Dreary state of precipitation in global models

Graeme L. Stephens,[1] Tristan L'Ecuyer,[1] Richard Forbes,[2] Andrew Gettlemen,[3] Jean-Christophe Golaz,[4] Alejandro Bodas-Salcedo,[5] Kentaroh Suzuki,[1] Philip Gabriel,[1] and John Haynes[6]

Received 25 May 2010; revised 7 September 2010; accepted 21 September 2010; published 21 December 2010.

[1] New, definitive measures of precipitation frequency provided by CloudSat are used to assess the realism of global model precipitation. The character of liquid precipitation (defined as a combination of accumulation, frequency, and intensity) over the global oceans is significantly different from the character of liquid precipitation produced by global weather and climate models. Five different models are used in this comparison representing state-of-the-art weather prediction models, state-of-the-art climate models, and the emerging high-resolution global cloud "resolving" models. The differences between observed and modeled precipitation are larger than can be explained by observational retrieval errors or by the inherent sampling differences between observations and models. We show that the time integrated accumulations of precipitation produced by models closely match observations when globally composited. However, these models produce precipitation approximately twice as often as that observed and make rainfall far too lightly. This finding reinforces similar findings from other studies based on surface accumulated rainfall measurements. The implications of this dreary state of model depiction of the real world are discussed.

Citation: Stephens, G. L., T. L'Ecuyer, R. Forbes, A. Gettlemen, J.-C. Golaz, A. Bodas-Salcedo, K. Suzuki, P. Gabriel, and J. Haynes (2010), Dreary state of precipitation in global models, *J. Geophys. Res.*, *115*, D24211, doi:10.1029/2010JD014532.

1. Introduction

[2] The increase in global mean accumulated precipitation that is expected to occur with global warming is primarily controlled by requirements for global energy balance [*Allen and Ingram*, 2002; *Stephens*, 2005, among others]. These energy balance controls on the planet's hydrological cycle apply only to the global mean accumulation of precipitation. It is not obvious how the character of precipitation, defined not only by accumulation but also in terms of frequency and intensity among other measures, is expected to change [e.g., *Stephens and Hu*, 2010] nor is it obvious how these changes manifest themselves on the regional scale. It is well documented that changes to both the frequency and intensity of precipitation do occur in climate warming experiments performed using present generation climate models [e.g., *Dai and Trenberth*, 2004; *Sun et al.*, 2006; *Pall et al.*, 2007; *Neelin et al.*, 2006], which raises serious questions about the common hydrological practice of assuming stationary statistics of precipitation [*Milly et al.*, 2008].

[3] The purpose of this paper is to provide a new evaluation of the character of oceanic precipitation from three different types of global prediction models. This evaluation defines the character of precipitation as a combination of three quantities: (1) the accumulation of precipitation $a_{\Delta T}$ integrated over some time period ΔT,

$$a_{\Delta T} = \sum_{\Delta T} f_i \times r_i,$$

and (2) the frequency of occurrence f_i over this same period ΔT of (3) precipitation of (instantaneous) intensity r_i.

[4] Although comparison of modeled $a_{\Delta T}$ with observations is commonplace, use of such comparisons over seasonal and longer time scales offers little in way of testing the realism of hydrological processes in Earth system models since this time-integrated precipitation is primarily controlled by the energy balance of the planet [e.g., *Stephens and Ellis*, 2008]. Joint assessment of model frequency f_i and intensity r_i of precipitation provide more insight on the validity of modeled moist processes, but such assessments have proven to be difficult to perform since the necessary global observations are generally lacking. Land surface observations of the daily-accumulated rainfall intensities of rates >1 mm/d were compiled from the Global Historical Climatology Network by *Sun et al.* [2006] and compared to analogous model accumulated precipitation. As in other studies [e.g., *Dai and Trenberth*, 2004], the Sun et al. com-

[1]Department of Atmospheric Sciences, Colorado State University, Fort Collins, Colorado, USA.
[2]European Centre for Medium-Range Weather Forecasts, Reading, UK.
[3]National Center for Atmospheric Research, Boulder, Colorado, USA.
[4]Geophysical Fluid Dynamics Laboratory, Princeton University, Princeton, New Jersey, USA.
[5]Met Office Hadley Centre, Exeter, UK.
[6]Monash University, Melbourne, Victoria, Australia.

0148-0227/10/2010JD014532

parison revealed a general overestimate in the frequency of modeled precipitation and an associated underestimate of intensity and this finding is also reproduced in this study. Using comparisons to oceanic precipitation data derived from satellite microwave observations, *Wilcox and Donner* [2007] reached a similar conclusion. Neither study include the contributions of lighter precipitation (<1 mm/d) that approaches 10% of the tropical accumulated rain [*Dai et al.*, 2007; *Berg et al.*, 2010] and represents an even a larger proportion of midlatitude precipitation where estimates of both frequency and intensity of precipitation from satellite observations has proven problematic [*Petty*, 1997; *Ellis et al.*, 2009; *Berg et al.*, 2010]. The assessment reported in this study, by contrast, uses new precipitation information available from CloudSat observations [e.g., *Haynes et al.*, 2009]. The frequency of precipitation detected by CloudSat is intrinsically more accurate than any previous satellite sources of data owing to the acute sensitivity of the CloudSat radar to the presence of drizzle and lighter precipitation missed by other satellite observations [e.g., *Berg et al.*, 2010; *Ellis et al.*, 2009].

[5] The characteristics of these new precipitation data are introduced in section 2 along with a discussion of likely observational error. A description of model data used in comparisons to the observations is provided in section 3. The main results of the study, presented in section 4, focus on instantaneous precipitation rather than daily or seasonal mean accumulations because (1) model errors revealed in this quantity provide a clearer indication of flaws in model physics and possible hints where improvements might be made to fix them (e.g., G. Stephens et al., The character of low clouds over the Earth's oceans, submitted to *Journal of the Atmospheric Sciences*, 2010) and (2) instantaneous precipitation directly couples to other physical processes of the Earth system. For example, precipitation influences the water budget of clouds and related radiative effects [*Stephens et al.*, 2008], the washout of aerosol [e.g., *Xian et al.*, 2009], and thus aerosol effects on the radiation balance, the latent heating of the atmosphere and its effects on storm genesis and intensity (e.g., P. J. Webster and C. D. Hoyas, Tropical cyclone number and intensity in an evolving warm pool, submitted to *Geophysical Research Letters*, 2010). The intensity and frequency of precipitation also greatly influences surface hydrological processes such as water filtration into soils, runoff, river discharge, and soil drying [*Trenberth and Dai*, 2007; *Dai et al.*, 2009]. Section 6 summarizes the results of the study.

2. Observations

[6] Although the main source of observational data employed in this study is the CloudSat spaceborne cloud-profiling radar (CPR) [*Stephens et al.*, 2008], precipitation data taken from other sources are also used for comparison. These other sources of data include (1) the Version 2 AMSR-E level-2B Ocean precipitation product [*Kummerow et al.*, 2001], (2) the Version 6 13.8 GHz Tropical Rainfall Measurement Mission (TRMM) Precipitation Radar (PR) product 2A25 [*Iguchi et al.*, 2000] and the Version 6 TRMM Microwave Imager (TMI) product 2A12 [*Kummerow et al.*, 1998, 2001], and (3) Global Precipitation Climatology Project (GPCP) Version 2.1 precipitation data that combines satellite IR data from Geostationary imagers, sounding data from the TIROS Operational Vertical Sounder (TOVS) and the Atmospheric Infrared sounder (AIRS), microwave imager data from the Special Sensor Microwave Imagers (SSMIs), and surface rain gage data [*Adler et al.*, 2003].

[7] The CloudSat data used in this study are from the 2C-COLUMN-PRECIP product. Although the full description of this product is provided by *Haynes et al.* [2009], a brief description of the data and related errors is given here. The CloudSat precipitation information derives from estimates of the path-integrated attenuation (PIA) of the CPR. PIA is determined for each footprint of the radar (approximately 1.7 km in scale) [*Tanelli et al.*, 2008]. At this stage in development, only precipitation over the oceans is produced given the estimate of PIA can be more accurately determined over ocean surfaces than over land surface. PIA is used in two ways in the product. The existence of PIA above some background level of attenuation indicates the existence of rain. The intensity of rain that is reported below is directly proportional to the rain-water path in the column which is also directly proportional to the magnitude of the PIA [e.g., *Haynes et al.*, 2009].

