Reconciling Simulated and Observed Views of Clouds: MODIS, ISCCP, and the Limits of Instrument Simulators

ROBERT PINCUS

University of Colorado, and NOAA/Earth System Research Laboratory, Physical Sciences Division, Boulder, Colorado

STEVEN PLATNICK

Earth Sciences Division, NASA Goddard Space Flight Center, Greenbelt, Maryland

STEVEN A. ACKERMAN

Department of Atmospheric and Oceanic Sciences, and Cooperative Institute for Meteorological Satellite Studies, University of Wisconsin—Madison, Madison, Wisconsin

RICHARD S. HEMLER

NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

ROBERT J. PATRICK HOFMANN

University of Colorado, and NOAA/Earth System Research Laboratory, Physical Sciences Division, Boulder, Colorado

(Manuscript received 17 May 2011, in final form 6 February 2012)

ABSTRACT

The properties of clouds that may be observed by satellite instruments, such as optical thickness and cloud-top pressure, are only loosely related to the way clouds are represented in models of the atmosphere. One way to bridge this gap is through ''instrument simulators,'' diagnostic tools that map the model representation to synthetic observations so that differences can be interpreted as model error. But simulators may themselves be restricted by limited information or by internal assumptions. This paper considers the extent to which instrument simulators are able to capture essential differences between the Moderate Resolution Imaging Spectroradiometer (MODIS) and the International Satellite Cloud Climatology Project (ISCCP), two similar but independent estimates of cloud properties. The authors review the measurements and algorithms underlying these two cloud climatologies, introduce a MODIS simulator, and detail datasets developed for comparison with global models using ISCCP and MODIS simulators. In nature MODIS observes less midlevel cloudiness than ISCCP, consistent with the different methods used to determine cloud-top pressure; aspects of this difference are reproduced by the simulators. Differences in observed distributions of optical thickness, however, are not captured. The largest differences can be traced to different approaches to partly cloudy pixels, which MODIS excludes and ISCCP treats as homogeneous. These cover roughly 15% of the planet and account for most of the optically thinnest clouds. Instrument simulators cannot reproduce these differences because there is no way to synthesize partly cloudy pixels. Nonetheless, MODIS and ISCCP observations are consistent for all but the optically thinnest clouds, and models can be robustly evaluated using instrument simulators by integrating over the robust subset of observations.

1. Evaluating simulations of present-day cloudiness with satellite observations

The fidelity of global climate models is frequently judged by comparing simulations of the present day with

Corresponding author address: Robert Pincus, Cooperative Institute for Research in Environmental Sciences, University of Colorado, 325 Broadway, R/PSD1, Boulder, CO 80305.
E-mail: robert.pincus@colorado.edu.

DOI: 10.1175/JCLI-D-11-00267.1

© 2012 American Meteorological Society

observations (e.g., Gleckler et al. 2008; Reichler and Kim 2008). Evaluating the distribution of clouds produced by these models (e.g., Pincus et al. 2008) is of particular interest because clouds have such strong and variable impacts on the earth's radiation budget. Satellite observations are central to this task because of the near-global view they provide.

The evaluation of cloud properties is more complicated than that of, say, temperature, because the cloud properties observable by remote sensing instruments are quite different from the variables used to represent clouds within a model. One way to narrow this gap is to build a diagnostic tool that converts the model state into synthetic observations. The "International Satellite Cloud Climatology Project (ISCCP) simulator" (Klein and Jakob 1999; Webb et al. 2001), for example, uses the internal model representation of cloudiness to produce estimates of the cloud-top pressure and optical thickness that would be reported by a particular satellite observing program. The simulator accounts for effects at the pixel scale, including the screening of clouds low in the atmosphere by clouds above them, the interpretation of measurements as if they arise from clouds in a single homogeneous layer, and the estimation of cloud-top pressure based on infrared brightness temperatures. The treatment of pixel-scale observations is similar to, though less rigorous than, the observation operators used in data assimilation. Because climate models are compared to observations on a statistical basis, the simulator also reproduces the averaging strategies adopted during the processing of ISCCP observations.

The ISCCP simulator has proved quite valuable in diagnosing GCM performance (see, e.g., Norris and Weaver 2001; Zhang et al. 2005; Williams and Webb 2009; among many others) This has inspired a number of other simulators for cloud-related instruments, including spaceborne radars (*CloudSat*; Haynes et al. 2007), lidars [*Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations* (*CALIPSO*; Chepfer et al. 2008)], and multiangle radiometers [Multiangle Imaging SpectroRadiometer (MISR; see Marchand and Ackerman 2010)]. *CALIPSO* and MISR have also produced summary datasets against which the results of the simulator may be compared (Chepfer et al. 2010; Marchand et al. 2010). What these efforts share is the desire to put model predictions and observations on the same footing so that differences between the two can be unambiguously interpreted as model error.

But not all ambiguities can be removed. Here we investigate the large impact that differing treatments of one class of difficult-to-interpret observations can have on the climatological distribution on cloud properties, demonstrate why this subset of observations cannot be modeled by instrument simulators, and show that comparisons between models and observations that exclude such observations are nonetheless robust. The next section describes long-term observations obtained from two similar but independent satellite records [ISCCP and Moderate Resolution Imaging Spectroradiometer (MODIS)], and section 3 explores the most prominent differences between these two views of the earth's clouds. Section 4 describes a method for emulating MODIS observations within a climate model, and section 5 illustrates the degree to which differences between these two sets of observations are captured by instrument simulators. We conclude by discussing the roles and limitations of instrument simulators in enabling model evaluation and diagnosis.

2. Global observations of cloudiness: ISCCP and MODIS

There are at least nine available climatologies of cloud properties based on satellite observations (Stubenrauch et al. 2009). We focus on two—ISCCP and MODIS, described below—with long data records, good spatial sampling, and the ability to detect clouds throughout the atmosphere. MODIS and ISCCP estimate many of the same quantities but, as we will show, often arrive at different answers. In this section we review the way these two cloud climatologies are constructed, including descriptions of individual pixel-level retrievals and the strategies used to aggregate the observations in space and time. Our goal is to highlight the similarities and differences in these datasets and understand the consequences for model evaluation, including motivating decisions made in constructing the MODIS simulator described in section 4. Readers seeking greater detail might start with Rossow and Schiffer (1999) for descriptions of ISCCP, Platnick et al. (2003) and King et al. (2003) for MODIS, and Marchand et al. (2010) for somewhat lengthier explanations and several illuminating case studies; many of the conclusions we reach in this section are also evident in that paper. We limit ourselves to observations made during daytime, as these contain a richer set of retrievals and better accuracy in cloud detection.

a. ISCCP observations

The ISCCP (see Rossow and Schiffer 1991) has produced a long (26 yr and growing), well-documented, and well-used record of cloudiness. ISCCP obtains and calibrates observations at two wavelengths from operational sensors aboard geostationary and polar-orbiting satellites, interprets these observations to determine which pixels are cloudy, then estimates the cloud-top pressure p_c of the highest cloudy layer and

column-integrated optical thickness τ within each (Rossow and Schiffer 1999).

ISCCP uses observations in the visible (approximately 0.6 μm, nominally 1-km resolution) and infrared (11 μm, 1–4-km resolution depending on the instrument) portions of the spectrum. Clouds are detected by comparing the observed values with estimates of the clear-sky contribution; pixels that are somewhat colder and/or brighter than the clear sky are expected to be are flagged as cloudy. Retrievals assume that each pixel is entirely covered with homogeneous clouds. Cloud-top temperature is estimated from the infrared brightness temperature, from which p_c is computed by comparing the temperature with a profile normally obtained from sounding instruments. Cloud optical thickness is inferred by assuming a thermodynamic phase based on the brightness temperature and then interpreting the observed visible reflectance as an optical thickness based on precomputed tables assuming constant particle size distributions. If the cloud is optically thin, the infrared emissivity is adjusted and a new cloud-top temperature and pressure inferred; if no self-consistent solution can be found, the cloud is assigned a pressure just above the tropopause.

Pixel-level retrievals are aggregated on an equal-area grid with roughly 250-km resolution 8 times per day. The ISCCP reports monthly averages for each of the eight daily observation times (i.e., the monthly average diurnal cycle) and for the daily mean (averaging uniformly over time is equivalent to assuming that the distribution of cloud properties is well observed at each hour of each month). The mean optical thickness reported is not the average of the individual values, but rather the "radiative mean" τ_r defined as the value of τ that produces the mean albedo using tables that map τ to albedo and vice versa. It is possible to reconstruct the linear-mean τ using the mean liquid and ice water paths and assumed particle sizes. ISCCP also provides a joint histogram of the cloud fraction c_f as a function of τ and p_c; this histogram has six categories of τ and seven categories of p_c in the 3-hourly observations, though this is reduced to three categories for each parameter in the monthly average files. Nighttime cloud amounts are adjusted to correct for systematic differences between the daytime and nighttime algorithms; these adjustments are applied to the monthly mean cloud amounts.

ISCCP processes observations from both geostationary and polar-orbiting satellites, preferring geostationary observations where they are available for more uniform temporal sampling and more stable calibration. But because cloud detection and cloud property retrievals depend on the zenith angle at which the scene is viewed (e.g., Rossow and Garder 1993; Maddux et al. 2010), the ISCCP record contains modest but significant spatially dependent artifacts (Norris 2000; Evan et al. 2007) visible as radial patterns centered on the subsatellite point below the geostationary platforms. These patterns are unique to ISCCP, since no other global satellite cloud climatology relies so heavily on geostationary observations.

ISCCP OBSERVATIONS FOR CLIMATE MODEL EVALUATION

We have constructed a special-purpose climatology of ISCCP cloud properties to facilitate comparisons with MODIS and with global models using the ISCCP simulator. We begin with the original 8-times-daily, gridded observations and average all daytime values within a month. We provide two estimates of mean cloud albedo, cloud-top pressure, and cloud-top temperature in each grid cell: one is a linear average over time and the second is weighted by the cloud fraction at each time, consistent with the MODIS averaging strategy described in the next section. Albedo is computed using an analytic approximation (described in the appendix) to the ISCCP lookup tables.

We have included the full-resolution joint histogram $c_f(\tau, p_c)$ to complement the reduced-resolution joint histogram distributed by ISCCP. Monthly mean files in network Common Data Form (netCDF) format may be obtained online (at http://climserv.ipsl.polytechnique.fr/cfmip-obs.html); we also plan to distribute them on the Earth System Grid alongside climate model results from the upcoming Coupled Model Intercomparison Project phase 5 (CMIP5) experiment (http://cmip-pcmdi.llnl.gov/cmip5/).

b. *MODIS observations*

Cloud observations from MODIS provide a useful complement to the ISCCP record. There are two MODIS instruments, one each aboard the *Terra* and *Aqua* satellites. Some of the 36 spectral observations available allow for different approaches for retrieving p_c and identifying thermodynamic phase than are taken by ISCCP, while others allow for the retrieval of complementary information, such as particle size. [Data products by the MODIS Science Team are available at http://ladsweb.nascom.nasa.gov. A separate set of retrievals is made in support of the Clouds and the Earth's Radiant Energy System (CERES) products, as described in Loeb et al. (2005) and Minnis et al. (2011), among others. Those retrievals follow somewhat different logic and have different error characteristics.]

MODIS pixel-scale retrievals, as described in detail by Platnick et al. (2003), proceed in steps:

- *Cloud detection/masking:* Clouds are initially identified by a cloud mask that uses tests incorporating

a large number of spectral bands (Frey et al. 2008). Different tests are used in different domains (over oceans, vegetation, ice, etc.). The cloud mask summarizes these tests by computing a likelihood that each pixel is clear (Ackerman et al. 1998), and then it assigns the result to one of four categories (confident clear, probably clear, uncertain/probably cloudy, and cloudy) based on probability boundaries 0.99, 0.95, and 0.66 for clear sky. The cloud mask is clear-sky conservative: the label ''cloudy'' means ''containing some amount of cloud (or maybe heavy aerosol).''

- *Cloud-top properties:* Cloud-top pressure retrievals are initially attempted using CO_2 slicing (Menzel et al. 1983), which infers the amount of CO_2 above an emitting cloud from brightness temperature measurements in several closely spaced bands near 15 μm (pressure can be inferred from the amount of carbon dioxide because the gas is well mixed). The method fails when the integrated amount of carbon dioxide becomes optically thick and clear-sky emission dominates the signal. This limits CO_2 slicing to values of p_c less than about 700 hPa; for clouds below this level, p_c is retrieved by matching the brightness temperature at 11 μm to a temperature profile (this is similar to the ISCCP algorithm). Cloud-top pressure is computed at 5-km scale based on 1-km pixels; it includes all such regions in which more than 4/25 ($\geq$16%) of the underlying 1-km pixels are deemed cloudy or probably cloudy by the cloud mask.
- *Cloud thermodynamic phase and optical properties:* Optical thickness and effective radius r_e are retrieved, following Nakajima and King (1990), by minimizing the difference between the observed reflectance in one nonabsorbing and one absorbing wavelength (nominally 0.86 and 2.13 μm, respectively, both at 1-km scale) and forward calculations tabulated across this parameter space. The thermodynamic phase of each pixel is required because the retrievals use separate forward calculations (i.e., lookup tables) for liquid and ice clouds. This determination is made based on a range of tests in the visible, near-infrared, and infrared portions of the spectrum (see King et al. 2010); where these tests disagree the pixel is labeled ''unknown/mixed phase'' and liquid water cloud libraries are used in the retrieval. When the simultaneous retrieval of τ and r_e fails (i.e., the observations cannot be fit to the forward calculations to within the desired accuracy), the pixel is not included in aggregated statistical summaries. For these pixels a retrieval of optical thickness alone is attempted using the fixed particle sizes assumed by ISCCP and is available in the pixel-level data.

