LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-JRNL-409248

Evaluation of tropical cloud and precipitation statistics of CAM3 using CloudSat and CALIPSO data

Y. Zhang, S. Klein, J. Boyle, G. G. Mace

December 5, 2008

Journal of Geophysical Research-Atmosphere

Disclaimer

Evaluation of tropical cloud and precipitation statistics of CAM3 using CloudSat and CALIPSO data

Y. Zhang[1], S. A. Klein[1], J. Boyle[1], and G. G. Mace[2]

1. Lawrence Livermore National Laboratory, Livermore, California

2. Department of Meteorology, University of Utah, Salt Lake City, Utah

Submitted to the Journal of Geophysical Research - Atmospheres

Abstract

The combined CloudSat and CALIPSO satellite observations provide the first simultaneous measurements of cloud and precipitation vertical structure, and are used to examine the representation of tropical clouds and precipitation in the Community Atmosphere Model Version 3 (CAM3). A simulator package utilizing a model-to-satellite approach facilitates comparison of model simulations to observations, and a revised clustering method is used to sort the subgrid-scale patterns of clouds and precipitation into principal cloud regimes.

Results from weather forecasts performed with CAM3 suggest that the model underestimates the horizontal extent of low and mid-level clouds in subsidence regions, but overestimates that of high clouds in ascending regions. CAM3 strongly overestimates the frequency of occurrence of the deep convection with heavy precipitation regime, but underestimates the horizontal extent of clouds and precipitation at low and middle levels when this regime occurs. This suggests that the model overestimates convective precipitation and underestimates stratiform precipitation consistent with a previous study that used only precipitation observations.

Tropical cloud regimes are also evaluated in a different version of the model, CAM3.5, which uses a highly entraining plume in the parameterization of deep convection. While the frequency of occurrence of the deep convection with heavy precipitation regime from CAM3.5 forecasts decreases, the incidence of the low clouds with precipitation and congestus regimes increases. As a result, the parameterization change does not reduce the frequency of precipitating convection that is far too high relative to observations. For both versions of CAM, clouds and precipitation are overly

65 reflective at the frequency of the CloudSat radar and thin clouds that could be detected by
66 the lidar only are underestimated.

1. Introduction

Although Global Climate Models (GCMs) are the primary tools to predict climate change, large uncertainties remain in projections of future climate after more than 30 years of GCM development (Houghton et al., 2001; Randall et al., 2007). The different representations of clouds and their feedback processes in GCMs have been identified as the major source of differences in model climate sensitivities (Cess et al. 1990; Soden et al. 2004; Zhang et al. 2005). These differences arise because contemporary GCMs cannot resolve clouds and highly simplified parameterizations are used to represent the interactions between clouds and radiation and the large-scale environment resolved by GCMs. It has been pointed out that improved present-day cloud simulations are needed to reduce the uncertainties in predicting future climate (Bony et al. 2006; Williams and Tselioudis, 2007). Widely collected observations are required to assess model performance and provide valuable information for the development of new parameterizations. However, the evaluation of GCM cloud simulations has long been hampered by the lack of suitable observations.

Field programs with intensive observations are not sufficient to solve the parameterization problem, because it is unlikely that a few cases will be representative enough. Traditional methods to obtain global perspective, such as the International Satellite Cloud Climatology Project (ISCCP; Rossow and Schiffer, 1999) and the Earth Radiation Budget Experiment (ERBE; Wielicki et al. 1996) rely on radiances observed by passive sensors on satellites. But because these radiances depend on the integrated effect of properties of the whole atmospheric column, they provide little information of the vertical structure of cloud fields. The lack of vertical structure information prevents an understanding of the hydrologic cycle and the modulation by clouds of the vertical

92 distribution of radiative heating rates; it also hinders the evaluation of GCM cloud
93 simulations. Launched in April 2006, the CloudSat and CALIPSO satellites, flying in the
94 A-Train constellation (Stephens et al. 2002), provide the first global survey of the vertical
95 distribution of cloud condensate and precipitation. The Cloud Profiling Radar (CPR) on
96 CloudSat (Im et al., 2006) is the first spaceborne millimeter-wavelength radar capable of
97 penetrating optically thick hydrometeor layers. The CALIPSO satellite carries a lidar
98 system (Winker et al., 2007) as its primary payload capable of detecting optically thin
99 clouds. The combined information from the two instruments is able to accurately
100 characterize the vertical as well as horizontal structure of hydrometeor layers (Mace et al.
101 2008). The only clouds missed by the combined dataset are low-level clouds with
102 reflectivity less than the detection threshold of the radar that are also beneath clouds
103 which completely attenuate the lidar pulse (Mace et al. 2008).

104 In this study, CloudSat and CALIPSO data are used to evaluate simulations of
105 cloud and precipitation statistics from CAM3 (Collins et al. 2006), a major United States
106 climate model. Traditional methods of GCM evaluation use maps of large spatial and
107 temporal means of cloud variables from both models and observations. However, this
108 method cannot provide an effective constraint on cloud simulations and cannot assess
109 cloud radiative feedback due to compensating errors (Norris and Weaver, 2001; Williams
110 et al., 2005). Another popular method is to investigate relationships between clouds and
111 other atmospheric parameters using compositing techniques (Ringer and Allan, 2004).
112 Atmospheric parameters, such as 500-hPa vertical velocity, sea surface temperature, and
113 lower tropospheric stability (Bony et al., 2004; Williams et al. 2006), have been used in
114 order to document the relationships between clouds and the parameters that are thought to
115 affect their evolution. However, it is difficult to identify a small set of key atmospheric

parameters (Williams et al., 2003; Bony et al., 2004), and there is a lack of reliable data for some atmospheric parameters. In this study, the cluster analysis method is used to objectively identify cloud regimes based on cloud observations alone without any knowledge of other meteorological parameters. By looking for distinctive cloud subgrid-scale patterns in ISCCP data, this method has been widely used to characterize cloud regimes and evaluate model simulations in recent years (Jakob and Tselioudis, 2003; Rossow et al., 2005; Gordon et al., 2005; Williams and Tselioudis, 2007; Chen and Del Genio, 2008). The clustering method has also been used to evaluate precipitation regimes from Tropical Rainfall Measurement Mission (TRMM) precipitation radar data (Boccippio et al., 2005) and cloud regimes in CloudSat data (Zhang et al., 2007; hereafter Zhang07), and to stratify TRMM latent-heating observations by ISCCP cloud regimes (Jakob and Schumacher, 2008). More recently, Marchand et al. (2009) presented the evaluation of modeled hydrometeor occurrence vertical profiles at the ARM Oklahoma site by clustering the large-scale dynamic and thermodynamic fields. This is the first study to use the cluster analysis method on the combined data from CloudSat and CALIPSO to evaluate cloud and precipitation statistics of a climate model.

