Evaluation of the Multiscale Modeling Framework Using Data from the Atmospheric Radiation Measurement Program

MIKHAIL OVTCHINNIKOV, THOMAS ACKERMAN, AND ROGER MARCHAND

Pacific Northwest National Laboratory, Richland, Washington

MARAT KHAIROUTDINOV

Colorado State University, Fort Collins, Colorado

(Manuscript received 30 November 2004, in final form 7 September 2005)

ABSTRACT

In a recently developed approach to climate modeling, called the multiscale modeling framework (MMF), a two-dimensional cloud-resolving model (CRM) is embedded into each grid column of the Community Atmospheric Model (CAM), replacing traditional cloud and radiation parameterizations. This study presents an evaluation of the MMF through a comparison of its output with the output from the CAM and with data from two observational sites operated by the Atmospheric Radiation Measurement Program, one at the Southern Great Plains (SGP) in Oklahoma and one at the island of Nauru in the tropical western Pacific (TWP) region.

Two sets of one-year-long simulations are considered: one using climatological sea surface temperatures (SSTs) and another using 1999 SST. Each set includes a run with the MMF as well as a CAM run with traditional or standard cloud and radiation treatments. Time series of cloud fraction, precipitation intensity, and downwelling solar radiation flux at the surface are analyzed. For the TWP site, the distributions of these variables from the MMF run are shown to be more consistent with observation than those from the CAM run. This change is attributed to the improved representation of convective clouds in the MMF compared to the conventional climate model. For the SGP, the MMF shows little to no improvement in predicting the same quantities. Possible causes of this lack of improvement are discussed.

1. Introduction

The sophistication of global climate models (GCMs) has increased dramatically over the last several decades naturally raising both the prospect and expectations of greater accuracy for their predictions. For the potential of latest advances to be fully realized, however, model development must go hand-in-hand with appropriate expansion of validation procedures to include more quantitative and strict tests of model performance. Such tests require detailed observations of the evolving atmospheric state.

The importance of comprehensive long-term local observations for climate studies has been long recognized and was the primary driving force behind the Atmospheric Radiation Measurement (ARM) program initiated by the U.S. Department of Energy (DOE) in 1991 (Ackerman and Stokes 2003). Since then, ARM has instrumented several operational sites worldwide and compiled an unprecedented archive of cloud, radiation, and general meteorological observations. Using these datasets for testing and improving cloud and radiation parameterizations in GCMs is not straightforward however.

GCM variables are subjected to implicit spatial and temporal averaging. A spatial scale is determined by horizontal dimensions of the model grid of order 100 km while a temporal scale is determined by the model time step, typically on the order of tens of minutes. In contrast, most ground observations provide either a time series of scalar measurements or a time series of vertically resolved profiles usually averaged over a much shorter period of time, typically a few seconds to a few minutes, and corresponding to scales of under a kilometer. Therefore, there is an enormous discrepancy

Corresponding author address: Mikhail Ovtchinnikov, Pacific Northwest National Laboratory, P.O. Box 999, Richland, WA 99352.
E-mail: mikhail@pnl.gov

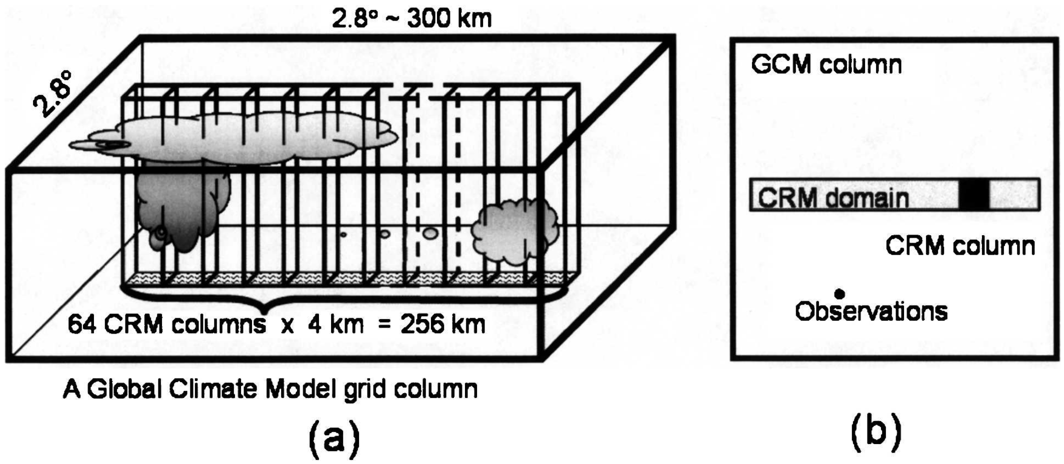

FIG. 1. (a) A sketch illustrating a two-dimensional CRM embedded within a grid column of a GCM. (b) An illustration of the horizontal projection (not to scale) shows inherently different spatial averaging of GCM and CRM variables and observations.

in time–space sampling that generally prevents direct model-to-observation comparison. An assumption of ergodicity [i.e., equivalence between temporal and spatial (or ensemble) statistics] and/or assumptions about subgrid-scale structure within each model grid cell can be used to convert the data into analogous formats, but these assumptions are not always accurate and are difficult to test.

An alternative approach to climate modeling called the multiscale modeling framework (MMF) employs an explicit treatment of cloud-scale processes. The MMF, also known as a cloud-resolving convection parameterization (Grabowski 2001) or superparameterization (Khairoutdinov and Randall 2001; Randall et al. 2003; Khairoutdinov et al. 2005), consists of a two-dimensional or small three-dimensional cloud-resolving model (CRM) embedded into each grid of the National Center for Atmospheric Research (NCAR) Community Atmospheric Model (CAM) or another GCM. As illustrated in Fig. 1, in the MMF the subgrid variability in cloud dynamics and cloud microphysics is explicitly resolved at spatial scales down to the resolution of the CRM. In addition, by applying the radiative transfer code to each CRM column, no arbitrary assumptions about cloud overlap are required, and the feedback of radiative heating on clouds is explicitly treated. The mean radiative heating for the CRM is applied to the heat budget of the large-scale grid cells. Thus, the CRM and the columnwise radiation code together replace the conventional cloud (stratiform and convective) and radiation parameterizations in the GCM. Detailed analysis of MMF-simulated global climate is ongoing. Several advantages of the approach have been already identified; including better defined Madden–Julian oscillation (MJO)-like systems and a more realistic diurnal cycle of precipitation (Khairoutdinov et al. 2005).

In this study, we evaluate the MMF by comparing its output with observations of cloud fraction, precipitation rate, and shortwave radiation flux from two ARM sites as well as with output from the CAM with its standard cloud and radiation parameterizations. Our approach to the comparative analysis is twofold. First, we look at the MMF output aggregated over the CRM domain. These domain statistics, which are representative of the CRM feedback to the parent GCM, are compared to statistics from traditional CAM parameterizations to test the potential advantage gained from the explicit treatment of cloud, precipitation, and radiation processes in the MMF. Second, we use the output from a single CRM column. This second approach allows for a more direct evaluation of the treatment of clouds and radiation in the MMF against ARM and other "local" observations, and to a limited degree lets us explore the opportunity that the MMF provides for simulating ground-based observing systems. A comparison of outputs from the CRM column and CRM domain can also help assess the relationship between temporal and spatial statistics at GCM grid scale to those at smaller scales.

The paper is organized as follows. Section 2 describes the observational datasets and model outputs used in the study and outlines the methodology of the comparison. Section 3 presents a comparative analysis of the cloud fraction, precipitation, and downwelling solar radiation flux at the surface. Finally, the results are summarized in section 4.