Uncertainty estimates are computed for each retrieval of τ and r_e. These include only the effects of uncertainties in instrument calibration and nominal plane-parallel forward models, surface spectral albedo, and spectral atmospheric correction (primarily due to above-cloud atmospheric moisture uncertainty). Other error terms can be important on a pixel-level basis (e.g., vertical and/or horizontal inhomogeneity) so the uncertainty estimates provided are a lower bound on the true uncertainty. Uncertainties are calculated from the cloud reflectance lookup tables used in the retrievals and depend on viewing and illumination geometry.

Beginning with the ''collection 5'' processing algorithms introduced in 2006, MODIS retrievals identify pixels that are not likely to be completely cloud covered (see King et al. 2006)—in other words, those pixels that are a poor fit to the plane-parallel, homogenous model used to interpret the observed reflectances in terms of optical properties—and use these to produce a cloud-conservative mask. This identification is based on tests for spectral and spatial uniformity to identify dust, smoke, snow or ice, sunglint, and some partly cloudy pixels, and on cloud-edge detection, which is also to remove partly cloudy pixels. In samples we have examined, the edge detection test is the trigger approximately 70% of the time. In collection 5, retrievals are not performed for pixels identified by this ''clear-sky restoral'' algorithm.

MODIS pixel-scale observations are aggregated on a 1° equal-angle grid (King et al. 2003). Observations from *Terra* and *Aqua* are aggregated separately. All observations from a given platform within a UTC day are included so the frequency of observations increases with distance from the equator, and observations from overpasses at different times (with potentially very different viewing and illumination geometries) may be included in the statistics. Statistics are weighted over time by the number of observations on each day; for cloud-related variables, this means that cloudier days count more than less-cloudy days when computing cloud properties (e.g., optical thickness, cloud-top pressure, particle size). Cloud fractions and cloud properties are summarized for all clouds and separately for ice- and liquid-phase clouds.

Two estimates of cloud fraction are produced: one counts the proportion of pixels deemed by the cloud mask to be cloudy or probably cloudy, and the other counts the proportion of available pixels for which cloud optical properties are successfully retrieved. We refer to these as the ''mask cloud fraction'' and ''retrieval cloud fraction'' below. Differences between these two estimates include the proportion of pixels for which retrievals are unsuccessful but are dominated by pixels

removed by clear-sky restoral. We have recently identified an error in the collection 5 processing: the number of failed retrievals is not included in the total number of observations used to determine the retrieval cloud fraction, so this quantity is slightly overestimated. The effect is not large because the number of failed retrievals is quite small.

Averages are reported for both optical thickness (i.e., $\overline{\tau}$) and its base-10 logarithm $\overline{\log_{10}(\tau)}$. Mean albedo $\overline{A}$ computed using, for example, two-stream methods can, to a good approximation, be computed as $\overline{A} \approx A[10^{\overline{\log_{10}(\tau)}}]$, making $\tau_l = 10^{\overline{\log_{10}(\tau)}}$ comparable to the radiative mean optical thickness τ_r reported by ISCCP (see the appendix for more details).

Liquid and ice water path are computed at the pixel scale, assuming that clouds are vertically homogeneous, and then averaged separately for liquid and ice clouds.

We estimate the uncertainty in the mean of τ and r_e from pixel-level uncertainties by assuming that errors are perfectly correlated within each grid cell on a given day but perfectly uncorrelated from day to day. This has the effect of greatly reducing typical daily uncertainties in a grid cell (by a factor of about $\sqrt{30}$ if roughly the same number of cloudy pixels is observed on each day of the month and cloud properties are reasonably consistent over the month). True errors are almost certainly not uncorrelated from day to day—calibration and modeling errors, in particular, are probably not well represented as random errors—but neither are errors within each day likely to be perfectly correlated. Uncertainty estimates are still the subject of active development (Platnick et al. 2004) and are intended as rough guidance.

MODIS OBSERVATIONS FOR CLIMATE MODEL EVALUATION

We have developed a dataset of MODIS observations targeted at the evaluation of clouds in climate models. We extract a small subset (listed in Table 1) of the many datasets available from the gridded monthly files produced by the MODIS project and reformat them as netCDF files that include the climate and forecast (CF, http://cf-pcmdi.llnl.gov/) metadata widely used in the climate community. The joint histogram of cloud-top pressure and cloud optical thickness included here uses the same bin definitions as does the analogous histogram produced by ISCCP. A second joint histogram with finer resolution is available in the original data. We produce separate files for the *Terra* (morning) and *Aqua* (afternoon) platforms along with an estimate that combines observations from the two platforms, weighted by the number of pixel-level observations. New files are added as data from the operational MODIS system become available and will be updated to reflect algorithmic updates as the MODIS record is reprocessed. The data are available online (at http://ladsweb.nascom.nasa.gov/) and are mirrored on another website (at http://climserv.ipsl.polytechnique.fr/cfmip-obs.html) and on the Earth System Grid.

Table 1. Parameters available in the MODIS dataset intended for climate model evaluation. All quantities are available as monthly averages on a 1° equal-angle grid. The difference between the total and (liquid + ice) retrieval cloud fractions is the fraction of detections for which phase is undetermined. Bin boundaries for the joint histogram of cloud top pressure and optical thickness follow the ISCCP definition.

Parameter	Subsets	Uncertainty provided
Cloud fraction – mask	Total	
	High, middle, low	
Cloud fraction – retrieval	Total	
	High, middle, low	
	Liquid, ice	
Mean cloud top pressure		
Mean optical thickness	Total	
	Liquid, ice	Yes
Mean $\log_{10}(\tau)$	Total	
	Liquid, ice	Yes
Mean particle size	Liquid, ice	Yes
Mean water path	Liquid, ice	Yes
Optical thickness/cloud top pressure joint histogram		

3. Understanding differences between ISCCP and MODIS climatologies of cloudiness and cloud properties

MODIS retrievals exploit the diversity of spectral information to offer a wider range of observations than ISCCP, but many MODIS observations (see Table 1) have analogs in the ISCCP climatology. In this section we compare the climatology of cloudiness, τ, and p_c, in these two datasets. Because cloud detection and the retrieval of optical thickness follow parallel approaches in the two datasets, one might expect MODIS and ISCCP climatologies of cloud fraction and the distribution of optical thickness to be similar. But, as we will show, differences between these two datasets can be large and arise in large part from quality-control decisions made during pixel-scale retrievals and aggregation. Here we compare long-term means (July 1983–June 2007 for ISCCP, July 2002–June 2010 for MODIS) using the datasets described in section 2.

a. Cloudiness (cloud fraction)

Cloud identification (masking) defines the starting population of pixels for which other cloud properties are

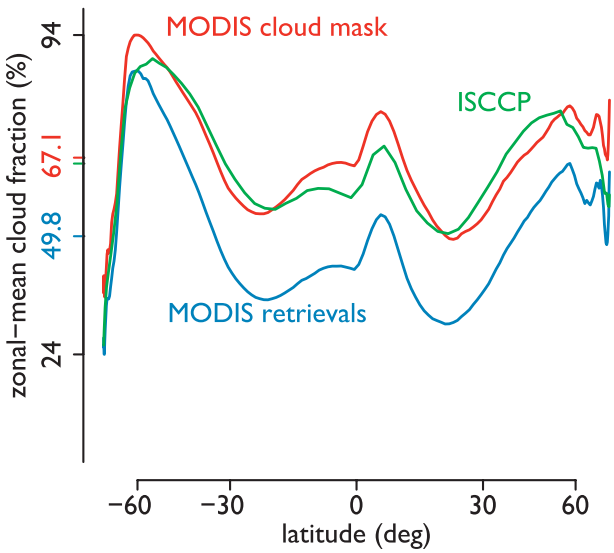

FIG. 1. Climatological zonal-mean distribution of daytime cloud fraction from ISCCP (green), the MODIS cloud mask (red), and MODIS cloud retrievals (blue, including all ice, liquid, and undetermined clouds). Tick marks on the ordinate show the corresponding global-mean cloud fraction and the range of zonally averaged cloud fractions. ISCCP and the MODIS cloud mask produce similar distributions of cloud occurrence, while the MODIS retrievals (which exclude pixels thought to be partly cloudy) are much more conservative.

determined, so comparisons of cloud properties between datasets can be expected to agree no better than does cloud fraction. The differentiation between clear and cloud-affected pixels is consistent between MODIS and ISCCP despite the much richer range of spectral information available to MODIS: daylight cloud fractions from ISCCP are quite similar to those from the MODIS cloud mask, as illustrated in Fig. 1. But the difference between the MODIS cloud mask and retrieval cloud fractions is striking, exceeding 20% in some subtropical latitude bands [see also Fig. 8 in Marchand et al. (2010)]. The retrieval cloud fraction includes all thermodynamic phases: ice, liquid, and undetermined, as described above.

The differences between the MODIS mask and retrieval cloud fractions, though large, are consistent with a high frequency of partly cloudy pixels at the ~1-km scale of the MODIS pixels. Partly cloudy pixels (as identified by clear-sky restoral) are counted as cloudy by the cloud mask, which seeks to identify fully clear pixels, and as clear by the retrieval process, which seeks to include only fully cloudy pixels. The large sensitivity of total cloud fraction to decisions about the fate of partly cloudy pixels has been known for as long as satellite observations of clouds have been available (see, e.g., Shenk and Salomonson 1972; Minnis and Wielicki 1988; Wielicki and Parker 1992; among many others); the magnitude of the difference between the two MODIS estimates (about 17% in the global, climatological daytime average) is quantitatively consistent with other estimates of this sensitivity.

The "Taylor diagram[1]" (Taylor 2001) in Fig. 2 shows the differences between these observational datasets in the way climate model simulations are normally evaluated against observations (see, e.g., Gleckler et al. 2008; Pincus et al. 2008): by summarizing the components of the total RMS difference between multiple datasets and a single reference. This example summarizes the agreement in the composite annual cycle in cloud fraction with our ISCCP daytime-only climatology (here and in Fig. 12 all datasets are mapped onto a 2.5° equal-angle grid). Although the diurnal cycle of cloudiness can be large in some locations (e.g., Cairns 1995), temporal sampling does not strongly affect cloudiness estimates at the global scale: both the ISCCP-provided diurnal average (labeled "D2" in the figure, and including corrections for nighttime cloud detection alluded to in section 2a) and the ISCCP daytime climatology sampled only at MODIS overpass times (labeled "Resamp," created by averaging only those observations that fall within 45°-wide longitude bands centered on the *Terra* and *Aqua* overpass times) are in very good agreement with the daytime-only dataset. We infer that the somewhat greater disagreement between ISCCP and the MODIS cloud mask reflects differences in performance in thin clouds, where the many spectrally dependent tests used by MODIS may have a different outcome than the two-threshold tests used by ISCCP, and other sampling differences, including the more uniform sampling of viewing and illumination angles by MODIS. Results for the MODIS retrieval fraction are omitted because the large bias dominates the figure, but the space–time standard deviation of this field and its correlation with ISCCP are about the same as the estimate from the MODIS cloud mask.

[1] Taylor diagrams indicate agreement with respect to a reference dataset. The radial axis denotes the standard deviation of each dataset and the azimuthal axis denotes the space–time correlation coefficient between each test dataset and the reference, and the size of each marker shows the bias with respect to the reference dataset; the magnitude of the bias can be measured against the radial axis. A light green line indicates the standard deviation of the reference dataset; perfect agreement would be shown as a point (zero bias) where the dashed line intersects the horizontal axis (correlation one and matching standard deviations). Arcs surrounding this point denote constant root-mean-square error. Points close to one another in a Taylor diagram are equally far from the reference dataset but not necessarily close to one another.

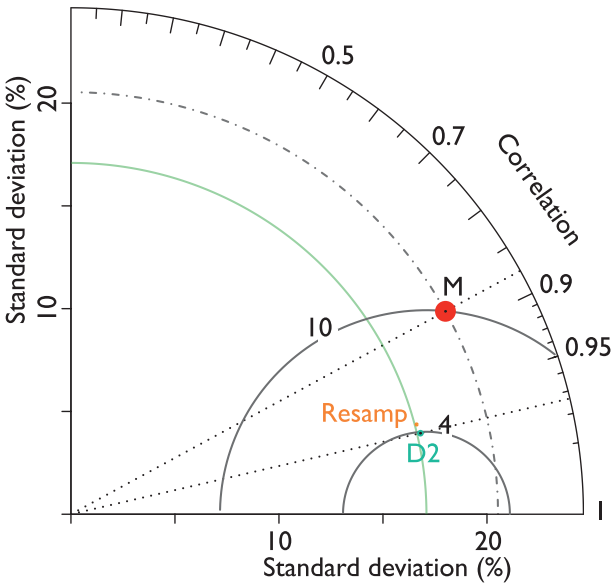

FIG. 2. Agreement in the composite annual cycle of daytime-mean climatology between daytime-only ISCCP observations (see subsection in section 2a) and other measures of cloud fraction. Standard deviation and correlation coefficients are computed over the 12 months of the composite annual cycle, weighting spatially according to area. Climatologies created using different strategies for averaging ISCCP data over the diurnal cycle ("D2" and "Resamp") are quite similar to the daytime-only climatology, implying that differences between ISCCP and the MODIS cloud mask ("M") reflect other sampling differences, including the sampling of viewing geometry.

