSDS bias (Figure 5a).

[33] In order to illustrate the effect of the underestimated cloud mass due to excluding the cloud mass from precipitation and convective core in the conventional GCMs, we draw a conceptual sketch of cloud-precipitation-radiation interactions for the real-world GCMs versus those for the conventional GCMs in Figure 15. The figure shows that the underestimated values of cloud mass in the CMIP3 and CMIP5 multimodel means might *directly, and in part*, lead to the overestimations of RSDS (Figure 5) and RLUT (Figure 8) and underestimation of RSUT (Figure 11) across the models and the multimodel mean in the heavily precipitating regions. This conjecture is supported in Figure 16, which shows that the latitude of the maximum zonal mean precipitation (180° to 360° ; 0°N to 15°N) is strongly correlated with the latitude of the maximum/minimum multimodel mean bias of RSDS, RLUT, and RSUT of the CMIP5 models. In other words, the RSDS, RLUT, and RSUT biases in CMIP3 and CMIP5 might be attributed to ignoring the radiative impacts from the precipitating clouds and convective core water mass. As cloud-climate feedback will undoubtedly represent a key uncertainty in the next Intergovernmental Panel on Climate Change assessment report, it is essential that cloud and radiation observations be utilized to their full extent and in concert to provide more complete constraints and that clouds, convection, precipitation, and radiation be treated in a consistent manner, as shown in Figure 15 (left).

[34] One GFDL GCM (CM3) and two GISS models as well as CESM1-CAM5 technically stand apart from the other GCMs as they attempt to model atmospheric ice/liquid water associated with convective clouds and precipitation and include their effects on radiation. The evaluations of RSDS (Figure 6a), RLUT (Figure 9a), and RSUT (Figure 12a) show that while NCAR CESM1-CAM5 has the best RLUT (Figures 9–11) among the models, GFDL CM3 slightly outperforms the other GCMs. The two GISS models do not illustrate a better performance than the other GCMs in terms of all the three radiative fluxes. The abilities of the CESM1-CAM5 and GFDL models to represent and output cloudy, precipitating ice, and convective profiles in GFDL-CM3 and to perform a more realistic radiation simulation, combined with the observational capabilities to roughly distinguish these types of ice mass, provide an additional means for constraining the model physics. Note that radiation biases from RSDS, RLUT, and RSUT in the two GISS models result from a substantial positive bias in CIWP/CIWC in the tropics. This positive bias in the tropics arises due to a compensation for a negative bias in clouds and CIWP/CIWC in the extratropics in conjunction with the need to have a globally averaged radiation budget consistent with observations [Li *et al.*, 2012].

[35] While it is beyond the scope of this study to probe the causes of the model-to-model differences and model-to-observation biases in radiation, based on the results, we hypothesize that the lack of an explicit representation of the cloudy, precipitating, and convective core components of the ice (and liquid) mass might play an important role for the biases in RSDS, RLUT, and RSUT. Our recent study [Waliser *et al.*, 2011] has shown that ignoring radiative effects of the precipitating components of the ice mass can result in

nontrivial biases in the shortwave and longwave radiation budgets at the surface and top of atmosphere and even more significant impacts on the vertical radiative heating profile. While more work needs to be pursued in this area, there is a strong suggestion from these studies that GCMs should strive to explicitly represent a broader range of ice and liquid hydrometeors, namely, the larger falling hydrometeors (rain, snow) as well as convective core mass, and include their effects in the radiative heating calculations which, for the moment, are largely ignored. Moreover, the evaluation results of this study show that the radiation balance in the CMIP class of GCMs is still underconstrained and, in many cases, is likely to have been achieved in unrealistic ways.

[36] Taken together, these points indicate the need for additional observational resources to adequately characterize and constrain cloud-precipitation-radiation interactions. Some potentially useful observational resources are a multichannel radar/lidar measurement to characterize the profile and spectrum of cloud and precipitation particle sizes, as well as a Doppler radar capability to provide information on cloud and precipitation dynamics. In addition, satellite observations are affected by spatiotemporal sampling, instrument sensitivity, and retrieval assumptions. Simulators are one method available to emulate these idiosyncrasies within a climate model and thus can be an invaluable tool for robust evaluation of model-simulated clouds. In the future, we plan to integrate these methodologies into our evaluation studies. These additional observational resources, used in conjunction with systematic model experimentation practices, will likely be a constructive strategy for improving the cloud-precipitation-radiation interactions alluded to above.