Due to the important role of tropical cloud system in global atmospheric circulation, our study will focus on the model simulations in tropical regions. The paper is organized as follows. In the next section, observational data, model simulations, and the cluster analysis method are briefly described. The simulator package that converts model output to observed variables is introduced in section 3, and cloud regimes from observational data are described in section 4. In section 5, model simulations are evaluated within the clustering framework, and changes resulting from the addition of new parameterizations to the CAM are shown. A summary is provided in section 6.

2. Data and Methodology

2.1 Observations

The CloudSat and CALIPSO satellites are maintained in tight orbital configuration to facilitate merging of data streams. The orbit is sun-synchronous with the overpass occurring around 1:30 am/pm local time. The ground track repeats every 16 days, and the orbital period is 99 minutes. The CPR on CloudSat is a 94-GHz nadir-pointing radar that records range-resolved profiles of backscattered power with a nominal footprint of 1.4 km across by 2.5 km along track. Due to the sensitivity of the radar to large particles, the CPR detects both clouds and precipitation. The estimated CPR minimum detectable signal is -30 dBZ, and contamination by surface reflection in the lowest 500m of the atmosphere renders the signal unusable for hydrometeor identification (Mace et al. 2007). Due to these limitations, CloudSat will miss some fraction of thin cirrus, mid-level liquid water clouds, and non-precipitating cumulus and stratocumulus clouds as well as all low-level clouds below 500m.

The two-wavelength (1064 nm and 532 nm) polarization lidar on CALIPSO provides high resolution vertical profiles of backscattered power from which clouds and aerosols may be identified. The lidar system, which has higher horizontal and vertical resolution than the CPR, has the capability to sense optically thin layers with optical depths of 0.01 or less (Winker et al. 2007), and other clouds such as non-precipitating stratocumulus whose reflectivity is below the detection threshold of the radar. On the other hand, the lidar quickly attenuates beyond optical depths of about 3 and cannot detect many clouds and precipitation identified by the radar (Zhang and Mace 2006; Mace et al. 2009). The CPR and the CALIPSO lidar complement each other in their capabilities to observe clouds.

165 In this study, two CloudSat standard data products are used to characterize cloud
166 vertical structures. The first is the Level 2 GEOPROF product (Mace 2004; Mace et al.
167 2007) which identifies the occurrence of hydrometeors with a masking algorithm and
168 provides the radar effective reflectivity factor, Z_e, expressed in dBZ($=10\log_{10}Z_e$). The
169 masking algorithm is described in more detail by Marchand et al. (2008). The second is
170 the Level 2 GEOPROF-LIDAR product (Mace et al. 2008) which contains the estimates
171 of lidar-determined cloud fraction within CPR sample volumes. The lidar information is
172 from the CALIPSO Level 2 Vertical Feature Mask which reports the location of aerosol
173 and cloud types.

174 In this study, tropical (23.5˚S − 23.5˚N) observations for the period June-
175 September in 2006 are used. Although not shown here, data for the same months in 2007
176 confirm the robustness of the results. Following the approach in Zhang07, a sequence of
177 200 adjacent profiles of satellite data (approximately 2° of latitude) define an individual
178 cloud region from which joint histograms of atmospheric pressure and signal strength are
179 computed to characterize the subgrid-scale patterns of cloud and precipitation. The
180 histograms contain the relative frequency of occurrence (RFO) of clouds and
181 precipitation in categories of seven signal bins and seven pressure levels; a sample
182 histogram is shown in Figure 1. To construct the joint histograms, radar reflectivity above
183 -30 dBZ with CPR cloud mask greater than or equal to 20, which means clouds with low
184 chance of a false detection (Marchand et al., 2008), is binned into six categories with a
185 bin interval of 10 dBZ. A seventh bin at the left side of the diagram displays the RFO of
186 lidar detected clouds which are not detected by the radar because the reflectivity is less
187 than -30 dBZ, the minimum detectable signal of the radar. The reported RFO is the
188 percentage of observations within a given pressure bin that have the reported signal

189 strength. Thus, if all volumes within a given pressure range for a 2° region had cloud or
190 precipitation identified by either CloudSat or CALIPSO, then the sum of RFOs over all 7
191 signal bins in the given pressure range would be 100%. To facilitate comparison with
192 previous cluster studies using ISCCP data, the boundaries of the seven pressure bins
193 coincide with those used by ISCCP, and the conversion from altitude to pressure is
194 attained by use of analysis data provided by the European Center for Medium-range
195 Weather Forecasts (ECMWF) in the ECMWF-AUX product released with CloudSat and
196 CALIPSO data. The characteristic patterns of this joint histogram will be used in the
197 cluster analysis technique to determine tropical cloud regimes. While these 200-profile
198 snapshots are created sequentially, a sensitivity study, which uses another set of
199 snapshots collected by taking a 100-profile step forward compared with the original set,
200 shows that these 200-profile snapshots are able to independently represent the tropical
201 cloud regimes.

202 ***2.2 CAM3 and model integrations***

203 In this study, simulations of cloud and precipitation statistics of the Community
204 Atmospheric Model version 3.1 (Collins et al., 2006) are examined. The version of
205 CAM3.1 used in this study employs the finite volume dynamical core with horizontal
206 resolution of 1.9º latitude by 2.5º longitude and 26 vertical levels. CAM 3.1 treats
207 stratiform cloud microphysics based on the prognostic cloud water formulation of Rasch
208 and Kristjansson (1998) with modifications made by Zhang et al. (2003). There are two
209 parameterizations of moist convection in the model: a shallow depth mixing
210 parameterization (Hack 1993) and a deep convection parameterization (Zhang and
211 McFarlane, 1995) which convects whenever the convective available potential energy
212 exceeds a small threshold of 70 J/kg.