Thus, MODIS cloud properties are derived from two distinct populations: those detected by the cloud mask, counted in the cloud mask fraction, and included in retrievals of cloud-top pressure, and the subset of these that remain after the removal of failed retrievals and filtering by clear-sky restoral, from which the retrieval fraction and statistics of optical thickness, particle size, and liquid and ice water path are computed. Because ISCCP estimates of cloudiness are consistent with MODIS cloud detection, differences between MODIS and ISCCP estimates of cloud-top pressure in the next section can be attributed to sensor capabilities and algorithmic differences, while differences in optical properties (section 3c) must have other roots, including, but not limited to, this filtering.

b. Cloud-top pressure

The estimation of cloud-top pressure is where MODIS and ISSCP rely most heavily on different information, with ISCCP using infrared window brightness temperature to infer pressure and MODIS using CO_2 slicing for higher clouds. Because the population of clouds for which cloud-top pressure is retrieved is roughly the same for MODIS and ISCCP, large differences in these estimates reflect errors in one or both of the retrievals. In practice, the largest differences between the two datasets are for midlevel clouds ($440 \leq p_c < 680$ hPa), which are much less frequent in the MODIS climatology than in ISCCP (see Fig. 3). This is likely due to ambiguities in the ISCCP algorithms: thermal emission techniques may interpret the emission from high, cold, thin clouds over low, warm, brighter clouds as being a single layer of midlevel clouds (see, e.g., Mace et al. 2006), while CO_2 slicing is less susceptible to this error. Low clouds are correspondingly more frequent in the MODIS cloud mask than in ISCCP, indicating that a majority of pixels from the large, cloud-affected population that are described by ISCCP as midlevel is assigned higher cloud-top pressures by MODIS. In the population remaining after clear-sky restoral, however, the proportion of low clouds is roughly similar in ISCCP and MODIS, while the amount of high cloudiness is unaffected, implying that clear-sky restoral is more likely to exclude low clouds (as identified by MODIS) than high clouds. Since edge detection is the primary trigger for clear-sky restoral, we infer that low clouds are more spatially inhomogeneous than high clouds on the scales observed by MODIS (see also Maddux et al. 2010).

c. Optical thickness

Figures 1 and 2 demonstrate that the population of pixels included in ISCCP retrievals is roughly the same population of clouds identified by the MODIS cloud mask, while the population for which MODIS performs and aggregates retrievals is smaller (and, in some regions, substantially smaller). Because retrieved values of τ do not depend strongly on the value of r_e (see, e.g., Nakajima and King 1990), filtering by clear-sky restoral would seem to be the largest algorithmic difference between ISCCP and MODIS estimates of optical thickness. Indeed, the largest differences between the two distributions of optical thickness (Fig. 4, top) is in the frequency of clouds with $0.3 \leq \tau < 3.6$, which cover 10.6% in the MODIS record and 33.1% in the ISCCP record. The very thinnest detectable clouds ($0.3 \leq \tau < 1.3$) are observed almost 10 times more frequently by ISCCP (14.2%) than MODIS (1.7%). As one result, the logarithmic-mean optical thickness τ_l reported by MODIS is substantially larger than the radiative mean τ_r reported by ISCCP (Fig. 5). This result is robust to differences in the definition of τ_l with τ_r, as the appendix demonstrates. When optically thin clouds are excluded, global-mean cloud fraction estimates from the two platforms are quite similar (Fig. 4, bottom). The amount of all but the thinnest clouds is identified consistently in both datasets: space–time correlations between MODIS and ISCCP

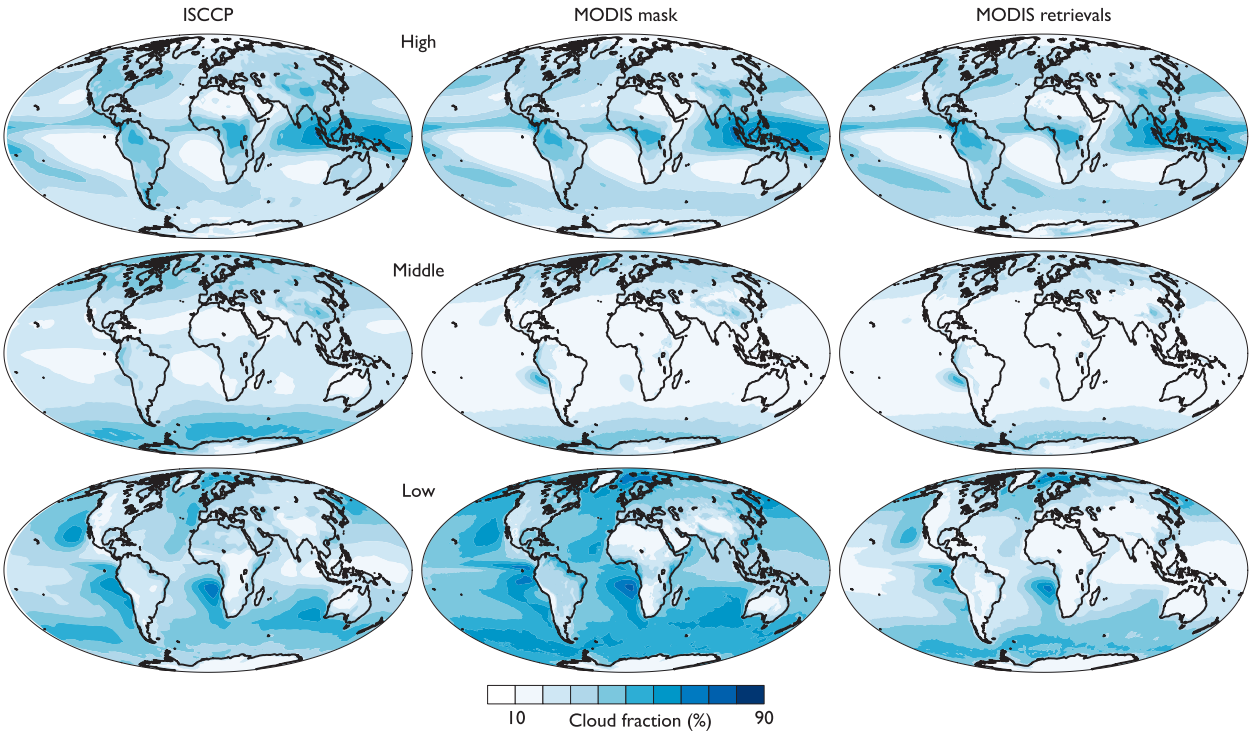

FIG. 3. Annual-mean distribution of (top) high ($p_c < 440$ hPa), (middle) middle ($440 < p_c < 680$ hPa), and (bottom) low ($p_c > 680$ hPa) clouds obtained from (left) ISCCP, (middle) the MODIS cloud mask, and (right) MODIS cloud retrievals. ISCCP depends on signals in the thermal infrared and assigns clouds to midlevels more frequently than MODIS (which uses CO_2 slicing) or other sensors. The clear-sky restoral that distinguishes the MODIS retrievals from the MODIS cloud mask is more likely to exclude low clouds than high clouds; we infer that low clouds are more spatially inhomogeneous.

estimates of $c_f(\tau \geq \tau_{\min})$ exceed 0.8 for $1.3 < \tau_{\min} \leq 23$ (above which sample sizes are quite small), consistent with case studies showing that MODIS, ISCCP, and MISR estimates of the presence of optically thick clouds agree quite well (Marchand et al. 2010).

Many of the tests used in clear-sky restoral, and particularly the frequently invoked edge detection trigger, are to identify and remove partly cloudy pixels from the population considered by MODIS. Our results are consistent with this goal having been achieved: the largest differences between the distributions derived from ISCCP (which does not filter for partly cloudy pixels) and MODIS are at low values of τ, and the optical thickness retrieved from partly cloudy pixels is systematically smaller than from fully cloudy pixels (Chang and Coakley 2007). Furthermore, the optically thinnest pixels are assigned cloud-top pressures throughout the atmosphere by ISCCP (Fig. 6) and categorized as low clouds by MODIS (cf. middle and right columns of Fig. 3). More than a third of these clouds are assigned the lowest possible cloud-top pressure by ISCCP, consistent with, though not proof of, ISCCP being unable to find self-consistent solutions for τ and p_c from visible and infrared observations. [That being said, neither case studies by the reviewers of this paper, among others, nor the spatial distribution of clouds in this category (Rossow et al. 2005) are consistent with broken low clouds being the only contributors to the lowest-τ, lowest-p_c bin.]

But partly cloudy pixels are not the whole story. Though the largest differences between ISCCP and MODIS estimates of the optical thickness distribution are for low values of τ, differences exist across the entire range. In particular, clouds with $\tau > 3.6$ are observed more frequently by MODIS (39.2%) than by ISCCP (32.9%). And though clear-sky restoral filters a sizable proportion of the MODIS observations, there are several other reasons why MODIS and ISCCP estimates of optical thickness might differ, including the following.

- *Differing viewing and illumination geometry:* Optical thickness retrievals are sensitive to both the viewing and illumination zenith angle (see Maddux et al. 2010; Loeb and Coakley 1998, respectively), which are sampled quite differently in the two datasets. MODIS observations are nominally made within about 1.5 h of local noon but include a wide range of viewing zenith angles. ISCCP observations are made at all sunlit times and so include low solar zenith angles, while viewing angles

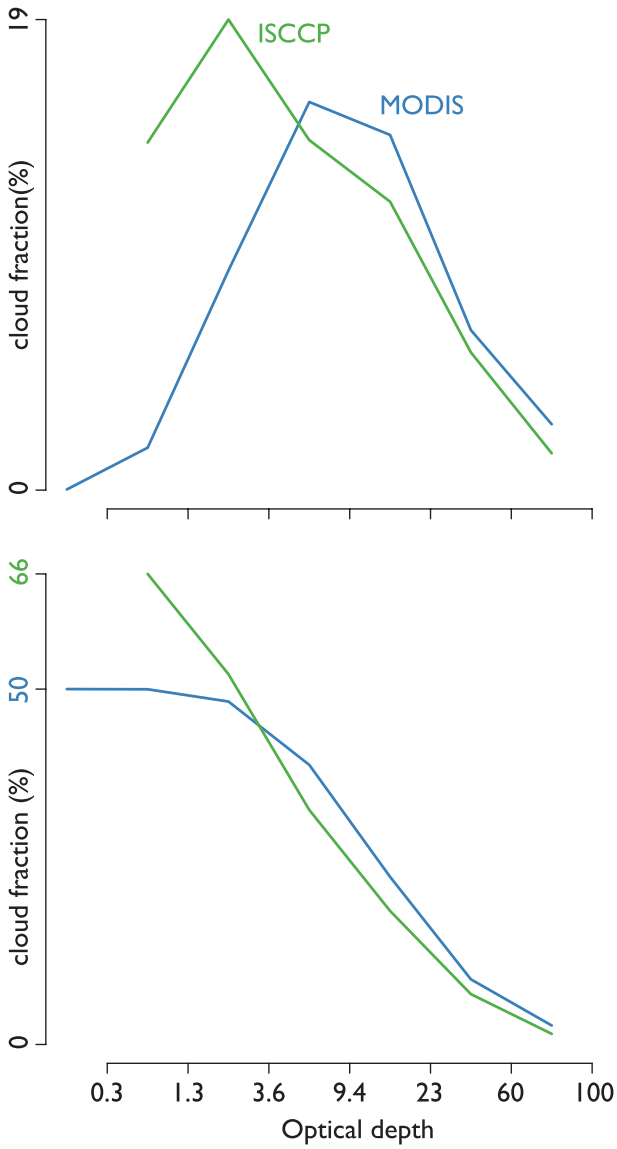

FIG. 4. (top) Global-mean climatological distribution of optical thickness from ISCCP and MODIS. The larger value of total cloudiness viewed by ISCCP is due to much higher values of optically thin ($\tau \leq 3.6$) clouds, while clouds of moderate-to-high optical thickness are more prevalent in the MODIS observations. (bottom) Global-mean climatological value of daytime cloud fraction $c_f(\tau \geq \tau_{\min})$ as a function of the minimum optical thickness. Though the two datasets disagree on the detailed distribution of τ, the overall amount of cloudiness for all but the thinnest clouds is similar in magnitude and in the sensitivity to $\tau_{\min}$.

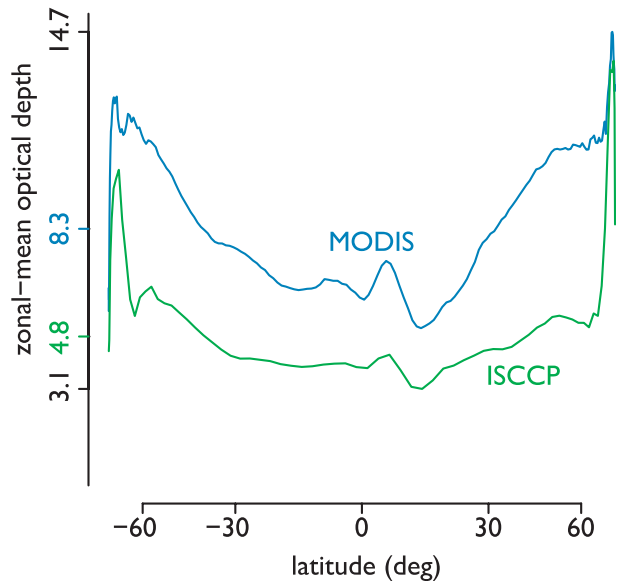

FIG. 5. (top) Climatological zonal-mean distribution of "radiatively effective" optical thickness (i.e., the optical thickness that produces the mean albedo, described in section 2) from ISCCP (τ_r, green) and MODIS cloud (τ_l, blue). Tick marks on the ordinate show the corresponding global mean for each dataset and the overall range. Temporal averaging is uniform: monthly means are weighted by the daily faction and then combined linearly over years. MODIS optical depths are substantially higher than ISCCP because the edge detection invoked by clear-sky restoral preferentially removes optically thin pixels.

remain constant for long periods in many locations equatorward of 60° due to the preference for geostationary observations. Marchand et al. (2010) offer this as a possible explanation for differences in optical thickness distributions between ISCCP, MISR, and MODIS.