[37] **Acknowledgments.** We would like to thank the Editor and two reviewers for giving very insightful and helpful comments and suggestions. We thank Prof. M-D Chou/NCU and Dr. W-L Lee/RCEC-Academia Sinica for useful comments; Dr. Anthony Del Genio/NASA GISS, Dr. Ken Lo/NASA GISS, Dr. Voldoir Aurore/CNRM, Dr. Masahiro Watanabe and Dr. Shingo Watanabe/MIROC, Dr. Leon Rotstajn/CSIRO, Dr. Knut von Salzen/CCCma, Dr. Gary Strand/NCAR, Dr. Alf Kirkevag/NCC, Dr. Seiji Yukimoto/MRI, Dufresne Jean-Louis/IPSL, Dr. Tongwen Wu/BCC-CMA, and many other colleagues from climate modeling centers for providing model information. Thanks also to Prof. W-T Anne Chen/NTU when at JPL and Gregory Huey/JPL with data. The contributions by DEW and JLL to this study were carried out on behalf of the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. The contribution of Hsi-Yen Ma to this work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344.

References

- Austin, R. T., A. J. Heymsfield, and G. L. Stephens (2009), Retrieval of ice cloud microphysical parameters using the CloudSat millimeter-wave radar and temperature, *J. Geophys. Res.*, *114*, D00A23, doi:10.1029/2008JD010049.
- Bodas-Salcedo, A., et al. (2008), Evaluating cloud systems in the Met Office global forecast model using simulated CloudSat radar reflectivities, *J. Geophys. Res.*, *113*, D00A13, doi:10.1029/2007JD009620.
- Bodas-Salcedo, A., et al. (2011), COSP: Satellite simulation software for model assessment, *Bull. Am. Meteorol. Soc.*, *92*, 1023–1043, doi:10.1175/2011BAMS2856.1
- Bony, S., et al. (2011), CFMIP: Towards a better evaluation and understanding of clouds and cloud feedbacks in CMIP5 models, *CLIVAR Exchanges*, *56*(2), 20–22.
- Chen, W.-T., C. P. Woods, J.-L. F. Li, D. E. Waliser, J.-D. Chern, W.-K. Tao, J. H. Jiang, and A. M. Tompkins (2011), Partitioning CloudSat ice water content for comparison with upper-tropospheric ice in global atmospheric models, *J. Geophys. Res.*, *116*, D19206, doi:10.1029/2010JD015179.

