COSP

Satellite simulation software for model assessment

BY A. BODAS-SALCEDO, M. J. WEBB, S. BONY, H. CHEPFER, J.-L. DUFRESNE, S. A. KLEIN, Y. ZHANG, R. MARCHAND, J. M. HAYNES, R. PINCUS, AND V. O. JOHN

By simulating the observations of multiple satellite instruments, COSP enables quantitative evaluation of clouds, humidity, and precipitation processes in diverse numerical models.

General circulation models (GCMs) of the atmosphere, including those used for numerical weather prediction (NWP) and climate projections, operate with resolutions from a few kilometers to hundreds of kilometers. Many atmospheric processes, such as turbulence and microphysical processes within clouds, operate at smaller scales and hence cannot be resolved by current model resolutions. These processes are included by means of parameterizations, which are semiempirical or statistical models that relate gridbox mean variables to these subgrid processes. For instance, some cloud parameterizations diagnose the amount of cloud condensate and the fraction of the grid box that a cloud occupies (cloud area fraction) as a function of the relative humidity (RH) of the grid box (Slingo 1980; Smith 1990). The formulation of these parameterizations is very important for the model evolution because they modify the three-dimensional structure of temperature and humidity directly (e.g., condensation/evaporation) or indirectly by interacting with other parameterizations (e.g., radiation) and the large-scale dynamics. Therefore, the evaluation of these parameterizations is crucial to improving our weather forecasts or increasing our confidence in climate projections.

Satellites have proven to be very helpful tools for this purpose because they provide global or near-global coverage, thereby giving a representative sample of all meteorological conditions. However, satellites do not measure directly those geophysical quantities of interest, such as the amount or phase of cloud condensate. They measure the intensity of radiation coming from a particular area and direction in a particular wavelength range (radiances). The range of wavelengths covered by past and current systems spans several orders of magnitude, from

AFFILIATIONS: BODAS-SALCEDO, WEBB, AND JOHN—Met Office Hadley Centre, Exeter, United Kingdom; BONY, CHEPFER, AND DUFRESNE—Laboratoire de Météorologie Dynamique/L'Institut Pierre-Simon Laplace, Centre National de la Recherche Scientifique, Université Pierre et Marie Curie, Paris, France; KLEIN AND ZHANG—Program For Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, Livermore, California; MARCHAND— Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, Seattle, Washington; HAYNES—School of Mathematical Sciences, Monash University, Clayton, Victoria, Australia; PINCUS—University of Colorado and NOAA/Earth System Research Laboratory, Boulder, Colorado
CORRESPONDING AUTHOR: Dr. A. Bodas-Salcedo, Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB United Kingdom
E-mail: alejandro.bodas@metoffice.gov.uk

The abstract for this article can be found in this issue, following the table of contents.

DOI:10.1175/2011BAMS2856.1

In final form 8 April 2011
©2011 American Meteorological Society

the ultraviolet (10^{-7} m) to radio frequencies (1 m). Information on the geophysical quantities of interest is inferred by inverse modeling, usually called satellite retrievals (e.g., Stephens and Kummerow 2007). A great deal of research has been conducted into producing satellite retrievals of many different geophysical variables, such as water vapor, atmospheric temperature, cloud properties, and land surface products (e.g., Randel et al. 1996; Chahine et al. 2006; Wylie and Menzel 1999; Schaaf et al. 2002). Satellite retrievals have been used in numerous studies to analyze the performance of NWP and climate models (e.g., Gates et al. 1999; Allan et al. 2007; Gleckler et al. 2008; Pincus et al. 2008). However, it is very difficult to extract quantitative information from these types of comparisons because part of the difference between model variables and retrievals stems from the different definitions of the variable under study (e.g., total cloudiness) among models and between the models and the satellite retrievals. This issue holds even for well-defined quantities—for example, ice water path—because of the inevitable assumptions required in retrievals. Additionally, satellite sensors have limitations—finite sensitivity, fixed viewing geometry, among others—that introduce uncertainties into the retrievals. Moreover, the retrievals may be sensitive to the first guess of the atmospheric state that is used in the inversion (Eyre 1987). Despite these difficulties, comparisons using retrievals have the advantage that the analysis is always carried out in the space of the geophysical variables, and hence the interpretation of the results can be easily linked with model deficiencies.

In the last two decades, a different avenue has been followed to exploit satellite data in model evaluation: the use of forward modeling of basic satellite measurements from model fields. Simulation of radiances in the longwave (LW) window and 6.7-μm channels have been widely used to evaluate models (e.g., Morcrette 1991; Salathé and Chesters 1995; Ringer et al. 2003; Iacono et al. 2003; Allan et al. 2003; Brogniez et al. 2005; Zhang et al. 2008). A different approach, although with the same aim of avoiding ambiguities in the comparisons between model variables and satellite retrievals, has also been pursued by developing simulators that mimic the observational process, including as many known biases as are practical (e.g., Yu et al. 1996). The simulator approach essentially acknowledges the issue that a retrieval produced by a satellite might not be directly comparable to a model variable, giving rise to multiple values of geophysical quantities from different sensors and retrieval algorithms. An example of this is the International Satellite Cloud Climatology Project (ISCCP; Rossow and Schiffer 1999) simulator that has been widely used in many efforts to evaluate the simulation of clouds in models (e.g., Klein and Jakob 1999; Webb et al. 2001; Zhang et al. 2005; Williams and Tselioudis 2007). More recently, several studies have followed this approach to exploit data from the active sensors (lidar and radar) (Haynes et al. 2007; Bodas-Salcedo et al. 2008; Chepfer et al. 2008; Wilkinson et al. 2008; Marchand et al. 2009; Zhang et al. 2010) and from the Multiangle Imaging Spectroradiometer (MISR) (Marchand and Ackerman 2010). The development of simulators is also an active area of research in the observational community (e.g., Voors et al. 2007; Masunaga et al. 2010), and it may also help address the question as to whether cloud-resolving models (CRMs) used to develop parameterizations are adequate for doing so. The main drawback of the simulator approach is that interpreting results in terms of physical processes may sometimes be problematic because the comparison variables are not trivially related to any single geophysical quantity.

The Cloud Feedback Model Intercomparison Project (CFMIP) community has developed an integrated satellite simulator, the CFMIP Observation Simulator Package (COSP). COSP is a flexible software tool that enables the simulation from model variables of data from several satelliteborne active and passive sensors. It facilitates the use of satellite data to evaluate models in a consistent way. The flexibility of COSP makes it suitable for use in many types of numerical models, from high-resolution models (~1-km resolution) to coarse-resolution models, such as the GCMs used in climate modeling, and the scales in between used in weather forecast and regional models. The fact that COSP includes several simulators under the same interface facilitates the implementation of a range of simulators in models. Another advantage of COSP—and in general, the simulator approach—is that it facilitates model intercomparison, not only model–satellite comparison (e.g., comparisons of cloud properties simulated by GCMs and CRMs).

The remainder of this paper is organized as follows: Section 2 describes the observations and presents a technical overview of COSP; section 3 presents the observational results, with special emphasis in the complementary nature of the different datasets; sections 4 and 5 show results from a model intercomparison using COSP; and conclusions and future plans are presented in section 6.

TECHNICAL DESCRIPTION OF COSP AND SATELLITE INSTRUMENTS. COSP is a modular

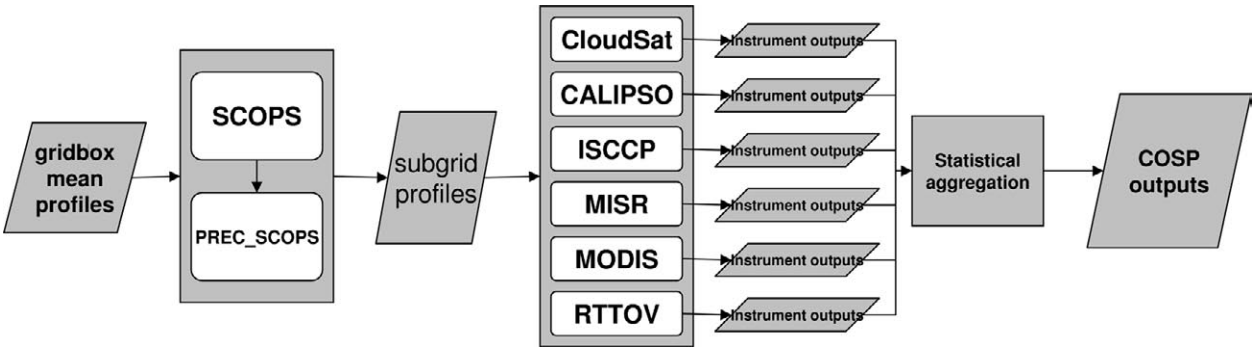


Fig. 1. COSP schematic.

piece of software written almost entirely in Fortran 90 and is distributed under Berkeley Software Distribution (BSD) open-source license via the CFMIP Web site (www.cfmip.net). Figure 1 depicts a schematic of its main components. COSP requires gridbox mean vertical profiles of temperature, humidity, hydrometeor (clouds and precipitation) mixing ratios, cloud optical thickness, and emissivity, along with surface temperature and emissivity from model output. COSP produces output comparable to satellite data in three broad steps. First, to address the mismatch in scale between that of a GCM grid box and that of a satellite pixel, the gridbox mean profiles for each grid box are broken into subcolumns, which may be thought of as representing an area much smaller than the grid box but commensurate with that of a satellite pixel. Second, the vertical profiles of individual subcolumns are passed to each of the instrument simulators, which apply models of differing complexity to simulate the signals and/or retrievals of various instruments. Finally, statistical modules, provided by COSP or the individual instrument simulators, gather the outputs from all the instruments and build diagnostics that can be compared to similar statistics from observations.

The current version of COSP includes simulators for datasets produced by the following instruments (Table 1): the cloud profiling radar (CPR) on board *CloudSat* (Stephens et al. 2002), the Cloud–Aerosol Lidar with Orthogonal Polarization (CALIOP) lidar onboard Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO; Winker et al. 2010), the ISCCP (Rossow and Schiffer 1999), the MISR (Diner et al. 2005), and the Moderate Resolution Imaging Spectroradiometer (MODIS) (King et al. 2003). The fast radiative transfer code Radiative Transfer for Television and Infrared Observation Satellite (TIROS) Operational Vertical Sounder (RTTOV) (Saunders et al. 1999) can also be linked to COSP to produce clear-sky brightness temperatures for many different channels of past and current infrared and passive microwave radiometers.

COSP instrument simulators operate on subcolumns (e.g., Klein and Jakob 1999; Räisänen et al. 2004; Pincus et al. 2006). Subcolumns are discrete samples, each of which is assumed to be internally homogeneous but constructed so that a large number of samples reproduces any internal inhomogeneity (including fractional cloudiness) in each grid cell as well as the model's overlap assumptions about vertical correlation. COSP can make use of subcolumns generated by the host model if they are already available [because, e.g., they are used in the model's radiative transfer algorithms (e.g. Pincus et al. 2003)]. COSP can also generate its own subcolumns using the Subgrid Cloud Overlap Profile Sampler (SCOPS; Webb et al. 2001) originally developed for the ISCCP simulator. Some instrument simulators also require knowledge about the distribution of precipitation within the columns; COSP employs a simple algorithm (Zhang et al. 2010; labeled PREC SCOPS in Fig. 1) to determine the subgrid distribution of precipitation fluxes from the subcolumn cloud distribution and gridbox mean precipitation fluxes. The use of subcolumns is only required in models with very large grid spacing (i.e., GCMs and NWP models). In cloud-resolving models, the generation of subcolumns is unnecessary, and the profiles from each grid box of the model can be directly passed to the instrument simulators that operate in step 2 of COSP.

The following subsections describe the satellite instruments and associated COSP modules that produce observations compatible with those instruments.