[8] The lower detection threshold of oceanic PIA (about 2 dB) translates to a lower threshold precipitation of approximately 0.02–0.05 mm/h, and hereafter we use 0.05 mm/h as a somewhat conservative estimate of the minimum observational threshold of an individual CPR footprint. The intensity of rain is proportional to the PIA but total attenuation of the radar and multiple scattering limit estimation of the heaviest rain to about 3–5 mm/h [e.g., *L'Ecuyer and Stephens*, 2002; *Haynes et al.*, 2009]. Although total attenuation occurs in only about 3% of all precipitating profiles and then primarily at low latitudes, this limitation leads to a bias in the PIA estimated tropical accumulation of about 10%–15% when compared to TRMM observations matched specifically to the CPR [*Berg et al.*, 2010]. These direct comparisons to the TRMM PR, provided in the CloudSat 2D-CLOUDSAT-TRMM product (http://cloudsat.cira.colostate.edu), are used here to correct the underestimated intensities for this relatively small number of cases of total attenuation removing a source of bias in the CPR seasonal mean precipitation. Analysis of the most intense rainfall regime in the *Berg et al.* [2010] study suggests that the overall low bias in CloudSat rain accumulation can be removed by simply increasing the PIA-estimated intensities of completely attenuated CloudSat pixels by a factor of 2.5. This correction has been applied to only that (small) fraction of completely attenuated profiles that occur in the TRMM region.

[9] Determination of the occurrence of precipitation combines the PIA with the radar reflectivity measured adjacent to the ocean surface to construct the equivalent unattenuated radar reflectivity at those levels. The larger the value of this unattenuated reflectivity in the near-surface range bins, the larger are the reflecting hydrometeors and the more likely it is that these particles fall producing precipitation at the surface. Threshold values of this near-surface reflectivity are used to define the likelihood of precipitation as summarized by *Haynes et al.* [2009]. In this study, we combine the thresholds that define the rain certain category (unattenuated near-surface reflectivity of 0 dBZ or higher)

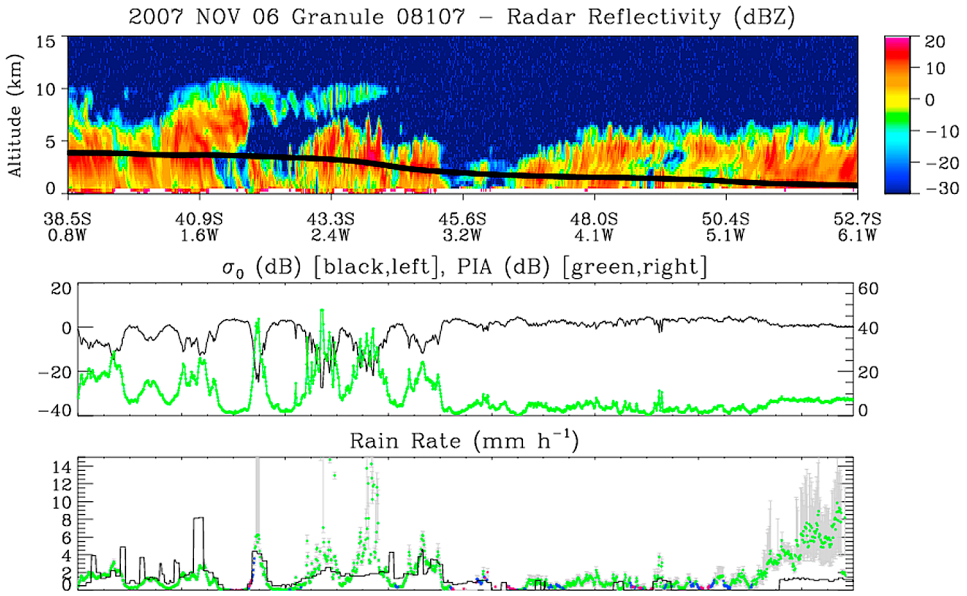

Figure 1. An approximately 400 km long portion of the CloudSat orbit showing the cloud and precipitation features of a frontal system observed on 8 November 2007. (top) The radar reflectivity observed by the CPR with the freezing level (0C isotherm) superimposed for reference. (middle) The measured surface reflectivity (black) and the inferred PIA (green). (bottom) The CPR-based intensity (green) with AMSR-E-based precipitation in black for reference.

and the rain probable category (unattenuated near-surface reflectivity is between −7.5 and 0 dBZ) of *Haynes et al.* [2009] to define the occurrence of precipitation. This largely excludes the likelihood of light drizzle (of rain rate less than about 0.02 mm/d) from the analysis.

[10] Assigning uncertainty to the detection of precipitation is not simple. *Ellis et al.* [2009] offer a limited assessment using surface observations. Agreement between the number of raining days observed at 9 out of 10 individual surface stations examined with CloudSat observations confined to a 2.5° × 2.5° box centered on the given stations was within 23% despite the obvious inherent sampling differences between such data. Another indication of uncertainty is the approximate 20% difference between rain certain and the combination of "rain certain" plus "rain probable" also reported by *Ellis et al.* [2009]. Taken together, it is reasonable to suggest that the uncertainty in detection is approximately 20%–25%, which is significantly less than the differences between models and observations shown below.

[11] Errors in precipitation intensity are more directly quantifiable (see *Haynes et al.* [2009] for details). It is reasonable to assume many retrieval errors are random with small contributions to total error when data are accumulated over time and space as done in this study. However, there are two potentially important sources of bias error that need to be accounted for and are well exemplified in the portion of a CloudSat orbit highlighted in Figure 1. The first is a bias error that occurs in the heaviest precipitation events that attenuate the CloudSat radar completely. As already mentioned, these circumstances are flagged in the algorithm, and although a mean multiplicative factor of 2.5 has been applied to these cases, it should be noted that the analysis of *Berg et al.* [2010] reveal that the scaling factor required to correct for the underestimation of tropical precipitation intensity by CloudSat PIA in cases of total attenuation actually varies between approximately 2 and 3 depending on water vapor and other environmental factors. We take this range as an upper and lower bound and use it to determine the error of tropical precipitation between 35°N-S for which the Berg et al. analysis is relevant and where most of the occurrence of total attenuation occurs.

[12] The second potential source of bias error arises from the uncertain assignment of the height of the precipitation water column. Precipitation intensity as derived from the column information in both the passive microwave observations and the CloudSat column PIA method is proportional to the precipitation water content defined as the water path divided by the depth of the column. The depth of the water column is typically assumed to be the thickness of the layer between the surface and the height of the freezing level neglecting the possible existence of supercooled water above this freezing level as often occurs in deeper convection and slantwise convection of higher-latitude frontal systems. When the freezing level touches the surface, as in the example of Figure 1, the estimated intensity becomes unrealistically too high and increasingly is sensitive to the specification of this column water height with errors exceeding 100% when freezing levels occur below 1.5 km [*Haynes et al.*, 2009]. The range in precipitation that results from a doubling and halving of the precipitation for cases that occur below the freezing level of 1.5 km is taken as indicative of error ranges for these cases.

3. Annual Oceanic Accumulations

[13] Figure 2a presents the oceanic rainfall accumulation (expressed as a time-mean precipitation rate) derived from

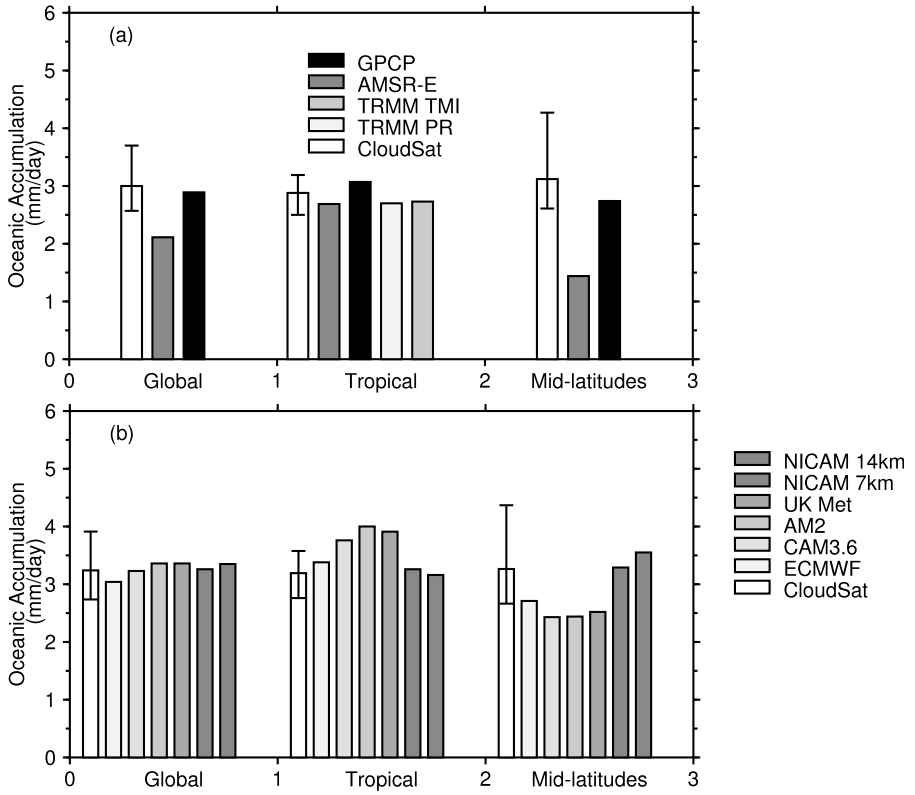

Figure 2. (a) The observed oceanic rainfall accumulation derived from different sources. (b) Comparison between CloudSat observed accumulation and model accumulation. The uncertainty ranges indicated on the CloudSat observations are discussed in the text.