- 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, *J. Geophys. Res. Lett.*, **35**, L15704, doi:10.1029/2008GL034207.
- Chou, M.-D., and M. J. Suarez (1999), A solar radiation parameterization for atmospheric studies, NASA TM-104606, *Technical Report Series on Global Modeling and Data Assimilation*, vol. 15, 40 pp.
- Colbo, K., and R. Weller (2009), Accuracy of the IMET sensor package in the subtropics, *J. Atmos. Oceanic Technol.*, **9**, 1867–1890, doi:10.1175/2009JTECHO667.1.
- 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. Clim.*, **24**, 2715–2727, doi:10.1175/2011JCLI3652.1
- Collins, W. D., P. J. Rasch, and Others (2004), Description of the NCAR Community Atmosphere Model (CAM 3.0), *Tech. Rep. NCARTN-464 + STR*, 210 pp., Natl. Cent. for Atmos. Res., Boulder, Colo.
- Delanoë, J., and R. J. Hogan (2008), A variational scheme for retrieving ice cloud properties from combined radar, lidar, and infrared radiometer, *J. Geophys. Res.*, **113**, D07204, doi:10.1029/2007JD009000.
- Delanoë, J., and R. J. Hogan (2010), Combined CloudSat-CALIPSO-MODIS retrievals of the properties of ice clouds, *J. Geophys. Res.*, **115**, D00H29, doi:10.1029/2009JD012346.
- Delanoë, J., R. J. Hogan, R. M. Forbes, A. Bodas-Salcedo, and T. H. M. Stein (2011), Evaluation of ice cloud representation in the ECMWF and UK Met Office models using CloudSat and CALIPSO data, *Q. J. R. Meteorol. Soc.*, **137**, 2064–2078, doi:10.1002/qj.882.
- Deng, M., G. G. Mace, Z. Wang, and H. Okamoto (2010), Tropical composition, cloud and climate coupling experiment validation for cirrus cloud profiling retrieval using CloudSat radar and CALIPSO lidar, *J. Geophys. Res.*, **115**, D00J15, doi:10.1029/2009JD013104.
- Deng, M., G. G. Mace, Z. Wang, and R. P. Lawson (2013), Evaluation of several A-Train ice cloud retrieval products with in situ measurements collected during the SPARTICUS campaign, *J. Appl. Meteor. Climatol.*, **52**, 1014–1030, doi:10.1175/JAMC-D-12-054.1.
- Donner, L. J., et al. (2011), The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component of the GFDL coupled model CM3, *J. Clim.*, **24**, 3484–3519, doi:10.1175/2011JCLI3955.1.
- Fouquart, Y., and B. Bonnel (1980), Computations of solar heating of the Earth's atmosphere: A new parameterization, *Beitr. Phys. Atmos.*, **53**, 35–63.
- Gent, P. R., et al. (2011), The Community Climate System Model version 4, *J. Clim.*, **24**(19), 4973–4991, doi:10.1175/2011JCLI4083.1.
- Gottelman, A., X. Liu, S. J. Ghan, H. Morrison, S. Park, A. J. Conley, S. A. Klein, J. Boyle, D. L. Mitchell, and J.-L. F. Li (2010), Global simulations of ice nucleation and ice supersaturation with an improved cloud scheme in the Community Atmosphere Model, *J. Geophys. Res.*, **115**, D18216, doi:10.1029/2009JD013797.
- Grant, K. E., and A. S. Grossman (1998), Description of a solar radiative transfer model for use in LLNL climate and atmospheric chemistry studies, UCRL-ID-129949, Lawrence Livermore Natl. Lab., Livermore, Calif., <http://www.llnl.gov/tid/lof/documents/pdf/233048.pdf>, 17 pp.
- Grant, K. E., C. C. Chuang, A. S. Grossman, and J. E. Penner (1999), Modeling the spectral optical properties of ammonium sulfate and biomass burning aerosols: Parameterization of relative humidity effects and model results, *Atmos. Environ.*, **33**, 2603–2620.
- Han, Q. Y., W. B. Rossow, J. Chou, K. S. Kuo, and R. M. Welch (1999), The effects of aspect ratio and surface roughness on satellite retrievals of ice-cloud properties, *J. Quant. Spectrosc. Radiat. Transfer*, **63**(2–6), 559–583.
- Jiang, J. H., et al. (2012), Evaluation of cloud and water vapor simulations in CMIP5 climate models using NASA A-Train satellite observations, *J. Geophys. Res.*, **117**, D1410, 24 PP, doi:10.1029/2011JD017237.
- Kato, S., S. Sun-Mack, W. F. Miller, F. G. Rose, Y. Chen, P. Minnis, and B. A. Wielicki (2010), Relationships among cloud occurrence frequency, overlap, and effective thickness derived from CALIPSO and CloudSat merged cloud vertical profiles, *J. Geophys. Res.*, **115**, D00H28, doi:10.1029/2009JD012277.
- Kato, S., et al. (2011), Computation of top-of-atmosphere and surface irradiance with CALIPSO, CloudSat, and MODIS-derived cloud and aerosol properties, *J. Geophys. Res.*, **116**, D19209, doi:10.1029/2011JD016050.
- Kato, S., N. G. Loeb, F. G. Rose, D. R. Doelling, D. A. Rutan, and T. E. Caldwell (2012a), Modeled surface irradiances consistent with CERES-derived top-of-atmosphere shortwave and longwave irradiances, submitted to *J. Clim.*
- Kato, S., N. G. Loeb, D. A. Rutan, F. G. Rose, S. Sun-Mack, W. F. Miller, and Y. Chen (2012b), Uncertainty estimate of surface irradiances computed with MODIS-, CALIPSO-, and CloudSat-derived cloud and aerosol properties, *Surv. Geophys.*, **33**(3–4), 395–412, doi:10.1007/s10712-012-9179-x.