Active instruments: CloudSat and CALIPSO. *CloudSat* and CALIPSO provide vertical information on clouds and aerosols. *CloudSat* CPR operates at a frequency of 94 GHz (Im et al. 2005). The CPR points in the nadir direction and its pulses sample a volume of 480 m in

Table 1. List of diagnostics from the COSP version 1.3.	
Simulator	**Output diagnostics**
CALIPSO	Lidar total backscatter (532 nm)
	Lidar molecular backscatter
	Height–scattering ratio histograms
	Low-level cloud fraction (CTP > 680 hPa)
	Midlevel cloud fraction (440 < CTP < 680 hPa)
	High-level cloud fraction (CTP < 440 hPa)
	3D cloud fraction
	Total cloud fraction
CloudSat	Radar reflectivity
	Height–reflectivity histograms
ISCCP	Mean cloud albedo
	Mean CTP
	Mean 10.5-μm T_B
	Mean clear-sky 10.5-μm T_B
	Mean cloud optical depth
	CTP in each subcolumn
	Cloud optical depth in each subcolumn
	CTP–τ histograms
	Total cloud fraction
MISR	CTH–τ histograms
MODIS	Total cloud fraction
	Liquid cloud fraction
	Ice cloud fraction
	High-level cloud fraction
	Midlevel cloud fraction
	Low-level cloud fraction
	Total cloud optical thickness
	Liquid cloud optical thickness
	Ice cloud optical thickness
	Total cloud optical thickness [Log_{10}(mean)]
	Liquid cloud optical thickness [Log_{10}(mean)]
	Ice cloud optical thickness [Log_{10}(mean)]
	Liquid cloud particle size
	Ice cloud particle size
	CTP-τ histograms
	Cloud liquid water path
	Cloud ice water path
	Cloud area fraction
PARASOL	Monodirectional reflectance
RTTOV	Clear-sky T_B
Combined	CALIPSO cloud fraction undetected by *CloudSat*
	Total cloud fraction from *CloudSat* and CALIPSO

the vertical, with a horizontal resolution of 1.4 km across track. We use the level 2 *CloudSat* geometric profile (2B-GEOPROF) dataset, which provides the radar reflectivity and identifies where hydrometeors occur (Marchand et al. 2008).

CALIOP operates at 532 nm and 1,064 nm. It is nadir pointing with a beam diameter of 70 m at the Earth's surface and a pulse repetition frequency that produces footprints every 333 m in the alongtrack direction. In this study, we use the GCM-Oriented CALIPSO Cloud Product (GOCCP; Chepfer et al. 2010). GOCCP processes CALIOP level 1 data following the same steps as the CALIPSO diagnostics in COSP. Instantaneous profiles of the lidar scattering ratio (SR) are first computed at the highest horizontal resolution of the data but at the vertical resolution typical of current GCMs, and then cloud diagnostics are inferred from these profiles. The lidar scattering ratio measures the effectiveness of the target in reflecting the lidar beam, and it is measured relative to the backscatter signal a molecular atmosphere (without clouds or aerosols) would have produced. Values of SR greater than five are taken as indications of volumes containing clouds.

For the simulation of *CloudSat* radar reflectivities, COSP uses QuickBeam (Haynes et al. 2007). QuickBeam simulates the vertical profile of radar reflectivities from a mixture of a flexible number of hydrometeor species (clouds and precipitation). Each hydrometeor type is described by a vertical profile of mixing ratios and a particle size distribution. The particle size distributions can be chosen from a library of built-in distributions. QuickBeam accounts for attenuation by atmospheric gases. By default, it is configured to reproduce *CloudSat*-like radar reflectivities, but it can also be used to simulate Tropical

Rainfall Measuring Mission precipitation radar reflectivities and vertically pointing ground-based radars. At the *CloudSat* radar frequencies, the attenuation by gases and liquid water is not negligible, and it has to be taken into account in the radar reflectivity computations. Also, for large cloud ice particles and precipitation, the wavelength is comparable to the size of the scatterers, and the Rayleigh approximation is no longer valid. QuickBeam accounts for these effects in its radar reflectivity computations. Multiple scattering is also present in a fraction of precipitating profiles (Battaglia et al. 2008), although this effect is neglected in the current version of the radar simulator used in COSP.

For the simulation of the CALIOP total backscatter, COSP uses software developed by Chepfer et al. (2008), based upon an earlier code known as active remote sensing simulator (ActSim) (Chiriaco et al. 2006; Chepfer et al. 2007). The lidar-attenuated backscatter signal is simulated for the 532-nm channel. From the backscatter profiles calculated with and without clouds, the lidar scattering ratio can be calculated. At the simulated wavelength, atmospheric cloud particles and gas molecules contribute to scattering but not to absorption. Cloud particles are considered to be spherical, and therefore the backscattering phase function is parameterized as a function of the effective radius using Mie theory. An option is included to account for the impact of nonsphericity of ice particles. The results presented here were derived using the spherical particles parameterization. Aerosols are not included in the simulations. A conservative cloud detection threshold is applied to the simulations and to the observations to minimize the impact of not including aerosols.

Passive imagers: ISCCP, MISR, and MODIS. ISCCP (Rossow and Schiffer 1999) collects visible and infrared radiances from operational weather satellites in sun-synchronous (near polar) and geostationary orbits. Cloud retrieval algorithms are then applied to produce a long-term climatology of cloud properties. The retrieval algorithm allows for the estimation of cloud-top pressure (CTP) and cloud optical depth (τ). Statistical summaries at lower resolution (2.5° × 2.5° latitude/longitude, daily and monthly means) are constructed in the form of joint CTP–τ histograms of cloud fraction. A simulator of Polarization and Anisotropy of Reflectances for Atmospheric Sciences coupled with Observations from a Lidar (PARASOL) monodirectional reflectances has also been implemented in COSP, although results from this simulator will be reported elsewhere.

MISR images Earth in nine different view directions to infer the angular variation of reflected sunlight and the physical characteristics of the observed scenes (Diner et al. 1998, 2005). It flies aboard the *Terra* satellite, sun-synchronous platform with the equator crossing at 1030 local time descending node. MISR determines cloud-top height (CTH) using a stereo-imaging technique (Moroney et al. 2002; Muller et al. 2002). MISR also retrieves cloud optical depth from the visible radiances, although only over ocean. These retrievals allow the computation of joint CTH–τ histograms similar to the ISCCP histograms.

MODIS is a 36-channel radiometer flying aboard the *Terra* and *Aqua* platforms. *Aqua* is also a sun-synchronous platform but with the equator crossing at 1330 local time ascending node. MODIS produces a wide range of cloud and aerosol products, including phase discrimination, cloud-top pressure estimation using CO_2 slicing (King et al. 2003; Platnick et al. 2003), and the determination of particle sizes based on near-infrared reflectance. A special purpose MODIS dataset has been constructed to facilitate a comparison with output from COSP and the MODIS simulator. This dataset contains only those quantities produced by the MODIS simulator. Uncertainty estimates are provided where they are available.

In COSP, the ISCCP simulator (Klein and Jakob 1999; Webb et al. 2001) outputs cloud fraction in cloud-top pressure and cloud optical depth bins, as well as mean cloud fraction, optical depth, and cloud-top pressure. The ISCCP simulator does not make full forward simulations of radiances; however, it takes into account the effects of overlapping clouds on the cloud fraction as seen from space, and it mimics the interpretation of brightness temperatures as coming from a single homogenous layer. Neither does the ISCCP simulator account for the calibration and view-angle-dependent biases that affect the observational database (Evan et al. 2007).

The MISR simulator used in COSP is described by Marchand and Ackerman (2010). It follows the same approach as the ISCCP simulator in the sense that it does not carry out comprehensive radiative transfer computations, but it does estimate the cloud-top height using a set of simple rules obtained from comparison of the MISR stereo height retrieval with ground-based radar and lidar. This approach allows the main characteristics of the retrieval to be captured in an efficient manner.

The MODIS simulator included in COSP emulates pixel-scale retrievals and monthly averages from MODIS (R. Pincus et al. 2011, manuscript submitted

to *J. Climate*). At the pixel (subcolumn) scale, the simulator uses simple integration to determine cloud detection and to compute cloud optical depth. Cloud-top pressure and cloud phase are estimated using extinction-weighted integration. Particle size is estimated using a pseudo-inversion: top-of-atmosphere near-infrared fluxes are computed using polynomial fits of optical properties as a function of (height dependent) cloud particle size; differences between the true fluxes and trial fluxes are then minimized by changing the particle size. Aggregation over time includes both linear and logarithmic averages of optical thickness and a joint histogram of cloud-top pressure and optical thickness.

Infrared and microwave sounders: RTTOV. RTTOV is a fast radiative transfer code that was originally developed to simulate radiances for the assimilation of data from the TIROS Operational Vertical Sounder (Saunders et al. 1999). RTTOV is developed by several European meteorological services within the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT)'s NWP Satellite Application Facility (http://research.metoffice.gov.uk/research/interproj/nwpsaf/rtm/). RTTOV has its own licensing terms, and therefore it is not distributed with COSP; however, it can be integrated and run within COSP. RTTOV is capable of computing brightness temperatures for most of the infrared sounders and passive microwave radiometers that are currently operational or have been operational in the past.

Diagnostic outputs. Table 1 lists the output diagnostics from the current version of COSP (version 1.3). The Climate Model Output Rewriter (http://www2-pcmdi.llnl.gov/cmor) library is used to write the outputs to network common data form (NetCDF) files that comply with the climate and forecast metadata convention and fulfill the requirements of the climate community's standard model experiments. The Working Group on Coupled Modelling has recommended the use of COSP in a subset of simulations for the Coupled Model Intercomparison Project (CMIP) Phase 5 that will be assessed by the fifth IPCC assessment report (http://cmip-pcmdi.llnl.gov/cmip5/experiment design.html). CMIP is a standard experimental protocol for studying the output of coupled ocean–atmosphere general circulation models. It provides a community-based infrastructure in support of climate model diagnosis, validation, intercomparison, documentation, and data access (Meehl et al. 2007).

OBSERVATIONAL RESULTS. Figure 2 shows observational diagnostics from these datasets for September–November (SON) 2006 over a region in the North Pacific (40°–60°N, 160°E–125°W). This region will be used to compare against model simulations in the following section. We chose this region because of its rich mixture of cloud types.

Vertical profiles of radar reflectivities are used to construct frequency histograms of radar reflectivity as a function of height (Fig. 2a). To interpret these diagrams, it is useful to remember that the radar reflectivity is not only sensitive to the amount of cloud or precipitation in the target volume but also to the particle size of the scatterers. In the Rayleigh scattering regime (wavelength much larger than particle diameter), the reflectivity is the sixth moment of the particle size distribution (e.g., Sauvageot 1992). This makes the reflectivity extremely sensitive to the particle size of the cloud and precipitation. The reflectivity uses a logarithmic scale, expressed in decibels (dB): an increase of 10 dB implies an increase of one order of magnitude in the intensity of the backscattered radiation. Although we show one region here, the reflectivity histograms have common characteristics, irrespective of the region or period analyzed. The observations sample a roughly triangular region in this 2D space, although they are mainly clustered in three distinct regions in the histogram that are related to different hydrometeor types. There seems to be a linear relationship between maximum reflectivity and height above the freezing level, at approximately 3 km in this region. In this first cluster, or ice branch of the histogram, the strength of the signal grows as height decreases because of higher water content and the effects of aggregation, which produces larger particles. The maximum values of the ice branch are bounded by the reflectivity produced by an ice particle size distribution with ice water content given by the difference between liquid water saturation and ice water saturation mixing ratios at that height (P. Field et al. 2011). Once the falling ice reaches the freezing level, it melts and falls as rain. Rainfall is the second distinct cluster of the histogram, and it is located in the bottom right corner of the histogram (cluster in yellow), with large reflectivities due to the large particle sizes. At low levels, below 2 km, the distribution seems to be slightly bimodal, with a second peak around −25 dB, corresponding to nonprecipitating clouds (third cluster).