213 In addition to CAM3.1, a later version of the model, CAM3.5, will also be
214 evaluated. While there are numerous differences between the two versions, the key
215 difference lies in two modifications to the parameterization of deep convection. The first
216 modification is the inclusion of a parameterization of cumulus momentum transport
217 (Richter and Rasch, 2008). The second modification uses a highly entraining (as opposed
218 to undilute) plume to calculate available potential energy and prohibits convection when
219 there is no available potential energy for this entraining plume (Neale et al., 2008). As a
220 result, deep convection will be suppressed if the troposphere is dry even if the convective
221 available potential energy for an undilute plume exceeds 70 J/kg.

222 Although CAM is a climate model, we examine simulations of CAM performed in
223 weather-forecast mode (Phillips et al., 2004) to better identify parameterization-related
224 deficiencies in the simulation of clouds and precipitation. With a weather-forecasting
225 approach, it is more likely that errors can be ascribed to the model parameterizations of
226 moist processes, because the large-scale atmospheric state in the early periods of a
227 forecast is relatively close to reality. In this study, a series of forecasts are performed
228 which commence every day in the time period from June to September 2006. Forecasts
229 are initialized from analyses of the National Center for Environmental Prediction (NCEP)
230 and we examine model data from day-2 forecasts. We analyze model output from this
231 forecast time-range because most of the fast-time scale spin-up issues are resolved by day
232 2 (Boyle et al. 2008).

233 Considering the overpass time of the A-Train constellation, the model simulations
234 at 1 am and pm local time are compared to observations. Tests show that the geographical
235 distribution of the RFO of cloud regimes significantly changes if simulator output at
236 other times is used while the joint histograms of atmospheric pressure and signal strength

237 are still similar to those from model output at 1 am and 1pm local time. This reminds the
238 reader that some of the geographical patterns shown below result from an incomplete
239 sampling of the diurnal cycle by CloudSat and CALIPSO (Liu et al., 2008).

240 ***2.3 Clustering method***

241 In this paper, the joint histograms of atmospheric pressure and signal strength are
242 used to characterize the vertical distributions of hydrometeors. In Zhang07, characteristic
243 patterns in these histograms of CloudSat data were identified using a K-means cluster
244 algorithm (Anderberg, 1973). The algorithm determined the patterns from a vector that
245 consisted of the 42 independent elements of the joint histogram. A drawback of this
246 method is that information on the distance in pressure or signal strength between
247 elements is not considered and thus results may be sensitive to the discretization of the
248 histogram (Williams and Webb, 2008). As an alternative, clustering is performed using a
249 7 element vector that equivalently illustrates the vertical profiles of signal strength. This 7
250 element vector which we call the normalized mean dBZ index is computed from the joint
251 histogram of cloud patterns in the following manner. As depicted in the upper abscissa of
252 Figure 1 (c), a dBZ index integer for each bin of signal strength is assigned. For example,
253 if the radar reflectivity dBZ is between -20 and -10, the dBZ index is 3. Likewise if the
254 hydrometeor is detected by the lidar only, the dBZ index is set to 1. The normalized mean
255 dBZ index at each of the 7 pressure levels is computed as the sum of RFO_{cld}*dBZ_ind,
256 where RFO_{cld} is the relative occurrence frequency of a certain dBZ range/lidar bin in all
257 the cloudy pixels at a given pressure level, and dBZ_ind is the dBZ index. If there are no
258 hydrometeors in a pressure level, then the normalized mean dBZ index is set to 0. In
259 Figure 1 (c), the line with diamonds shows the vertical profile of the normalized mean
260 dBZ index for this cloud pattern.

There are two major benefits to expressing the vertical structure of a hydrometeor pattern in this way. First, the vertical profile of the normalized mean dBZ index describes the dominant hydrometeor system in a region. This is because higher radar reflectivity roughly corresponds to larger particle sizes and cloud water contents. Rain and drizzle is indicated by dBZ larger than ~ -15 (Frisch et al., 1995; Stephens and Woods, 2007) whereas liquid clouds without rain or drizzle will have dBZ less than -15 and often less than -30, in which case only the lidar can detect the cloud. For ice, thin cirrus clouds typically have dBZ of -50 to -20 dBZ, whereas larger ice particles exhibit dBZ larger than -20. Second, the use of a normalized mean dBZ index facilitates the comparison of observations with model simulations. This is because the model only predicts the grid-box mean cloud and precipitation condensate and thus assumptions would be necessary to reproduce the spread of dBZ often observed in clouds. Although we could use assumptions to generate the subgrid-scale variability in the simulator, the current version of the simulator distributes the model's cloud condensate and precipitation uniformly among the subgrid-scale columns designated to have cloud or precipitation, with the result that the histograms of signal strength are more narrow than is typically observed. A negative consequence of using the normalized mean dBZ index is that cloud coverage, a variable used in previous clustering analyses (Jakob 2003, Williams and Webb, 2008), is unused. Note that while clustering is performed using the normalized mean dBZ index, all results in this paper are displayed using the joint histogram of atmospheric pressure and signal strength.

The clustering method iteratively searches for a predefined number of clusters starting with initial seeds. These seeds, used to create the initial cluster centroids, are selected randomly from the dataset with the only restriction being low correlation

between any two seeds. The cluster centroids represent specific patterns in the vertical profile of the mean dBZ index. Every 2 degree CloudSat curtain is assigned to the cluster whose centroid has the minimum Euclidian distance in the vertical profile of the mean dBZ index. There are two ways to calculate the cluster centroids during the iterations. One is to recalculate the centroids after all elements are assigned to a cluster, and the other is to recalculate the cluster centroid each time an element is assigned to a cluster. The latter way is used here because results depend less on the initial seeds chosen and the algorithm converges faster. To simply test the sensitivity of clustering results to initial seeds, the algorithm was repeated 30 times and a dominant set of cloud clusters is obtained in at least 75% of tests.

A limitation of the K-means algorithm is that the number of clusters needs to be subjectively specified in advance. Here the number of clusters is determined following the empirical criteria of Rossow et al. (2005). The correlation coefficients among the vertical profiles of the normalized mean dBZ index of the centroids and the geographical distributions of the frequency of occurrence of each cluster are used to judge the outcome. If the correlation between any two resulting clusters in both the centroid and the geographical distribution exceeds 0.7, the two clusters are designated as belonging to the same principal cloud regime. Although we did not find it necessary in this study, other studies have made subjective decisions to combine as a final step some of the resulting clusters into a set of principal cloud regimes (Williams and Tselioudis 2007; Williams and Webb 2008).