- Kay, J. E., et al. (2012), Exposing global cloud biases in the Community Atmosphere Model (CAM) using satellite observations and their corresponding instrument simulators, *J. Clim.*, **25**, 5190–5207, doi:10.1175/JCLI-D-11-00469.1.
- Klein, S. A., and C. Jakob (1999), Validation and sensitivities of frontal clouds simulated by the ECMWF model, *Mon. Weather Rev.*, **127**(10), 2514–2531.
- Klein, S. A., Y. Zhang, M. D. Zelinka, R. N. Pincus, J. Boyle, and P. J. Gleckler (2012), Are climate model simulations of clouds improving? An evaluation using the ISCCP simulator, *J. Geophys. Res.*, **118**, 1329–1342, doi:10.1002/jgrd.50141.
- Kodama, C., A. T. Noda, and M. Satoh (2012), An assessment of the cloud signals simulated by NICAM using ISCCP, CALIPSO, and CloudSat satellite simulators, *J. Geophys. Res.*, **117**, D12210, doi:10.1029/2011JD017317
- Li, J. (2002), Accounting for unresolved clouds in a 1d infrared radiative transfer model. Part I: Solution for radiative transfer, including cloud scattering and overlap, *J. Atmos. Sci.*, **59**, 3302–3320, doi:10.1175/1520-0469(2002)059<3302:AFUCIA>2.0.CO;2.
- Li, J.-L. F., D. E. Waliser, and J. H. Jiang (2011), Correction to “Comparisons of satellites liquid water estimates to ECMWF and GMAO analyses, 20th century IPCC AR4 climate simulations, and GCM simulations,” *Geophys. Res. Lett.*, **38**, L24807, doi:10.1029/2011GL049956.
- Li, J.-L. F., et al. (2012), An observationally-based evaluation of cloud ice water in CMIP3 and CMIP5 GCMs and contemporary analyses, *J. Geophys. Res.*, **117**, D16105, doi:10.1029/2012JD017640.
- Lin, W. Y. and M. H. Zhang (2004), Evaluation of clouds and their radiative effects simulated by the NCAR Community Atmospheric Model CAM2 against satellite observations, *J. Clim.*, **17**, 3302–3318, doi:10.1175/1520-0442.
- 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. Clim.*, **22**(3), 748–766, doi:10.1175/2008JCLI2637.
- Loeb, N. G., J. M. Lyman, G. C. Johnson, R. P. Allan, D. R. Doelling, T. Wong, B. J. Soden, and G. L. Stephens (2012), Observed changes in top-of-the-atmosphere radiation and upper-ocean heating consistent within uncertainty, *Nat. Geosci.*, **5**, 110–113, doi:10.1038/NNGEO1375.
- Ma, H.-Y., M. Köhler, J.-L. F. Li, J. D. Farrara, C. R. Mechoso, R. M. Forbes, and D. E. Waliser (2012), Evaluation of an ice cloud parameterization based on a dynamical-microphysical lifetime concept using CloudSat observations and the ERA-Interim reanalysis, *J. Geophys. Res.*, **117**, D05210, doi:10.1029/2011JD016275.
- Mace, G. G., Q. Zhang, M. Vaughan, R. Marchand, G. Stephens, C. Trepte, and D. Winker (2009), A description of hydrometeor layer occurrence statistics derived from the first year of merged CloudSat and CALIPSO data, *J. Geophys. Res.*, **114**, D00A26, doi:10.1029/2007JD009755.
- Medovaya, M., D. E. Waliser, R. A. Weller, M. McPhaden (2002), Assessing ocean buoy shortwave observations using clear-sky model calculations, *J. Geophys. Res.*, **107**, C2, doi:10.1029/2000JC000558.
- Minnis, P., W. L. Smith, Jr., D. P. Garber, J. K. Ayers, D. R. Doelling (1994), Cloud Properties Derived From GOES-7 for Spring 1994 ARM Intensive Observing Period Using Version 1.0.0 of ARM Satellite Data Analysis Program, NASA Ref. Publ. 1366, 59 pp.
- Minnis, P., et al. (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**, 11, doi:10.1109/TGRS.2011.2144601.
- Mlawer E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, S. A. Clough (1997), Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-*k* model for the longwave, *J. Geophys. Res.*, **102**, 16,663–16,682.
- Morcrette, J.-J., and Y. Fouquart (1986), The overlapping of cloud layers in shortwave radiation parameterizations, *J. Atmos. Sci.*, **43**, 321–328.
- Morcrette, J. J., Smith L., Fouquart Y. (1986), Pressure and temperature dependence of the absorption in longwave radiation parameterizations, *Contrib. Atmos. Phys.*, **59**(4), 455–469.
- Nam, C. C. W., and J. Quaas (2012), Evaluation of clouds and precipitation in the ECHAM5 general circulation model using CALIPSO and CloudSat satellite data, *J. Clim.*, **25**, 4975–4992, doi:10.1175/JCLI-D-11-00347.1.
- Norris, J. R., and C. P. Weaver (2001), Improved techniques for evaluating GCM cloudiness applied to the NCAR CCM3, *J. Clim.*, **14**, 2540–2550.
- Rasch, P. J., and J. E. Kristjánsson (1998), A comparison of the CCM3 model climate using diagnosed and predicted condensate parameterizations, *J. Clim.*, **11**, 1587–1614.
- Randall, D. A., et al. (2007), Climate models and their evaluation, in *Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon et al., pp. 589–662, Cambridge Univ. Press, Cambridge, U. K.
- Rienecker, M. M., et al. (2008), The GEOS-5 Data Assimilation System—Documentation of Versions 5.0.1, 5.1.0, and 5.2.0. *Technical Report Series on Global Modeling and Data Assimilation*, 27.