The lidar observations from CALIPSO provide a complementary perspective to *CloudSat* (Delanoë and Hogan 2010). Figure 2b shows the scattering ratio histograms as a function of height as derived

by the CALIPSO-GOCCP observations. Only the nighttime measurements have been used to minimize the impact of instrumental noise in the scattering ratio computation (Chepfer et al. 2010). The histogram shows a main cluster of points in the mid and high troposphere that are classified as cloud (SR >

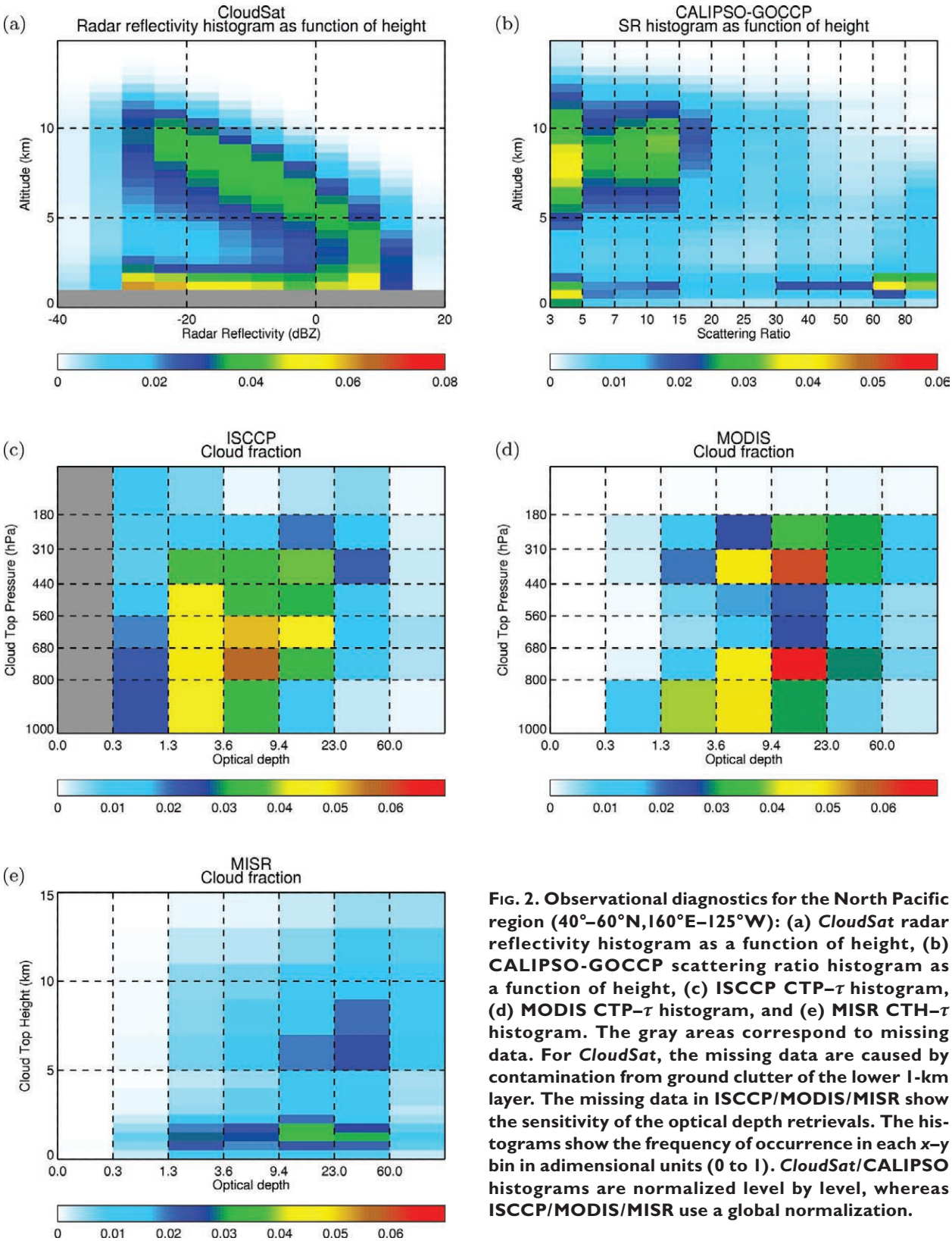


Fig. 2. Observational diagnostics for the North Pacific region (40°–60°N,160°E–125°W): (a) *CloudSat* radar reflectivity histogram as a function of height, (b) CALIPSO-GOCCP scattering ratio histogram as a function of height, (c) ISCCP CTP–τ histogram, (d) MODIS CTP–τ histogram, and (e) MISR CTH–τ histogram. The gray areas correspond to missing data. For *CloudSat*, the missing data are caused by contamination from ground clutter of the lower 1-km layer. The missing data in ISCCP/MODIS/MISR show the sensitivity of the optical depth retrievals. The histograms show the frequency of occurrence in each *x*–*y* bin in adimensional units (0 to 1). *CloudSat*/CALIPSO histograms are normalized level by level, whereas ISCCP/MODIS/MISR use a global normalization.

5), with contributions from cirrus and deeper frontal clouds. The lidar signal is quickly attenuated by liquid cloud, and it becomes fully attenuated for clouds with optical depth $\tau > 5$. This makes the number of fully attenuated samples (SR < 0.01, not shown) increase drastically for levels below 5 km. Low-level cloud (below 3 km) spans SR values from 5 to greater than 80. The histogram also shows some population for very high SR values (SR > 80), implying liquid water, above 2 km and up to 6 km high. Congestus clouds and midlevel mixed-phase layered clouds (Hogan et al. 2003) may contribute to populate this region in the histogram, as well as the transition from ice to mixed phase in deep frontal clouds.

While the form of the ISCCP, MISR, and MODIS histograms (Figs. 2c–e) are conceptually similar, the satellite sensors and the algorithms differ, with the result that the joint histograms can differ quite significantly even when viewing the same clouds. The spatiotemporal sampling of these datasets is also different. ISCCP uses data from geostationary satellites and sun-synchronous satellites, which provide a good spatial and temporal sampling that are able to resolve the diurnal cycle. MISR and MODIS fly on sun-synchronous satellites, and therefore they sample each latitude band at the same local time. However, the amplitude of the cloud fraction and optical depth diurnal cycle in the extratropical oceans is very small (Rossow and Schiffer 1999; Warren et al. 1988), so its impact on the results presented in Fig. 2 is negligible. All three instruments report clouds across a broad range of optical depths. ISCCP (Fig. 2c) reports a substantial amount of optically thin-to-moderate clouds with cloud-top pressures in the middle of the troposphere. Whereas, MODIS (Fig. 2d) and MISR (Fig. 2e) produce a bimodal distribution of cloud elevations, with larger frequencies of low clouds across a range of optical depths and frequent high, optically thick clouds. The CTP retrieval from ISCCP uses the cloud-top temperature obtained from the measured radiances and collocated atmospheric temperature profiles produced operationally. However, MODIS uses a CO_2 slicing method for clouds with top pressures less than about 700 hPa. The CO_2 slicing technique uses several channels around the 15-μm CO_2 band with different absorption properties to estimate cloud-top pressure. The ISCCP method is susceptible to reporting a high, semitransparent cloud over a low-level cloud as a midlevel cloud due to the contribution of the warm low cloud to the measured radiance (Marchand and Ackerman 2010). The CO_2 slicing method is less susceptible to being confused by this type of multilayer cloud situations. Hence, these types of clouds will be reported as midlevel by ISCCP and as high-level by MODIS. At low levels (CTP > 700 hPa), MODIS reverts to an algorithm similar to ISCCP for low clouds. The differences in these observational retrievals highlight the importance of the simulator approach in model evaluation. The simulators will be able to reproduce the apparent midlevel clouds in ISCCP and high-/low-level clouds in MODIS and MISR if the models produce optically thin clouds over low-level clouds.

The combination of these datasets tells us more about the observed cloud fields than the three datasets individually. The different behavior of the satellite retrievals in the presence of a high-level thin cloud above a low-level cloud can be used to extract additional information. In these situations, which we refer to as multilayer cloud, ISCCP tends to report a cloud-top height in the midtroposphere, whereas the MISR stereo-imaging technique is capable of "seeing" through the optically thin cloud. That means that these situations will be reported as midlevel cloud by ISCCP and low-level cloud by MISR (i.e., ISCCP reported high + mid = optically thick high + thick mid + optically thin high with low, while MISR reported high + mid = thick high + thick mid). Therefore, the amount of multilayer cloud can be estimated by subtracting the fraction of high-level cloud + midlevel cloud retrieved by ISCCP from the fraction of high-level cloud + midlevel cloud retrieved by MISR (Marchand et al. 2010). Table 2 shows the results of these estimates for clouds with cloud-top pressure smaller than 680 hPa. Marchand et al. (2010) remap the MISR data onto pressure layers to provide an ISCCP-like histogram and make an accurate estimate of MISR cloud above 680 hPa.

Table 2. High-level plus midlevel ISCCP and MISR cloud fractions (%) over the North Pacific region for the models and the observational datasets. The difference (ISCCP minus MISR) is an estimate of the amount of multilayer cloud (Marchand et al. 2010).

	ISCCP	MISR	Difference
HadGEM1	47.9	36.2	11.7
LMDZ4	33.8	31.4	2.4
MMF 1kmL52	56.5	41.3	15.2
MMF 4kmL26	47.8	38.9	8.9
Observations	55.5	32.0	23.5

Because this mapped dataset was not available for this study, we have added all MISR clouds with tops above 3 km as an estimate of high-level MISR + midlevel MISR cloud. This height is an estimate of the average height of the 680-hPa pressure surface. The amount of multilayer cloud is on the order of 23%, suggesting that a substantial amount of high-level thin cloud above low-level cloud is misclassified by ISCCP as midlevel cloud.

MODEL INTERCOMPARISON OF CLOUDS AND PRECIPITATION. We have applied COSP to four different atmospheric GCMs run with prescribed sea surface temperatures (SSTs) for the season of September–November 2006. Table 3 lists the models, resolutions, and references to the model description papers. The Hadley Centre Global Environmental Model version 1 (HadGEM1) and the Laboratoire de Météorologie Dynamique (LMD) model with zoom capability (LMDZ) are standard GCMs, whereas the multiscale modeling framework (MMF) consists of the National Center for Atmospheric Research Community Atmosphere Model in which a two-dimensional cloud-resolving model has been embedded in each GCM grid box. Two MMF simulations have been run with different resolutions of the cloud-resolving model: i) 64 columns and 26 vertical levels, approximately 4-km horizontal resolution (MMF 4kmL26); and ii) 256 columns and 52 vertical levels, approximately 1-km horizontal resolution (MMF 1kmL52). We will refer to these MMF simulations as MMF 4 km and MMF 1 km. COSP was run on high-frequency instantaneous fields, and then 3-month average statistics of all the COSP outputs were produced. In the following subsections we show the results of this intercomparison for selected COSP diagnostics.

CloudSat reflectivities. The models' histograms of radar reflectivity are shown in Figs. 3a–d. Generally, the models exhibit some characteristics similar to the *CloudSat* observations, but there are many noticeable differences. The height–reflectivity relationship of the ice branch is present in the simulations, showing that models capture some basic microphysical relationships. The standard GCMs show simulated reflectivities larger than in the observations for levels in the midtroposphere. All models tend to overestimate the frequency of occurrence. The overestimation of reflectivities and frequency of occurrence is largest in LMDZ. This may be due to an overestimation of ice water content, effective particle size, or a combination of both.