- *Differing observation times:* MODIS observations are taken near midday, while ISCCP accumulates observations over all daylight hours, which could lead to systematic differences where the diurnal cycle of τ is not roughly symmetric about local noon.
- *Differing approaches to thermodynamic phase detection:* Phase detection affects the retrieval of optical thickness because it determines which forward calculations are matched to observed intensities. ISCCP separates ice from liquid clouds based on a single threshold of infrared brightness temperature where MODIS uses a wide range of spectral tests with potentially higher accuracy.
- *Differing assumptions about cloud single-scattering properties:* ISCCP uses fixed particle size distributions (and hence fixed single-scattering properties) for each thermodynamic phase, while MODIS sizes can vary from pixel to pixel. The larger impact, however, is likely to be the choice of particle shapes used to represent ice crystals: the polycrystal habit used by ISCCP (Macke 1993) produces lower values of asymmetry parameter (and hence lower optical depths for a given observed reflectance) than does the physically based model used by MODIS (Baum et al. 2005).

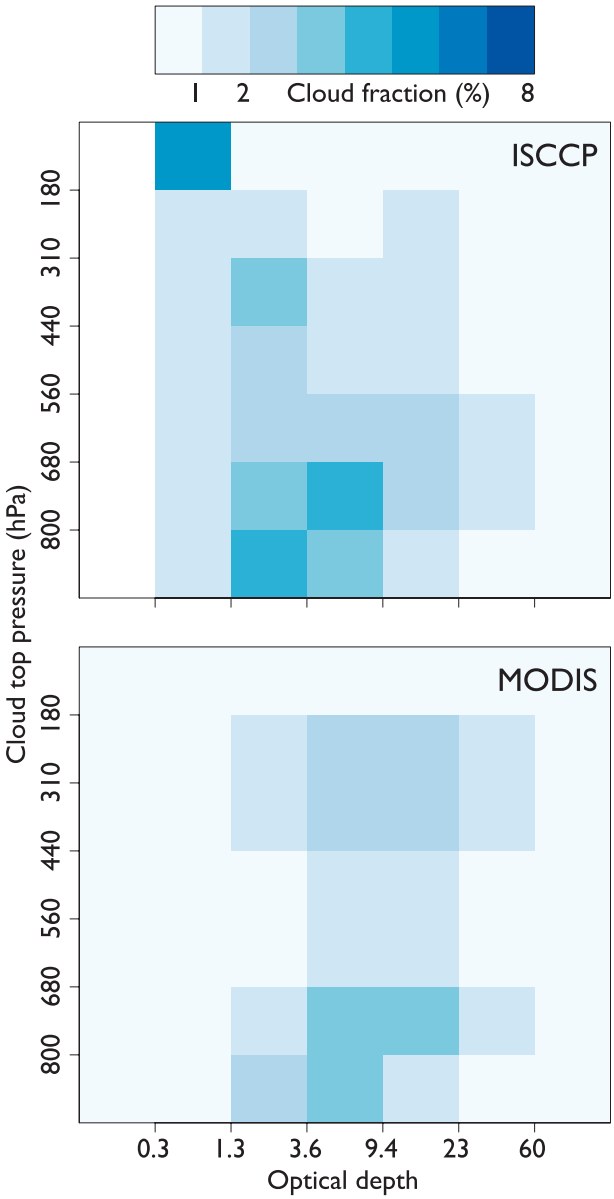

FIG. 6. Climatological distribution of global-mean cloud fraction as a function of cloud top pressure (vertical axis) and cloud optical thickness (horizontal axis) observed by (top) ISCCP and (bottom) MODIS. ISCCP provides retrievals for all pixels identified as cloud affected, while MODIS excludes those identified as marginal by the clear-sky restoral process described in the text. The partly cloudy pixels most frequently removed during the clear-sky restoral represent almost all the optical thinnest clouds ($\tau < 1.3$) observed by ISCCP. The frequent observation of very high, very thin clouds by ISCCP may indicate that many of these observations are difficult to interpret self-consistently.

We isolate the role of filtering decisions by creating two alternate MODIS datasets for April 2005. We begin with a special-purpose pixel-level dataset in which retrievals are attempted for all pixels labeled as being cloud affected by the cloud mask. We then create an ''omit-failed'' dataset (Fig. 7, dotted line) that includes statistics for all pixels for which simultaneous estimates of τ and r_e are available, and an ''ISCCP like'' dataset (Fig. 7, dashed line) that also aggregates optical thickness from pixels, in which the simultaneous retrievals failed but single-parameter retrievals using assumed particle sizes were successful. Global-mean cloud cover from the ISCCP-like dataset is within 1% of the ISCCP estimate for this month.

Both filtering strategies (omitting failed retrievals, evidenced by the difference between the ISCCP-like and omit-failed datasets, and the addition of clear-sky restoral, as the difference between the omit-failed and default results) are most active at low optical thickness: both remove essentially all observations with $\tau < 0.3$, most with $\tau < 1.3$, and have essentially no effect on the population of clouds with $\tau \geq 9.4$. Clear-sky restoral has the largest relative impact on the population of clouds with $1.3 \leq \tau < 9.4$. But aggregation strategies alone do not explain the much higher frequency of optically thin ($\tau < 3.6$) clouds in the ISCCP dataset relative to MODIS, nor the correspondingly lower frequency of optically thick clouds. Even after moving to an ISCCP-like filtering strategy, in fact, half the discrepancy between ISCCP and MODIS estimates of cloud fractions for $\tau < 3.6$ remains.

The precise distribution of optical thickness may, in fact, be quite difficult to determine even in the best of circumstances. Figure 8 shows the marginal and cumulative distributions of optical thickness for April 2005 over the western Pacific (10°N–10°S, 170°W–130°E) obtained from ISCCP, three sets of MODIS observations filtered with various levels of stringency, and, over ocean only, the MISR instrument using the retrievals described in Marchand et al. (2010). During this month ISCCP observations in this domain are obtained from a geostationary satellite sitting at its center so that ISCCP viewing angles are near nadir. The character of the differences between ISCCP and MODIS, and the sensitivity to filtering strategy, are similar to those for the entire globe (Fig. 7), in that the distribution of optical thickness derived from MODIS observations contains fewer optically thin clouds and correspondingly more optically thick clouds. This suggests that differences in the global distribution are not driven primarily by the distribution of ISCCP viewing angles. The MISR retrieval technique is quite similar to the one used by ISCCP, especially in the fixed assumptions made regarding particle distributions, and both are restricted to near-nadir views in this comparison. Nonetheless, the two disagreements in cloud amount outside the range $3.6 \leq \tau < 60$ are considerable.

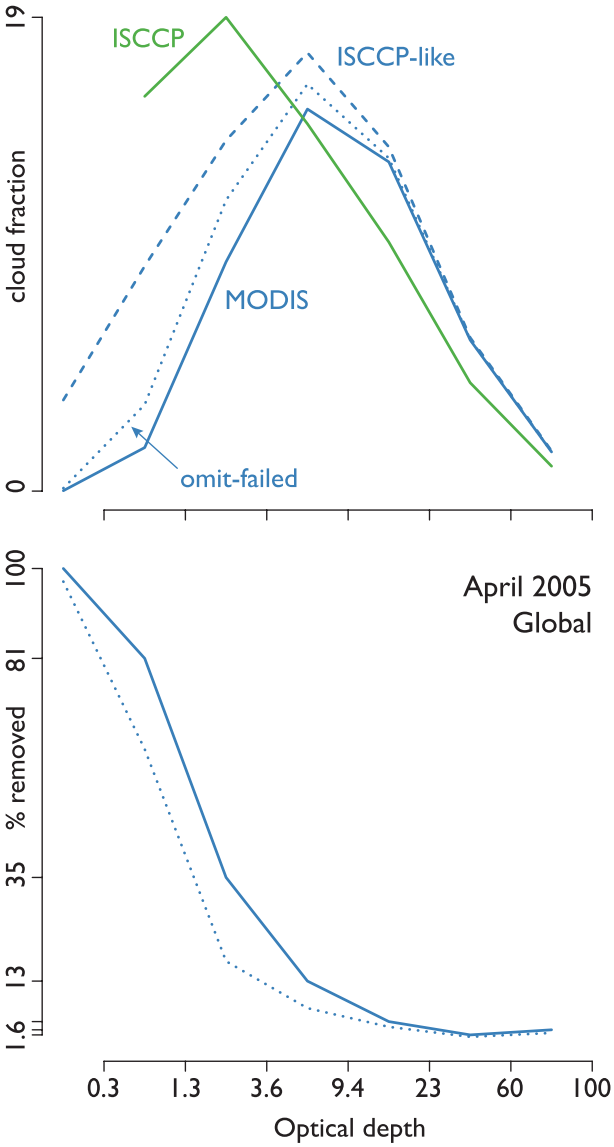

FIG. 7. (top) Impact of filtering strategies on the global-mean distribution of optical thickness for April 2005. Solid lines indicate the distributions obtained by ISCCP (green) and the operational MODIS algorithm (solid blue). The dotted line (omit-failed) shows the distribution of MODIS retrievals when clear-sky restoral is not invoked but failed retrievals (i.e., those for which self-consistent values of τ and r_e cannot be found) are discarded. The dashed line (ISCCP like) shows results from a liberal strategy that ignores clear-sky restoral and also aggregates optical thickness retrievals using assumed particle sizes that are attempted when the normal retrievals fail to find self-consistent values of τ and r_e. (bottom) Fraction of ISCCP-like cloudiness removed by omit-failed and operational filtering. Both strategies are intended to remove pixels that do not fit the plane-parallel, homogenous model used to derive τ, and both provide similar results: many pixels flagged by clear-sky restoral also fail to produce simultaneous retrievals, and these represent most of the optically thinnest clouds.

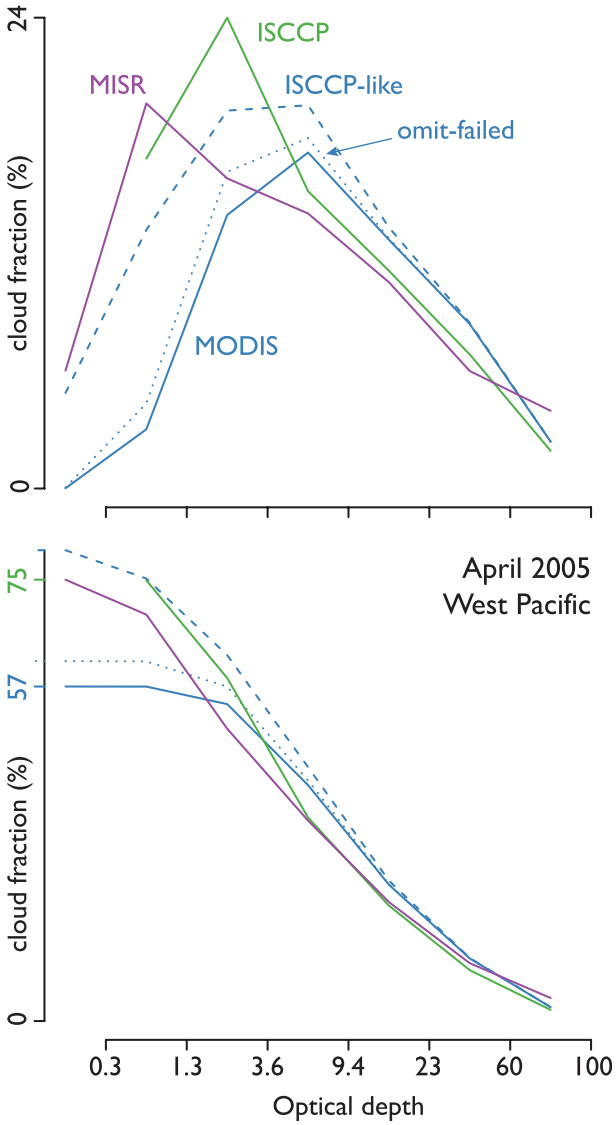

FIG. 8. (top) Marginal and (bottom) cumulative distributions of optical thickness as observed by ISCCP (green), MODIS (blue), and MISR (purple, ocean only) under good viewing and illumination conditions: during the month these data were obtained (April 2005), ISCCP relied on a geostationary satellite almost perfectly centered over this domain, the western Pacific. Dotted and dashed lines show results for MODIS using the increasingly liberal omit-failed and ISCCP-like filtering strategies described in the text, respectively. The MISR retrieval technique is quite similar to the one used by ISCCP, but the two estimates disagree at high and low optical thickness.

d. *Interpreting observations from partly cloudy pixels*

All satellite retrievals, including those made by MODIS and ISCCP, relate reflectance measurements to physical properties using simple models of the atmosphere. In particular, optical property retrievals assume that each pixel is completely covered by a single

vertically and horizontally uniform cloud layer. One-kilometer (pixel size) regions that fit this model precisely are rare in the earth's atmosphere if they exist at all, and there are many ways in which a given pixel may fail to match this model. For some combinations of model failure and retrieval, this may not be relevant: top-of-atmosphere visible-wavelength reflectance for a given optical thickness, and hence the value of τ inferred from that reflectance, is essentially unaffected by the vertical distribution of particle size (Zhang et al. 2010), for example. But the ability to determine when the model is failing in relevant ways depends in part on how much information is available. Figure 7, for example, shows that requiring that consistency in simultaneous retrievals of optical thickness and particle size alone significantly reduces the population of pixels considered useful, while studies based on angular consistency of optical thickness retrievals (Horváth and Davies 2004; Liang et al. 2009; Di Girolamo et al. 2010) suggest that only roughly 30% of 1-km pixels in marine boundary layer clouds fit the interpretive model within 5% reflectance across the viewing angle space sampled by MISR.

Clear-sky restoral was introduced into MODIS processing because it was thought that the large values of effective radius frequently produced by MODIS might be the result of retrievals in partly cloudy pixels. Excluding cloud edges did not, as it turns out, make a substantial difference to the average particle size, in part because many retrievals at cloud edges failed to produce consistent retrievals of τ and r_e and so were ignored during aggregation. Pixels excluded by clear-sky restoral are quite frequent (about 17% as a global average); though the majority of these pixels was removed for being near cloud edges, it is also the great majority of optically thin ($\tau \leq 1.3$) pixels observed by MODIS (Fig. 7). This distribution, along with the high frequency of the cloud edge trigger, suggests that many of the pixels identified by clear-sky restoral are only partly cloudy. We argue that it is misleading to interpret retrievals of optical thickness from these clouds literally.