3. CFMIP observation simulator package (COSP)

To facilitate a meaningful comparison of the model with CloudSat and CALIPSO measurements, we use version 1.1 of a simulator package which has been developed

311 through international collaborations under the framework of the Cloud Feedback Model
312 Intercomparison Project (CFMIP, http://cfmip.metoffice.com/COSP.html). To avoid
313 significant ambiguities in the direct comparison of model simulations with retrievals from
314 observations, the CFMIP Observation Simulator Package (COSP) converts model clouds
315 into pseudosatellite observations with a model to satellite approach that mimics the
316 satellite view of an atmospheric column with model-specified physical properties. The
317 approach accounts for observational limitations of the instruments as described below.

318 COSP has three major parts: 1) the generation of a subgrid-scale distribution of
319 cloud and precipitation, 2) the simulation of radar and lidar signals from this distribution,
320 and 3) the computation of statistical summaries from the subgrid-scale distribution of
321 simulated signals which can then be compared to similar statistical summaries computed
322 from observations. In the first part, each GCM grid box is equally divided into a number
323 of vertical columns (50 in this case) and clouds are assigned to these columns in a manner
324 consistent with the model's grid-box average stratiform and convective cloud amounts
325 and its cloud overlap assumption. The scheme which produces a subgrid distribution of
326 clouds is the Subgrid Cloud Overlap Profile Sampler (SCOPS) which is also used in the
327 ISCCP simulator (Klein and Jakob, 1999; Webb et al. 2001). Note that the grid-box mean
328 cloud condensate is divided equally among all columns that SCOPS designates as cloudy.

329 The next step is to determine which of the columns generated by SCOPS contain
330 rain and snow. The scheme used is called SCOPS_PREC and is similar to that of
331 Chevallier and Bauer (2003) and O'Dell et al. (2007). The inputs to SCOPS_PREC
332 include the column distribution of large-scale and convective clouds from SCOPS and the
333 model's grid-box mean precipitation flux of large-scale and convective rain and snow.
334 Note that this scheme currently ignores any parameterization of precipitation area

fraction that some models have (Jakob and Klein, 2000). To allow a close match between clouds allocated by SCOPS and precipitation produced by the clouds, precipitation is assigned to columns with the following algorithm which starts at the top-of-atmosphere and proceeds downward to the surface. There are in total five possibilities for the assignment of precipitation to columns, and they are used with different priorities. First, large-scale precipitation is assigned to all columns that either have stratiform clouds in the current level (possibility one) or large-scale precipitation in the level above (possibility two). These two possibilities account for the overwhelming majority of cases. However, there may be rare instances where precipitation is not assigned after applying these possibilities. For these rare instances, the following possibilities are applied. The third possibility is to assign large-scale precipitation to all columns that have stratiform clouds in the level below. If precipitation is not assigned with the third possibility, then large-scale precipitation is assigned to all columns that have stratiform clouds anywhere in the vertical column (possibility four). If after this possibility, precipitation is still not assigned then it is assumed that large-scale precipitation covers 100% of the area and every column is filled with precipitation (possibility five). Possibility five is only used in the pathological case where the grid box has stratiform precipitation but no stratiform clouds. The same method is used to assign convective precipitation to columns using the convective clouds apportioned by SCOPS. The only difference is that convective precipitation is assumed to cover 5% of the area in possibility five. Following this assignment, the gridbox mean precipitation flux is, for lack of a better method, divided equally among all of the columns assigned to have precipitation. Then, the local precipitation flux is converted to a mixing ratio following Khairoutdinov and Randall

(2003) who assume a Marshall-Palmer size distribution for precipitation and make a set of assumptions for particle terminal velocity.

In the second part of COSP, the radar and lidar signals are calculated using the column distribution of cloud and precipitation. The QuickBeam code (Haynes et al., 2007) is used to simulate the radar signal and calculates the vertical profiles of radar reflectivity accounting for attenuation of the radar beam from intervening hydrometeors, the atmospheric profiles of temperature and humidity, and assumptions for the particle size distributions of each hydrometeor. The ACTSIM code (Chiriaco et al. 2006; Chepfer et al. 2007) is used to simulate the lidar signal and calculates the vertical profile of lidar backscatter from the same set of modeling variables excluding precipitation hydrometeors which contribute negligibly to the lidar backscatter. The simulated signals are considered valid where cloud optical depth is lower than about 2.5 and saturated if cloud optical depth exceeds this value. Aerosols are not currently included in the lidar simulator.

In the third part of COSP, statistical summaries are generated from these simulated signals in a manner similar to that used to derive the hydrometeor mask from the CloudSat and CALIPSO observations (Mace et al., 2009). In particular, we compute the joint histogram of atmospheric pressure and signal strength taking into account the radar sensitivity of -30 dBZ, surface contamination effects (Mace et al. 2007), and saturation of lidar signals. When the lidar detects cloud using a threshold value of normalized backscatter ratio of 3 and radar reflectivity is less than -30 dBZ, the occurrence frequency will contribute to the first column of the histogram. Volumes with radar reflectivity less than -30 dBZ that are beneath the level of complete attenuation of the lidar beam will be considered as clear. In these ways, the cloud and precipitation fields from model

382 simulations are diagnosed in a manner as close as possible as the diagnosis with real
383 observations.