- Rossow, W. B., and R. A. Schiffer (1999), Advances in understanding clouds from ISCCP, *Bull. Am. Meteorol. Soc.*, *80*(11), 2261–2287.
- Rossow, W. B., and Y. C. Zhang (1995), Calculation of surface and top of atmosphere radiative fluxes from physical quantities based on ISCCP data sets: 2. Validation and first results, *J. Geophys. Res.*, *100*(D1), 1167–1197.
- Schmidt, G. A., et al. (2006), Present day atmospheric simulations using GISS ModelE: Comparison to in-situ, satellite and reanalysis data, *J. Clim.*, *19*, 153–192, doi:10.1175/JCLI3612.1.
- Stephens, G. L. (2005), Cloud feedbacks in the climate system: A critical review, *J. Clim.*, *18*, 237–273, doi:10.1175/JCLI-3243.1.
- Sundqvist, H. (1988), Parameterization of condensation and associated clouds in models for weather prediction and general circulation simulation, in *Physically-Based Modeling and Simulation of Climate and Climate Change*, vol. 1, edited by M. E. Schlesinger, pp. 433–461, Kluwer Academic Publishers, Dordrecht, the Netherlands.
- Taylor, K. E. (2001), Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res.*, *106*, 7183–7192.
- Trenberth, K. E., and J. T. Fasullo (2009), Simulation of present day and 21st century energy budgets of the southern oceans, *J. Clim.*, *23*(2), 440–454, doi:10.1175/2009JCLI3152.1.
- Waliser, D. E., R. A. Weller, R. D. Cess (1999), Comparisons between buoy-observed, satellite-derived and modeled surface shortwave flux over the subtropical North Atlantic during the subduction experiment, *J. Geophys. Res.*, *104*, 31,301–31,320.
- Waliser, D. E., et al. (2009), Cloud ice: A climate model challenge with signs and expectations of progress, *J. Geophys. Res.*, *114*, D00A21, doi:10.1029/2008JD010015.
- Waliser, D. E., J.-L. F. Li, T. S. L'Ecuyer, and W.-T. Chen (2011), The impact of precipitating ice and snow on the radiation balance in global climate models, *Geophys. Res. Lett.*, *38*, L06802, doi:10.1029/2010GL046478.
- Webb, M., C. Senior, S. Bony, and J. J. Morcrette (2001), Combining ERBE and ISCCP data to assess clouds in the Hadley Centre, ECMWF and LMD atmospheric climate models, *Clim. Dyn.*, *17*(12), 905–922.
- Wielicki, B. A., B. R. Barkstrom, E. F. Harrison, R. B. Lee, G. L. Smith, and J. E. Cooper (1996), Clouds and the Earth's Radiant Energy System (CERES): An Earth observing system experiment. *Bull. Am. Meteorol. Soc.*, *77*, 853–868.
- Zhang, Y.-C., W. B. Rossow, A. A. Lacis, V. Oinas, and M. I. Mishchenko (2004), Calculation of radiative fluxes from the surface to top of atmosphere based on ISCCP and other global data sets: Refinements of the radiative transfer model and the input data, *J. Geophys. Res.*, *109*, D19105, doi:10.1029/2003JD004457.
- Zhang, M. H., et al. (2005), Comparing clouds and their seasonal variations in 10 atmospheric general circulation models with satellite measurements. *J. Geophys. Res.*, *110*, D15S02, doi:10.1029/2004JD005021.
- Zhang, Y., S. A. Klein, J. Boyle, and G. G. Mace (2010), Evaluation of tropical cloud and precipitation statistics of Community Atmosphere Model version 3 using CloudSat and CALIPSO data, *J. Geophys. Res.*, *115*, D12205, doi:10.1029/2009JD012006.