2. Models, data, and method of analysis

a. The CAM

The CAM is the atmospheric component of the Community Climate System Model (CCSM; Blackmon et al. 2001). The CAM configuration used in this study includes the semi-Lagrangian dynamical core (Williamson and Olson 1994), 26 layers in the vertical (stretched grid with the top at 3.5 mb), and T42 horizontal (spectral) resolution, which approximately corresponds to the 2.8° × 2.8° grid. Thus, each grid cell represents an area of approximately 300 × 300 km^2 in the Tropics. The time step for the CAM is 1 h. The model is forced with prescribed sea surface temperatures (SSTs).

The CAM physics package predicts cloud condensate (liquid plus ice) for layered (stratiform) clouds and diagnoses it for convective clouds. The model also diagnoses precipitation in the form of rain or graupel-like snow (Rasch and Kristjánsson 1998). This bulk microphysical parameterization is similar to that used in earlier cloud-resolving models. The diagnosis of cloud fraction is a generalization of the scheme by Slingo (1987). Cloud fraction depends on relative humidity, vertical velocity, atmospheric stability, and convective mass fluxes. Three types of clouds are diagnosed by the scheme with each having its own vertical profile of cloud fraction: low-level marine stratus ($\boldsymbol{C}_{\mathrm{st}}$), convective cloud with coupled anvils ($\boldsymbol{C}_{\mathrm{cir}}$), and layered cloud ($\boldsymbol{C}_c$). The total cloud fraction $\boldsymbol{C}_{\mathrm{tot}}$ at each level is then diagnosed as $\boldsymbol{C}_{\mathrm{tot}} = \max(\boldsymbol{C}_{\mathrm{st}}, \boldsymbol{C}_{\mathrm{cir}}, \boldsymbol{C}_c)$, which is equivalent to a maximum overlap assumption of cloud types within each grid box. The condensate value is assumed uniform within any and all types of clouds within each grid box.

The shortwave radiation is treated using the δ–Eddington approximation adopted by Briegleb (1992). The method employed to represent longwave radiative transfer is based on an absorptivity/emissivity formulation (Ramanathan and Downey 1986). The cloud overlap for radiative calculations is maximum random, that is, clouds in adjacent layers are maximally overlapped, and groups of clouds separated by one or more clear layers are randomly overlapped. The treatment of cloud vertical overlap follows Collins (2001).

b. The MMF

The MMF configuration used in this study is described in detail by Khairoutdinov and Randall (2001) and Khairoutdinov et al. (2005), and is only briefly summarized here. It consists of the CAM described above and an individual CRM running in each CAM grid column, resulting in a total of 8192 CRMs operating in parallel. The CRM grid includes 64 columns at 4-km spacing and 24 layers in the vertical coinciding with the lowest 24 CAM levels. Lateral boundary conditions are cyclic. The CRM predicts all three components of wind, the liquid water/ice moist static energy, and mixing ratios of total nonprecipitating water (vapor + cloud water + cloud ice) and total precipitating water (rain + snow + graupel). Partitioning of the two predicted water categories into vapor and hydrometeor mixing ratios is done diagnostically every time step using prescribed temperature dependencies. The CRM is forced by the temperature and humidity tendencies of the large-scale grid cells and feeds the response back to the large scale as a heating and moistening term in the large-scale budget equations for heat and moisture. The CRM–CAM exchange occurs every CAM time step (i.e., hourly), but the CRM runs continuously using a 20-s time step. The CRM orientation is east–west. Because the orientation is arbitrary and independent of the large-scale flow, the feedback of the CRM on the CAM momentum budget is neglected. In all cases, the location of the CRM within a CAM grid column is undefined because the surface conditions are the same for each CRM column. The CRM replaces the conventional cloud (stratiform and convective) parameterization in the CAM. By applying the CAM radiative transfer code to each CRM column, no arbitrary assumptions about cloud overlap are required and the subgrid feedback of radiative heating on clouds is explicitly treated. The grid cell mean of the radiative heating is applied to the heat budget of the large-scale grid cells.

As described, the MMF is computationally about 200 times more intensive compared to the "traditional" CAM. Consequently, only a handful of year-long simulations have been conducted to date, and two of these are analyzed here, as described below.

c. Observations

Observations used in this study come from two ARM sites: the tropical western Pacific (TWP) site located on the island of Nauru (0.521°S, 166.916°E) and the Southern Great Plain (SGP) Central Facility site in north-central Oklahoma (36.617°N, 97.50°W).

Cloud fraction statistics are derived from vertically pointing millimeter wave (35 GHz) cloud radar and lidar observations using the cloud masking algorithm of Clothiaux et al. (2000). The radar detects hydrometeors with reflectivities in the range of approximately −50 to +20 dBZ with good accuracy up to heights of 10–15 km above ground level with 90-m vertical resolution. The lidar is used to detect the cloud-base height with 30-m resolution.

Precipitation rates on the ground are measured by an optical rain gauge at the TWP Nauru site and by an electrically heated, tipping-bucket precipitation gauge at the SGP site. The optical rain gauge records the rainfall rate every minute with the uncertainty of ±0.1 mm h^{-1}. For the tipping-bucket gauge the sampling uncertainty is one full bucket (0.254 mm) per sample period, or 0.254 mm h^{-1} for hourly rain rate.

Each ARM site is equipped with a pyranometer to measure broadband solar irradiance on a planar surface. Time series analyzed in this study are composed of 1-min averages of downward shortwave flux that has an uncertainty of ~15 W m^{-2} or 3%, whichever is larger.

d. Time series and data analysis

Two sets of model runs are analyzed in this study: one using climatological SSTs and another using the observed SSTs for the year of 1999. All simulations are initialized on 1 September and last for about 500 days until the end of the following year (e.g., simulations for the year of 1999 start from 1 September 1998). To prevent seasonal bias and to eliminate model spinup effects, only one full year's worth of data (January through December) is analyzed, meaning that the first fall season is excluded from consideration.

Observations cover a period from 1998 through 2002. The complete time series is taken to represent climatological conditions while a subset for 1999 is compared with the model runs forced by that year's SSTs. For the SGP site, we use four years' worth of data from 1 September 1998 through 31 August 2002. During this period, the radar data are available 84%, precipitation measurements 93%, and broadband solar radiation measurements about 96% of the time. For the TWP site, we use three years' worth of data from 1 November 1998 through 31 October 2001. For that period, the radar data are available 64% of the time, the precipitation record is nearly continuous, and the broadband shortwave radiation measurements have very short interruption periods and cover 98.5% of the time. There could be a small bias introduced by missing observations, especially for cloud fraction at the TWP site.

All the time series used in the analysis are composed of 1- or 3-h averages as indicated in the text. The spatial averaging is inherently different for the CAM, MMF, and observations (Fig. 1b). The effect of these differences is discussed in the following sections, but in general one should expect progressively smaller extremes and less variability in moving from local (point) measurements to CRM grid column to CRM strip domain to CAM grid column. The long-term averages are expected to be independent of the spatial averaging.

TABLE 1. Parameters of the simulated and observed climatological total cloud fraction distributions for the TWP and SGP. For MMF the spatial (over CRM domain) cloud fraction is given. Zero median values indicate that clear-sky conditions are predicted more than 50% of the time.