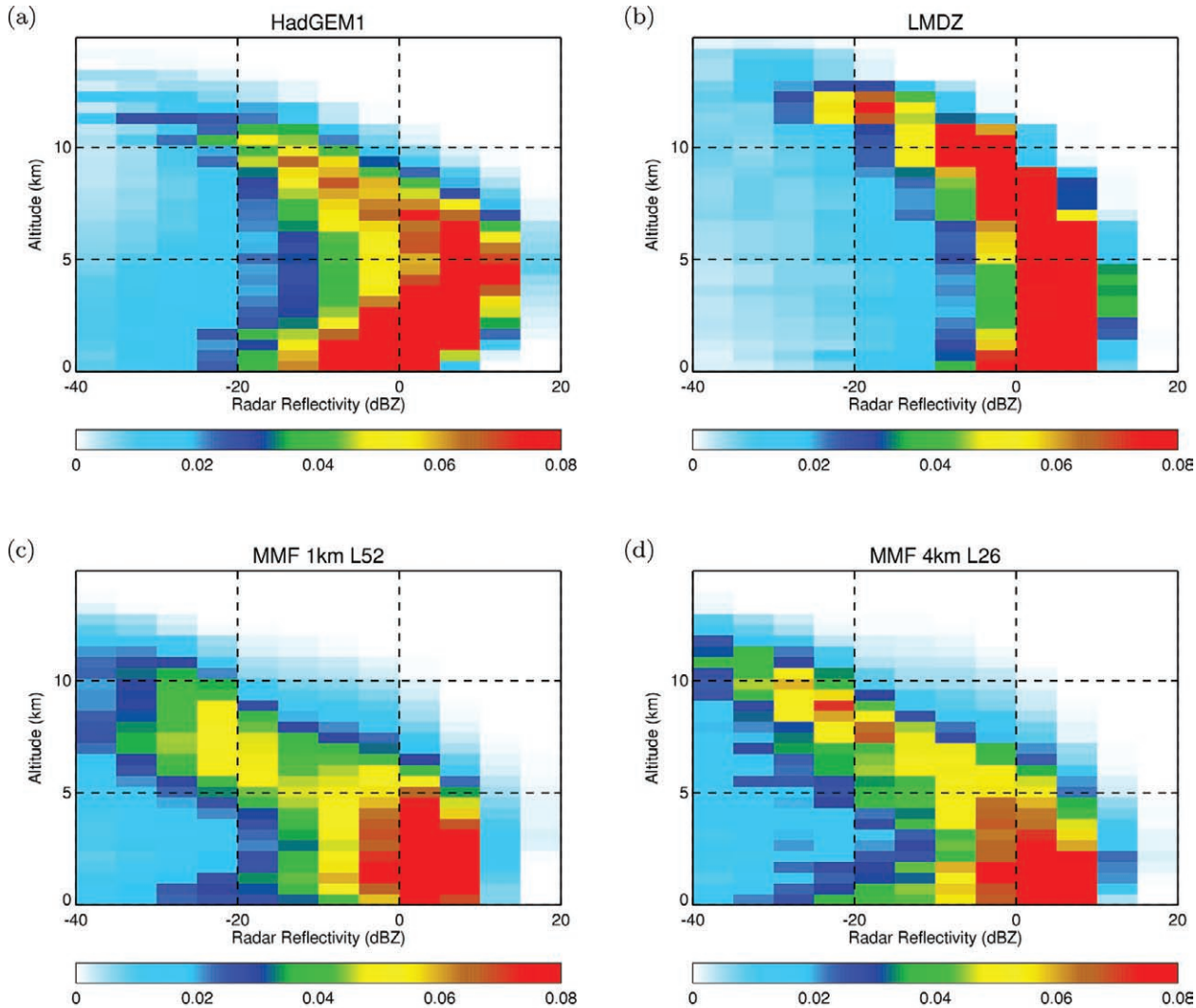
The proportion of nonprecipitating clouds (bottom left region of the histogram) is underestimated by all models. The precipitation cluster (bottom right corner, dB > −5) is overpopulated in all models, which indicates that the proportion of drizzling/raining cloud is overestimated compared to the observations. This behavior was also observed in the global forecast configuration of the Met Office model (Bodas-Salcedo et al. 2008) and other models (Stephens et al. 2010).

CALIPSO scattering ratios. Figure 4 shows the simulated histograms of scattering ratio as a function of height. The high-level cloud cluster of the observational histogram is broadly represented by all models, although with most of the population concentrated within a smaller range of SR values. The MMF 1 km does the best job in reproducing the frequency of occurrence of the high cloud mode. Figure 4b shows scattering ratios much larger in the LMDZ model than in the observations at high levels, consistent with the overestimation of the reflectivity in the ice branch (Fig. 3b). Because the lidar scattering ratio is influenced less by particle size than the radar reflectivity, this suggests an excess of condensate in the LMDZ model between 9 and12 km.

At levels below 3 km, models span a wide range of SR, although less than the observations. Also, the

Table 3. Models used in this study. In the case of the two versions of the MMF simulations, the number of columns refers to those used by the CRM embedded in each grid box of the parent GCM. CMMAP stands for Center for Multiscale Modeling of Atmospheric Processes. Two MMF simulations have been run with different resolutions of the CRM embedded in each GCM grid box: 4kmL26 and 1kmL52 (refer to text for details).

Model	Institution	References	Resolution
HadGEM1	Met Office	Martin et al. (2006)	1.25° lat × 1.875° lon, 38 levels
LMDZ4	LMD	Hourdin et al. (2006)	2.5° lat × 3.75° lon, 19 levels
MMF 4kmL26	CMMAP	Marchand and Ackerman (2010)	2° lat × 2.5° lon, 26 levels, 4 km
MMF 1kmL52	CMMAP	Marchand and Ackerman (2010)	2° lat × 2.5° lon, 52 levels, 1 km



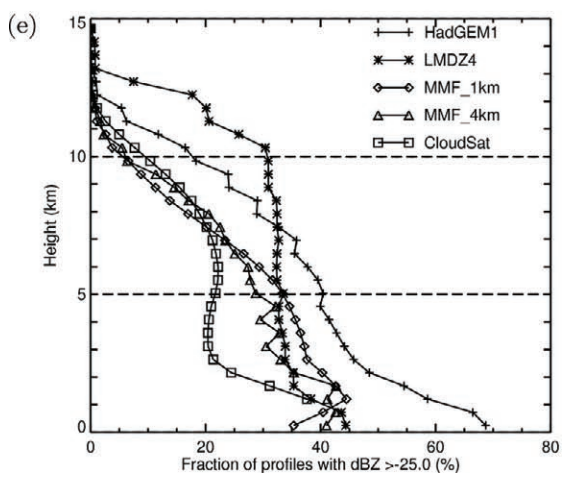
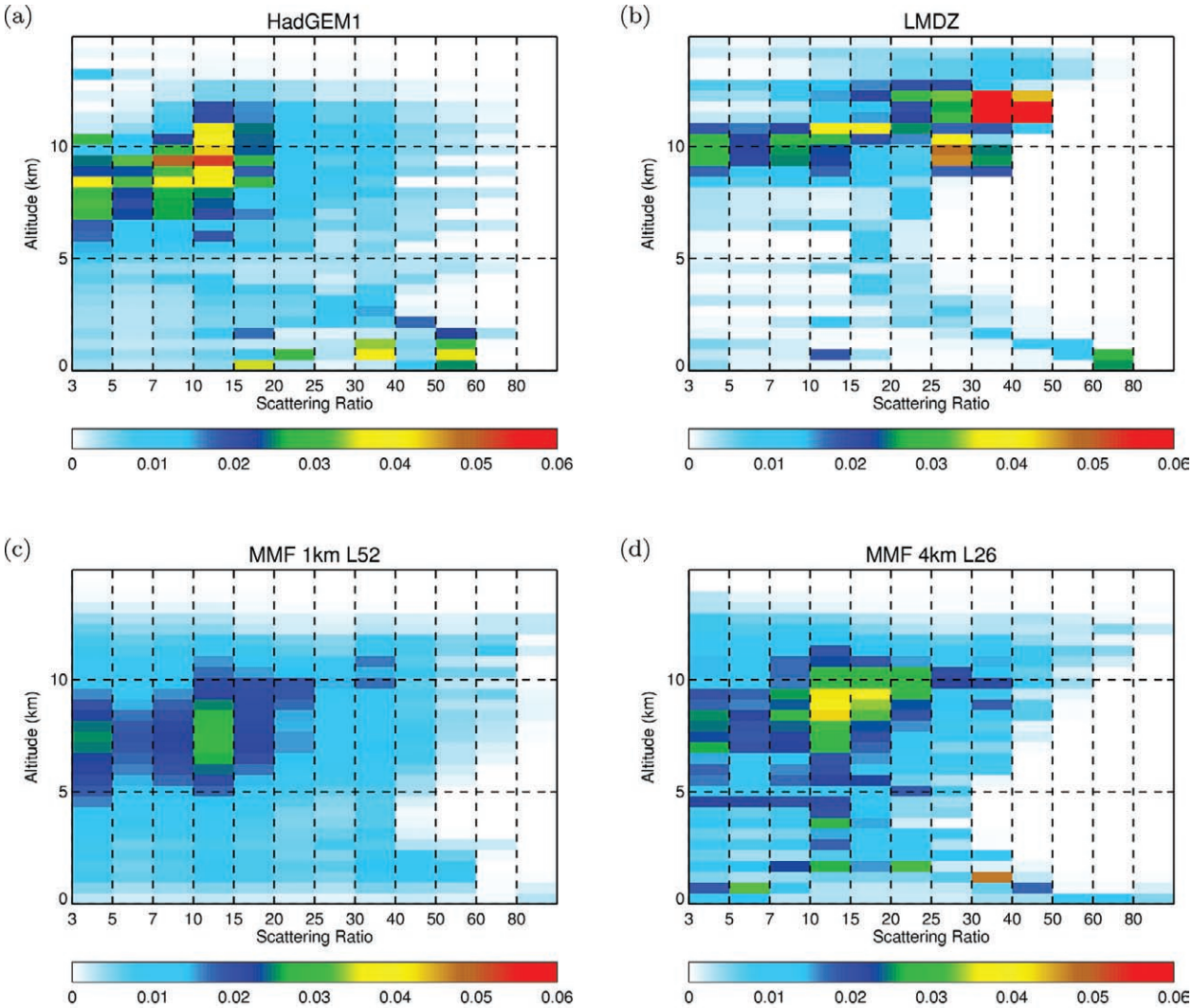


Fig. 3. Model histograms of radar reflectivity as a function of height (North Pacific region): (a) HadGEM1, (b) LMDZ, (c) MMF 1 km , (d) MMF 4 km , and (e) hydrometeor fraction as a function of height (fraction of samples with $Z > -25$ dB). The color scale represents the frequency of occurrence in adimensional units, from 0 to 1.

structure of the histogram at these levels is often noisier than in the observations.

The MMF 1 km is an exception, showing a closer agreement to the observations, which suggests that the increase in vertical resolution has a positive effect in the simulation of boundary layer cloud. The midlevel cloud with large SR is missing in all models, suggesting that they lack a good representation of congestus or mixed-phase layered clouds.