384 While many sources of uncertainty can affect the output of COSP, two major
385 uncertainties arise from the assumed particle size distributions for different hydrometeors
386 and the methods used to generate subgrid-scale inhomogeneity in cloud condensate and
387 precipitation. For example, Bodas-Salcedo et al. (2008) examined the role of the shape of
388 the ice particle size distribution and found that the calculated radar reflectivity can
389 change by around 5 dBZ from increasing or decreasing the intercept of the assumed
390 exponential distribution by a factor of 5. Since the signal bin width we select is 10 dBZ,
391 an uncertainty of this magnitude will not significantly change our conclusions. Further
392 exploration of uncertainties can be made by using the different distribution models
393 available in the radar simulator. The applicability of homogenous horizontal distribution
394 of cloud condensate and precipitation in subgrid scale and the cloud and precipitation
395 overlap are two important issues for an accurately simulated signal. Zhang et al. (2005)
396 found little sensitivity of model biases in comparison with ISCCP observations to the
397 replacement of randomly overlapped horizontally homogenous clouds with exponentially
398 decaying overlapped horizontally inhomogeneous clouds following the method reported
399 on in Pincus et al. (2006). For COSP, the signals will also be sensitive to the assumption
400 that the entire cloud generates precipitation and that the precipitation area does not
401 decrease beneath the cloud unless all of the precipitation evaporates. Testing the
402 sensitivity of the simulated signals to these assumptions will require future work. In the
403 context of this study, we will partially address the possible bias caused by distribution
404 assumptions by artificially homogenizing the observations to GCM gridbox scale as a
405 sensitivity study (see section 5.1).

406 Figure 2 displays a sample comparison between simulator output from CAM3.1
407 day-2 forecasts and the observations. It shows the east-to-west distribution of clouds in
408 the tropics formed as an average over tropical latitudes for June through September 2006.
409 CAM3.1 is able to capture some aspects of clouds related to the large-scale circulation
410 such as the abundance of clouds in the Asian Monsoon (70E through the dateline) and the
411 predominance of low clouds to the west of South America (200˚E to 280˚E) and Africa
412 (320˚E to 360˚E). However, it is clear that the model has too frequent high clouds
413 particularly in the Asian Monsoon region. One interesting detail is that many CAM3.1
414 clouds have cloud water contents too small to be detected by either the radar or the lidar
415 (This is the so-called "empty cloud" problem where cloud fraction is non-zero but cloud
416 condensate is zero. These "empty clouds" are not included in the figure.). However,
417 because these figures display averages over large temporal and spatial scales, they cannot
418 indicate the exact disparities in cloud types between simulated and observed cloud
419 systems. More detailed comparison is required to investigate whether the model can
420 simulate specific clouds with the correct frequency in the right location. This motivates
421 the following analysis of cloud regimes.
422
423 **4. Clustering of tropical CloudSat and CALIPSO data**
424
425 The results of applying the clustering method to CloudSat and CALIPSO
426 observations are shown in Figures 3 and 4. These figures depict the cluster centroids in
427 terms of the joint histogram of atmospheric pressure and signal strength (Figure 3) and
428 the occurrence frequency maps of different cloud regimes (Figure 4). The different
429 locations of maximum RFO for different cloud regimes is indicative of the association of
430 cloud regimes with specific characteristics of the large-scale atmospheric circulation and
431 thermodynamic states (Del Genio and Kovari, 2002; Rossow et al. 2005). Table 1

displays the tropical average relative frequency of occurrence and total cloud cover for each cloud regime.

Six cloud regimes are able to describe the variations of tropical cloud systems. Cloud regimes are given names based on the qualitative assessment of the joint histograms of atmospheric pressure and signal strength for each cluster (Figure 3). The first regime with an occurrence frequency of 35% (Table 1) is the most common cluster of the six, and is named as low cloud with less precipitation. Most of the clouds are detected by the lidar, and only a small fraction of clouds is detected by the radar. The second regime is named low cloud with precipitation due to the greater fraction of dBZ values in excess of -15, which is an approximate threshold that distinguishes cloud from drizzle and rain (Frisch et al., 1995; Stephens and Wood, 2007). These two regimes are found with concentrations in the large subsidence regions of the tropical oceans. The first regime has the highest RFO at the west coasts of continents where marine stratocumulus clouds are known to be prevalent (Klein and Hartmann 1993). The second regime happens more frequently in regions where trade cumulus are predominant. Over higher ocean temperatures than the first regime, the low clouds and precipitation extend deeper with increased clouds and precipitation occurring in the 680 to 800 hPa bin. The third regime is named thin cirrus and is characterized by clouds at high levels with low dBZ and sometimes only detectable by the lidar. This regime is most common in the Caribbean, the African Monsoon and the Asian Monsoon regions of India and South Asia. The fourth regime consists of clouds and precipitation over a wide range of dBZ below 440 hPa. This regime is suggestive of isolated convection that reaches the middle troposphere and will be named cumulus congestus. It often occurs as an important regime in the transition from shallow cumulus to deep convection. This regime is most common

over the northwestern Pacific on the eastern edge of the Asian monsoon and with lesser frequency over the Inter-Tropical Convergence Zones of the Atlantic and Pacific oceans and the African and Asian monsoons. It also has a high RFO over the high topography of the west coast of South America, east central Africa, and South Asia. The fifth regime is named cirrus anvils and has a higher RFO at larger dBZ and occurs over a wider range of pressure as compared to the thin cirrus regime. This cloud type is generally produced by outflow from deep cumulus or synoptic and mesoscale disturbances (Sassen and Mace, 2002; Mace et al. 2006) and preferentially occurs over land areas in the monsoons of Asia, Africa, and Central America. The sixth and last regime is named deep convection with heavy precipitation. It occurs most frequently in the west Pacific warm pool and the Asian Monsoon region (Zipser et al., 2006; Liu and Zipser, 2005).

By comparing these cloud regimes to those determined from an analysis of only CloudSat data (Figure 1 in Zhang07), the value of combining the radar and lidar data is readily apparent. First, the increase relative to Zhang07 of cloud RFO in the highest pressure level for most regimes illustrates the capability of the lidar to sense tenuous cirrus whose radar reflectivity is less than the radar detection threshold. Second, a large portion of non-drizzling cumulus or stratocumulus are detected only by lidar as indicated prominently by the two low-cloud regimes. Third, the lidar is capable of detecting thin mid-level liquid water clouds particularly in the thin cirrus, congestus, and cirrus anvil regimes. As a result, the occurrence of clear sky decreases from 30% in Zhang07 to 8% in this study (Table 1). Note that clear sky is defined as when fewer than 5% of adjacent 200 profiles of satellite data have cloud or precipitation; obviously this number is dependent on the number of profiles in the samples.