	CAM	MMF	Observations
	Tropical western Pacific (Nauru)		
Mean	0.85	0.28	0.54
Median	0.99	0.25	0.51
Standard deviation	0.26	0.19	0.32
	Southern Great Plains		
Mean	0.30	0.24	0.48
Median	0.00	0.00	0.43
Standard deviation	0.42	0.34	0.43

3. Results

a. Cloud fraction

Cloud fractions (CFs) analyzed here are determined as follows. For any model column, the CAM CF is taken as predicted by the cloud parameterization (section 2a) and averaged over 1 or 3 h, as indicated below. In the MMF, the domain (or spatial) CF at any time is defined as the ratio of the number of cloudy CRM columns to the total number of columns. A CRM column is considered cloudy if its liquid/ice water path exceeds 20 g m^{-2}. Similar to the CAM CF, the MMF domain CF is averaged in time over the same period. Although time averaging is used in the analysis, both these CFs are instantaneously defined for any CAM grid point. In contrast, CF derived from the vertically pointing cloud radar is defined only for a time series as a ratio of the number of observations with a return signal exceeding a specified threshold to the total number of observations over that period. To mimic the radar observations, we introduce a third type of model CF. An MMF temporal CF is also computed using output from a single column from each CRM 2D strip. In this case, the MMF CF is either zero or one at a given instant (depending on whether the liquid/ice water content exceeds a prescribed threshold) much like a radar observation. Unfortunately, the CRM column data are available for the 1999 MMF run only.

A comparison of model-predicted CF from the climatological SST runs with radar-derived CF at the TWP site is presented in Table 1. The 3-h mean CFs differ greatly, ranging from 28% for the MMF (CRM domain) to 54% for observations, to 85% for CAM. Figure 2 shows the corresponding distributions of CF frequency of occurrence. Compared to observations at the TWP site, the CAM greatly overpredicts overcast conditions, while the MMF simulation lacks this regime (Fig. 2a). The latter discrepancy could be due, at least

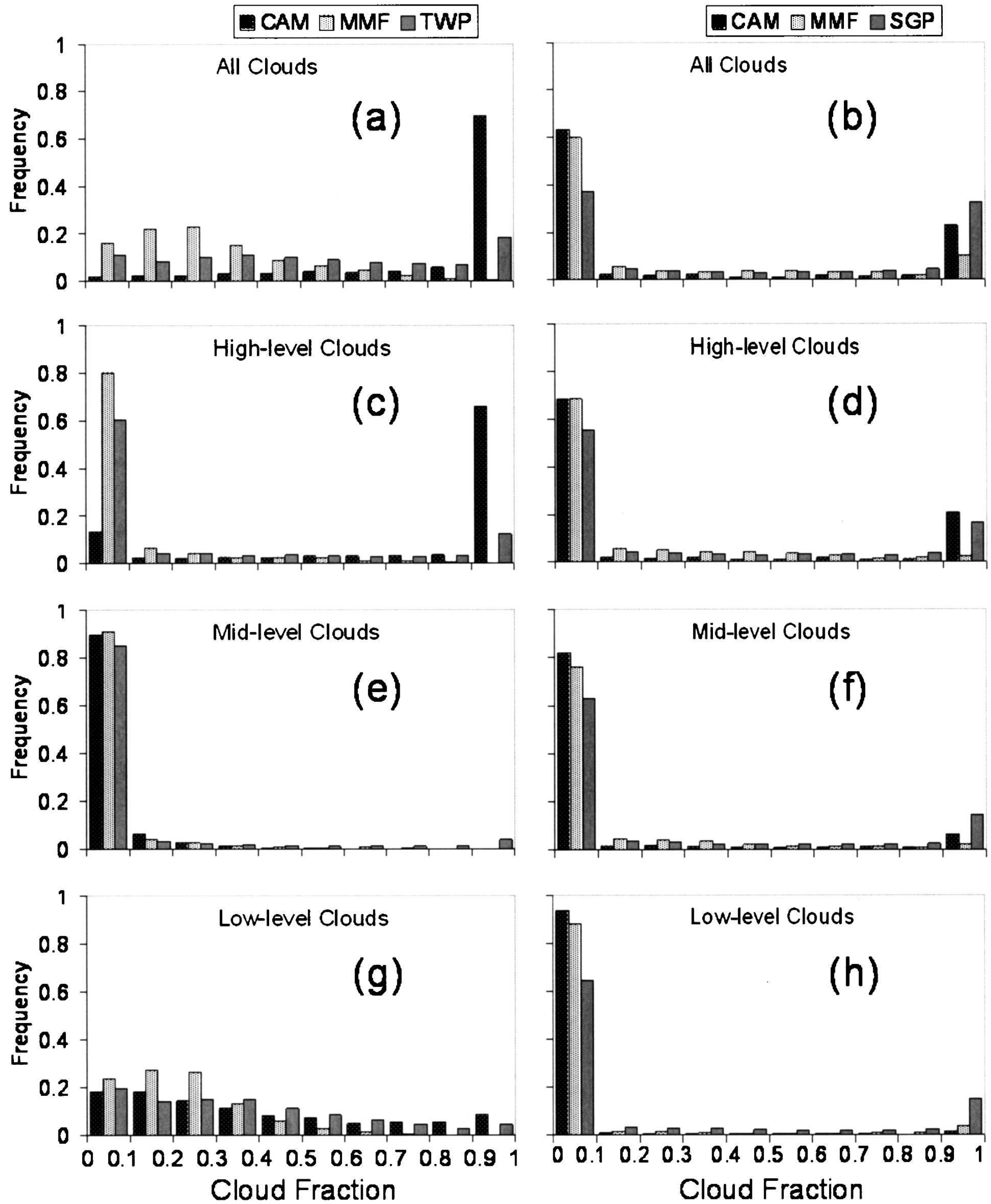

FIG. 2. Histograms of cloud fractions from CAM, MMF, and cloud radar observations. Four rows from top to bottom show cloud fractions for all, high-, middle-, and low-level clouds, respectively. The histograms are constructed by grouping all 3-h average values into 10 cloud fraction intervals. Model cloud fractions are taken at the grid columns covering the (left) TWP and (right) SGP locations.

partially, to sampling because the radar (with its narrow vertically pointing beam) can observe cloud continuously (for 1 to 3 h) and yet the true large-scale aerial cloud fraction can be less than 100%. The CAM CF on the other hand is clearly unrealistic since an overcast condition over a larger area (as in a CAM grid cell) must also be observed at any location within this area.

Comparison of cloud fractions for low- (below 700 mb), middle- (between 700 and 400 mb), and high- (above 400 mb) level clouds reveals that the frequent complete overcast conditions in the CAM run are primarily due to high-level (cirrus) clouds (Fig. 2c). This known shortcoming of the traditional CAM cloud scheme (Lin and Zhang 2004) is corrected, and perhaps overcorrected in the MMF run, in which overcast conditions are extremely rare in any 3-h period.

For the low-level cloud cover, CAM predictions are closer to observations than the MMF (Fig. 2g). There are at least two factors contributing to the negative MMF CF bias relative to observations. First, small cumulus clouds, which are a prominent feature in the Tropics, are not resolved on the 4-km CRM grid. Second, radar-derived cloud statistics at the TWP site are known to be contaminated by the presence of the island of Nauru, which acts as a local hot spot and preferentially generates boundary layer clouds during clear-sky or thin cloud periods. McFarlane et al. (2005) found that the absolute increase in low cloud frequency over the ARM site from the island-generated cloud plume can be as much as 10% during daytime under prevailing easterly trade winds. A long-term average effect is expected to be smaller but could contribute to the underprediction of low clouds by the MMF.

At the SGP site, both MMF and CAM underpredict the cloud amount at all levels (Figs. 2b,f,h), the only exception being the overprediction of complete overcast by high-level clouds in CAM (Fig. 2d). The positive bias in high-level cloudiness in CAM, although small in this case, partially compensates for the lack of low clouds and results in a total CF that is closer to observations than the CF from the MMF (Table 1).