ISCCP, MISR and MODIS histograms. All models reproduce the vertically bimodal nature of the MISR histogram, as can be seen in Fig. 5. The high-top cloud maximum in optical depth is shown by all models but with model-specific differences. The HadGEM1 and MMF overestimate cloud fraction with cloud



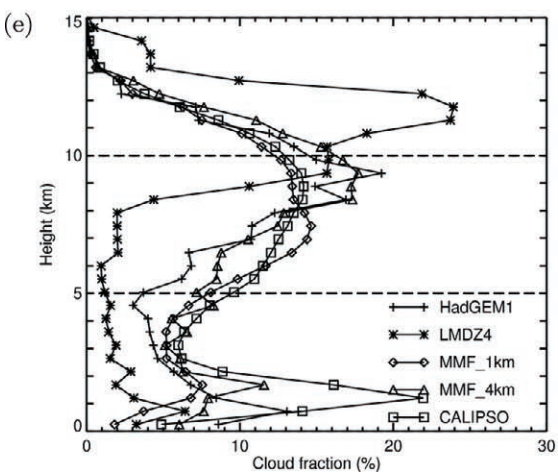
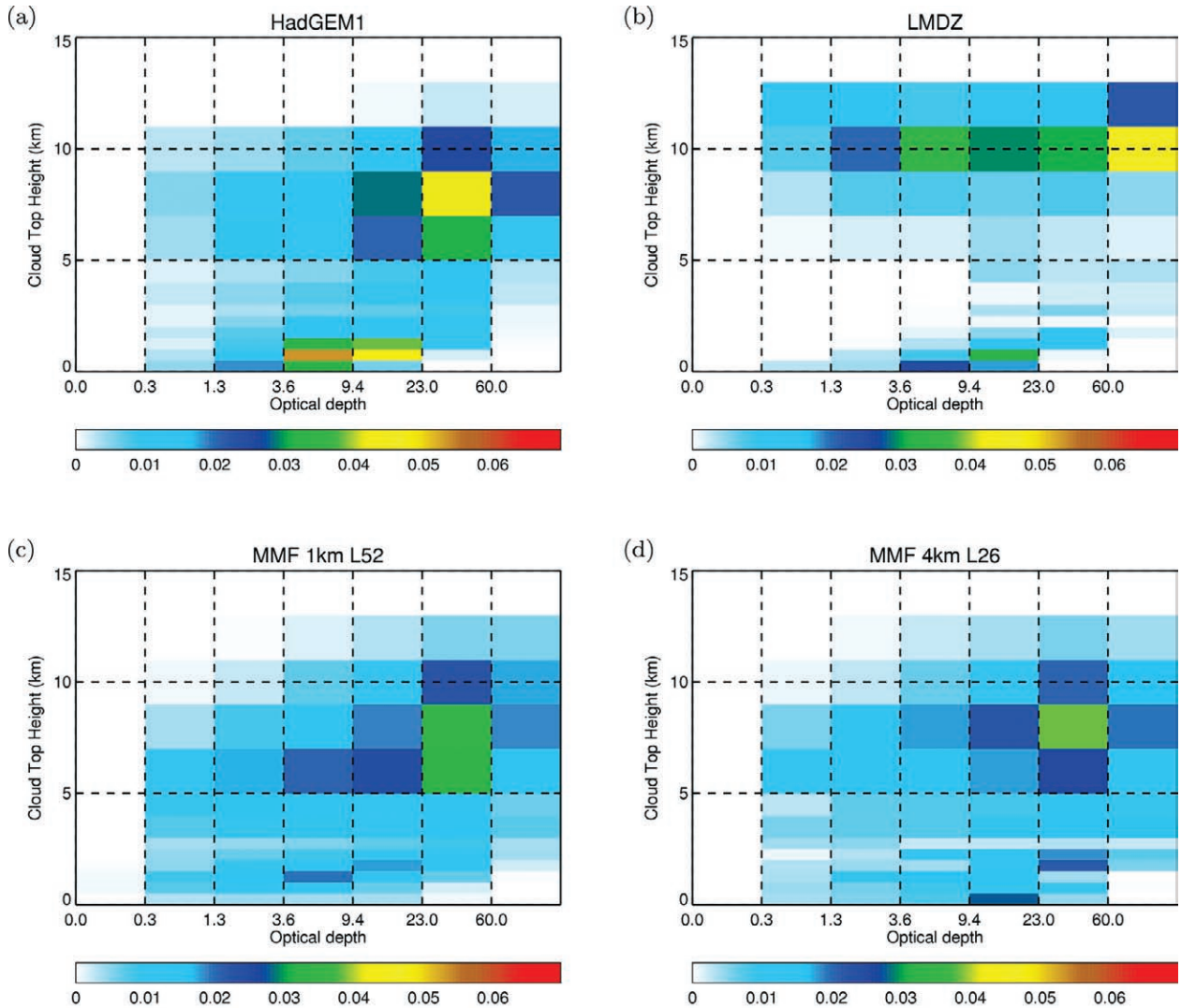


Fig. 4. As in Fig. 3, but for the histograms of scattering ratio as a function of height. (e) Fraction of samples with SR > 5 as a function of height. The color scale represents the frequency of occurrence in adimensional units, from 0 to 1.

tops between 5 and 10 km and underestimate cloud fraction with tops above 10 km. LMDZ overestimates the optical depth and the height of the high cloud (Figs. 5b and 6b). Both histograms show high cloud with optical depth greater than 23 in both cases, in line with the previous interpretation of the SR histograms. The total amount of cloud at those levels is also overestimated, as shown in Figs. 3e–6e.

Low-level cloud occurs less frequently than in the observations in all models. The optical depth distribution is narrower than in the observations, with a lack of thinner cloud and a complete absence of thicker cloud. The effect of small broken clouds is not represented in the simulators, so the distribution of optical depths for low clouds in the models are very similar for both MISR and ISCCP (Figs. 5 and 6), being concentrated between 3.6 and 23. The MMF simulations also show some significant population



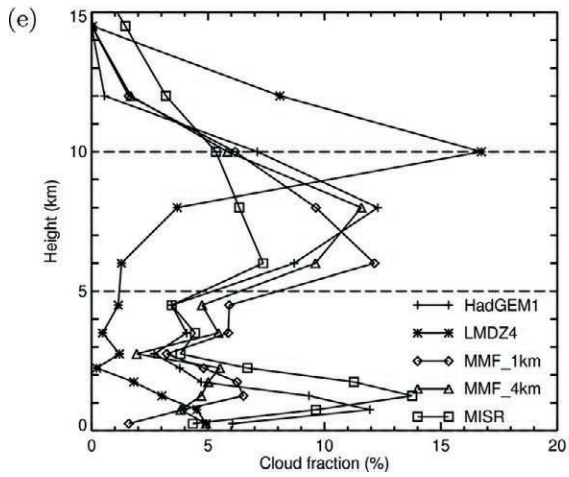


Fig. 5. As in Fig. 3, but for the MISR CTH–τ histograms. (e) Cloud fraction as a function of CTH (sum of all optical depths greater than 0.3). The color scale represents the cloud fraction in adimensional units, from 0 to 1.

outside this range. This is on the low side with respect to MISR and on the high side with respect to ISCCP. In the observations, the low optical depth region of these histograms often denotes the presence of pixels partially filled by clouds (Marchand et al. 2010). The reported optical depth will likely be biased low because the retrievals assume a pixel completely filled by a plane-parallel cloud. Because of the technical characteristics of the instruments and retrievals, the ISCCP retrievals seem to be affected more by this bias than MISR and MODIS (Marchand et al. 2010). These results then imply that models have insufficient cloud with optical depths greater than 23.

MISR and ISCCP report very different fractions of cloud with tops at midlevels, between 3 and 7 km or between 700 and 400 hPa (Figs. 5e and 6e). HadGEM1 and the MMF simulations show more cloud than

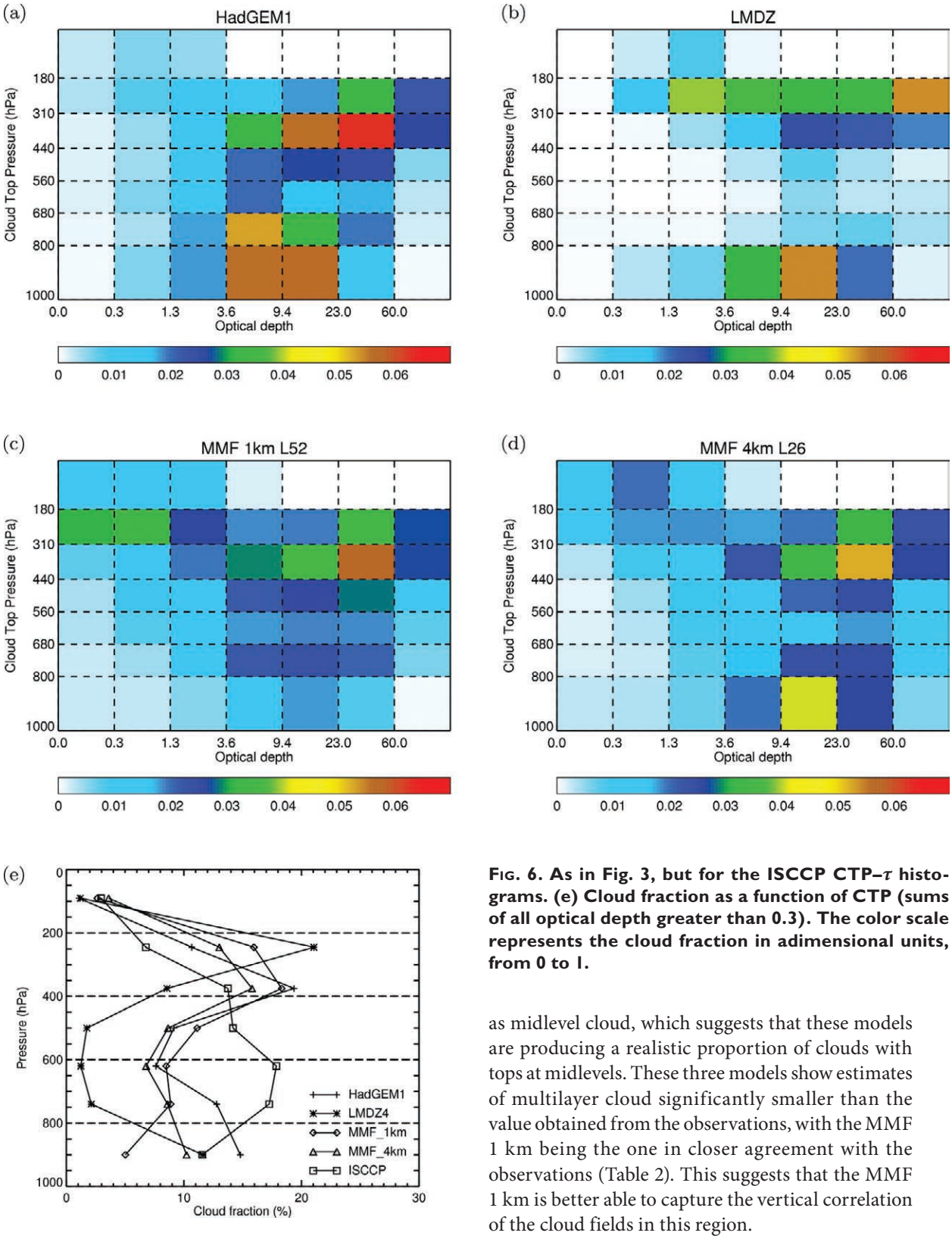


Fig. 6. As in Fig. 3, but for the ISCCP CTP–τ histograms. (e) Cloud fraction as a function of CTP (sums of all optical depth greater than 0.3). The color scale represents the cloud fraction in adimensional units, from 0 to 1.

MISR and less cloud than ISCCP. As we have shown in section 3, a substantial amount of high-level thin cloud above low-level cloud is misclassified by ISCCP as midlevel cloud, which suggests that these models are producing a realistic proportion of clouds with tops at midlevels. These three models show estimates of multilayer cloud significantly smaller than the value obtained from the observations, with the MMF 1 km being the one in closer agreement with the observations (Table 2). This suggests that the MMF 1 km is better able to capture the vertical correlation of the cloud fields in this region.

The results from LMDZ can be used as an example of the complementary information that the different datasets provide. This model shows a robust lack of cloud with tops at midlevels, according to MISR,

ISCCP, and CALIPSO. The estimate of multilayer cloud in LMDZ is very small (≈2%, Table 2). This is the result of having consistent MISR and ISCCP histograms, which show a strongly bimodal distribution of clouds: a mode of high-level thick cloud and a second mode of low-level cloud, with very little midlevel cloud. In addition to this, *CloudSat* simulations (Fig. 3e) show a nearly constant fraction of profiles with reflectivities greater than –25 dB below 10 km. This suggests that the entire high cloud in this model is precipitating, and there is little extra cloud at midlevels.