5. Evaluation of CAM simulations

5.1 CAM3.1

Model data can be either clustered independently or assigned into the observational cluster with the minimum Euclidian distance between the modeled and observed normalized dBZ index. However, if model data are clustered independently, a different number of clusters may result (Williams and Tselioudis, 2007). In this case, both the joint histograms and the geographic distributions may differ substantially from the observations leading to an ambiguous evaluation of model deficiencies. To reduce complexity, model simulations are assigned to the cluster centroids determined from observations, and the joint histograms formed by averaging the modeled elements in each cluster are shown in Figure 5 with their corresponding RFO geographic distributions in Figure 6. Tropical averages of the RFO and total cloud cover for each model regime are reported in Table 1. Projecting model simulations onto the observed clusters allows one to compare a common set of regimes.

Results indicate that the two modeled low level cloud regimes have much less hydrometeor fraction than observed in their joint histograms of atmospheric pressure and signal strength. In particular, the model strongly underestimates low level clouds that are detectable only by the lidar. In contrast, the low cloud with less precipitation regime has more precipitating cloud than observed, and the intensity of drizzle for the two low cloud regimes is too high compared with observations, similar to results reported recently elsewhere for other models (Bodas-Salcedo et al., 2008; O'Connor et al., 2009). At the same time, the modeled RFO of the low cloud with less precipitation regime is more frequent than observed in the oceanic subsidence regions, but too infrequent in ascent regions. Compared with observations, the oceanic peaks of modeled RFO of the low cloud with precipitation regime are shifted westward. For both the thin cirrus and cirrus

506 anvil regimes, the model has a reasonable vertical profile of cloud fraction in the upper
507 troposphere, and low clouds overlapped by high clouds are simulated well in the model.
508 However, the simulated RFOs of the two cirrus regimes are much lower than those
509 observed over the Americas, the central Pacific Ocean, and the Asian monsoon region.
510 This may be because cirrus clouds cooccur with deep convection too often in the model.
511 While the model has a reasonable occurrence frequency of cumulus congestus except for
512 an underestimate over the tropical western Pacific, the model overestimates the
513 occurrence of radar reflectivity above 10 dBZ suggesting that the simulated mid-level
514 clouds precipitate too heavily. For the deep convection with heavy precipitation regime,
515 the model simulates an occurrence frequency of 33%, more than 2 times the observed
516 occurrence frequency of 13%. At the same time, the hydrometeor coverage is lower than
517 observed at levels beneath 440 hPa, particularly for the range from -10 to 20 dBZ.
518 Regardless of the regime, a very prominent problem evident from the joint histograms is
519 that the model strongly underestimates the occurrence of clouds with reflectivity less than
520 -10, particularly clouds which are only detectable by the lidar. A separate comparison
521 between the cloud optical thickness from model data and those derived from Moderate-
522 Resolution Imaging Spectroradiometer (MODIS; Salomonson and Toll, 1991)
523 measurements also illustrates the modeled clouds are too optically thick (not shown).
524 These results indicate that the model clouds are too reflective, both at the frequency of
525 the CloudSat radar but probably also at visible wavelengths (Zhang et al., 2005).

526 To investigate the effect of homogenous distribution of cloud condensate and
527 precipitation used in the simulator package, we create another set of the joint histograms
528 for the observed cloud regimes by replacing the radar reflectivity at each level by the
529 grid-box (200 profiles) mean reflectivity and then calculate the joint histograms from the

means using the cluster number determined by the original joint histograms without grid-box averaging (not shown). The comparison between this recalculated set and the simulations supports the conclusion that the model clouds are still too reflective for the two low clouds and two cirrus regimes, and that the intensity of modeled precipitation for the low clouds with precipitation and congestus regimes are too high. Note that this test probably overestimates the impact of the homogeneity assumption, because the averaged histograms mix the cloud and precipitation together, while the model has a separate representation of cloud and precipitation.

In order to explore the relationship of model parameterizations to the discrepancies between models and observations, the simulator package is run for convective and stratiform components of cloud systems separately and the resulting cloud patterns (not shown) are constructed using the assigned cloud regimes determined from the simulator output created from the complete cloud systems. For the low clouds with less precipitation regime, most model clouds are stratiform, while those for the low clouds with precipitation regime are both convective and stratiform but the mean convective dBZ is larger than the stratiform dBZ which unsurprisingly indicates stronger precipitation. The high clouds overlapped with low clouds are generated from stratiform component. For the thin cirrus and anvil cloud regimes, model clouds are predominantly stratiform, while the clouds of the cumulus congestus regime are characterized by intense convective systems. The cloud coverage of the deep convection with heavy precipitation regime results comparably from convective and stratiform systems. Unsurprisingly the dBZ of the convective clouds and precipitation are greater than that of the stratiform clouds and precipitation, and the modeled stratiform precipitation is less frequent beneath 800 hPa than above which is suggestive of precipitation evaporation in the lower

554 troposphere. Considering that the model's precipitation area is too low and that the
555 model's RFO is far greater than observed for this regime, it suggests that the model
556 produces too much convective precipitation but too little stratiform precipitation. This
557 result would be consistent with that of Dai (2006) who found that this model (as well as
558 most conventional climate models) underestimate/overestimate the accumulated
559 stratiform/convective precipitation in the tropics based on TRMM observations.

560 ***5.2 CAM3.5***

561 Applying the same analysis approach to cloud simulations from CAM3.5 yields
562 joint histograms for the six regimes (Figure 7) that are similar to those from CAM3.1.
563 The most noticeable changes are that the cloud fraction at the highest pressure level is
564 lower than that from CAM3.1, and that the hydrometeor fractions at low levels increase,
565 particularly in the deep convection regime. However, the differences in the RFO and
566 spatial distributions are more significant (Figure 8 and Table 1). For example, the
567 occurrence frequency of deep convection and thin cirrus regimes decreases. In particular,
568 although the RFO of deep convection regime remains a factor of two too large, the
569 decrease in the total RFO of the three high cloud regimes (thin cirrus, cirrus anvil and
570 deep convection) from 39% to 31% corrects an overestimate of the observed occurrence
571 frequency of 30%. Additionally, the occurrence frequency of congestus increases from
572 6% to 13% leading to an overestimate of the observed occurrence frequency which is 9%.
573 At the same time, the occurrence frequency of low clouds with precipitation increases in
574 many oceanic regions. Also worthy of mention is that the total cloud coverage of all
575 regimes decreases when compared with that of CAM3.1