The MMF allows us to explore the effects of various definitions of cloud fraction. First, we look at the effect of the time period over which the temporal CF is determined and the spatial CF is averaged. Figure 3 shows that the spatial (CRM domain) CFs averaged over 1 h (dom1) and 3 h (dom3) are virtually indistinguishable. Thus, for the 1999 MMF run, the change in statistics of the TWP/Nauru cloud field in the domain, which is 264-km long, is negligible over these intervals. Increasing the interval over which the temporal (CRM column) CF is computed from 1 h (col1) to 3 h (col3) predictably decreases the frequency of complete overcast and clear conditions and increases the frequency of intermediate CF, most notably in the range from 0.3 to 0.5.

Figure 3 illustrates that the distribution of the CF is affected very strongly by whether CF is calculated spatially or temporally (e.g., compare dom1 and col1 in Fig. 3a). The change is almost exclusively due to variations in low-level CF (Fig. 3g). Thus, for example, the CRM column contains low-level clouds for the whole hour about 10% of the time (col1; Fig. 3g), while the probability of the domain CF being equal to one is close to zero (dom1; Fig. 3g). The influence of the change in CF definition on the mid- and high-level statistics is negligible (Figs. 3c,e). The result is to be expected since the horizontal extent of cirrus clouds is usually much larger than that of low-level clouds.

The difference between the spatial/domain and temporal/column CF diminishes substantially when the sampling/averaging time interval increases from 1 to 3 h. It is certainly not surprising that averaging the temporal (column) CF for a longer averaging interval makes the distribution closer to the spatial (domain) CF. However, there are still clear differences between the 3-h temporal (single column, time averaged) and spatial (domain and time averaged) CFs even when using 3-h averages and a year of model output.

We also note that overall the simulated sensitivity of the temporal CF to the sampling interval (i.e., the change in the temporal CFs when interval is increasing from 1 to 3 h) reproduces the observed change very well (Fig. 3).

For the SGP site, the differences between observations and simulation results are larger but sensitivity of the simulated and observed temporal CFs to the averaging period is similar. The difference between simulated spatial and temporal CF are smaller in this case than is seen for the TWP site, however. Time averaging also has a smaller effect on CF distributions in simulations and observations. The likely reason is a generally larger horizontal extent of clouds including an apparent lack of small cumulus clouds in the region.

A lingering question in model-to-observation comparison is the adequacy of a definition of a cloud. Because the CRM provides explicitly the spatial distribution of condensate it allows us to explore this issue in more detail than could be done using, for example, CAM-diagnosed CF. To this end, we compare the sensitivity of the simulated and observed distributions of the temporal 1-h CF at the TWP site to a selected threshold used to define clouds. Specifically, we calculate the model CF using two different thresholds specified in terms of liquid water content (LWC) and ice water content (IWC) as

$$\mathrm{IWC} > 0.0165\ \mathrm{g\ m^{-3}}\ \text{or}\ \mathrm{LWC} > 0.136\ \mathrm{g\ m^{-3}}, \quad (1)$$

and

$$\mathrm{IWC} > 0.0042\ \mathrm{g\ m^{-3}}\ \text{or}\ \mathrm{LWC} > 0.038\ \mathrm{g\ m^{-3}}. \quad (2)$$

These thresholds correspond roughly to radar reflectivities of −30 and −40 dBZ, respectively, which are the values determining the observed CF. While the previously discussed sensitivity to the time interval affected the distribution but not the mean CF, lowering the radar reflectivity threshold from −30 to −40 dBZ results in detecting more clouds and therefore increases the mean CF by shifting the distributions toward larger

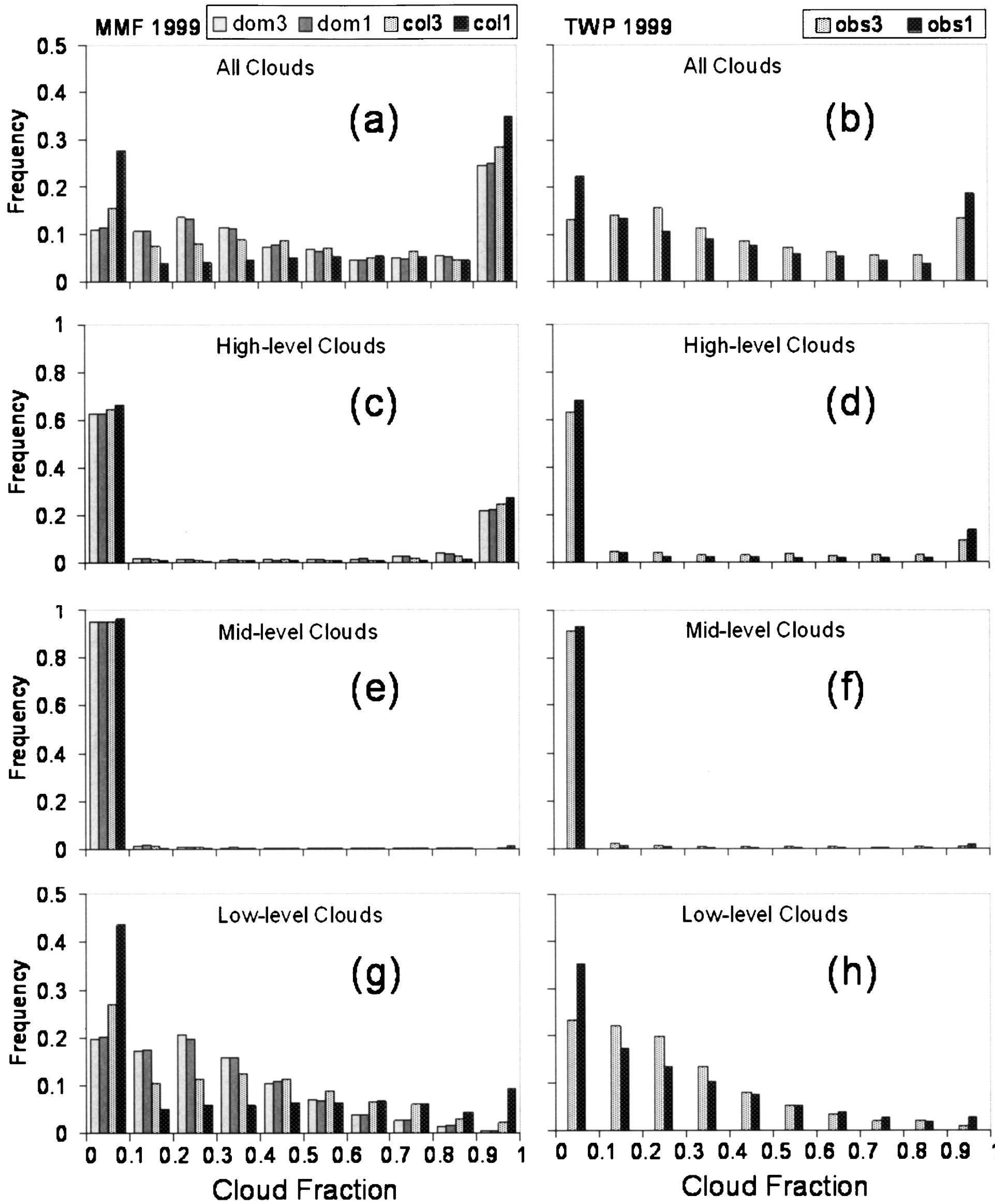

FIG. 3. Histograms of (left) simulated and (right) observed cloud fractions for TWP in 1999 for 1- and 3-h intervals. As in Fig. 2, four rows from top to bottom show cloud fractions for all, high-, middle-, and low-level clouds, respectively. Temporal cloud fractions (both simulated and observed) are for a radar reflectivity threshold of −40 d*BZ*.