33 months of CloudSat data taken from June 2006 to March 2009. The AMSR-E Goddard PROFiling algorithm (GPROF) accumulations also shown are for the same 33 month period, and the annual mean GPCP data are also a multiyear average. The precipitation data averaged between latitudes of 30°N-S are referred to as tropical, data averaged between 30°N-S to 60°N-S are similarly referred to as midlatitudes and data averaged between 60°N-S are referred to as global. Averages of 11 years of TRMM PR and TMI data are also shown for the tropics. The 60°N-S "global" oceanic mean accumulation of 3.0 +0.07/−0.04 mm/d determined from CloudSat compares well with the average GPCP value of 2.9 mm/d but not to the AMSR-E data that substantially underestimates precipitation in the midlatitudes. This lack of agreement between CloudSat and AMSR-E is expected given that the version of the AMSR-E data available for this study is known to miss significant amounts of midlatitude precipitation [e.g., *Ellis et al.*, 2009; *Petty*, 1997]. The agreement among the different observations for the tropics is also noteworthy with all estimates of accumulation falling within the CloudSat observational error range.

[14] The extent of the error bar on the CloudSat results of Figure 2a combines the two high and low ranges of each bias. This error range, however, omits a potentially important source of error in the CloudSat data composites. Neglected are representativeness errors due to the incomplete spatial and temporal sampling that is inherent to these asynoptically sampled data [e.g., *Salby and Callahan*, 2007]. One interpretation of the agreement shown in Figure 2a between the different data sources for the tropical oceans in particular is that the more limited space-time sampling of CloudSat has little influence on these accumulated annual statistics (as well as on seasonal statistics, data not shown), and thus, it is reasonable to assume the representativeness errors are small over oceans [see also *Berg et al.*, 2010]. Since the AMSR-E flies on the A-Train, these data have the same temporal but very different spatial sampling than CloudSat and the asynoptic nature of both intrinsically differ from the data sampling of TRMM. Since the TRMM TMI and the AMSR-E accumulations shown are based on the same GPROF algorithm [*Kummerow et al.*, 1998], the physical assumptions of the algorithms are thus the same. The small differences between these two different sources of precipitation data are most likely due to small sampling differences, which is consistent with the above interpretation that representativeness errors due to the asynoptic sampling of the A-Train sensors are small compared to the TRMM sample. The slightly larger accumulations observed by CloudSat compared to the TRMM and AMSR-E (about 7%) is due to the contribution of lighter rain detected by CloudSat but missed by the TRMM PR and by implication AMSR-E and TMI [e.g., *Berg et al.*, 2010]. The agreement between CloudSat and GPCP in midlatitudes also suggests the influence of the CloudSat sparse asynoptic sampling on

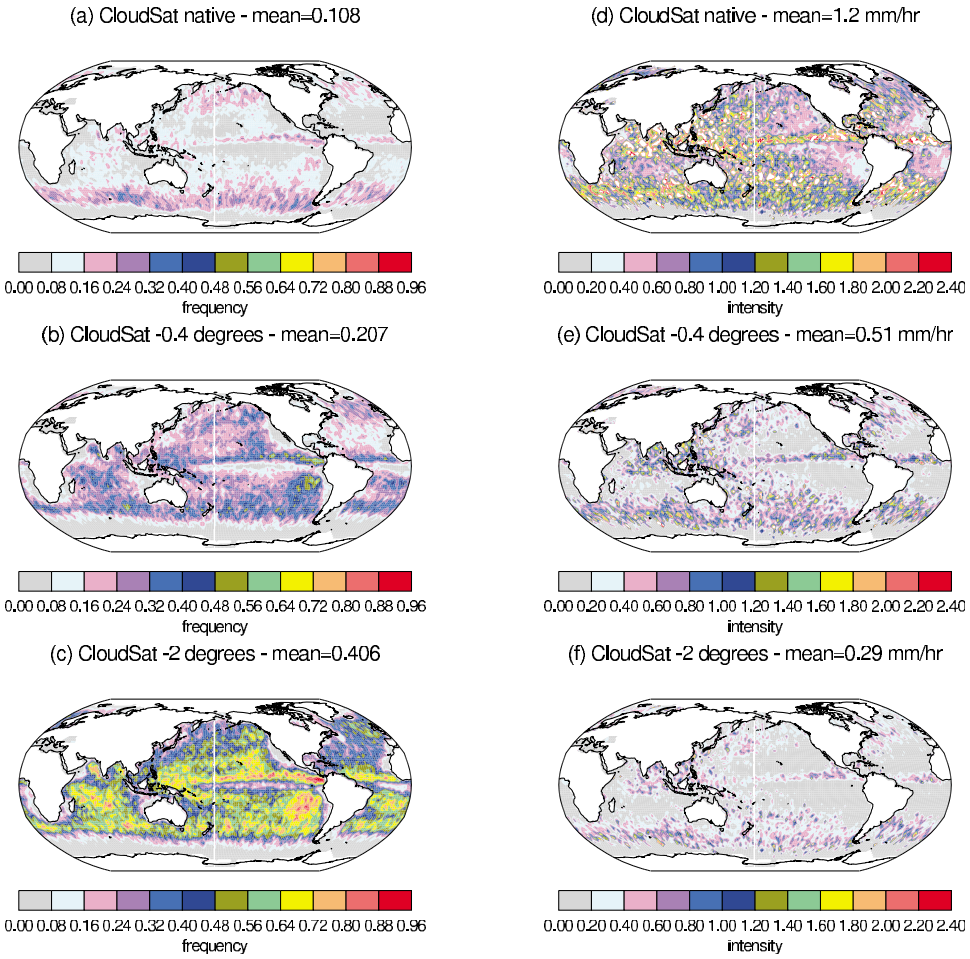

Figure 3. The CloudSat fractional occurrence of (left; liquid) precipitation and rain intensity inferred at the native resolution of the CloudSat measurements (1.75 km) and for CloudSat observations averaged up in scale to 0.4°S and 2°S. The oceanic mean values are also given above the images.

the space-time accumulated statistics presented in Figure 2a is also small.

4. Model Data

4.1. Data Sources

[15] Instantaneous (i.e., singe time step) model precipitation, sampled at single time steps every 3 h, was accumulated from the simulations of five different global models. While all model simulations apply to the same season, the data analyzed for each model correspond to different time periods. We applied different sampling approaches to one model. One approach sampled the data asynoptically along the projected satellite orbit whereas the second adopted the synoptic uniform space-time sampling across the globe. The results reported in this study for that model were unchanged regardless of the approach, thus suggesting that the results of this study are also generally robust to different sampling strategies.

[16] The models that are compared to the observations include as follows: (1) the version CY35R1 of the European Centre for Medium-Range Weather Forecasts (ECMWF) integrated forecast system (IFS) global model with a grid resolution of 40 km (T511) (The data are from a series of daily 12–36 h forecasts from operational analyses and were analyzed for the JJA season of that year. Values are extracted from three hourly model output along the A-Train track and are therefore always within 1.5 h of the overpass time of CloudSat.); (2) a 5 year JJA AMIP climatology of the CAM3 climate model [*Collins et al.*, 2006] with an approximate 2° resolution; (3) an equivalent JJA AMIP climatology of the AM2 climate model [*Geophysical Fluid Dynamics Laboratory Global Atmosphere Model Development Team* (*GFDL-GAMDT*), 2004] also with an approximate 2° resolution; (4) a 5 year JJA climatology from the atmospheric only version of the unified global model of the UK Meteorological Office (HadGEM1) [*Johns*, 2006]; and (5) a more restricted 4 day July simulation of the global cloud-resolving model NICAM (Nonhydrostatic ICosahedral Atmospheric Model) [e.g., *Satoh et al.*, 2008] in the form reported by *Suzuki et al.* [2008] for versions of the model with grid resolutions of 7 and 14 km. The data from the latter four models (CAM3, AM2, HadGem1, and NICAM) were sampled synoptically.