Cloud detection algorithms require thresholds to separate clear and cloudy skies; these thresholds may be adjusted depending on the purpose of the cloud mask (Yang and Di Girolamo 2008). Values for ISCCP were selected in part to minimize the differences between ISCCP-derived cloud amounts and independent estimates of this quantity (Rossow et al. 1993; Rossow and Schiffer 1999). In partly cloudy pixels, the desired threshold is that which balances "overestimates due to low spatial resolution offset by underestimates due to finite radiance threshold" (Rossow et al. 1993, p. 2394). MODIS thresholds were tuned differently (i.e., to provide a probability that the field of view contains some cloud; see Ackerman et al. 1998) but produce results quite similar to ISCCP. Still, MODIS cloud fraction is known to be biased in fields of small (subpixel-scale) clouds (Zhao and Di Girolamo 2006), suggesting that perfect thresholds are not achievable.

But even cloud masks tuned to produce unbiased estimates of cloud fraction will not produce unbiased cloud retrievals. Masks for cloud fraction attempt to balance over- and underestimations of cloud occurrence; this would yield unbiased retrievals in partly cloudy pixels only if retrievals were proportional to the product of the cloud fraction and the underlying property. This seems unlikely even in plane-parallel clouds, since strict linearity would require 1) that the measured reflectance is proportional to the product of cloud optical thickness and cloud fraction, and 2) that the retrieval of cloud properties from reflectance measurements is linear in reflectance. Given the nonlinear dependence of radiation fields on cloud properties and surface reflectance this is unlikely to be true in general.

Partly cloudy pixels are also those in which three-dimensional radiative transfer effects may be expected to affect the distribution of reflectance observed from satellites, especially if the subpixel cloud fraction is small (Yang and Di Girolamo 2008; Evans et al. 2008). The underlying pixel size for both MODIS and ISCCP retrievals of optical thickness is approximately 1 km near nadir and increases with viewing zenith angle. Three-dimensional effects can be expected where the scale of the clouds is commensurate with or smaller than the "radiative smoothing scale" (Marshak et al. 1995); this depends on the local value of optical extinction but is of order 300 m—commensurate, in other words, with the scale of broken fields of cumulus (e.g., Koren et al. 2008).

Fundamentally, then, the interpretation of optical thickness retrievals from partly cloudy pixels is ambiguous, and MODIS and ISCCP treat this ambiguity differently. ISCCP identifies these pixels as cloudy and retrieves optical properties retrieved from them consistently, even though more than a third (5.1% of 14.2%) of the clouds assigned $0.3 \leq \tau \leq 1.3$ are assigned $p_c <$ 180 hPa, consistent with ISCCP retrievals that did not produce self-consistent solutions for τ and p_c. The two primary cloud properties reported by ISCCP (τ and p_c) are most relevant to studies of the earth's radiation budget and the project's success can be judged, in part, by noting that the top-of-atmosphere radiation budget can be closed to good accuracy using the cloud properties retrieved by the project (Zhang et al. 2004); even very numerous clouds with $\tau \leq 1.3$ have a modest effect on this budget. MODIS takes a more conservative approach, identifying these pixels as cloud affected but excluding them from cloud retrievals. This approach

accepts a truncation error in lieu of introducing unknown biases (caused by applying algorithms assuming plane-parallel, homogeneous clouds outside that range) that might be expected to differ by retrieved variable. Either choice can be rationalized but, as we discuss below, the decision can have a strong influence on model evaluation.

4. Synthesizing MODIS observations of clouds from model states

The cloud properties observable from space are only indirectly related to the internal representation of those clouds in global models. In general, clouds are represented by a time-evolving probability distribution of cloud condensate in each grid cell; the most frequent representation is the cloud fraction and mean liquid and ice water contents. Instrument simulators combine the model's internal description of clouds with information about the retrieval process to produce the statistics available in the observational record. Simulators can be divided into three parts: 1) a treatment of subgrid-scale variability, 2) the simulation of pixel-scale retrievals, and 3) the calculation of statistics.

The first step is required because any given profile of cloudiness, in all but the simplest cases, implies a distribution of possible retrievals. The ISSCP simulator introduced the idea of drawing Monte Carlo samples, known as "subcolumns," from this distribution (see also Yu et al. 1996). The samples are constructed so that each subcolumn can be considered homogeneous, while a large ensemble reproduces the input statistics (e.g., fractional cloudiness in each layer). The rules for doing so can be arbitrarily simple or complex according to the assumptions made by the host model; readers seeking more details on the construction of subcolumns should see Fig. A1 of Klein and Jakob (1999) as well as Räisänen et al. (2004). This step is not necessary in cloud-scale models with sufficient resolution that subgrid-scale homogeneity can be assumed (e.g., Marchand and Ackerman 2010).

Steps 2 and 3 have direct analogies to the steps used to process observations (see section 2). Step 2 is also conceptually similar to the forward operators used to predict observations of, for example, satellite radiance during data assimilation. Instrument simulators in climate models tend to be less detailed than the operators used in data assimilation, however, in part because comparisons to observations are done statistically, rather than one by one.

We have developed a MODIS simulator in this framework. We assume that subcolumns are available. The software first operates on these to produce synthetic pixel-level retrievals, and then a second procedure aggregates collections of retrievals to produce temporal averages and joint histograms. The simulator is integrated into the Cloud Feedback Model Intercomparison Project (CFMIP) Observation Simulator Package (COSP) software suite (Bodas-Salcedo et al. 2011) though it can easily be adapted to other contexts (e.g., cloud-resolving models). The MODIS simulator is available on the COSP website (http://cfmip-obs-sim.googlecode.com).

a. Simulating ISCCP observations

To provide context for the MODIS simulator, it may be useful to review the ISCCP simulator that inspired it [see also the discussion in Mace et al. (2011)]. In each subcolumn, the ISCCP simulator computes τ by summing the optical depth (integrated extinction) in each layer. Clear- and all-sky infrared zenith intensities are then computed using a simple radiative transfer model and user-provided cloud emissivity in each layer. Cloud-top temperature T_c is inferred from τ, the top-of-atmosphere infrared intensity, and the surface temperature and emissivity. Two adjustments are made at this point. First, if a cloud-top temperature cannot be found that reproduces the top-of-atmosphere infrared intensity (i.e., if the retrieval fails), then the cloud-top temperature is set to 5 K less than the temperature at the tropopause (effectively placing the cloud-top pressure at the tropopause itself). Second, the optical thickness of very thin clouds may be slightly increased to account for "IR-only" detection. Finally, the temperature profile of the troposphere is searched to find the pressure corresponding to the cloud-top temperature, and the cloud albedo determined by an analytic approximation (described in the appendix) to the tables used in processing ISCCP observations.

The ISCCP simulator aggregates pixel-scale retrievals to report grid-mean values of c_f (the fraction of subcolumns with $\tau \geq 0.3$), $\overline{A}$, τ_r, $\overline{T_c}$, $\overline{p_c}$, and the joint histogram $c_f(\tau, p_c)$; the latter contains a category for subcolumns with $\tau < 0.3$ that are present in the model but would be too thin to be observed.

b. Simulating pixel-scale MODIS observations

In keeping with the wider diversity of retrievals produced by MODIS instruments, the MODIS simulator requires a greater diversity of inputs than does the ISCCP simulator, including profiles of particle size for liquid and ice clouds, $r_{e,l}$ and $r_{e,i}$, and the corresponding liquid and ice optical depths at 0.67 μm within each layer of each subcolumn as a function of the model's vertical coordinate z. Users may opt to provide a single value of optical depth and the mixing ratios of cloud ice and liquid, in which case optical depth is partitioned by phase, assuming that particles are in the geometric optics limit. Mixed-phase clouds may be represented by

providing values of particle size and optical depths for both phases within a grid cell. The value of cloud-top pressure in each subcolumn as diagnosed by the ISCCP simulator is also required.

For each subcolumn we return a binary cloud mask and, for subcolumns deemed cloudy, the phase P (liquid, ice, or undetermined), p_c, τ, and retrieved particle size r_e.

As with the ISCCP simulator, optical thickness is derived by integrating the sum of the extinction σ due to liquid and ice clouds through the depth of the atmosphere as follows:

$$\tau = \int_{\mathrm{TOA}}^{\mathrm{sfc}} \sigma_l(z) + \sigma_i(z)\, dz, \tag{1}$$

where the extinction is computed by dividing the optical depth in each layer by the physical thickness of the layer. This is implemented as a sum of the optical depths in each layer and, like the ISCCP simulator, assumes that optical thickness can be retrieved with no error.

We mimic the daytime cloud mask as specifying a minimum detectable optical thickness $\tau_{\min}$. In the present version, we set $\tau_{\min} = 0.3$ based on comparisons between MODIS and high-sensitivity lidars (Ackerman et al. 2008). Note that the same threshold is used in the ISCCP simulator.

Proxies for the phase and cloud-top pressure retrievals are fairly rough. Comparisons of CO_2 slicing retrievals of p_c with profiles of extinction from spaceborne lidar, suggest that the pressure reported is insensitive to cloud beyond the first optical depth (Holz et al. 2006). We approximate the retrieval of cloud-top pressure by reporting the mean extinction-weighted pressure of the first optical depth, that is,

$$p_c = \frac{1}{\tau}\int_{\mathrm{TOA}}^{\tau} p(z)[\sigma_l(z) + \sigma_i(z)]\, dz,$$
$$\tau = \min\left[1, \int_{\mathrm{TOA}}^{\mathrm{sfc}} \sigma_l(z) + \sigma_i(z)\, dz\right]. \tag{2}$$

The retrieved value of cloud-top pressure is not constrained to fall at a model interface level. We mimic the upper limit of CO_2 slicing (see section 2b) by reporting the cloud-top pressure from the ISCCP simulator for subcolumns where Eq. (2) provides a value greater than 700 hPa.

Similarly, calculations performed by Wind et al. (2010) suggest that MODIS phase determination is sensitive to the phase in the highest portion of the cloud. We model this simply by computing

$$P = \frac{1}{\tau}\int_{\mathrm{TOA}}^{\tau} P_l(z)\sigma_l(z) + P_i(z)\sigma_i(z)\, dz,$$
$$\tau = \min\left[1, \int_{\mathrm{TOA}}^{\mathrm{sfc}} \sigma_l(z) + \sigma_i(z)\, dz\right], \tag{3}$$

where P is an integer indicating liquid P_l or ice ($P_i = P_l + 1$). Where less than 70% of the extinction in the first optical depth arises from a single phase (i.e., where $0.3 < |P - P_l| < 0.7$), the phase is considered undetermined; otherwise, P is rounded to identify the phase as entirely liquid or ice.

Particle size is estimated using a simplified pseudoinversion. During the development of the operational MODIS algorithms we built two large lookup tables, one each for liquid and ice clouds, summarizing the optical properties of cloud particle size distributions as a function of effective radius. For use in the MODIS simulator, we approximate the lookup tables for wavelength $\lambda = 2.1\ \mu$m with polynomial fits for the dependence of single-scattering albedo ω_0 and asymmetry parameter g on r_e. We use these fits to determine the total (liquid plus ice) cloud single-scattering albedo $\omega_0[r_e(z)]$ and asymmetry parameter $g[r_e(z)]$, which we combine with the total extinction $\sigma_l(z) + \sigma_i(z)$ to compute a predicted top-of-atmosphere near-infrared albedo R_p for each subcolumn using the two-stream approximation in each layer and adding methods to compute the reflectance of the entire column. We compare this albedo to a set of albedos $R_t(r_e)$ made for homogenous clouds of phase P [as determined in Eq. (3)] using the value of optical thickness obtained in Eq. (1). We compute albedo at a fixed set of trial sizes and linearly interpolate to find the value of r_e that minimizes $R_p - R_t(r_e)$. As in the processing of MODIS observations, subcolumns for which phase is undetermined, assume liquid drops in the retrieval of particle size. If the pseudoretrieval fails, the values of all cloud properties are set to missing values. Uncertainties due to the clear atmosphere and surface are neglected.

c. *Simulating grid-scale MODIS observations*

Synthetic MODIS pixels are aggregated to produce the statistics available in the custom observational dataset (see Table 1), although only a single set of cloud fractions is produced. Liquid and ice water paths for each pixel are inferred from τ and r_e assuming a vertically uniform cloud. In practice, statistics are normally computed for each model column at each diagnostic time step. This slightly complicates time averaging: MODIS observations represent the average over the number of pixels observed, so a time series of aggregate results from the MODIS simulator must be weighted by the appropriate

cloud fraction at each time in order to produce a time mean comparable to the observations (the ISCCP simulator shares this trait).

5. Understanding differences in ISCCP-like and MODIS-like views of a climate model's cloudiness

To what extent can the differences between the MODIS and ISCCP views of the earth's cloudiness described in section 3 be captured by instrument simulators like those described in section 4? We address this question using results from the Atmospheric Model, version 3 (AM3; Donner et al. 2011), a new global model developed by the Geophysical Dynamics Laboratory, in which we have implemented the full COSP. The description of the subgrid-scale structure in AM3 is fairly involved, and includes a method for diagnosing internal variability based on cloud condensate amounts and cloud fraction in each layer as well as an overlap assumption that accounts for correlations in condensate concentration (Pincus et al. 2005, 2006). The model also predicts aerosol concentration and cloud drop sizes; MODIS is one of just a few publicly distributed global observational datasets available for evaluation of the latter [others include the Advanced Very High Resolution Radiometer Pathfinder Atmospheres Extended dataset (PATMOS-x, available from http://cimss.ssec.wisc.edu/patmosx/) and the retrievals from the MODIS CERES team]. Results here are taken from a 26-yr run (1980–2006) using observed sea surface temperatures and greatly simplified sulfate chemistry.