CF values (Fig. 4). Once again, the effect is primarily due to changes in detection of low-level clouds (not shown). The simulated sensitivity (Fig. 4a) is small compared to the observed one (Fig. 4b). The sensitivity of observations appears to be due to small clouds for which a relatively large fraction of their volumes has water content below the first threshold. The modeled sensitivity is diminished because clouds under 4 km are not resolved.

b. Precipitation

Table 2 summarizes the total annual precipitation amount from observations at the TWP site and MMF and CAM runs with both climatological and 1999 SSTs.

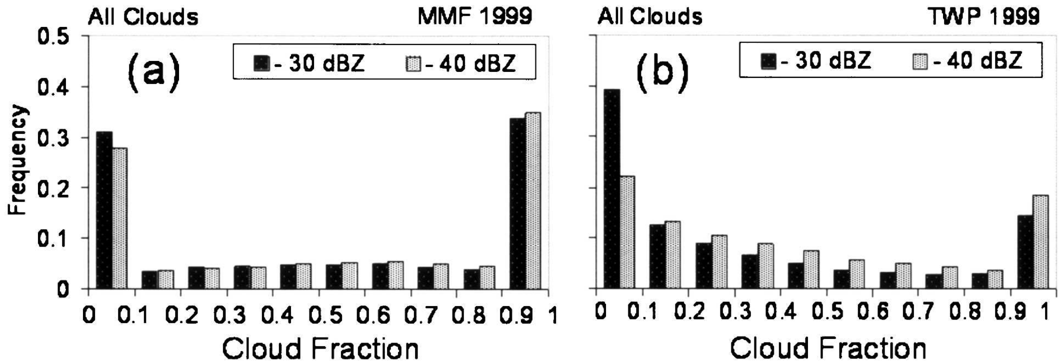

Fig. 4. Histograms of temporal (over 1 h) (a) simulated (CRM column) and (b) observed (cloud radar) cloud fractions for the TWP in 1999 using two different radar reflectivity thresholds for defining cloud boundary.

The year 1999 was a La Niña year of medium strength. At Nauru, this phase of ENSO is usually associated with a reduction in precipitation (negative precipitation anomaly) during the November through March period and near-normal precipitation for the May through September period. Therefore, the observed annual precipitation in 1999 is only 358 mm or 55% of the 1998–2002 (climatological) average of 637 mm. The MMF simulations are in close agreement with observations both in annual precipitation amount and in predicting a relatively dry year in 1999. In both CAM simulations, the precipitation is overestimated by a factor of 2 or more although the relative reduction for the year 1999 compared to climatology is still captured (Table 2).

The TWP region is characterized by a very strong north–south gradient in the total annual precipitation (Table 2 and Figs. 5a,b), and values at any point are very sensitive to the location of the intertropical convergence zone (ITCZ). In fact, for the year of 1999 the mean of the four MMF grid points surrounding the TWP/Nauru location (286 mm) is closer to the observations (358 mm) than is precipitation at the grid point closest to Nauru (216 mm). However, the grid point differences are much smaller than the difference between CAM and MMF, so that the MMF superiority in this case looks robust.

Table 2. Total annual precipitation (mm) at the ARM TWP Nauru site. For the CAM and MMF, values for the nearest four model grid points are shown and their means are given in bold.

	CAM			MMF			TWP (Nauru)
	1429		1315	700		613	
Climate		**1309**			**677**		**637**
	1235		1258	781		615	
	906		764	393		308	
Year 1999		**826**			**286**		**358**
	884		750	216		227	

For the northern Oklahoma region, La Niña years generally correlate with drier than normal conditions. In 1999, however, a notable positive precipitation anomaly was observed at the SGP site (Table 3) with total precipitation for the year exceeding the multiyear average by 15%. This is not a local effect: the increased precipitation for this region in 1999 is also confirmed by Oklahoma Mesonet stations surrounding the SGP Central Facility site, where observations used in this study were taken. The model fails to produce enough precipitation in the area, regardless of the cloud treatment (Table 3 and Figs. 5c,d). Both MMF and CAM runs predict only about half of the observed annual precipitation amount in all simulations. Moreover, in contrast to observations, both models show 1999 as the drier year compared to climatology by 20%–25%. Although the local gradients among the nearest four grid points are nearly the same as they were for the TWP region (Table 3), the closest regions with the simulated annual precipitation amount comparable to that observed at the SGP are found as far as the Gulf of Mexico (Fig. 5d) and northern Florida (Fig. 5c).

Recalling that the main benefit of the MMF is in its explicit treatment of convective clouds, it is instructive to examine if the lack of improvement in the SGP precipitation prediction can be attributed to a different type of precipitation in that region compared to the TWP. Figure 6 shows the relative contribution of the convection to the total annual precipitation around the two sites. These data from the 1999 CAM run, in which convective (subgrid) and stratiform (resolved) components of precipitation are predicted separately, are expected to be qualitatively indicative of the frequency of occurrence of the convection-prone environment. Predictably, around the TWP the fraction of convection-driven precipitation is near 100%. In the SGP region, however, the convection is still the primary source of

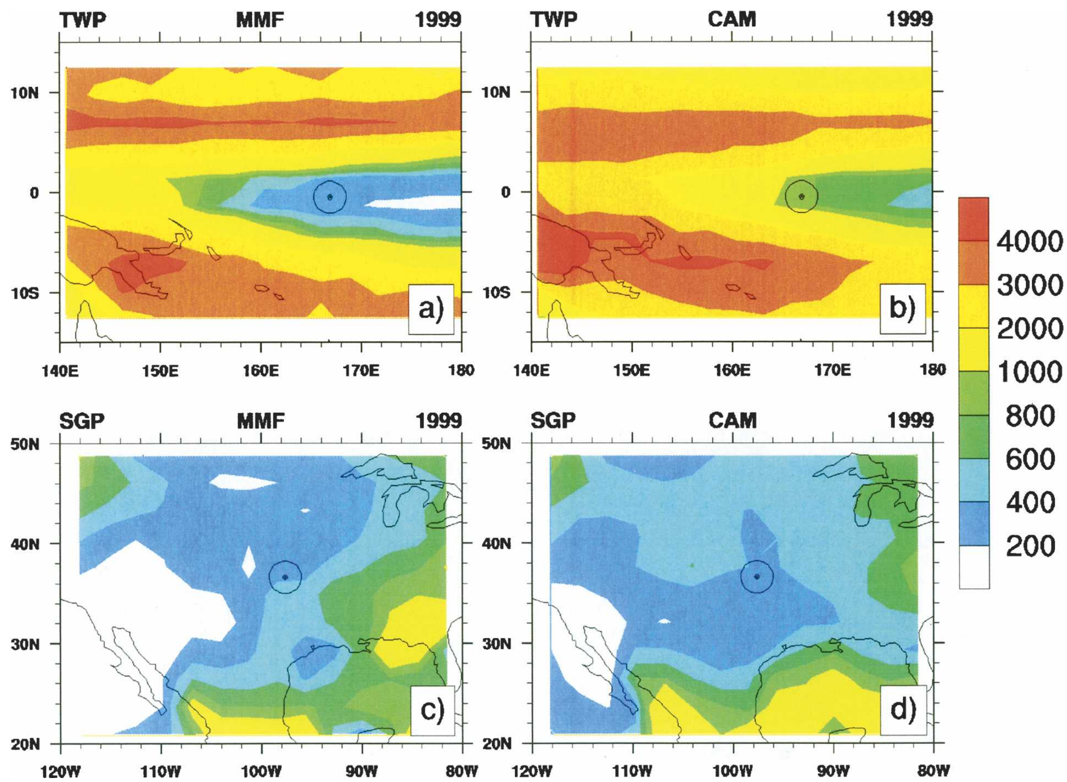

FIG. 5. Total annual precipitation amount (mm) as predicted by the (a), (c) MMF and (b), (d) CAM for the (a), (b) TWP and (c), (d) SGP regions for 1999. Locations of the two ARM sites are encircled. The observed precipitation amounts are 358 and 1031 mm for the TWP and SGP sites, respectively.

precipitation being responsible for over 70% of the annual amount.