[17] The global, tropical, and midlatitude JJA accumulations derived from these model data are compared to the mean of the combined 2006 and 2007 JJA seasons observed by CloudSat in Figure 2b. All model results lie within the

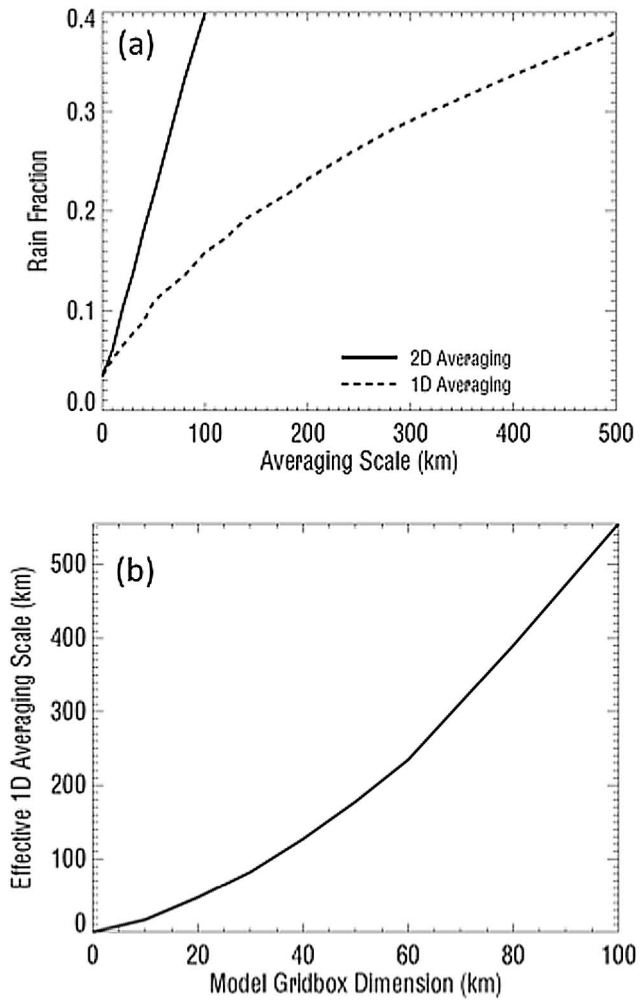

Figure 4. (a) Rainfall fraction from 35°S to 35°N derived from TRMM PR data with different degrees of spatial averaging. (b) Effective linear (1D) averaging scale required to reproduce rain fraction for various degrees of areal (2D) averaging.

observational uncertainties of the observed global, seasonal accumulations. Larger differences exist, however, between model and observations when considering tropical and midlatitude accumulations separately. The accumulations from the four general circulation models that employ more traditional parameterization of moist physics tend to produce too much precipitation over the tropical oceans compared to observations. This excess is then compensated for in the global averages by too little midlatitude precipitation that falls below the estimated lower error of observation. The tropical and midlatitude precipitation of the NICAM model, by contrast to these other models, lies within the range of observational uncertainty for both the tropics and midlatitudes.

4.2. Averaging Methods

[18] The comparisons between model and observed accumulated precipitation shown in Figure 2b are unaffected by differences in spatial resolution between models and observations assuming that the representativeness errors of the observations are negligible. Both the frequency and intensity of precipitation, however, are scale dependent, and some adjustment to the observations is required when making direct comparisons between 2D model fields at different resolutions and the 1D like CloudSat data obtained on a finer resolution. Upscaling such 1D-like data to compare to model areal grid point data is problematic in many ways. Precipitation is fundamentally nonuniform (or intermittent) and the statistical properties of its spatial and temporal averages are known to display a multifractal dimensional nature [e.g., *Lovejoy et al.*, 2008; *Hubert*, 2001]. Multidimensional phenomena-like precipitation obey a power law relation on the scales over which they are averaged [e.g., *Field and Shutts*, 2009]. Averaging precipitation over dimensions less than that of the full space-time in which the phenomena are embedded implies that in general some of the more intense regions (characterized by a lowest dimension) will be undersampled resulting in a shift in the probability distribution function (pdf) of precipitation in favor of lighter rains. This shift in the pdf does not generally result in any bias to the mean of the distribution itself.

[19] In this study we apply a procedure to upscale the CloudSat precipitation to the appropriate model resolution using the following simple averaging approach that is similar to the simple averaging employed in the *Sun et al.* [2006] study. Each granule of data (one orbit) is divided into a number of segments of given length that is predetermined by the model resolution in question. The number of individual CPR profiles within each segment of data is then determined. The existence of precipitation above the specified threshold in any one profile within such a segment is counted as an occurrence for the entire segment and the precipitation rate is then averaged across all profiles within the given segment. The threshold intensity, taken to be 0.05 mm/h at the native resolution of the observations, is similarly reduced by a factor that equals the number of profiles that fall within each segment. The approximate precipitation thresholds used to sample each model were 0.05, 0.01, 0.01, 0.017, 0.3, and 0.6 mm/d for the IFS, CAM3,AM2, UKMO, NICAM 7 km, and NICAM 14 km models, respectively.

[20] Figure 3 illustrates the effects of upscaling on both f_i and r_i. Figure 3 (left) depicts the 2007 JJA frequencies of occurrences defined at the native resolution of the observations accumulated into 2 × 2 latitude/longitude bins, as well as the frequencies of occurrences averaged up to 0.4° and 2° also accumulated into 2 × 2 latitude/longitude bins. Figure 3 (right) shows the matching intensities of precipitation. Figure 3 clearly demonstrates the effect of averaging over successively larger scales with the oceanic mean frequency increasing from 11% to 20% and 40% and the oceanic-wide intensities decreasing from 1.2 mm/h to 0.51 and 0.29 mm/h as the scale of observations increases from the native 1.7 km resolution of the observations to 0.4° and 2°, respectively.

[21] In an effort to place bounds on the impact of this simple upscaling in one as opposed to two dimensions, TRMM PR rainfall data were used to establish approximate 1D averaging scales that preserve the overall frequency of precipitation at a variety of areal averaging scales. PR data of 1 month were aggregated to a variety of spatial scales employing both linear along-track averaging and 2D areal averaging. Resulting estimates of rain fraction over the