576 It is tempting to attribute most of the changes in the regime occurrence frequencies
577 to the elimination of undilute plumes in the deep convection parameterization of

578 CAM3.5. Indeed, this is confirmed by examination of a separate integration of CAM3.5
579 modified to permit undilute plumes according to the formulation that was used in
580 CAM3.1 (Table 1). Physically, dilute plumes have detrainment levels in the middle and
581 lower troposphere and the inclusion of the dilute plumes likely explains the increase in
582 the occurrence frequency of low clouds with precipitation and congestus, and the
583 decrease in the occurrence frequency of deep convection. Indeed, in the simulation of
584 CAM3.5 with undilute plumes, the occurrence frequency of low clouds with precipitation
585 decreases from 21% to 19% and the occurrence frequency of congestus decreases from
586 13% to 6%, confirming that the change in the dilution of convective plumes is
587 responsible for most of the increase of these regimes from CAM3.1 to CAM3.5. The
588 reduction in the occurrence frequency and total cloud coverage of thin cirrus that results
589 from dilute plumes (Table 1) may be the result of decreased condensate and water vapor
590 detrainment from deep convection in the upper troposphere. This interpretation is
591 consistent with the strong decrease in the occurrence frequency of the deep convection
592 with heavy precipitation regime in CAM3.5 (Table 1).

593 ***5.3 The association of cloud regimes with large-scale dynamics***

594 To explore the coupling between cloud regimes and the large-scale dynamics that is
595 supportive of different cloud types, the occurrence frequency of cloud regimes from both
596 observations and model simulations over ocean are sorted by the value of monthly mean
597 vertical pressure velocity at 500 hPa (ω_{500}). Although cloud systems may be associated
598 with other large-scale parameters, such as sea surface temperature or lower tropospheric
599 stability (Klein and Hartmann, 1993; Weaver, 1999; Williams et al., 2003; Ringer and
600 Allan, 2004), we choose to examine ω_{500} because of its recent widespread use in the
601 analysis of tropical clouds following the pioneering approach of Bony et al. (2004).

602 NCEP vertical velocities are sorted into 8 bins such that the occurrence frequency of each
603 bin is equal. The compositing of observed cloud regimes into vertical velocity bins is
604 performed in two ways (Figure 9). In the first way, the fraction of elements of a given
605 regime which occur in a given vertical velocity bin is displayed in Figure 9a. If there
606 were no relationship between a cloud regime and ω_{500}, the occurrence frequency of a
607 regime in each velocity bin would be equal to 0.125 apart from random fluctuations. In
608 the second way, the fraction of elements in a given vertical velocity bin which belong to a
609 given regime is displayed in Figure 9b. In this way, the sum of the frequencies for the six
610 regimes in each velocity bin is 1. As expected, the two low cloud regimes are much more
611 common in subsidence regions and the remaining regimes are more common in ascent
612 regions. The association of cloud regimes with large-scale dynamics provides
613 quantitative targets for model simulations.

614 Figure 10 displays differences between observations and CAM3.1 and CAM3.5
615 simulations. Compared with observations (Figure 10a) CAM3.1 strongly overestimates
616 the occurrence frequency of deep convection with heavy precipitation in the three
617 dynamic regimes with the strongest upward motion. For example, in the strongest upward
618 motion bin, CAM3.1 simulates an occurrence frequency of 0.86 whereas the observed
619 occurrence frequency is only 0.35. As result, in the strongly ascending regimes the model
620 underestimates the occurrence frequency of all other cloud regimes. In the regimes with
621 moderate descending or ascending air motion, the model produces too many low clouds
622 with precipitation but too few low clouds with less precipitation.

623 The impact of the model changes between CAM3.1 and CAM3.5 on the frequency of
624 cloud regimes in different dynamical regimes is displayed in Figure 10c. The occurrence
625 frequency of congestus in each dynamic regime rises with stronger increases in the

626 ascending regimes, while the occurrence frequency of thin cirrus slightly decreases in all
627 regimes. At the same time, the occurrence frequency of deep convection with heavy
628 precipitation decreases in the bins with upward motion although the reduction does not
629 cancel the model overestimate especially in the strongest upward motion bin. In many
630 regimes, the occurrence frequency of low clouds with precipitation increases. As a result
631 of these changes, the simulation of cloud regimes from CAM3.5 compares somewhat less
632 favorably to observations in their occurrence frequencies (Figure 10b). Although the
633 overestimate of deep convection is reduced in ascending regimes, it is replaced with
634 overestimates of congestus and low clouds with precipitation in weakly ascending and
635 descending regimes. In general this suggests that the occurrence of precipitating
636 convection remains distressingly high. A possible concern is that the observed occurrence
637 frequency of the deep convection regime may be underestimated because CloudSat only
638 collects a curtain of data instead of a wide area along the satellite flight track. While this
639 may partly contribute to the remarkable difference between the modeled and observed
640 RFO of the deep convection with heavy precipitation regime, the comparison of CAM3
641 with ground-based as well as satellite observations in some previous studies (Xie et al.
642 2004, Dai and Trenberth 2004) also suggests that the model greatly overestimates
643 precipitation frequency.

644 ***5.4 Comparison between CAM3's forecasts and its climate***

645 In order to examine the consistency between cloud regimes of CAM3's climate and
646 its forecasts, data from 'climate' integrations of CAM3 using only observed sea-surface
647 temperatures and sea ice for June-September 2006 are analyzed following the method
648 used for the forecasts. The joint histograms of atmospheric pressure and signal strength of
649 the six regimes from the climate integrations are similar to those of the forecasts, but the

RFO of the individual cloud regimes have several noticeable differences. In the climate integrations of both CAM3.1 and CAM3.5, the low clouds with precipitation regime is more common in the subsidence regions, and the congestus regime occurs more frequently in ascending regimes. In contrast, the deep convection with heavy precipitation regime is less frequent in the climate integrations than that from the forecasts. These differences show that drifts in the large-scale atmospheric state lead to drifts in the population of cloud regimes.

To show the relationship between cloud regimes and the 500 hPa pressure vertical velocity, frequency differences similar to Figure 10 are created for the climate integrations (Figure 11). Although the differences with observations are not the same, the differences between CAM3.1 and CAM3.5 for almost all cloud regimes in climate integrations are similar to those of the forecasts but with much smaller magnitude. For example, low clouds with precipitation increase at the expense of low clouds with less precipitation, and the congestus clouds occur more frequently in ascending regions. However, the greater similarity of biases with observations between model versions indicates a compensation between the atmospheric state and the parameterization changes in the model's climate simulation.