Accurate prediction of the total annual amount is a necessary but not a sufficient condition for a successful simulation of local climate. Predicting extreme events or more generally the frequency distribution of the precipitation rate is as important as predicting the mean (Groisman et al. 1999). Cumulative probability distributions of 3-h average precipitation rates from meteorological observations and from CAM and MMF are shown in Fig. 7. Direct comparison of precipitation rate distributions is difficult because they are dominated by different spatial averaging (Fig. 1b). Indeed, observations represent point measurements with no spatial averaging, MMF/CRM column statistics represent an area of 4×4 km^2, MMF/CRM domain statistics correspond to a strip of 64 columns (area of 256×4 km^2), and the standard CAM parameterization provides mean values for the large-scale grid (roughly 300×300 km^2). In general, we should expect larger extremes and more variability in local (point) measurements than in the CRM strip, which in turn should be more variable than the parameterization representing CAM grid cell properties. In the Tropics, with precipitation dominated by convection, it is likely that the precipitation rate at points 100 km apart are statistically independent (or at best weakly correlated), although the precipitation distributions at points 4 km apart (as in the adjacent CRM columns) could be highly correlated. Thus, while the probability distributions in Fig. 7 look very different they are not necessarily inconsistent with each other. To illustrate the point, we construct a hypothetical domain containing 2, 4, 8, and 16 statistically independent simulated records such that the distribution properties for each record are the same as those taken from measurements (Fig. 8). [Numerically, we construct new time series by averaging the required number of randomly rearranged original (observed) time series,

TABLE 3. Same as in Table 2, but at the ARM SGP site.

	CAM			MMF			SGP
	480		520	468		603	
Climate		**431**			**478**		**878**
	328		397	395		444	
	392		390	294		302	
Year 1999		**319**			**407**		**1031**
	217		278	510		522	

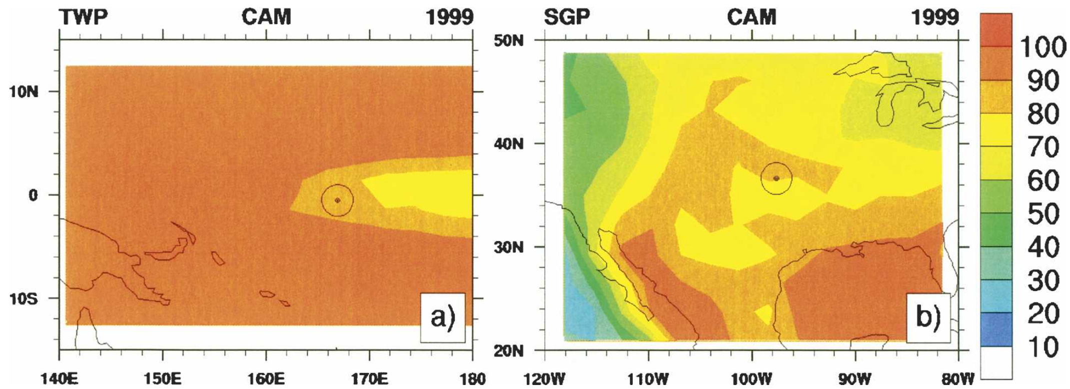

FIG. 6. Fraction (%) of the total annual precipitation amount originated from the CAM convective parameterization for the (a) TWP and (b) SGP regions. Circled points indicate locations of the ARM observation sites.

which mathematically represent convolving the original probability distribution function 2, 4, 8, and 16 times, respectively.] This simple example nicely illustrates the consequences of different sampling areas, even if only qualitatively. In reality, various precipitation events will have different correlation scales in time and space, and the result may not be well approximated by a simple averaging of independent samples.

Figure 7b shows that the SGP distributions are much flatter than those from TWP. (Note the stretched *y* axis in Fig. 7b.) The SGP also has much lower frequency of any precipitation event than TWP. In the CAM simulations, there is no precipitation at the SGP site location about 70% of the time, when for the TWP site location it is only 5% of the time.

The sensitivity of the cumulative probability distribution of the precipitation rate to the time-averaging interval is illustrated in Fig. 9. The difference between the distributions of 1- and 3-h mean rates decreases as the area of spatial averaging increases. Thus, the sensitivity of CAM is less than that of the MMF/CRM domain, which in turn is less than the sensitivity of the MMF/CRM column.

c. Downwelling solar radiation flux at the surface

A quantitative comparison performed here is for the 1999 runs when the downwelling shortwave flux was added to the saved statistics. From the climatological runs, only the net solar radiative flux at the surface was saved, which is not directly comparable with the measured quantity of the downwelling flux. Table 4 summarizes parameters of observed and modeled distributions for the TWP Nauru site. The distributions are composed of 3-h averages around local noon and the

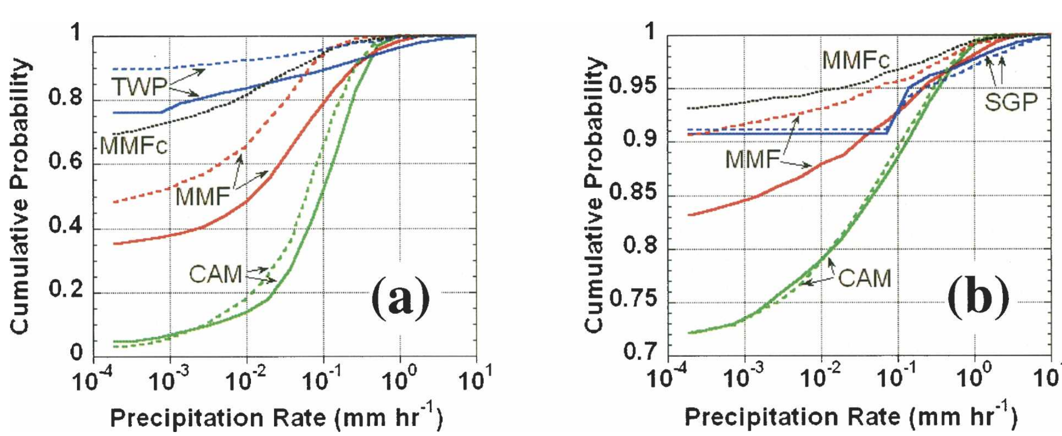

FIG. 7. Climatological (solid lines) and 1999 (dashed lines) cumulative probability distributions of 3-h mean precipitation rates from observations and models for the (a) TWP and (b) SGP sites. Note the stretched *y* axis in (b). The CRM column statistics (MMFc, black lines) are available for the 1999 run only. A projected intersection of each curve with the *y* axis indicates a fraction of time without precipitation.