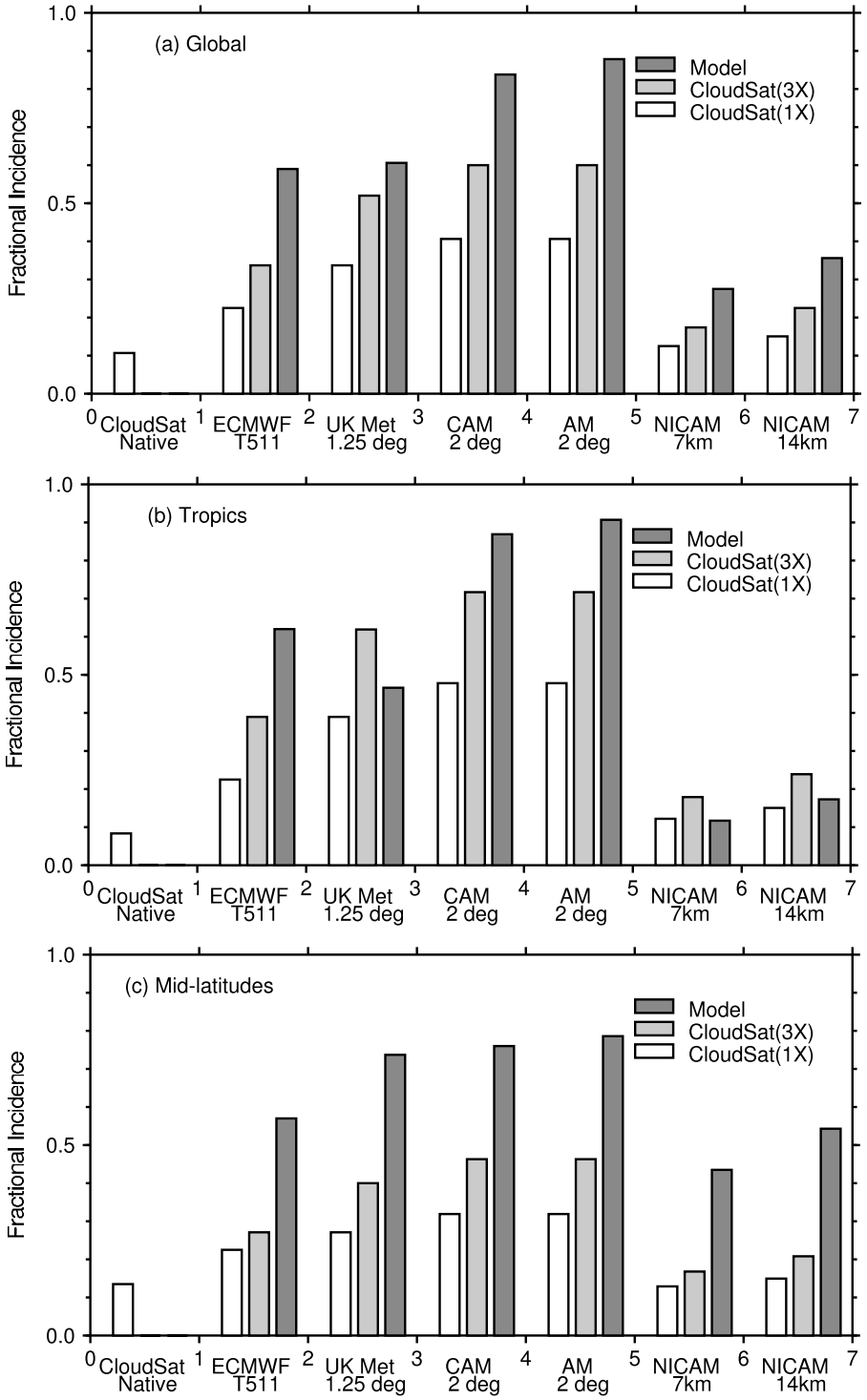

Figure 5. The frequency of precipitation from model simulations and CloudSat observations for global oceans, tropical oceans, and midlatitude oceans as defined in the text.

TRMM domain are presented as a function of averaging scale in Figure 4a. Since areal averaging covers a much larger area than averaging over a linear segment with the same dimension, significantly larger rain fractions are obtained with 2D averaging. We define the length of the 1D averaging segment that most closely approximates the rain fraction at areal averaging scale (or model resolution) to be the "effective" 1D averaging scale that is most representative of that resolution. Using the TRMM PR rainfall data, effective 1D averaging scales for model resolutions ranging from 5 to 100 km are presented in Figure 4b, confirming the expectation that a 1D averaging scale larger than the resolution of the given model is required to obtain the corresponding equivalent 2D rain fraction. To match the 40 km

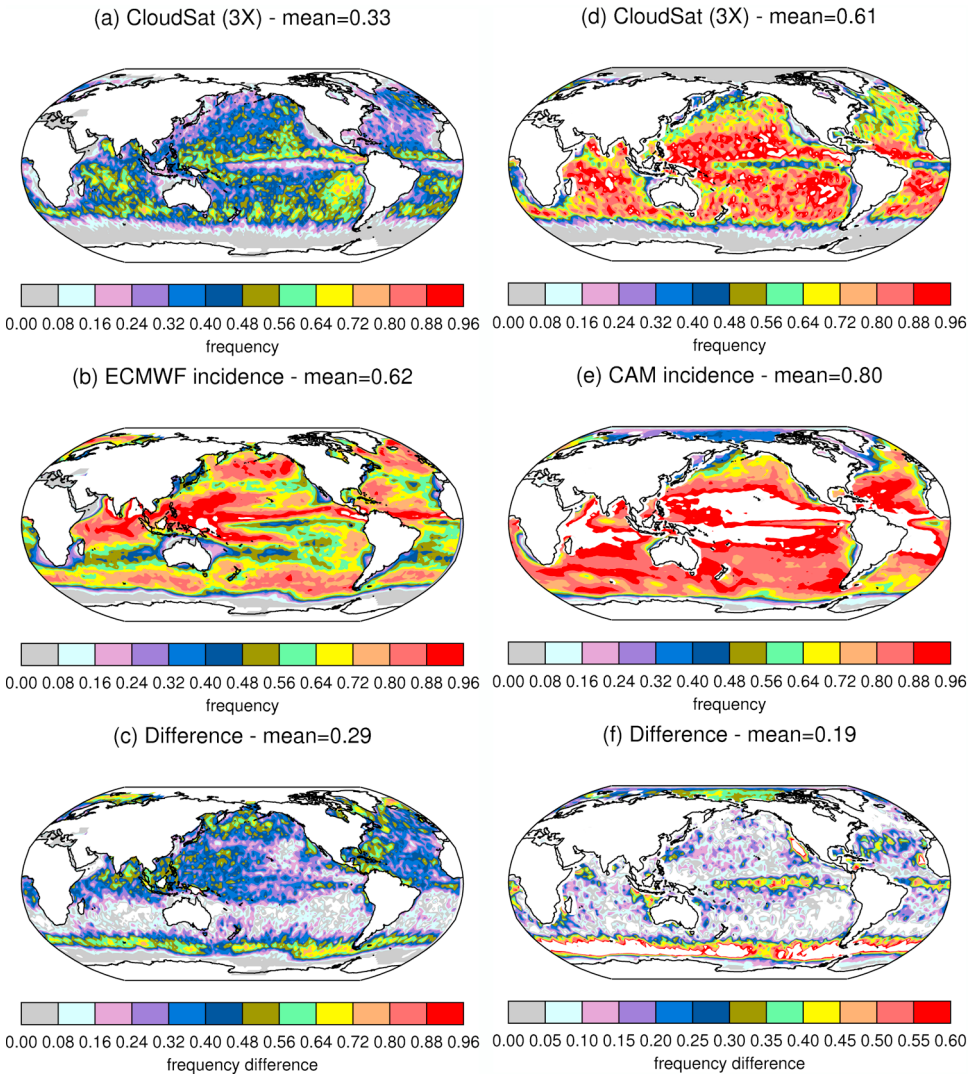

Figure 6. Comparison of the observed global distributions of frequency of occurrence of (top) precipitation to those of from three different models, (left) the 7 km NCAM model, (center) version CY35R1 of the ECMWF integrated forecast system, and (right) the CAM3 climate model. Also shown are zonally averaged differences between model and observations.

resolution of the ECMWF product, for instance, requires CloudSat data to be averaged along a 120 km averaging segment, a factor of three larger than the dimension of the ECMWF grid box. In practice, the scaling between 1D and 2D averaging will depend strongly on the organization of the rainfall being observed and should be expected to vary with location and rainfall regime. In light of the global focus of the current study, however, we adopt this simple approximate upscaling relationship acknowledging that it may introduce some uncertainty in the results. In an effort to represent these uncertainties, results are shown for upscale averaging to two different resolutions, one that matches the model resolution (1×) and one that is a factor of 3 larger than the corresponding model resolutions (3×).

5. Results

[22] The observed and modeled frequencies of occurrence of precipitation are summarized in Figures 5a–5c. The observations presented in Figure 5 are a result of the simple averaging procedures applied to the native CloudSat observations collected for both the 2006 and 2007 JJA seasons. These averaged data are then presented in Figure 5 in the form of composite averages for the global oceans, the tropical oceans, and the oceans of the midlatitudes as previously defined. The global distributions of the (3×) observed and modeled frequencies as well as the global and zonal differences between the model and observations are also shown in Figure 6 for two of the model studies. Presented on Figure 7 are the zonally averaged differences between (3×) observations of precipitation frequency and the frequency of occurrence of precipitation from three of the models. On the whole, the differences between modeled and observed frequency of occurrence of precipitation exceed that explained by observational errors and/or the amount of averaging of the data (Figure 5). The general tendency is for models to produce precipitation that is far too frequent, especially in midlatitudes (e.g., Figures 6 and 7). The occurrences averaged over