Figure 7 shows a comparison of simulated and observed ISCCP and MODIS CTP-τ histograms over a region in the South Pacific (60°–40°S, 180°–105°W). Diagnostics from the MODIS simulator were only available from HadGEM1, so only results for this model are shown. We choose this region because it highlights most of the differences between the retrievals. As shown in Fig. 7, ISCCP reports a large fraction of midlevel cloud. MODIS places most of this cloud in the high-level category (CTP < 440 hPa, Fig. 7e), because of the higher sensitivity of the MODIS CO_2 slicing method in these conditions. MODIS shows less low-level cloud (CTP > 680 hPa) than ISCCP, especially in the optically thin bins ($\tau < 3.6$). MODIS cloud properties retrievals are not applied to all the pixels identified as cloudy by the cloud mask. Additional screening tests are applied to filter out broken clouds, cloud edges, and heavy aerosol or sunglint situations. The cloud fraction reported by the MODIS cloud mask is very similar to the one reported by ISCCP in this region (0.83). However, the additional screening in the MODIS retrievals reduces the fraction to 0.71. This reduction is largely due to the MODIS histogram containing less optically thin cloud (Fig. 7f), as previously reported by Marchand et al. (2010). MODIS reports the same amount of cloud as ISCCP with optical depths between 3.6 and 9.4, but the vertical distribution is very different. In addition to reporting more high-level cloud, MODIS also reports more low-level cloud. This suggests that part of the cloud reported as midlevel in ISCCP is reported as low-level cloud by MODIS. This may be due to a systematic difference in the cloud-top pressure estimates between ISCCP and MODIS for clouds with tops in the lower 3–4 km of the troposphere. If there are clouds with tops around 3 km, then small differences in CTP retrievals may place clouds above or below the pressure level that separates low-level cloud from midlevel cloud (680 hPa). MISR CTH-τ histograms for the same region in the South Pacific (not shown) agree with MODIS. The simulators (Figs. 7a and 7b) capture some of the specific characteristics of the retrievals but not all. The ISCCP simulator reports more midlevel cloud because the cloud-top pressure estimates of high clouds put clouds at lower levels, consistent with the observations. The effect of broken cloud and the additional cloud screening applied by the MODIS retrievals is not represented in the simulators, so the distribution of optical depths from both simulators is nearly identical (Fig. 7f). This analysis illustrates again the synergy between different observations and how different simulators may reveal different aspects of the model simulations.

UPPER-TROPOSPHERIC HUMIDITY. As an example of RTTOV diagnostics, we use upper-tropospheric humidity (UTH; Fig. 8). UTH is a key climate variable that plays a very important role in Earth's radiation budget (e.g., Kiehl and Briegleb 1992; Ingram 2010). However, UTH is not well simulated by current climate models (Pierce et al. 2006; John and Soden 2007). Observed UTH is derived using the 183.31 ± 1.00-GHz channel of the Advanced Microwave Sounding Unit-B (AMSU-B), which is a scanning microwave radiometer on board National Oceanic and Atmospheric Administration (NOAA) operational satellites. This channel is sensitive to layers approximately between 500 and 200 hPa. Soden and Bretherton (1993) demonstrate that the natural logarithm of UTH is linearly related to the clear-sky brightness temperature (T_B) measured at a strong water vapor absorption line:

$$ln\,(\mathrm{UTH}) = a + b \times T_B \qquad (1)$$

The constants a and b are regression coefficients. These coefficients depend on the viewing angle and were derived by Buehler and John (2005) using a diverse atmospheric profile dataset (Chevallier et al. 2006). UTH data are limb corrected (John et al. 2006), and the data have been validated using high-quality radiosonde and satellite measurements (Milz et al. 2009). At this frequency, microwave measurements are compromised only by precipitating/deep convective clouds, and measurements contaminated by such clouds are filtered out. Because cloudiness is correlated with humidity, filtering out the cloudy/high humidity values could introduce a 2%–3% RH bias in the observed UTH in convectively active areas (Buehler et al. 2008).

Observed UTH by *NOAA-18* for SON 2006 is shown in Fig. 8a. Figure 8b shows the modeled UTH, which is computed using simulated RTTOV brightness temperatures following the methodology

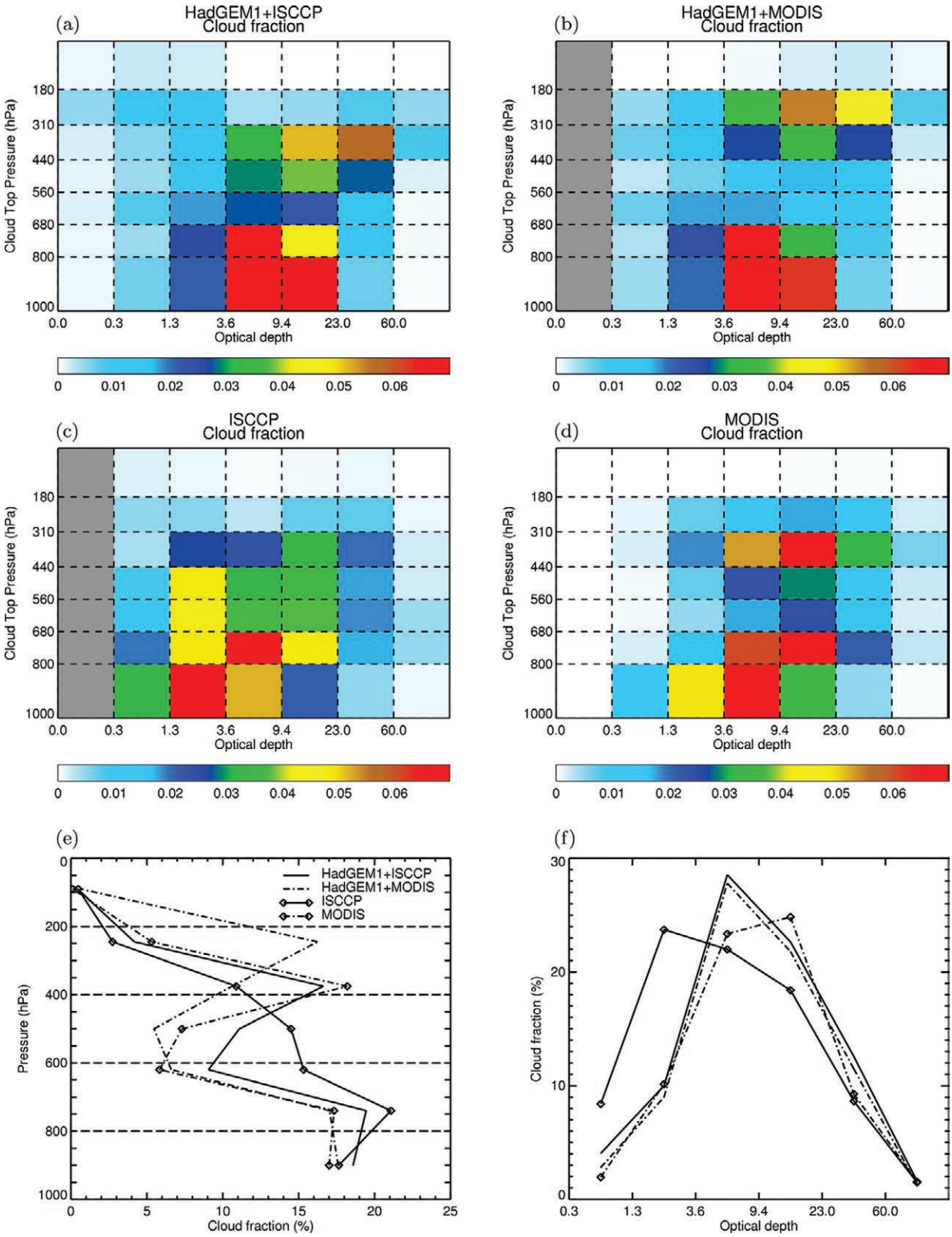


Fig. 7. Simulated (HadGEM1) and observed ISCCP and MODIS diagnostics over a region in the South Pacific (60°–40°S, 180°–105°W). CTP–τ histograms for (a) ISCCP simulator results from HadGEM1, (b) MODIS simulator results from HadGEM1, (c) ISCCP observations, and (d) MODIS observations. (e) Cloud fraction as a function of CTP (sums of all optical depth greater than 0.3), and (f) fraction as a function of τ (sum of all the CTP layers for each optical depth bin). The color scale represents the cloud fraction in adimensional units, from 0 to 1.

applied to the observations but with two differences. First, the model brightness temperatures are only simulated for nadir angle, so only one pair of *a* and *b* coefficients is used. Second, the precipitating and deep convective events are not filtered out, which will introduce a small moist bias with respect to the observations in convectively active regions. Over elevated areas—for example, the Himalayas—the atmosphere is so dry that the measured or modeled UTH can be contaminated by the surface. To filter out surface contaminated measurements, we use a radiance difference (radiance of 183.31 ± 3.00-GHz channel, which is a midtropospheric humidity channel, minus the radiance of 183.31 ± 1.00 GHz) that would be positive under normal conditions but negative for surface-contaminated scenes. This again highlights the utility of the simulator approach, as the same filtering process can be applied to the observed and modeled datasets.

There are a few noticeable differences between the observed and simulated UTH. In the model the Indo-Pacific warm-pool high-humidity region is more humid and is shifted eastward, and the model is more humid in the midlatitudes. Figure 8c shows the difference in modeled UTH compared to the observed UTH. In the tropics, in general, the differences are larger; for example, over southern India, the model has a dry bias of more than 20% RH and is up to 15% more humid over the midlatitude oceans. The differences may be partly due to mismatches of dynamical regimes in the model and observation. For example, Brogniez and Pierrehumbert (2007) have shown by partitioning modeled and measured UTH based on dynamical regimes that the models are in better agreement with the observations.

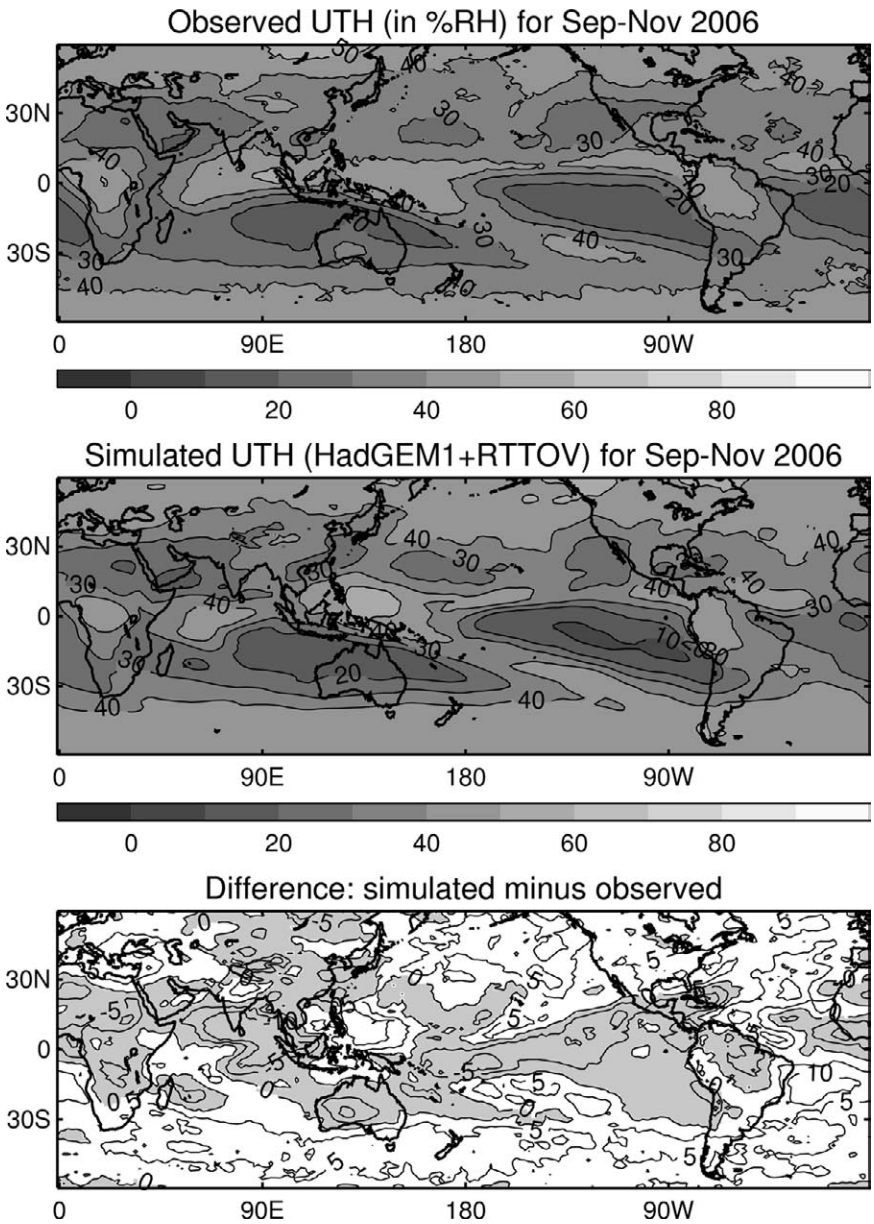


Fig. 8. UTH (RH, in %) from Sep to Nov 2006: (a) retrieved by *NOAA-18*, (b) simulated by HadGEM1 using RTTOV, and (c) absolute difference between modeled and observed UTH (negative values are shaded gray).