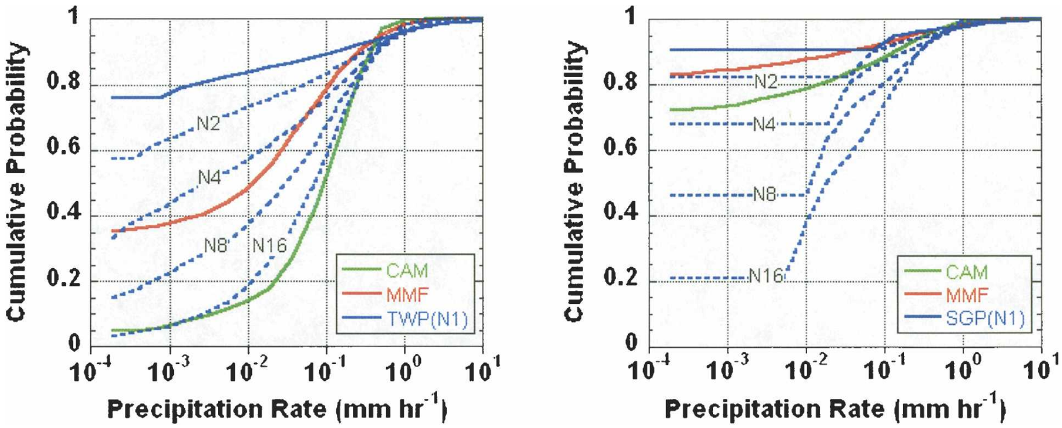

FIG. 8. Climatological cumulative probability distributions of 3-h mean precipitation rates from observations and models for the (left) TWP and (right) SGP sites (solid lines). Dashed lines represent distributions for four hypothetical domains (marked N2, N4, N8, and N16) containing, respectively, 2, 4, 8, and 16 statistically independent records for which distributions are identical to the observed one (marked N1).

fluxes are normalized by the cosine of the zenith solar angle to account for the seasonal change in insolation. Table 4 shows that the widths of both original and normalized distributions from the CAM as characterized by the standard deviations are significantly larger than those from either the MMF or observations, primarily because of the increase in frequency of low values (Fig. 10a). The CAM distribution also has the smallest (by about 100 W m^{-2}) mean, which is consistent with the highest cloud fraction in the CAM. The corresponding cloud fraction statistics for 3-h periods around local noon (not shown) are a little noisier (due to shorter time series) but generally differ only slightly from the statistics for all of the periods (Fig. 2 and Table 1). The distribution from a single CRM column is the closest to the observed distribution, while the distribution from a CRM domain is narrower and closer to the mean, as expected for the domain averages. The observed mean solar flux is larger than the one predicted by the MMF, despite the observed CF being almost twice as large as the MMF CF. This implies that model clouds are generally optically thicker than the real ones.

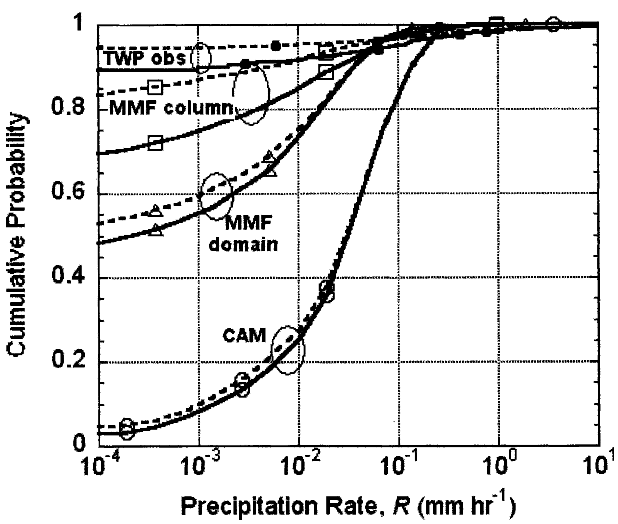

FIG. 9. Cumulative probability distributions of 1- (dashed lines) and 3- (solid lines) h mean precipitation rates from observations and models for 1999 runs for the TWP.

At the SGP site, both models overestimate the amount of solar radiation reaching the surface (Table 5), consistent with the models' underestimate of cloud amount (Table 1 and Fig. 2). The most notable discrepancy is in large fluxes (greater than 850 W m^{-2}), which are frequently predicted by the two models but rarely observed. The widths of all three distributions for the absolute downward flux are nearly identical because at this latitude they are dominated primarily by the seasonal variation of insolation rather than cloud parameters. The clear-sky flux at the SGP varies between the

TABLE 4. Parameters of simulated and observed frequency distributions of the downward solar radiation flux at the surface at the ARM TWP Nauru site for 1999. The distributions are composed of 3-h averages around local noon and normalized by the cosine of the zenith angle. For the MMF both CRM domain average and CRM column values are given.

	CAM	MMF domain/column	TWP Nauru
Mean (W m^{-2})	758	860/860	894
Minimum (W m^{-2})	128	527/453	125
Maximum (W m^{-2})	1029	1043/1039	1086
Median (W m^{-2})	834	868/890	925
Standard deviation (W m^{-2})	226	94/129	141

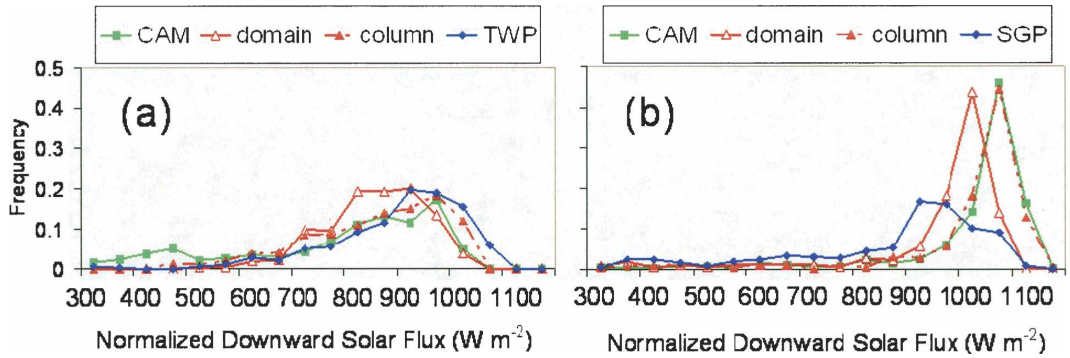

FIG. 10. Frequency distributions of observed and simulated downward solar radiation fluxes at the surface for (a) TWP and (b) SGP sites for 1999. The fluxes are normalized by the cosine of the zenith angle to account for seasonal changes in insolation. The distributions are constructed using 3-h averages around local noon grouped in bins, which are 50 $\mathrm{Wm^{-2}}$ wide. The MMF statistics are shown for both CRM domain and CRM column.

minimum of 400 W $\mathrm{m^{-2}}$ in the winter and the maximum near 950 W $\mathrm{m^{-2}}$ in the summer, while at the TWP the range is much narrower (between the minimum of 850 W $\mathrm{m^{-2}}$ and the maximum of 1000 W $\mathrm{m^{-2}}$). When the seasonal insolation variations are removed, the models' shortcomings are exposed much more clearly and the widths of the distributions of the normalized flux from both models are vastly different from the observed (Table 5; Fig. 10b).

4. Discussion and conclusions

We have tested the MMF approach to global modeling by comparing the simulated distributions of cloud fraction, precipitation rate, and downwelling shortwave radiation at the surface to both ARM observations and CAM simulations. The cloud fraction is intended to be the most basic categorization of the cloud amount, while the other two quantities are important components of the water and energy cycle.

We note that the cloud fraction in traditional GCM parameterizations is an intermediate parameter generated by a cloud scheme for injection into radiation calculations and other parameterizations, which in turn compute the effects of the cloud fields. For the model to perform well, the cloud fraction does not have to be realistic as long as its effect on the energy and water cycle (e.g., latent heat release, precipitation rate, and cloud radiative forcing) is correct. The proper energy and water balances are often achieved by tuning model's parameters. We want the cloud fraction to be realistic, however, not only because cloud amount is one of the more easily observable parameters but also because clouds represent a physical link between many thermodynamic, radiative, chemical, and other processes. Thus, it is imperative to test the model ability to represent real cloud fields in addition to the effects of the clouds.