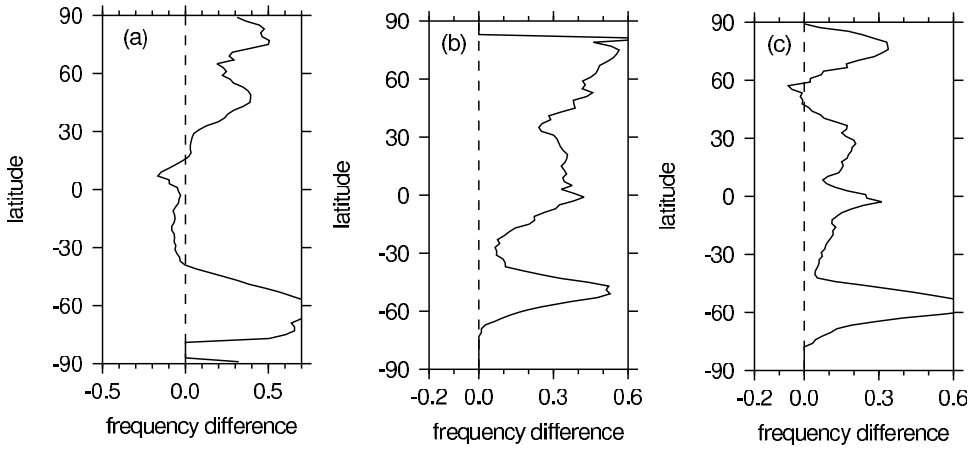

Figure 7. Differences between observed and modeled zonally averaged frequency of occurrence of precipitation: (a) the 7 km NCAM model, (b) version CY35R1 of the ECMWF integrated forecast system, and (c) CAM.

all six model data sets for the entire oceans (60°N-S) are 1.7 and 2.4 times more frequent than the 3× and 1× global CloudSat averages, respectively. The behavior of the NICAM model notably differs from that of the more conventional GCMs. The differences between the latter models and observations tend to be largest at the rainiest latitudes evident in the zonal averaged profiles of Figure 7 whereas the differences between the NICAM model and observations are largest in midlatitudes and there is a general agreement with observations over the tropical oceans. While the biases of NICAM are small over tropical oceans, *Sato et al.* [2009] note that these biases are larger over land.

[23] Comparison of observed and modeled intensity of precipitation is provided in Figure 8. Since the time-integrated accumulations of precipitation closely match the observations (Figure 2b), the consequence of precipitation that is predicted too frequently (Figures 5–7) is the intensity of model precipitation must then also occur too lightly. The mean-model bias of intensity is 1.3 and 1.9 times less than the 3× and 1× averaged observations. Differences in intensity also exist across models. The NICAM model generally agrees with the observations in the tropics as noted in reference to Figure 3 but contains substantial biases in rainfall intensity at higher latitudes. This agreement between the global cloud model and observations in the tropics warrants further comment and further analysis. *Demott et al.* [2007] contrasted the 3–6 hourly precipitation from the community atmospheric model (CAM) and the "multiscale modeling framework" (MMF) version of CAM in which cumulus parameterization is replaced with a cloud resolving model. They found that the CAM produces too much light rainfall that occurs as soon as the boundary layer energizes whereas the MMF precipitation, like observations, lags the build up of energy by several hours. This result suggests that there may be some reason to expect that the NICAM model may capture a more delayed and realistic triggering of convection that perhaps implies a more realistic representation of the frequency of convective precipitation and its intensity. The different behavior of the NICAM model precipitation in the tropics compared to GCMs, reinforced with the suggestions of the Demott et al. study, suggests that the reason for the different behavior of the tropical rainfall of the NICAM model compared to other models requires deeper analysis than can be provided in this study and is a topic of ongoing research.

[24] The results of Figure 9 offer further evidence that the model biases illustrated in Figures 5–8 are not merely a consequence of inadequate upscaling of observations but indicative of a systemic problem of models more generally. Figure 9 illustrates the comparison of model-derived radar reflectivity profiles and CloudSat observed profiles that have been constructed in the form of cumulative frequency in altitude display (CFAD) diagrams. These model reflectivity profiles are calculated using model condensate and model precipitation applied to a CloudSat radar simulator [e.g., *Haynes et al.*, 2009]. Each model grid box is split into subcolumns to represent the subgrid cloud/precipitation fraction and overlap profile explicitly and the reflectivity is then calculated from the "in-cloud" condensate for each subcolumn. These profiles are thus meant to be equivalent to the in-cloud profiles analogous to those obtained by the CloudSat CPR at its native resolution and therefore free of explicit effects of cloud and precipitation grid-scale fractions that affect the model grid box average values. The model CFADs shown are calculated from ECMWF model data extracted along the satellite track for July 2007 over tropical oceans between 30°N-S and the second is a similar composite of UKMO model forecasts for the same month. In both cases, the frequency of the drizzle/rain reflectivity mode (i.e., the low level reflectivity maximum above about −10 dBZ) is clearly overestimated compared to the observed profiles of reflectivity. The observations tend to exhibit a more predominant cloud reflectivity peak around −25 dBZ than is apparent in either model simulated reflectivity profiles. Other differences between model and observed CFADs, such as the differences at higher altitude associated with cloud ice particles, have been discussed in more detail by *Bodas-Salcedo et al.* [2008] and also by *Masunaga et al.* [2008] and *Satoh et al.* [2010]. *Satoh et al.* [2010], for example, produced CFADs that more closely match observations when grau-

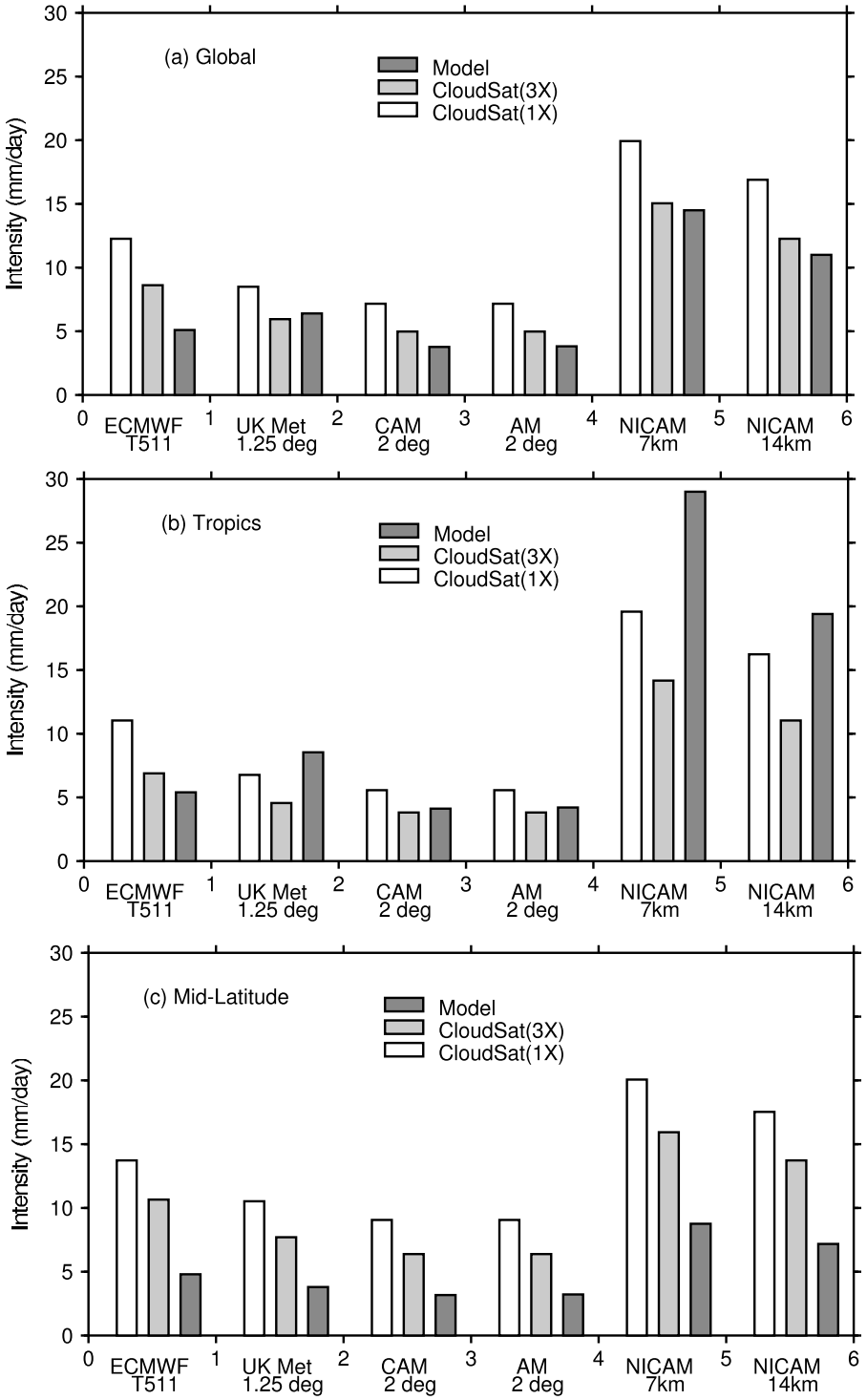

Figure 8. As in Figure 5, but for the instantaneous precipitation intensity.

pel and faster sedimentation of ice particles are introduced into the model.