Comparing the descriptions of the ISCCP and MODIS simulators (sections 4a and 4b, respectively), one can see that the determination of optical thickness (and hence cloud detection) in individual subcolumns is nearly the same in the MODIS and ISCCP simulators. There are two (presumably rare) exceptions: subcolumns containing very optically thin clouds, for which the ISCCP may increase the optical thickness to flag the subcolumn as cloudy, and subcolumns that are excluded because the MODIS simulator's particle size pseudoretrieval fails. Indeed, total cloud cover produced by the ISCCP simulator running in AM3 (56.2%, counting only "detectable" clouds with $\tau > 0.3$) is only slightly larger than that produced by the MODIS simulator (54%). Almost half this difference is due to clouds with $1.3 \leq \tau < 3.6$ (see Fig. 9), although simulated ISCCP cloud fractions exceed simulated MODIS values slightly across the range of optical thickness, which we attribute to failures of the MODIS simulator particle size pseudoretrieval in mixed-phase clouds for which liquid cloud lookup tables were used.

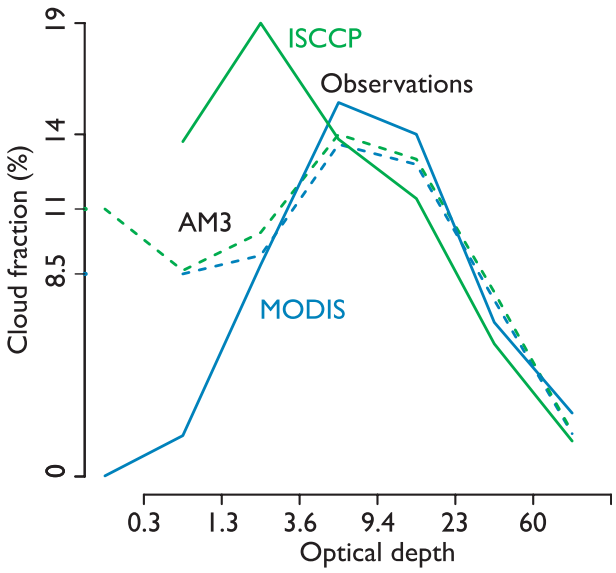

FIG. 9. Marginal histogram of global-mean cloud fraction as a function of optical thickness in observations (solid) and from instrument simulators running in the AM3 climate model (dashed): ISCCP and the ISCCP simulator are in green, MODIS and the MODIS simulator are in blue. The instrument simulators produce very similar distributions, though some MODIS simulator cloud fractions are slightly smaller at most values of τ due to failed retrievals of particle size in mixed-phase clouds. Neither simulator accounts for retrievals in scattered clouds where the observations disagree most sharply (see section 3). Clouds with $\tau \leq 0.3$ are quite common in AM3 though they are not detectable in passive observations.

The degree to which the two simulators produce different distributions of cloud-top pressure, on the other hand, depends more strongly on the joint distribution of τ and p_c produced by the host model. If, for example, all high clouds in the model are optically thick, then infrared-based estimates should agree well with the proxy for CO_2 slicing [Eq. (2)]. This is not the case with AM3, however: the largest differences between synthetic MODIS and ISCCP observations is related to high clouds (Fig. 10), which the ISCCP simulator puts in the highest bin ($p_c \leq 180$ hPa) and the MODIS simulator places somewhat lower in the atmosphere. This result is somewhat surprising at face value, since both simulators assume perfect knowledge of τ, but it may arise from inconsistencies between the clouds' true emissivity, as computed by AM3, and that assumed by the ISCCP simulator based on visible wavelength optical thickness; the joint histograms of cloud-top pressure and optical thickness (Fig. 11) for the two simulators are quite similar for clouds with $\tau > 3.6$, while the ISCCP simulator puts essentially all clouds with $\tau \leq 1.3$ in the highest cloud-top pressure category, consistent with the failure to find consistent τ, p_c pairs.

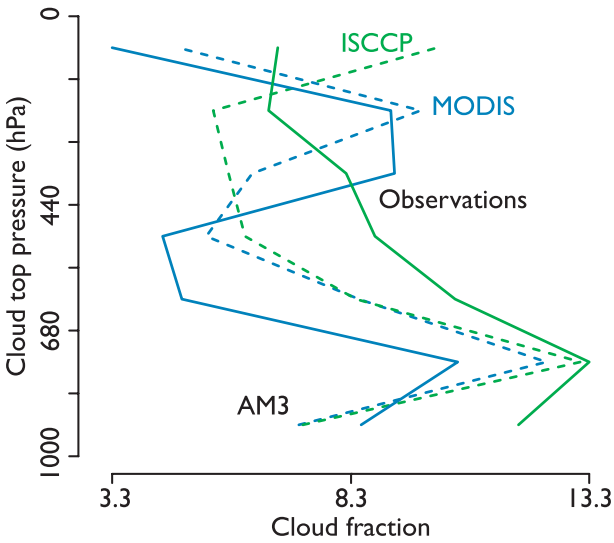

FIG. 10. Marginal histogram of global-mean cloud fraction as a function of cloud top pressure in observations (solid) and from instrument simulators running in the AM3 climate model (dashed): ISCCP and the ISCCP simulator (excluding clouds with $\tau \leq 0.3$) are in green, MODIS and the MODIS simulator are in blue. Total observed cloud fractions are lower in MODIS than ISCCP primarily because MODIS does not aggregate marginal pixels, which tend to be optically thin and low in the atmosphere. ISCCP thermal algorithms assign far more cloudiness to the middle of the atmosphere ($p_c \leq 440 < 680$ hPa) than does MODIS's CO_2 slicing method. The MODIS and ISCCP simulators do not capture this detail, in part because the joint distribution of clouds in AM3 is not the same as the distribution in nature.

Figures 9 and 10 also include verifying observed distributions of cloudiness, and Fig. 11 can be compared with Fig. 6. Robust differences between the model and observations, such as the lack of modeled clouds with $p_c > 800$ hPa, can be confidently attributed to model error. For clouds with $\tau \leq 1.3$, however, the difference between the observations is commensurate with the difference between the model and either set of observations.

6. Interpreting differences among observations and model estimates of cloud properties

a. *How well can the distribution of optical thickness be measured?*

Instrument simulators codify knowledge about observational uncertainties. The ISCCP and MODIS simulators assume that optical thickness can be determined perfectly [as does the MISR simulator described by Marchand and Ackerman (2010)], based in part on comparisons between optical thickness retrievals and independent measurements in very carefully controlled situations (e.g., Platnick and Valero 1995). But comparisons of the distribution of optical thickness obtained by

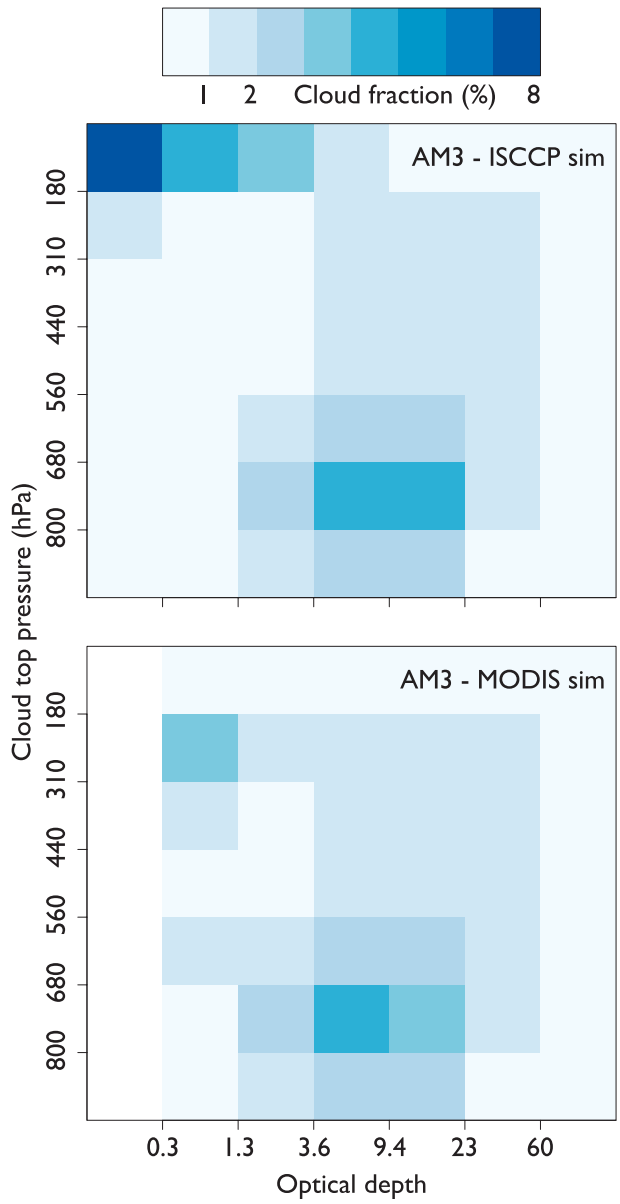

FIG. 11. Climatological distribution of global-mean cloud fraction as a function of cloud top pressure (vertical axis) and cloud optical thickness (horizontal axis) from the (top) ISCCP and (bottom) MODIS simulators running in the AM3 climate model. Very optically thin clouds ($\tau \leq 0.3$) cover about 11% of the earth in AM3; these are not detectable in either set of observations. The simulators reproduce the higher cloud-top pressures observed by ISCCP (Fig. 6) but not the divergent treatments of partly cloudy pixels. Robust differences between the model and observations, such as the lack of modeled clouds with $p_c > 800$ hPa, can be confidently attributed to model error.

these programs (section 3c) disagree substantially, indicating that the true uncertainty is far from zero.

Some of the differences between MODIS and ISCCP are likely to be explained by the sensitivity of cloud

detection and optical thickness to viewing and illumination geometry, and quantifying how this sensitivity impacts climatological estimates would be useful. Others differences are due to different assumptions and ancillary data, and these will continue to be refined. As one example, recent comparisons between the optical thickness derived from MODIS observations in the visible with those in the infrared and those from active sensors suggest that the model of ice particles used by MODIS is too strongly forward scattering, which will lead to systematic overestimates of optical thickness in ice clouds. Based on this evidence, the MODIS science team has plans to change the ice model used in our retrievals; ice cloud optical thicknesses will change accordingly.

Some of the uncertainty, too, no doubt arises from interpreting top-of-atmosphere measurements using plane-parallel, homogeneous models. Mace et al. (2011, and references therein), for example, show that ISCCP and similar algorithms routinely underestimate cloud optical thickness relative to simultaneously obtained surface measurements. They attribute this difference to small (subpixel)-scale variability in cloud properties within fully cloudy pixels; a similar explanation has been offered to explain counterintuitive relationships between the strength of absorption by water and the particle size retrieved by MODIS as a function of wavelength (Zhang and Platnick 2011). Since the observed radiation field is only rarely consistent with plane-parallel, homogeneous clouds (Liang et al. 2009; Di Girolamo et al. 2010), it seems useful to quantify the impacts of realistic small-scale variability on cloud retrievals and, where possible, to find ways to identify observations strongly affected.

The largest differences between distributions of optical thickness derived by MODIS and ISCCP, though, stem from the treatment of partly cloudy pixels (section 3c). These account for almost all pixels with $\tau < 1.3$ in the MODIS record (Fig. 7) and for 17% of all MODIS observations. This suggests that the ISCCP distribution contains a significant proportion of retrievals from partly cloudy pixels.

Partly cloudy pixels are a dramatic failure of the plane-parallel, homogeneous model—such a dramatic failure, we argue, that it is not clear what is meant by "optical thickness" in broken clouds: though the quantity is formally defined as the vertical integral of extinction at any location, there is little utility in using this measure when it does not describe the interaction of clouds with the radiation field. Rossow et al. (2002, p. 560) made the same point a decade ago: "There is a minimum horizontal scale at which the optical thickness is a meaningful quantity. This scale is imposed by the interaction of the radiation field with the inhomogeneous extinction field: *when 3D radiative effects are important, then optical thickness is not a meaningful parameter*" (italics in the original). We argue that it is misleading to interpret optical thickness retrievals from partly cloudy pixels in the same way as measurements from pixels that more closely fit the homogeneous, plane-parallel model.

b. *On the limits of instrument simulators*

Forecast or climate model evaluation relies on the comparison of simulations with observations, and where some aspect of a model's state is not directly observable a further model of the observation process must be invoked. Instrument simulators are one class of such interpretive models. Using retrievals of physical quantities (e.g., water and ice concentrations from radar reflectivity observations) is, in many ways, less direct. Users are stuck with whatever mapping between physical and observable properties is used in the retrievals, even when this is inconsistent with the model being evaluated: there can be no accounting for missing observations in, for example, very thick or thin clouds, and so on. Some mapping between model state and observations is always necessary; instrument simulators can be thought of as the most direct mapping for satellite observations of clouds.

We stress that the term "instrument simulator" is a misnomer, since all three steps described in section 4 are relevant for global models: assumptions made about the vertical coherence of clouds can have a profound impact on model predictions of column-integrated quantities, such as optical thickness (Morcrette and Jakob 2000; Pincus et al. 2006), and comparisons of optical thickness must account for the averaging strategy (e.g., radiative or linear weighting).

But, like all models, instrument simulators are incomplete. Our results suggest at least five circumstances in which even perfect instrument simulators may not reproduce artifacts in the observations.

1) Some observational errors depend on the state of the clouds and atmosphere. Adjustments to cloud-top pressures derived from infrared intensity, for example, are large only when the cloud is optically thin. If the distribution of cloud parameters in nature differs from those in the model being evaluated, then even ideal simulators will reproduce a different mixture of observations and observational artifacts than is observed. As one example, it appears that the large values of midlevel cloudiness provided by ISCCP are primarily observational artifacts that arise from the relatively frequent occurrence in nature of optically thin high clouds over optically thick low clouds, so the amount derived from models using simulators will be quite sensitive to the frequency and optical thickness distribution of high clouds.