TABLE 5. Same as in Table 4, but for the ARM SGP site.

	CAM	MMF domain/column	SGP
Mean (W $\mathrm{m^{-2}}$)	985	931/978	750
Minimum (W $\mathrm{m^{-2}}$)	34	46/48	17
Maximum (W $\mathrm{m^{-2}}$)	1151	1097/1151	1142
Median (W $\mathrm{m^{-2}}$)	1064	1013/1059	912
Standard deviation (W $\mathrm{m^{-2}}$)	220	213/222	330

Our analysis indicates that the cloud fraction being the simplest one-parameter measure of cloud amount is not a good predictor for either the total annual precipitation or the surface shortwave radiation flux. For the TWP site, for example, MMF predicts the observed precipitation amount well (Table 2) but only half the observed cloud fraction (Table 1). Furthermore, both MMF and observations give similar values for climatological and 1999 CFs, but the annual precipitation amount changes by a factor of 2. The highest cloud fraction for the TWP predicted by CAM (Table 1) agrees with the lowest downwelling radiation flux in that run (Table 4). However, despite the observed TWP CF being nearly twice the MMF CF (Table 1), the observed solar flux is larger than the MMF prediction (Table 4). Some of these CF differences can be attributed to differences in the time and space scales.

The analysis shows that the frequency distributions of the cloud fraction, precipitation rate, and solar flux depend strongly on the sampling/averaging in both time and space. This time- and space-scale gap, which is responsible, at least partly, for differences between the observations and simulations, is greatly lessened in comparisons of the CRM column fields to the observations. Comparison of cloud fraction is further complicated by some arbitrariness in defining cloud boundaries. This points to the potential utility of simulating radar reflectivity fields from the CRM fields and deriving cloud fractions from these simulated radar data, rather than using simple water mixing ratio thresholds as have been applied here.

By explicitly resolving clouds, the MMF (in this 2D configuration) notably improves realism of prediction of precipitation and cloud amount, as well as the net solar radiation flux in the TWP region but not in the SGP region. Let us consider some of the possible reasons behind these regional differences. The TWP site is in an equatorial ocean while the SGP is in the middle of an extratropical continent. Many of the cloud formation mechanisms in these locations are different, and the MMF performance is determined by whether its configuration, including the structure of the CRM and the way it is coupled to the parent global model, is able to simulate the dominant processes. Convection is the primary source of precipitation in both the TWP and SGP regions (Fig. 6). However, in the MMF used here, the CRM–surface interaction is very limited since an exchange between a CRM and the surface is handled by a large-scale component of the MMF (i.e., by the CAM). Because all of the model runs are driven by the prescribed SSTs, the forcing over the ocean (TWP) is much more direct and faithful to reality than it is over the continents (SGP) where the lack of direct CRM–surface coupling plays a larger role.

Large-scale dynamical forcing is also known to be much more complicated and important in midlatitudes compared to the Tropics. Given the fact that the CRMs have periodic lateral boundary conditions and cannot maintain large-scale horizontal gradients (and hence baroclinicity) and that the exchange among CRMs is handled through tendencies of large-scale variables, the MMF ability to propagate synoptic-scale disturbances, such as fronts or mesoscale convective systems, may not be much different from that of the CAM.

Finally, with much colder winter temperatures at the SGP, part of the precipitation in the region is in a form of snow. Frozen precipitation introduces more uncertainty in both simulations and observations, but we do not believe this effect is nearly strong enough to explain the discrepancies.

The comparison of model simulations and ARM observations presented in this study shows both MMF promise and the need for additional work. In the future, we plan to improve evaluation of the model's performance by analyzing other variables at multiple sites as well as regionally and globally using satellite observations. This will enable us to better identify and correct deficiencies. Because of its higher resolution and more explicit treatment of cloud and radiation processes, the MMF is better suited for simulations of observing systems than traditional GCMs, thus making MMF evaluation more robust and the path to its improvement more straightforward.

Acknowledgments. The research described in this paper was conducted under the Laboratory Directed Research and Development (LDRD) Program at the Pacific Northwest National Laboratory, a multiprogram national laboratory operated by Battelle for the U.S. Department of Energy (DOE) under Contract DE-AC06-76RL01830. The data were provided by the DOE Atmospheric Radiation Measurement Program Climate Research Facility. Stimulating discussions with Chuck Long about the data analysis are acknowledged. Two anonymous reviewers are thanked for the comments that helped to improve the clarity of the paper.

REFERENCES

Ackerman, T. P., and G. M. Stokes, 2003: The Atmospheric Radiation Measurement Program. *Phys. Today,* **56,** 38–44.

Blackmon, M., and Coauthors, 2001: The Community Climate System Model. *Bull. Amer. Meteor. Soc.,* **82,** 2357–2376.

Briegleb, B. P., 1992: Delta-Eddington approximation for solar radiation in the NCAR Community Climate Model. *J. Geophys. Res.,* **97,** 7603–7612.

Clothiaux, E. E., T. P. Ackerman, G. G. Mace, K. P. Moran, R. T. Marchand, M. A. Miller, and B. E. Martner, 2000: Objective determination of cloud heights and radar reflectivities using a combination of active remote sensors at the ARM CART sites. *J. Appl. Meteor.,* **39,** 645–665.

Collins, W. D., 2001: Parameterization of generalized cloud overlap for radiative calculations in general circulation models. *J. Atmos. Sci.,* **58,** 3224–3242.

Grabowski, W. W., 2001: Coupling cloud processes with the large-scale dynamics using the cloud-resolving convection parameterization (CRCP). *J. Atmos. Sci.,* **58,** 978–997.

Groisman, P. Y., and Coauthors, 1999: Changes in the probability of heavy precipitation: Important indicators of climatic change. *Climatic Change,* **42,** 243–283.

Khairoutdinov, M. F., and D. A. Randall, 2001: A cloud resolving model as a cloud parameterization in the NCAR Community Climate Model: Preliminary results. *Geophys. Res. Lett.,* **28,** 3617–3620.

——, ——, and C. DeMott, 2005: Simulations of the atmospheric general circulation using a cloud-resolving model as a superparameterization of physical processes. *J. Atmos. Sci.,* **62,** 2136–2154.

Lin, W. Y., and M. H. Zhang, 2004: Evaluation of clouds and their radiative effects simulated by the NCAR Community Atmospheric Model against satellite observations. *J. Climate,* **17,** 3302–3318.

McFarlane, S. A., C. N. Long, and D. M. Flynn, 2005: Impact of island-induced clouds on surface measurements: Analysis of the ARM Nauru Island Effect Study data. *J. Appl. Meteor.,* **44,** 1045–1065.

Ramanathan, V., and P. Downey, 1986: A nonisothermal emissivity and absorptivity formulation for water vapor. *J. Geophys. Res.,* **91,** 8649–8666.

Randall, D., M. Khairoutdinov, A. Arakawa, and W. Grabowski, 2003: Breaking the cloud parameterizations deadlock. *Bull. Amer. Meteor. Soc.,* **84,** 1547–1564.

Rasch, P. J., and J. E. Kristjánsson, 1998: A comparison of the CCM3 model climate using diagnosed and predicted condensate parameterizations. *J. Climate,* **11,** 1587–1614.

Slingo, J. M., 1987: The development and verification of a cloud prediction scheme for the ECMWF model. *Quart. J. Roy. Meteor. Soc.,* **113,** 899–927.

Williamson, D. L., and J. G. Olson, 1994: Climate simulations with a semi-Lagrangian version of the NCAR community climate model. *Mon. Wea. Rev.,* **122,** 1594–1610.