6. Summary and Discussion

[25] CloudSat's depiction of the character of precipitation over the global oceans [e.g., *Haynes et al.*, 2009] is significantly different from the character of the precipitation produced by global weather and climate models. The five different models used include version CY35R1 of the ECMWF integrated forecast system (IFS) global model with a grid resolution of 40 km (T511), the CAM3 and AM2 climate models each at an approximate 2° resolution, the atmospheric-only version of the unified global model of the UK Meteorological Office (HadGEM1) at 1.25° resolution and the NICAM global cloud-resolving model NICAM with grid resolutions of 7 and 14 km. These represent both

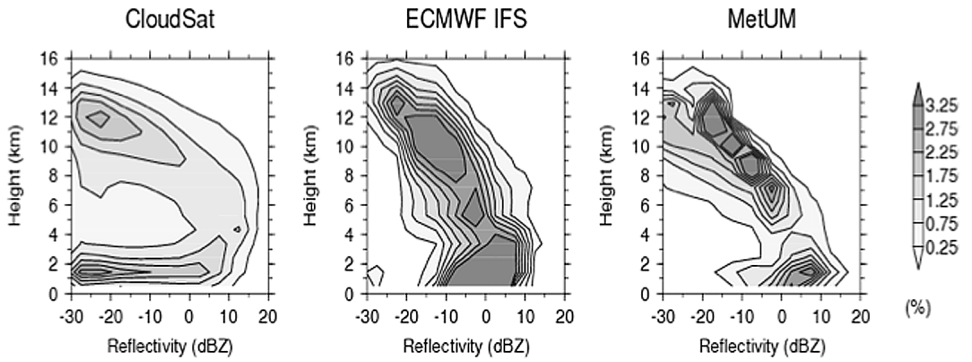

Figure 9. Cumulative frequency in altitude display of 94 GHz radar reflectivities measured by CloudSat and simulated from cloud and precipitation data obtained from two different global prediction models. The data apply to July 2007 and to the tropical region confined between latitudes 30°N/S.

state-of-the-art weather prediction models and climate models and a state-of-the-art emerging high-resolution global cloud model.

[26] The focus of the comparison is on the character of oceanic precipitation defined as a combination of the accumulation of precipitation $a_{\Delta T}$ integrated over some time period ΔT, the frequency of occurrence of precipitation f_i over this same period ΔT and the (instantaneous) intensity r_i of the precipitation. The comparisons presented are restricted to oceanic precipitation primarily because the observations are more reliable there and the analogous results for land precipitation based on surface observations have been described elsewhere [e.g., *Sun et al.*, 2006].

[27] The main findings of the research reported in this paper are as follows:

[28] 1. The oceanic rainfall accumulations derived from a composite of 33 months of CloudSat observations from June 2006 to March 2009, compared to analogous multiyear accumulations derived from AMSR-E, TRMM, and GPCP, reveal a high level of agreement amongst these different data in tropical latitudes (between 30°N-S). The agreement of the CloudSat observations with these other data sources over tropical oceans suggests that the more limited space-time sampling of CloudSat compared to these other observations has little influence on the accumulated annual (and seasonal) statistics. The accumulations in the extra tropical latitudes (between 30°N-S and 60°N-S) exhibit less agreement with the Version-2 AMSR-E data falling significantly below both the CloudSat and GPCP estimates and below the estimated error range of CloudSat. This result was expected given past studies that have pointed to detection issues in AMSR-E precipitation algorithms [*Petty*, 1997; *Ellis et al.*, 2009]. The CloudSat-based global mean oceanic accumulation is 3.0 +0.7/−0.4 mm/d that agrees with GPCP derived from entirely independent data. A surprising result, however, is the observed large accumulation that occurs in midlatitudes (3.12 +1.15/−0.51 mm/d) that independently supports and adds confidence to the more empirical results of GPCP in these latitudes.

[29] 2. The comparison of the observed global mean accumulated precipitation with simulated data from five different models reveals a remarkable level of agreement (Figure 2b). While not entirely unexpected given the basic energy balance control on accumulation, this result nevertheless suggests that the global mean atmospheric energy budget of the models studied is close to reality. The agreement between model and observations begins to break down when regional accumulations are considered. The accumulations from the four general circulation models that employ more traditional parameterization of moist physics tend to produce too much precipitation over the tropical oceans compared to observations. This overestimate is then compensated for in the global averages by too little midlatitude precipitation. The tropical and midlatitude precipitation of the NICAM model, by contrast, lies within the range of observational uncertainty for both the tropics and midlatitudes. The extent that these differences between models relate to the different convective parameterizations used in the GCMs compared to the more explicit convection of NICAM warrants further study.

[30] 3. The differences between the modeled and observed frequencies of precipitation are larger than can be explained by observational errors and by procedures to average the observations to model resolution. The tendency is for models to produce precipitation that is too frequent, as already noted in past studies. The occurrences averaged over the six model data sets are approximately twice the frequency of observations. This bias is partially alleviated in the higher resolution model that treats convection explicitly and whose resolution is closer to the native resolution of the observations where agreement between this model and observations is better in the tropics but significantly worse in midlatitudes. Since the time-integrated accumulations of precipitation more closely match the observations (Figure 2b), it follows that precipitation is predicted too frequently and too lightly. The mean model intensity lies between 1.3 and 1.9 times less than the averaged observations.

[31] 4. As further evidence of the model biases in precipitation character, model reflectivity profiles calculated using model condensate and precipitation applied to a CloudSat radar simulator [e.g., *Haynes et al.*, 2009] were compared to observed profiles. This comparison is meant to represent in-cloud properties free of the uncertainties associated with upscale averaging. The model reflectivity profiles show the influence of excessive drizzle/rain on reflectivity with a maximum low-level frequency of reflectivity occurring low in the atmosphere and above −10 dBZ in contrast to the observed, more predominant low-level cloud

reflectivity peak at −25 dBZ. The extent these differences are merely an effect of subgrid scale variability not captured by the models versus errors in the way precipitation physics is included in models is understudy.

[32] The differences in the character of model precipitation are systemic and have a number of important implications for modeling the coupled Earth system as discussed above. It is also well known that the ability of a numerical model for resolving wave-like fields that vary continuously in time and space is several times the grid resolution [e.g., *Williamson*, 2008]. Our results suggest this is also true of intermittent fields like precipitation. Since the tendency is for increased frequency of precipitation as the averaging scale of observations increases (e.g., Figures 3 and 5), the much higher frequency of occurrences of model grid point precipitation implies that this precipitation is more representative of a scale that is many times the model grid resolution. *Roberts and Lean* [2008], for example, demonstrated that an acceptable measure of skill in precipitation forecasts from high-resolution models of 1 and 12 km resolution occurs at scales of 45–60 and 50–80 km, respectively. This suggests that the real resolution of model precipitation is several times that of the model grid resolution. This implies little skill in precipitation calculated at individual grid points, and thus applications involving downscaling of grid point precipitation to yet even finer-scale resolution has little foundation and relevance to the real Earth system.

[33] **Acknowledgments.** Aspects of this work have been supported both under NASA grant NAG5-11475 and NOAA grant NA17RJ1228 193. A. Bodas-Salcedo was supported by the Joint DECC and Defra Integrated Climate Programme-DECC/Defra (GA01101). NICAM simulations have been performed on the Earth Simulator with financial supports by MEXT/RR2002 project, MEXT/Innovative Program of Climate Change Projection for the 21st Century and JST/CREST.

References

Adler, R. F., et al. (2003), The version 2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present), *J. Hydrometeorol.*, *4*, 1147–1167.

Allen, M. A., and W. J. Ingram (2002), Constraints on future changes in climate and the hydrologic cycle, *Nature*, *419*, 224–232, doi:10.1038/nature01092.

Berg, W., T. L'Euyer, and J. M. Haynes (2010), The distribution of rainfall over oceans from spaceborne radars, *J. Clim. Appl. Meteorol.*, *49*, 535–543.

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

Collins, W. D., et al. (2006), The Community Climate System Model version 3 (CCSM3), *J. Clim.*, *19*, 2122–2143.

Dai, A., and K. E. Trenberth (2004), The diurnal cycle and its depiction in the Community Climate System Model, *J. Clim.*, *17*, 930–951.

Dai, A., X. Lin, and K.-L. Hsu (2007), The frequency, intensity, and diurnal cycle of precipitation in surface and satellite observations over low and midlatitudes, *Clim. Dyn.*, *29*, 727–744.