2) Some observational errors are caused by erroneous ancillary data. As one example, ISCCP relates cloud-top temperature to pressure using a temperature profile derived from sounding instruments, but these profiles rarely contain the sharp inversions that cap many boundary layer clouds. The resulting underestimation of cloud-top pressure (Wang et al. 1999; Del Genio et al. 2005) cannot be reproduced by instrument simulators using perfect soundings. MODIS estimates for low clouds exhibit a similar bias (Holz et al. 2008). Similarly, errors caused by imperfect forward models, such as the possible overestimation of optical thickness caused by ice models used in MODIS retrievals, cannot be captured by the instrument simulator.
3) Simulators may assume no error in some part of the retrieval process, as both the MODIS and ISCCP simulators do for the retrieval of cloud optical thickness, when errors may be present in the observations. Making this assumption, even if there are no reasonable alternatives, may have further implications. In particular, we expect that assuming perfect knowledge of τ, even for small τ, means that the ISCCP and MISR MODIS simulators are more likely to find self-consistent retrievals than are the corresponding observations.
4) Observational artifacts may be caused by factors that are neglected by the simulators. The modest but clear geographic patterns in ISCCP observations (section 2a), for example, are due to the dependence of retrieved cloud properties on view angle and ISCCP's preference for geostationary observations. Including this effect in global models is possible but has not yet been considered important enough to warrant the considerable effort.
5) Observational artifacts may be caused by factors that are conceptually incompatible with some aspect of the simulators. As we showed in section 3, differences in the way ISCCP and MODIS treat partly cloudy pixels have a dramatic effect on the climatology of cloudiness and the distribution of optical thickness in thin clouds. The frequency of these pixels depends directly on the relative spatial scales of the satellite instruments and the clouds themselves. But the Monte Carlo generation of subcolumns (section 4) is agnostic with respect to scale and cannot produce partly cloudy pixels. One might, for example, develop a model of cloud spatial scale that depends on cloud type (e.g., Alexandrov et al. 2010), but the synthetic observational artifacts introduced would depend entirely on this model and so might be better addressed independently.

Given that the ISCCP simulator predicts cloud-top pressure quite accurately given accurate inputs (Mace et al. 2011), we suspect the large discrepancies between the observed and simulated distributions of cloudiness with height (Fig. 10) is due to some combination of the first three factors in addition to model errors being sought, while the simulation for optical thickness retrieval is affected by the second (e.g., erroneous models for ice particle scattering in one or both sets of retrievals), fourth (no accounting for artifacts caused by viewing and illumination geometry), and fifth (no treatment of partly cloudy pixels) factors.

c. *Designing robust measures for model evaluation*

Although instrument simulators cannot entirely close the gap between models and observations of clouds, there are broad regimes in which they work well: we showed in section 3 that MODIS and ISCCP agree when integrating over all but the optically thinnest pixels and, in section 5, that the corresponding instrument simulators share this behavior (see also Mace et al. 2011). This suggests that model evaluation can be made more robust by focusing on that part of parameter space where instrument simulators capture the bulk of the observational artifacts, and by focusing on measures that integrate over this space rather than the detailed distribution of clouds within it (e.g., Williams and Webb 2009).

One specific implication is that total cloud fraction is too fragile a measure to be useful in model evaluation. Cloud fraction has historically been attractive because it has been considered easier to make a binary decision about the presence or absence of cloud than to infer the value of some continuous quantity (see, e.g., Pincus et al. 2011). But, as the results of section 3 highlight, total cloud fraction is quite sensitive to the spatial scale at which it is measured (see also Wielicki and Parker 1992; Di Girolamo and Davies 1997), and at the kilometer scales of current satellite instruments, different assumptions can have dramatic impacts on the climatology. The agreement between MODIS and ISCCP is greatly improved when clouds with $\tau < 1.3$ are excluded, suggesting that applying a similar (entirely empirical) filter to the models before evaluation would make the comparisons more robust. Figure 12 demonstrates how this impacts the evaluation of AM3. The top panel shows an evaluation of cloud fraction (ISCCP and MODIS retrieval fraction) against the respective simulators; the biases between the two observational datasets (pink) is much larger than either model estimate and a substantial fraction of the variability in cloudiness. Restricting the comparison to pixels and subcolumns with $\tau \geq 1.3$ brings agreement between the two sets observations to nearly the levels seen in comparing the cloud-affected area (cf. with Fig. 2). Using more restrictive thresholds (as in Marchand et al. 2010,

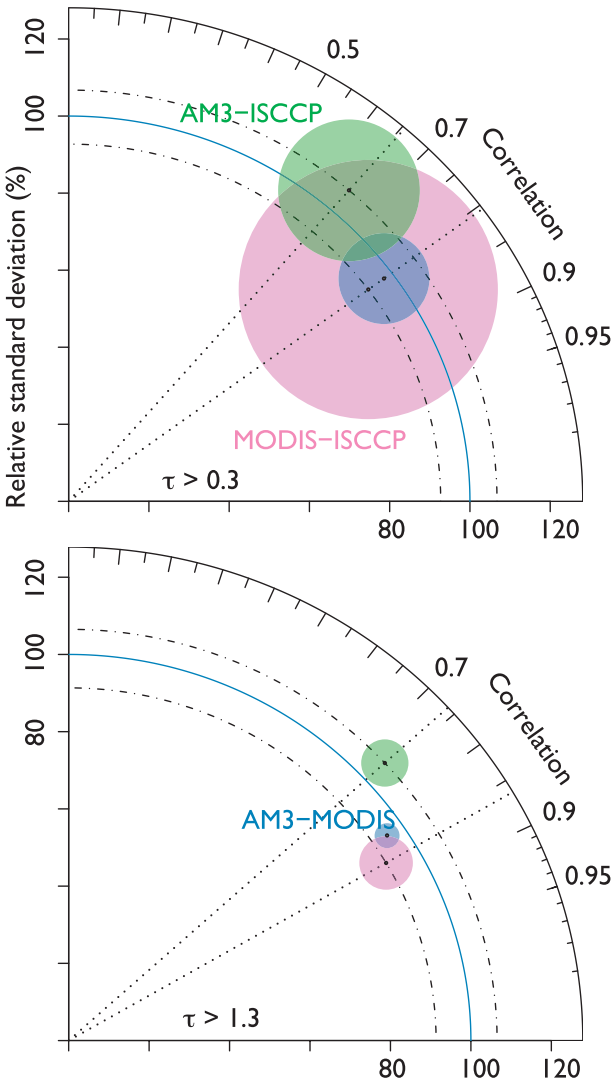

FIG. 12. Taylor diagrams showing agreement between measures of (top) total cloudiness and (bottom) clouds with greater than 1.3 between AM3's ISCCP simulator and ISCCP observations (green), AM3's MODIS simulator and MODIS observations (blue), and the two sets of observations (pink). Statistics are computed globally over the composite seasonal cycle. Observations are the reference for comparison with the model; because the reference varies among the comparisons, the Taylor diagram has been nondimensionalized. Excluding optically thin clouds from the comparison greatly reduces the bias and improves the correlation between the two sets of observations, making model evaluation more robust.

e.g.) reduces bias differences between the MODIS and ISCCP datasets but degrades other measures, including the bias relative to the remaining cloud fraction and the correlations between the two climatologies.

The stark differences between MODIS and ISCCP estimates of cloudiness suggest more flexible strategies for aggregating observations. The treatments of marginal pixels in the present versions of the observations lie at two ends of a spectrum. The next revision of the MODIS data will include separate aggregation of one or more categories of pixels flagged by the clear-sky restoral algorithm, allowing users to compute statistics for larger populations at the cost of increasing uncertainty. ISCCP could follow suit by separately aggregating those pixels for which self-consistent pairs of τ and p_c could not be found. Even a small number of categories would allow users to explore the consequences of making less drastic decisions and to more sharply delineate more substantive disagreements.

Acknowledgments. We appreciate advice from Stephen A. Klein on the design of the MODIS simulator and with helping us sharpen the arguments in the initial manuscript. We thank Tony Del Genio, Larry Di Girolamo, and Roger Marchand for a terrifically helpful set of reviews. Gang Ye and Gary Fu of the MODIS Data Processing System went to great lengths to produce our custom dataset from the original MODIS files, and Gala Wind processed heroic amounts of data to help us understand the impact of clear-sky restoral. We are grateful to Yuying Zhang for help in the production of the ISCCP dataset and to Paul Hubanks for sharing essential knowledge about the details of the MODIS level 3 products. RP thanks the Max Planck Institute for Meteorology (Hamburg) for its gracious hospitality and lively discussions during portions of this project. This work was supported by NASA under Grant NNX08AD65G.

APPENDIX

Computing "Radiatively Effective" Optical Thickness

One of the fundamental properties retrieved by MODIS and ISCCP at the pixel scale is the optical thickness τ measured at some wavelength in the visible portion of the spectrum. Optical thickness is the primary control on albedo, but the relationship between these quantities is nonlinear. For this reason, both datasets provide ways to estimate the optical thickness corresponding to the mean albedo: ISCCP provides τ_r (see section 2a), while MODIS provides τ_l (section 2b). This appendix briefly explores the comparability of these two measures.

The ISCCP simulator uses an analytic approximation [$A = \tau^{\alpha}/(c + \tau^{\alpha})$, where $\alpha = 0.895$ and $c = 6.82$] to the lookup tables used to relate albedo A to τ during operational processing. We used this approximation to compute temporal averages when building the ISCCP dataset described in section 2a(1). The approximation to the lookup tables is quite good (Fig. A1). Albedo is linear

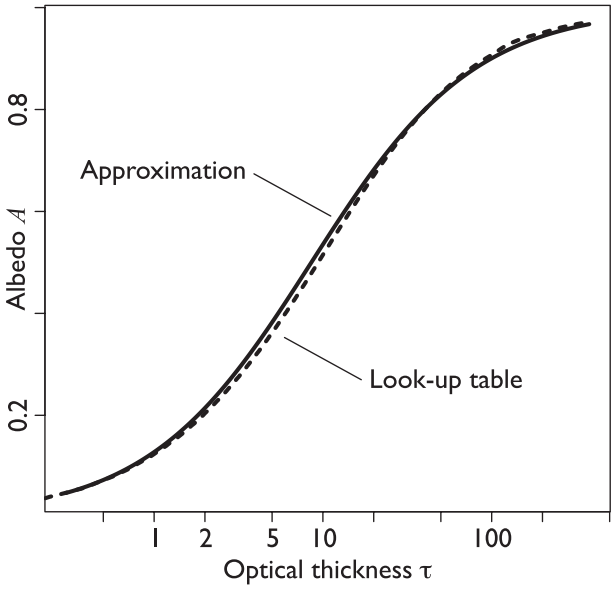

FIG. A1. Relationship between cloud albedo and optical thickness used in the ISCCP simulator (solid) and from the lookup tables used during operational processing of the ISCCP data (dashed). Values on the abscissa are logarithmically spaced. Over much of the range of optical thickness, A is linear in $\log(\tau)$, so that the logarithmically averaged optical thickness used by MODIS is comparable to the albedo-averaged optical thickness used by ISCCP.

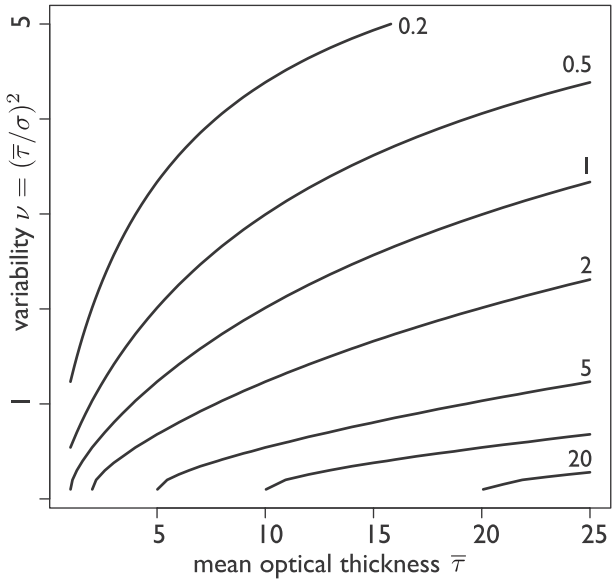

FIG. A2. Difference between logarithmically and radiative-averaged optical thickness $\tau_r - \tau_l$ for gamma distributions of optical thickness as a function of the distribution mean $\overline{\tau}$ and shape parameter $\nu = (\overline{\tau}/\sigma)^2$. The difference is in the opposite sense than is observed (see Fig. 5), while observations (discussed in the text) of ν support values of $\tau_r - \tau_l$ in the range 1–2. Thus, differences between ISCCP and MODIS mean optical depths indeed arise from differences in the populations of observations being averaged.

in $\log(\tau)$ over a large range of optical thickness (roughly 2–50); within this range τ_l and τ_r are comparable.

When integrated over a distribution of optical thickness (i.e., when using τ_l or τ_r to represent averages over space and/or time), the difference between the two measures depends on the details of the underlying distribution. Figure A2 shows this difference as a function of the parameters of a gamma distribution [as suggested by the observations of Barker et al. (1996)]. Note that $\tau_r > \tau_l$ for every choice of parameters for this distribution, while τ_l derived from MODIS is significantly larger, on average, than τ_r derived from ISCCP (see Fig. 5). Typical values of the shape parameter ν at daily time scales in 1° areas are roughly 2–4 (see Table 3b in Oreopoulos and Cahalan 2005), for which values of $\tau_r - \tau_l$ are less than 2 for almost all values of $\overline{\tau}$. The monthly averages in sections 2 and 5 will be broader because they include temporal variability (Pincus et al. 1999), but this will be mitigated by the weighting by cloud fraction at each observation time. This suggests that the result in Fig. 5 is indeed due to averaging over different populations of observations.