CONCLUSIONS AND FUTURE PLANS. The Cloud Feedback Model Intercomparison Project (CFMIP) Observation Simulator Package (COSP) is a flexible software tool that enables the simulation of data from several satelliteborne active and passive sensors from model variables. The flexibility of COSP and a common interface for all sensors facilitates its use in any type of numerical model, from high-resolution cloud-resolving models to

the coarser-resolution GCMs. This paper describes the main capabilities of COSP and presents an intercomparison of four different models. It demonstrates that COSP facilitates the evaluation of models against observations and comparisons between them in a more consistent manner. Two of the models used in this study are GCMs used in the latest IPCC assessment reports, and the others are two versions of the multiscale modeling framework (MMF) with different horizontal and vertical resolutions of the embedded cloud-resolving model. These models have been run in atmosphere-only mode, and COSP outputs have been extracted for one season, from September to November 2006. We focus the intercomparison on a region in the North Pacific, characterized by a rich mixture of different cloud types. This intercomparison shows how COSP can be applied to different types of models in a transparent way.

The *CloudSat* radar has the capability of providing information about the vertical distribution of clouds, and because of its sensitivity to large particles, it provides information on the cloud and precipitation microphysics. Frequency histograms of *CloudSat* radar reflectivity as a function of height show that all models capture the basic features of these observational histograms, but they also show some common biases that seem to be present across models. They lack a mode of low-level nonprecipitating cloud.

The CALIPSO lidar also provides information on the vertical distribution of clouds. The lidar is able to observe clouds with small particle sizes (characteristic of some boundary layer clouds and some thin cirrus) that go undetected by the *CloudSat* radar. Because of its visible wavelength, the signal is fully attenuated for liquid clouds or thick ice clouds. This makes this information complementary to that provided by *CloudSat* (Delanoë and Hogan 2010). Frequency distributions of scattering ratio as a function of height suggest that all models lack a good representation of congestus and mixed-phase layered clouds.

Cloud-top pressure/height versus optical depth, as derived from passive imagers (ISCCP, MISR, and MODIS), show noticeable differences that arise from the different algorithms used to derive the cloud properties (pressure/height, optical depth). The impact of the different cloud property retrievals is illustrated by the comparisons. The differences in these observational retrievals highlight the importance of the simulator approach in model evaluation.

In general, the MMF simulations seem to perform better than the other climate models in all the diagnostics, showing a distribution of hydrometeors in the vertical and also in optical depth closer to the observations. This suggests that a more complex microphysical scheme and resolving finer scales help when simulating cloud processes.

The combination of all these diagnostics tells us more about the representation of clouds and precipitation in models than any of them separately. This demonstrates the value of multiple simulators when making quantitative comparisons of model outputs with observations.

The comparison undertaken here comprises only a single 3-month period and is intended only to highlight the type of comparison that can be undertaken using COSP. Comparisons with longer datasets and more models will be undertaken as part of CFMIP2 and will make clear to what degree the differences between models and between models and observations identified here are robust.

One important aspect that we have not considered in these comparisons is the impact of mismatch between meteorological regimes in the observations and models due to positional errors (e.g., the displacement of the storm tracks). This type of error may be responsible for some of the differences between the models and the observations, or may it be compensating for cloud physical errors. The fact that we have applied this methodology to atmosphere-only runs, where the SSTs are prescribed, helps to minimize this concern. There are methodologies that can be used to minimize this problem, that stratify the data into regimes by using compositing (e.g., Bony et al. 2004; Field et al. 2008) or clustering techniques (e.g., Jakob and Tselioudis 2003; Williams and Webb 2009; Zhang et al. 2010). Under CFMIP2, there are plans to apply these methodologies to a wide range of models using COSP.

COSP is a community tool that is distributed under the BSD license. It can be downloaded from the CFMIP Web site (www.cfmip.net). We welcome the involvement of the external community in the development of new capabilities for COSP, such as new diagnostics for existing instruments and the development of new modules for other instruments.

There are already plans for the development of new diagnostics, such as information on the vertical correlation of hydrometeors from the profiles of radar reflectivity. Future versions of the MISR simulator will include pixel alignment and resolution capability. The radar simulator will include support for two-moment microphysics as well as better support for different microphysical settings.

ACKNOWLEDGMENTS. A. Bodas-Salcedo, M. J. Webb, and V. O. John were supported by the Joint DECC/Defra

Met Office Hadley Centre Climate Programme (GA01101). Support for S. A. Klein and Y. Zhang was provided by the Regional and Global Climate Modeling and Earth System Modeling Programs of the Office of Science at the U.S. Department of Energy with additional support from the NASA Modeling, Analysis, and Prediction Program. The contribution of S. A. Klein and Y. Zhang to this work was performed under the auspices of the U.S. Department of Energy by the Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. This work benefited from the financial support of the French LEFE project MISSTERRE. This work was partly funded by the Seventh Framework Programme project EUCLIPSE. The authors would like to thank the Department of Energy Pacific Northwest National Laboratory (PNNL), John Helly and Dave Randall at the Center for Multiscale Modeling of Atmospheric Processes (CMMAP), and the San Diego Supercomputing Center (SDSC) for providing the computer time and support needed to run the MMF climate model. Funding for this research was supported in part by the NASA Jet Propulsion Laboratory (Contract NMO710860). R. Pincus was supported by NASA Grant NNX08AD65G. *CloudSat* data were obtained from the *CloudSat* Data Processing Center (http://cloudsat.cira.colostate.edu). CALIPSO-GOCCP data were obtained from the IPSL ClimServ/ICARE data center (http://climserv.ipsl.polytechnique.fr/cfmip-obs.html). ISCCP data were obtained from the NASA Langley Research Center Atmospheric Science Data Center (http://eosweb.larc.nasa.gov). Thanks to those members of the NASA Langley Research Center and NASA Jet Propulsion Laboratory who are generating the MISR CTH-OD dataset, including Jeff Walters, Mike Smyth, Alex Menzies, and Catherine Moroney. MODIS data were obtained from the level 1 and Atmosphere Archive and Distribution System (LAADS), NASA Goddard Space Flight Center. We thank T. H. Stein, M. Woodage, R. Forbes, C. Nam, M. Satoh, and M. Ringer for their helpful comments and feedback. We thank the three anonymous reviewers, who helped to improve the quality of the paper.

REFERENCES

Allan, R. P., M. A. Ringer, and A. Slingo, 2003: Evaluation of moisture in the Hadley Centre climate model using simulations of HIRS water-vapour channel radiances. *Quart. J. Roy. Meteor. Soc.,* **129,** 3371–3389, doi:10.1256/qj.02.217.

——, A. Slingo, S. F. Milton, and M. E. Brooks, 2007: Evaluation of the Met Office global forecast model using Geostationary Earth Radiation Budget (GERB) data. *Quart. J. Roy. Meteor. Soc.,* **133,** 1993–2010, doi:10.1002/qj.166.

Battaglia, A., J. M. Haynes, T. L'Ecuyer, and C. Simmer, 2008: Identifying multiple-scattering-affected profiles in CloudSat observations over the oceans. *J. Geophys. Res.,* **113,** D00A17, doi:10.1029/2008JD009960.

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

Bony, S., J.-L. Dufresne, H. Le Treut, J.-J. Morcrette, and C. A. Senior, 2004: On dynamic and thermodynamic components of cloud changes. *Climate Dyn.,* **22,** 71–86, doi:10.1007/s00382-003-0369-6.

Brogniez, H., and R. T. Pierrehumbert, 2007: Intercomparison of tropical tropospheric humidity in GCMs with AMSU-B water vapor data. *Geophys. Res. Lett.,* **34,** L17812, doi:10.1029/2006GL029118.

——, R. Roca, and L. Picon, 2005: Evaluation of the distribution of subtropical free tropospheric humidity in AMIP-2 simulations using METEOSAT water vapor channel data. *Geophys. Res. Lett.,* **32,** L19708, doi:10.1029/2005GL024341.

Buehler, S. A., and V. O. John, 2005: A simple method to relate microwave radiances to upper tropospheric humidity. *J. Geophys. Res.,* **110,** D02110, doi:10.1029/2004JD005111.

——, M. Kuvatov, V. O. John, M. Milz, B. J. Soden, D. L. Jackson, and J. Notholt, 2008: An upper tropospheric humidity data set from operational satellite microwave data. *J. Geophys. Res.,* **113,** D14110, doi:10.1029/2007JD009314.

Chahine, M. T., and Coauthors, 2006: AIRS: Improving weather forecasting and providing new data on greenhouse gases. *Bull. Amer. Meteor. Soc.,* **87,** 911–926.

Chepfer, H., M. Chiriaco, R. Vautard, and J. Spinhirne, 2007: Evaluation of MM5 optically thin clouds over Europe in fall using ICESat lidar spaceborne observations. *Mon. Wea. Rev.,* **135,** 2737–2753.

——, 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.

——, ——, ——, G. Cesana, J.-L. Dufresne, P. Minnis, C. J. Stubenrauch, and S. Zeng, 2010: The GCM-oriented CALIPSO Cloud Product (CALIPSO-GOCCP). *J. Geophys. Res.,* **115,** D00H16, doi:10.1029/2009JD012251.

Chevallier, F., S. Di Michele, and A. P. McNally, 2006: Diverse profile datasets from the ECMWF 91-level short-range forecasts. NWP SAF Document NWP-SAF-EC-TR-010, 16 pp.

Chiriaco, M., R. Vautard, H. Chepfer, M. Haeffelin, J. Dudhia, Y. Wanherdrick, Y. Morille, and A. Protat,

2006: Ability of MM5 to simulate ice clouds: Systematic comparison between simulated and measured fluxes and lidar/radar profiles at the SIRTA atmospheric observatory. *Mon. Wea. Rev.,* **134,** 897–918.

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.

Diner, D. J., J. C. Beckert, T. H. Reilly, C. J. Bruegge, J. E. Conel, and R. Kahn, 1998: Multi-Angle Imaging Spectroradiometer (MISR) description and experiment overview. *IEEE Trans. Geosci. Remote Sens.,* **36,** 1072–1087.

——, and Coauthors, 2005: The value of multiangle measurements for retrieving structurally and radiatively consistent properties of clouds, aerosols, and surfaces. *Remote Sens. Environ.,* **97,** 495–518, doi:10.1016/j.rse.2005.06.006.

Evan, A. T., A. K. Heidinger, and D. J. Vimont, 2007: Arguments against a physical long-term trend in global ISCCP cloud amounts. *Geophys. Res. Lett.,* **34,** L04701, doi:10.1029/2006GL028083.