Dai, A., T. Qian, K. E. Trenberth, and J. D. Milliman (2009), Changes in continental freshwater discharge from 1948–2004, *J. Clim.*, *22*, 2773–2791.

Demott, C. A., D. A. Randall, and M. Khairoutdinov (2007), Convective precipitation variability as a tool for general circulation model analysis, *J. Clim.*, *20*, 91–112.

Ellis, T. D., T. L'Ecuyer, J. M. Haynes, and G. L. Stephens (2009), How often does it rain over the global oceans? The perspective from CloudSat, *Geophys. Res. Lett.*, *36*, L03815, doi:10.1029/2008GL036728.

Field, P. R., and G. J. Shutts (2009), Properties of normalized rain-rate distributions in the tropical Pacific, *Q. J. R. Meteorol. Soc.*, *135*, 175–186.

Geophysical Fluid Dynamics Laboratory Global Atmosphere Model Development Team (GFDL-GAMDT) (2004), The new GFDL global atmosphere and land model AM2-LM2: Evaluation with prescribed SST simulations, *J. Clim.*, *17*, 4641–4673.

Haynes, J. M., T. S. L'Ecuyer, G. L. Stephens, S. D. Miller, C. Mitrescu, N. B. Wood, and S. Tanelli (2009), Rainfall retrieval over the ocean with spaceborne W-band radar, *J. Geophys. Res.*, *114*, D00A22, doi:10.1029/2008JD009973.

Hubert, P. (2001), Multifractals as a tool to overcome scale problems in hydrology, *Hydrol. Sci. J.*, *46*(6), 897–905.

Iguchi, T., T. Kozu, R. Meneghini, J. Awaka, and K. Okamoto (2000), Rain-profiling algorithm for the TRMM precipitation radar, *J. Appl. Meteorol.*, *39*, 2038–2052.

Johns, T. C. (2006), The New Hadley Centre Climate Model (HadGEM1): Evaluation of coupled simulations, *J. Clim.*, *19*, 1327–1353.

Kummerow, C., Y. Hong, W. S. Olson, S. Yang, R. F. Adler, J. McCollum, R. Ferraro, G. Petty, D. B. Shin, and T. T. Wilheit (2001), The evolution of the Goddard Profiling Algorithm (GPROF) for rainfall estimation from passive microwave sensors, *J. Appl. Meteorol.*, *40*, 1801–1820.

Kummerow, C., W. Barnes, T. Kozu, J. Shiue, and J. Simpson (1998), The Tropical Rainfall Measuring Mission (TRMM) sensor package, *J. Atmos. Oceanic Technol.*, *15*, 809–817.

L'Ecuyer, T. S., and G. L. Stephens (2002), An estimation-based precipitation retrieval algorithm for attenuating radars, *J. Appl. Meteorol.*, *41*, 272–285.

Lovejoy, S., D. Schertzer, and V. Allaire (2008), The remarkable wide range spatial scaling of TRMM precipitation, *Atmos. Res.*, *90*, 10–32, doi:10.1016/j.atmosres.2008.02.016.

Masunaga, H., M. Satoh, and H. Miura (2008), A joint satellite and global cloud-resolving model analysis of a Madden-Julian Oscillation event: Model diagnosis, *J. Geophys. Res.*, *113*, D17210, doi:10.1029/2008JD009986.

Milly, E. C. D., J. Betancourt, M. Falkenmark, R. M. Hirsch, Z. W. Kudzewicz, D. R. Lettenmaier, and R. J. Stouffer (2008), Climate change: Stationarity is dead: Whither water management?, *Science*, *319*, 573–574.

Neelin, J. D., M. Munnic, H. Su, J. E. Meyerson, and C. E. Holloway (2006), Tropical drying trends in global warming models and observations, *Proc. Natl. Acad. Sci. U. S. A.*, *103*, doi:10.1073/pnas.06017981093.

Pall, P., M. R. Allen, and D. A. Stone (2007), Testing the Clausius-Clapeyron constraint on changes in extreme precipitation under CO_2 warming, *Clim. Dyn.*, *28*, 351–363, doi:10.1007/s00382-006-0180-2.

Petty, G. W. (1997), An intercomparison of oceanic precipitation frequencies from 10 special sensor microwave/imager rain rate algorithms and shipboard present weather reports, *J. Geophys. Res.*, *102*(D2), 1757–1777, doi:10.1029/96JD03000.

Roberts, N. M., and H. W. Lean (2008), Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events, *Mon. Weather Rev.*, *136*, 78–97.

Salby, M., and P. Callahan (2007), Sampling error in climate properties derived from satellite measurements: Consequences of under sampled diurnal variability, *J. Clim.*, *10*, 18–36.

Sato, T., H. Miura, M. Satoh, Y. N. Takayabu, and Y. Wang (2009), Diurnal cycle of precipitation in the tropics simulated in a global cloud-resolving model, *J. Clim.*, *22*, 4809–4826, doi:10.1175/2009jcli2890.1.

Satoh, M., T. Matsuno, H. Tomita, H. Miura, T. Nasuno, and S. Iga (2008), Nonhydrostatic icosahedral atmospheric model (NICAM) for global cloud resolving simulations, *J. Comput. Phys.*, *227*, 3486–3514.

Satoh, M., T. Inoue, and H. Miura (2010), Evaluations of cloud properties of global and local cloud system resolving models using CALIPSO/CloudSat simulators, *J. Geophys. Res.*, *115*, D00H14, doi:10.1029/2009JD012247.

Stephens, G. L. (2005), Cloud feedbacks in the climate system: A critical review, *J. Clim.*, *18*, 237–273.

Stephens, G. L., et al. (2008), CloudSat mission: Performance and early science after the first year of operation, *J. Geophys. Res.*, *113*, D00A18, doi:10.1029/2008JD009982.

Stephens, G. L., and T. D. Ellis (2008), Controls of global mean precipitation increases in global warming GCM experiments, *J. Clim.*, *21*, 6141–6155.

Stephens, G. L., and Y. Hu (2010), Are climate-related changes to the character of global precipitation predictable?, *Environ. Res. Lett.*, *5*, 025209, doi:10.1088/1748-9326/5/2/025209.

Sun, Y., S. Solomon, A. Dai, and R. W. Portmann (2006), How often does it rain?, *J. Clim.*, *19*, 916–934.

Suzuki, K., T. Nakajima, M. Satoh, H. Tomita, T. Takemura, T. Y. Nakajima, and G. L. Stephens (2008), Global cloud-system-resolving simulation of aerosol effect on warm clouds, *Geophys. Res. Lett.*, *35*, L19817, doi:10.1029/2008GL035449.

Trenberth, K. E., and A. Dai (2007), Effects of Mount Pinatubo volcanic eruption on the hydrological cycle as an analog of geoengineering, *Geophys. Res. Lett.*, *34*, L15702, doi:10.1029/2007GL030524.

Tanelli, S., S. L. Durden, E. Im, K. S. Pak, D. G. Reinke, P. Partain, J. M. Haynes, and R. T. Marchand (2008), CloudSat's cloud profiling radar after two years in orbit: Performance, calibration, and processing, *IEEE Trans. Geosci. Remote Sens.*, *46*, 3560–3573.

Wilcox, E. M., and L. J. Donner (2007), The frequency of extreme rain events in satellite rain-rate estimates and an atmospheric General Circulation Model, *J. Clim.*, *20*, 53–69.

Williamson, D. L. (2008), Equivalent finite volume and Eulerian spectral transform horizontal resolutions established from aqua-planet simulations, *Tellus, Ser. A*, *60*, 839–847, doi:10.1111/j.1600-0870.2008.00340.x.

Xian, P., J. S. Reid, J. F. Turk, E. J. Hyer, and D. L. Westphal (2009), Impact of modeled versus satellite measured tropical precipitation on regional smoke optical thickness in an aerosol transport model, *Geophys. Res. Lett.*, *36*, L16805, doi:10.1029/2009GL038823.

A. Bodas-Salcedo, Met Office Hadley Centre, Exeter, UK.

R. Forbes, European Centre for Medium-Range Weather Forecasts, Reading, UK.

A. Gettlemen, National Center for Atmospheric Research, Boulder, CO 80307-3000, USA.

J.-C. Golaz, Geophysical Fluid Dynamics Laboratory, Princeton University, P.O. Box 308, Forrestal Campus, Princeton, NJ 08542, USA.

J. Haynes, Monash University, Alfred Hospital Commercial Road, Melbourne, VIC 3004, Australia.

T. L'Ecuyer, P. Gabriel, G. L. Stephens, and K. Suzuki, Department of Atmospheric Sciences, Colorado State University, Fort Collins, CO 80523, USA.