Top-of-atmosphere reflectance depends on surface albedo in addition to cloud properties, so that τ_l and τ_r provide the optical thickness corresponding to the average albedo only over dark surfaces. Users seeking more refined estimates must combine the distribution of optical thickness with temporally and spatially varying estimates of surface albedo.

REFERENCES

Ackerman, S. A., K. I. Strabala, W. P. Menzel, R. A. Frey, C. Moeller, and L. E. Gumley, 1998: Discriminating clear sky from clouds with MODIS. *J. Geophys. Res.,* **103** (D24), 32 141–32 157.

——, R. E. Holz, R. Frey, E. W. Eloranta, B. C. Maddux, and M. McGill, 2008: Cloud detection with MODIS. Part II: Validation. *J. Atmos. Oceanic Technol.,* **25,** 1073–1086.

Alexandrov, M. D., A. Marshak, and A. S. Ackerman, 2010: Cellular statistical models of broken cloud fields. Part I: Theory. *J. Atmos. Sci.,* **67,** 2125–2151.

Barker, H. W., B. A. Wielicki, and L. Parker, 1996: A parameterization for computing grid-averaged solar fluxes for inhomogeneous marine boundary layer clouds. Part II: Validation using satellite data. *J. Atmos. Sci.,* **53,** 2304–2316.

Baum, B. A., P. Yang, A. J. Heymsfield, S. Platnick, M. D. King, Y.-X. Hu, and S. T. Bedka, 2005: Bulk scattering properties for the remote sensing of ice clouds. Part II: Narrowband models. *J. Appl. Meteor.,* **44,** 1896–1911.

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

Cairns, B., 1995: Diurnal variations of cloud from ISCCP data. *Atmos. Res.,* **37,** 133–146, doi:10.1016/0169-8095(94)00074-N.

Chang, F.-L., and J. A. Coakley, 2007: Relationships between marine stratus cloud optical depth and temperature: Inferences from AVHRR observations. *J. Climate,* **20,** 2022–2036.

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

——, ——, ——, 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.

Del Genio, A. D., A. B. Wolf, and M.-S. Yao, 2005: Evaluation of regional cloud feedbacks using single-column models. *J. Geophys. Res.,* **110,** D15S13, doi:10.1029/2004JD005011.

Di Girolamo, L., and R. Davies, 1997: Cloud fraction errors caused by finite resolution measurements. *J. Geophys. Res.,* **102** (D2), 1739–1756.

——, L. Liang, and S. Platnick, 2010: A global view of one-dimensional solar radiative transfer through oceanic water clouds. *Geophys. Res. Lett.,* **37,** L18809, doi:10.1029/2010GL044094.

Donner, L. J., and Coauthors, 2011: The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component of the GFDL global coupled model CM3. *J. Climate,* **24,** 3484–3519.

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.

Evans, K. F., A. Marshak, and T. Várnai, 2008: The potential for improved boundary layer cloud optical depth retrievals from the multiple directions of MISR. *J. Atmos. Sci.,* **65,** 3179–3196.

Frey, R. A., S. A. Ackerman, Y. Liu, K. I. Strabala, H. Zhang, J. R. Key, and X. Wang, 2008: Cloud detection with MODIS. Part I: Improvements in the MODIS cloud mask for collection 5. *J. Atmos. Oceanic Technol.,* **25,** 1057–1072.

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

Holz, R. E., S. A. Ackerman, P. Antonelli, F. Nagle, R. O. Knuteson, M. McGill, D. L. Hlavka, and W. D. Hart, 2006: An improvement to the high-spectral-resolution CO_2-slicing cloud-top altitude retrieval. *J. Atmos. Oceanic Technol.,* **23,** 653–670.

——, ——, F. W. Nagle, R. Frey, S. Dutcher, R. E. Kuehn, M. A. Vaughan, and B. Baum, 2008: Global Moderate Resolution Imaging Spectroradiometer (MODIS) cloud detection and height evaluation using CALIOP. *J. Geophys. Res.,* **113,** D00A19, doi:10.1029/2008JD009837.

Horváth, Á, and R. Davies, 2004: Anisotropy of water cloud reflectance: A comparison of measurements and 1D theory. *Geophys. Res. Lett.,* **31,** L01102, doi:10.1029/2003GL018386.

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.

——, S. Platnick, P. A. Hubanks, G. T. Arnold, E. G. Moody, G. Wind, and B. Wind, 2006: Collection 005 change summary for the MODIS cloud optical property (06_OD) algorithm. NASA Rep., Version 3.1, 23 pp. [Available online at http://modis-atmos.gsfc.nasa.gov/C005_Changes/C005_CloudOpticalProperties_ver311.pdf.]

——, ——, G. Wind, G. T. Arnold, and R. T. Dominguez, 2010: Remote sensing of radiative and microphysical properties of clouds during TC^4: Results from MAS, MASTER, MODIS, and MISR. *J. Geophys. Res.,* **115,** D00J07, doi:10.1029/2009JD013277.

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

Koren, I., L. Oreopoulos, G. Feingold, L. A. Remer, and O. Altaratz, 2008: How small is a small cloud? *Atmos. Chem. Phys.,* **8,** 3855–3864, doi:10.5194/acp-8-3855-2008.

Liang, L., L. Di Girolamo, and S. Platnick, 2009: View-angle consistency in reflectance, optical thickness and spherical albedo of marine water-clouds over the northeastern Pacific through MISR-MODIS fusion. *Geophys. Res. Lett.,* **36,** L09811, doi:10.1029/2008GL037124.

Loeb, N. G., and J. A. Coakley, 1998: Inference of marine stratus cloud optical depths from satellite measurements: Does 1D theory apply? *J. Climate,* **11,** 215–233.

——, S. Kato, K. Loukachine, and N. Manalo-Smith, 2005: Angular distribution models for top-of-atmosphere radiative flux estimation from the clouds and the earth's radiant energy system instrument on the *Terra* satellite. Part I: Methodology. *J. Atmos. Oceanic Technol.,* **22,** 338–351.

Mace, G. G., and Coauthors, 2006: Cloud radiative forcing at the Atmospheric Radiation Measurement Program Climate Research Facility: 1. Technique, validation, and comparison to satellite-derived diagnostic quantities. *J. Geophys. Res.,* **111,** D11S90, doi:10.1029/2005JD005921.

——, S. Houser, S. Benson, S. A. Klein, and Q. Min, 2011: Critical evaluation of the ISCCP simulator using ground-based remote sensing data. *J. Climate,* **24,** 1598–1612.

Macke, A., 1993: Scattering of light by polyhedral ice crystals. *Appl. Opt.,* **32,** 2780–2788.

Maddux, B. C., S. A. Ackerman, and S. Platnick, 2010: Viewing geometry dependencies in MODIS cloud products. *J. Atmos. Oceanic Technol.,* **27,** 1519–1528.

Marchand, R., and T. P. 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.

——, ——, M. Smyth, 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.

Marshak, A., A. D. Davis, W. J. Wiscombe, and R. F. Cahalan, 1995: Radiative smoothing in fractal clouds. *J. Geophys. Res.,* **100** (D12), 26 247–26 261.

Menzel, W. P., W. L. Smith, and T. R. Stewart, 1983: Improved cloud motion wind vector and altitude assignment using VAS. *J. Climate Appl. Meteor.,* **22,** 377–384.

Minnis, P., and B. A. Wielicki, 1988: Comparison of cloud amounts derived using GOES and Landsat data. *J. Geophys. Res.,* **93** (D8), 9385–9403.

——, and Coauthors, 2011: CERES edition-2 cloud property retrievals using TRMM VIRS and Terra and Aqua MODIS data—Part I: Algorithms. *IEEE Trans. Geosci. Remote Sens.,* **49,** 4374–4400, doi:10.1109/TGRS.2011.2144601.

Morcrette, J.-J., and C. Jakob, 2000: The response of the ECMWF model to changes in the cloud overlap assumption. *Mon. Wea. Rev.,* **128,** 1707–1732.

Nakajima, T., and M. D. King, 1990: Determination of the optical thickness and effective particle radius of clouds from reflected solar radiation measurements. Part I: Theory. *J. Atmos. Sci.,* **47,** 1878–1893.

Norris, J. R., 2000: What can cloud observations tell us about climate variability? *Space Sci. Rev.,* **94,** 375–380.

——, and C. P. Weaver, 2001: Improved techniques for evaluating GCM cloudiness applied to the NCAR CCM3. *J. Climate,* **14,** 2540–2550.

Oreopoulos, L., and R. F. Cahalan, 2005: Cloud inhomogeneity from MODIS. *J. Climate,* **18,** 5110–5124.

Pincus, R., S. A. McFarlane, and S. A. Klein, 1999: Albedo bias and the horizontal variability of clouds in subtropical marine boundary layers: Observations from ships and satellites. *J. Geophys. Res.,* **104** (D6), 6183–6191.

——, C. Hannay, S. A. Klein, K.-M. Xu, and R. S. Hemler, 2005: Overlap assumptions for assumed probability distribution function cloud schemes in large-scale models. *J. Geophys. Res.,* **110,** D15S09, doi:10.1029/2004JD005100.

——, R. S. 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. P. Hofmann, K. E. Taylor, and P. J. Gleckler, 2008: Evaluating the present-day simulation of clouds, precipitation, and radiation in climate models. *J. Geophys. Res.,* **113,** D14209, doi:10.1029/2007JD009334.

——, R. J. P. Hofmann, J. L. Anderson, K. Raeder, N. Collins, and J. S. Whitaker, 2011: Can fully accounting for clouds in data assimilation improve short-term forecasts? *Mon. Wea. Rev.,* **139,** 946–957.

Platnick, S., and F. P. J. Valero, 1995: A validation of a satellite cloud retrieval during ASTEX. *J. Atmos. Sci.,* **52,** 2985–3001.

——, M. D. King, S. A. Ackerman, W. P. Menzel, B. A. Baum, J. C. Riédi, 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.

——, R. Pincus, B. Wind, M. D. King, M. A. Gray, and P. Hubanks, 2004: An initial analysis of the pixel-level uncertainties in global MODIS cloud optical thickness and effective particle size retrievals. *Passive Optical Remote Sensing of the Atmosphere and Clouds IV,* S. C. Tsay, T. Yokota, and M.-H. Ahn, Eds., International Society for Optical Engineering (SPIE Proceedings, Vol. 5652), doi:10.1117/12.578353.

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.

Reichler, T., and J. Kim, 2008: How well do coupled models simulate today's climate? *Bull. Amer. Meteor. Soc.,* **89,** 303–311.

Rossow, W. B., and R. A. Schiffer, 1991: ISCCP cloud data products. *Bull. Amer. Meteor. Soc.,* **72,** 2–20.

——, and L. C. Garder, 1993: Validation of ISCCP cloud detections. *J. Climate,* **6,** 2370–2393.

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

——, A. W. Walker, and L. C. Garder, 1993: Comparison of ISCCP and other cloud amounts. *J. Climate,* **6,** 2394–2418.

——, C. Delo, and B. Cairns, 2002: Implications of the observed mesoscale variations of clouds for the earth's radiation budget. *J. Climate,* **15,** 557–585.

——, G. Tselioudis, A. Polak, and C. Jakob, 2005: Tropical climate described as a distribution of weather states indicated by distinct mesoscale cloud property mixtures. *Geophys. Res. Lett.,* **32,** L21812, doi:10.1029/2005GL024584.

Shenk, W. E., and V. V. Salomonson, 1972: A simulation study exploring the effects of sensor spatial resolution on estimates of cloud cover from satellites. *J. Appl. Meteor.,* **11,** 214–220.

Stubenrauch, C., and Coauthors, 2009: Assessment of global cloud climatologies. *GEWEX News,* No. 1, International GEWEX Project Office, Silver Spring, MD, 6–7.

Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.,* **106** (D7), 7183–7192.

Wang, J., W. B. Rossow, T. Uttal, and M. Rozendaal, 1999: Variability of cloud vertical structure during ASTEX observed from a combination of rawinsonde, radar, ceilometer, and satellite. *Mon. Wea. Rev.,* **127,** 2484–2502.

Webb, M. J., 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.

Wielicki, B. A., and L. Parker, 1992: On the determination of cloud cover from satellite sensors: The effect of sensor spatial resolution. *J. Geophys. Res.,* **97** (D12), 12 799–12 823.

Williams, K. D., 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.

Wind, G., S. Platnick, M. D. King, P. A. Hubanks, M. J. Pavolonis, A. K. Heidinger, P. Yang, and B. A. Baum, 2010: Multilayer cloud detection with the MODIS near-infrared water vapor absorption band. *J. Appl. Meteor. Climatol.,* **49,** 2315–2333.

Yang, Y., and L. Di Girolamo, 2008: Impacts of 3-D radiative effects on satellite cloud detection and their consequences on cloud fraction and aerosol optical depth retrievals. *J. Geophys. Res.,* **113,** D04213, doi:10.1029/2007JD009095.

Yu, W., M. Doutriaux, G. Seze, 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.

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., 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, Z., and S. Platnick, 2011: An assessment of differences between cloud effective particle radius retrievals for marine water clouds from three MODIS spectral bands. *J. Geophys. Res.,* **116,** D20215, doi:10.1029/2011JD016216.

——, ——, P. Yang, A. K. Heidinger, and J. M. Comstock, 2010: Effects of ice particle size vertical inhomogeneity on the passive remote sensing of ice clouds. *J. Geophys. Res.,* **115,** D17203, doi:10.1029/2010JD013835.

Zhao, G., and L. Di Girolamo, 2006: Cloud fraction errors for trade wind cumuli from EOS-Terra instruments. *Geophys. Res. Lett.,* **33,** L20802, doi:10.1029/2006GL027088.