Eyre, J. R., 1987: On systematic and their climatological mean values. *Quart. J. Roy. Meteor. Soc.,* **113,** 279–292, doi:10.1002/qj.49711347516.

Field, P. R., A. Gettelman, R. B. Neale, R. Wood, P. J. Rasch, and H. Morrison, 2008: Midlatitude cyclone compositing to constrain climate model behavior using satellite observations. *J. Climate,* **21,** 5887–5903.

——, A. Bodas-Salcedo, and M. E. Brooks, 2011: Using model analysis and satellite data to assess cloud and precipitation in midlatitude cyclones. *Quart. J. Roy. Meteor. Soc.,* in press.

Gates, W. L., and Coauthors, 1999: An overview of the results of the Atmospheric Model Intercomparison Project (AMIP I). *Bull. Amer. Meteor. Soc.,* **80,** 29–55.

Gleckler, P. J., K. E. Taylor, and C. Doutriaux, 2008: Performance metrics for climate models. *J. Geophys. Res.,* **113,** D06104, doi:10.1029/2007JD008972.

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.

Hogan, R. J., A. J. Illingworth, E. J. O'Connor, and J. P. V. Poiares Baptista, 2003: Characteristics of mixed-phase clouds. II: A climatology from ground-based lidar. *Quart. J. Roy. Meteor. Soc.,* **129,** 2117–2134, doi:10.1256/qj.01.209.

Hourdin, F., and Coauthors, 2006: The LMDZ4 general circulation model: Climate performance and sensitivity to parametrized physics with emphasis on tropical convection. *Climate Dyn.,* **27,** 787–813, doi:10.1007/s00382-006-0158-0.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, and S. A. Clough, 2003: Evaluation of upper tropospheric water vapor in the NCAR Community Climate Model (CCM3) using modeled and observed HIRS radiances. *J. Geophys. Res.,* **108,** 4037, doi:10.1029/2002JD002539.

Im, E., C. Wu, and S. L. Durden, 2005: Cloud profiling radar for the CloudSat mission. *IEEE Trans. Aerosp. Electron. Syst.,* **20,** 15–18, doi:10.1109/MAES.2005.1581095.

Ingram, W., 2010: A very simple model for the water vapour feedback on climate change. *Quart. J. Roy. Meteor. Soc.,* **136,** 30–40, doi:10.1002/qj.546.

Jakob, C., and G. Tselioudis, 2003: Objective identification of cloud regimes in the tropical western Pacific. *Geophys. Res. Lett.,* **30,** 2082, doi:10.1029/2003GL018367.

John, V. O., and B. J. Soden, 2007: Temperature and humidity biases in global climate models and their impact on climate feedbacks. *Geophys. Res. Lett.,* **34,** L18704, doi:10.1029/2007GL030429.

——, S. A. Buehler, and N. Courcoux, 2006: A cautionary note on the use of Gaussian statistics in satellite-based UTH climatologies. *IEEE Geosci. Remote Sens. Lett.,* **3,** 130–134, doi:10.1109/LGRS.2005.859350.

Kiehl, J. T., and B. P. Briegleb, 1992: Comparison of the observed and calculated clear sky greenhouse effect: Implications for climate studies. *J. Geophys. Res.,* **97,** 10 037–10 049.

King, M. D., and Coauthors, 2003: Cloud and aerosol properties, precipitable water, and profiles of temperature and water vapor from MODIS. *IEEE Trans. Geosci. Remote Sens.,* **41,** 442–458, doi:10.1109/TGRS.2002.808226.

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

Marchand, R., and T. Ackerman, 2010: An analysis of cloud cover in multiscale modeling framework global climate model simulations using 4 and 1 km horizontal grids. *J. Geophys. Res.,* **115,** D16207, doi:10.1029/2009JD013423.

——, G. G. Mace, T. Ackerman, and G. L. Stephens, 2008: Hydrometeor detection using Cloudsat—An earth-orbiting 94-GHz cloud radar. *J. Atmos. Oceanic Technol.,* **25,** 519–533.

——, J. Haynes, G. G. Mace, T. Ackerman, and G. Stephens, 2009: A comparison of simulated cloud radar output from the multiscale modeling

framework global climate model with CloudSat cloud radar observations. *J. Geophys. Res.,* **114,** D00A20, doi:10.1029/2008JD009790.

——, T. Ackerman, M. Smythe, and W. B. Rossow, 2010: A review of cloud top height and optical depth histograms from MISR, ISCCP and MODIS. *J. Geophys. Res.,* **115,** D16206, doi:10.1029/2009JD013422.

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.

Masunaga, H., and Coauthors, 2010: Satellite Data Simulator Unit: A multisensor, multispectral satellite simulator package. *Bull. Amer. Meteor. Soc.,* **91,** 1625–1632.

Meehl, G. A., and Coauthors, 2007: Global climate projections. *Climate Change 2007: The Physical Science Basis,* S. Solomon et al., Eds., Cambridge University Press, 747–845.

Milz, M., S. A. Buehler, and V. O. John, 2009: Comparison of AIRS and AMSU-B monthly mean estimates of upper tropospheric humidity. *Geophys. Res. Lett.,* **36,** L10804, doi:10.1029/2008GL037068.

Morcrette, J. J., 1991: Evaluation of model-generated cloudiness: Satellite-observed and model-generated diurnal variability of brightness temperature. *Mon. Wea. Rev.,* **119,** 1205–1224.

Moroney, C., R. Davies, and J.-P. Muller, 2002: Operational retrieval of cloud-top heights using MISR data. *IEEE Trans. Geosci. Remote Sens.,* **40,** 1532–1540, doi:10.1109/TGRS.2002.801150.

Muller, J.-P., A. Mandanayake, C. Moroney, R. Davies, D. J. Diner, and S. Paradise, 2002: MISR stereoscopic image matchers: Techniques and results. *IEEE Trans. Geosci. Remote Sens.,* **40,** 1547–1559, doi:10.1109/TGRS.2002.801160.

Pierce, D. W., T. P. Barnett, E. J. Fetzer, and P. J. Gleckler, 2006: Three-dimensional tropospheric water vapor in coupled climate models compared with observations from the AIRS satellite system. *Geophys. Res. Lett.,* **33,** L21701, doi:10.1029/2006GL027060.

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:10.1029/2002JD003322.

——, R. Hemler, and S. A. Klein, 2006: Using stochastically generated subcolumns to represent cloud structure in a large-scale model. *Mon. Wea. Rev.,* **134,** 3644–3656.

——, C. P. Batstone, R. J. Patrick-Hofmann, K. E. Taylor, and P. E. Gleckler, 2008: Evaluating the present-day simulation of clouds, precipitation and radiation in climate models. *J. Geophys. Res.,* **133,** D14209, doi:10.1029/2007JD009334.

Platnick, S., M. D. King, S. A. Ackerman, W. P. Menzel, B. A. Baum, J. C. Riedi, and R. A. Frey, 2003: The MODIS cloud products: Algorithms and examples from Terra. *IEEE Trans. Geosci. Remote Sens.,* **41,** 459–473, doi:10.1109/TGRS.2002.808301.

Räisänen, P., H. W. Barker, M. F. Khairoutdinov, J. Li, and D. A. Randall, 2004: Stochastic generation of subgrid-scale cloudy columns for large-scale models. *Quart. J. Roy. Meteor. Soc.,* **130,** 2047–2067, doi:10.1256/qj.03.99.

Randel, D. L., T. H. Vonder Haar, M. A. Ringerud, G. L. Stephens, T. J. Greenwald, and C. L. Combs, 1996: A new global water vapor dataset. *Bull. Amer. Meteor. Soc.,* **77,** 1233–1246.

Ringer, M. A., J. M. Edwards, and A. Slingo, 2003: Simulation of satellite channel radiances in the Met Office Unified Model. *Quart. J. Roy. Meteor. Soc.,* **129,** 1169–1190, doi:10.1256/qj.02.61.

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

Salathé, E. P., and D. Chesters, 1995: Variability of moisture in the upper troposphere as inferred from TOVS satellite observations and the ECMWF model analyses in 1989. *J. Climate,* **8,** 120–132.

Saunders, R., M. Matricardi, and P. Brunel, 1999: An improved fast radiative transfer model for assimilation of satellite radiance observations. *Quart. J. Roy. Meteor. Soc.,* **125,** 1407–1425.

Sauvageot, H., 1992: *Radar Meteorology.* Artech House Inc., 366 pp.

Schaaf, C. B., and Coauthors, 2002: First operational BRDF, albedo nadir reflectance products from MODIS. *Remote Sens. Environ.,* **83,** 135–148.

Slingo, J. M., 1980: A cloud parametrization scheme derived from GATE data for use with a numerical model. *Quart. J. Roy. Meteor. Soc.,* **106,** 747–770.

Smith, R. N. B., 1990: A scheme for predicting layer clouds and their water content in a general circulation model. *Quart. J. Roy. Meteor. Soc.,* **116,** 435–460.

Soden, B. J., and F. P. Bretherton, 1993: Upper tropospheric relative humidity from the GOES 6.7 μm channel: Method and climatology for July 1987. *J. Geophys. Res.,* **98** (D9), 16 669–16 688.

Stephens, G. L., and C. D. Kummerow, 2007: The remote sensing of clouds and precipitation from space: A review. *J. Atmos. Sci.,* **64,** 3742–3765.

——, and Coauthors, 2002: The CloudSat mission and the A-Train. *Bull. Amer. Meteor. Soc.,* **83,** 1771–1790.

——, and Coauthors, 2010: Dreary state of precipitation in global models. *J. Geophys. Res.,* **115,** D24211, doi:10.1029/2010JD014532.

Voors, R., and Coauthors, 2007: ECSIM: The simulator framework for EarthCARE. *Sensors, Systems, and Next-Generation Satellites XI,* R. Meynart et al., Eds., International Society for Optical Engineering (SPIE Proceedings, Vol. 6744), doi:10.1117/12.737738.

Warren, S. G., C. J. Hahn, J. London, R. M. Chervin, and R. L. Jenne, 1988: Global distribution of total cloud cover and cloud type amounts over the ocean. U.S. DOE Tech. Rep. DOE/ER-0406 and NCAR Tech. Note NCAR/TN-317-STR, 170 pp.

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. *Climate Dyn.,* **17,** 905–922.

Wilkinson, J. M., R. J. Hogan, A. J. Illingworth, and A. Benedetti, 2008: Use of a lidar forward model for global comparisons of cloud fraction between the ICESat lidar and the ECMWF model. *Mon. Wea. Rev.,* **136,** 3742–3759.

Williams, K. D., and G. Tselioudis, 2007: GCM intercomparison of global cloud regimes: Present-day evaluation and climate change response. *Climate Dyn.,* **29,** 231–250, doi:10.1007/s00382-007-0232-2.

——, and M. J. Webb, 2009: A quantitative performance assessment of cloud regimes in climate models. *Climate Dyn.,* **33,** 141–157, doi:10.1007/s00382-008-0443-1.

Winker, D. M., and Coauthors, 2010: The CALIPSO mission: A global 3D view of aerosols and clouds. *Bull. Amer. Meteor. Soc.,* **91,** 1211–1229.

Wylie, D. P., and W. P. Menzel, 1999: Eight years of high cloud statistics using HIRS. *J. Climate,* **12,** 170–184.

Yu, W., M. Doutriaux, G. Sèze, H. Treut, and M. Desbois, 1996: A methodology study of the validation of clouds in GCMs using ISCCP satellite observations. *Climate Dyn.,* **12,** 389–401, doi:10.1007/BF00211685.

Zhang, M. H., and Coauthors, 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., and Coauthors, 2008: On the diurnal cycle of deep convection, high-level cloud, and upper troposphere water vapor in the multiscale modeling framework. *J. Geophys. Res.,* **113,** D16105, doi:10.1029/2008JD009905.

——, 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.