6. Summary

This paper uses tropical measurements of cloud fields from CloudSat and CALIPSO to evaluate simulated cloud and precipitation statistics from the CAM3. Although several prior studies assess model performance using CloudSat or CALIPSO data (Bodas-Salcedo et al. 2008; Chepfer et al. 2008; Marchand et al. 2009), this is the first study to assess a model using both data streams, which is beneficial due to the different sensitivities of the radar and lidar for hydrometeor detection (Mace et al. 2009). The

merged CloudSat and CALIPSO dataset provides the most comprehensive description of the vertical structure of hydrometeor fields currently possible on a global basis. It has the potential to advance our understanding of cloud processes and improve model evaluations. Observations are analyzed in terms of cloud regimes using a clustering technique applied to tropical data for the period June to September 2006, so that model simulations can be evaluated as a function of characteristic cloud type. Six cloud regimes with distinctive cloud subgrid-scale patterns to the vertical profiles of signal strength are identified, and the geographical distributions of the occurrence frequencies of these principal cloud regimes illustrate the association with the large-scale atmospheric circulation.

A satellite simulator package is applied to the model to aid quantitative evaluation of model performance using the new data. The joint histograms of atmospheric pressure and signal strength generated by the simulator package are used to assess model performance under the clustering framework. Assigning model histograms to the observed cloud regimes facilitates comparison in terms of both the occurrence frequency and properties of cloud regimes.

The comparison of the geographical distributions between model simulations and observations shows that CAM3.1 overestimates the area coverage of high clouds especially in the Tropical Western Pacific, east central Africa, and northern South America, and underestimates the area coverage of low clouds in subsidence regions. More insightful are the differences in the joint histograms of atmospheric pressure and signal strength that are able to expose model deficiencies in the simulated vertical structure of hydrometeor properties. It is found that cloud coverage of the two low cloud regimes and congestus regimes are significantly lower than observed. Low- and mid-level

clouds may precipitate too heavily. The biases in the joint histogram and occurrence frequency for the deep convection with heavy precipitation regime suggest that the model overestimates convective precipitation but underestimates stratiform precipitation. Particularly striking is the model overestimate of the occurrence frequency of deep convection and the complete absence of cirrus anvils. In general, the modeled clouds are too reflective in all regimes, which is consistent with that seen by Bodas-Salcedo et al. (2008) who used CloudSat data to evaluate clouds and precipitation in the Unified Model of the United Kingdom Meteorological Office. Also, it is particularly prominent in the fact that the model is unable to simulate hardly any clouds with radar reflectivity less than -30 but still detectable by CALIPSO, and a similar result was found with the French climate model (Chepfer et al. 2008).

It has been reported that the CALIPSO Vertical Feature Mask version 2 used to create the GEOPROF-LIDAR product used in this study has some error (D. Winker, personal communication). The error is expected to cause an overestimate in the occurrence of isolated low-level clouds on the order of 5-10% in the maritime trade cumulus regions and have little effect elsewhere. This error will not explain the lack of model clouds in the lidar-only bin of the two low cloud regimes. Thus while a future version of the data may change the cloud coverage and/or RFO of the two low cloud regimes, there will be less impact on the other four cloud regimes. For these other regimes, the occurrence frequency at the lowest level for the lidar-only bin may decrease in Figure 3, but the RFO of the regimes will likely not change.

To examine the impact of model parameterizations on the simulated clouds, we also evaluate CAM 3.5. The cloud subgrid-scale patterns of CAM3.5 are similar to those from CAM 3.1, but the geographical distributions of the RFO are significantly different. The

724 new version of the model reduces deep convection and high clouds but increases
725 congestus and low clouds with precipitation. These changes are primarily due to
726 implementation of dilute plumes in the deep convection parameterization which leads to
727 greater detrainment in the middle troposphere and less detrainment in the upper
728 troposphere.

729 The cloud regimes are also sorted by the monthly mean vertical wind at 500 hPa to
730 show the relationship between tropical cloud systems and the large-scale environment
731 that influences the evolution of cloud systems. It is shown that, relative to CAM3.1,
732 CAM3.5 suppresses deep convection with heavy precipitation and generates more
733 congestus in ascending regions and low clouds with precipitation in subsidence regions.
734 However, deep convection is still too frequent in strongly ascending regions, and low
735 clouds with less precipitation are still too infrequent.

736 Although results from climate integrations of CAM show different geographical
737 distributions of the occurrence frequencies for the individual cloud regimes relative to
738 those of the forecasts, the changes from CAM3.1 to CAM3.5 are identical but smaller in
739 magnitude for all the regimes. The differences of simulated cloud statistics between
740 forecasts and climate integrations may imply that the feedback processes are partly
741 responsible for the climatological biases. However, the details about how the feedback
742 processes generate these differences will need more sensitivity studies. Our result is not
743 consistent with that in Williams and Brooks (2007), which found the cloud regimes are
744 similar for the forecasts and the climate integrations with the Met Office Unified
745 Forecast-Climate Model. The lack of differences in their case for cloud regimes between
746 forecasts and climate integrations may be partly due to the fact that the analysis used to
747 initialize their climate model is from a data-assimilation system with the same physical

748 model. Furthermore, larger differences between clouds in the climate and forecast
749 integration of CAM3 may occur because differences between the tropical state of
750 CAM3's climate and the NCEP analysis are larger. Additional investigation of the spin-
751 up of model clouds and precipitation in CAM3 is warranted.

752 Some of the conclusions from this study echo those of previous studies such as the
753 overabundance of deep convection (Xie et al. 2004), the near absence of anvil cirrus
754 (Williams et al. 2005), the overestimate of convective precipitation (Dai 2006), and
755 overly reflective clouds (Zhang et al. 2005; Bodas-Salcedo et al. 2008). However, some
756 new perspectives are provided, including an underestimation of thin clouds that can only
757 detected by the lidar and an overestimation of precipitation frequency from CAM3.5. The
758 fact that the CPR can see precipitation allows one to diagnose errors in model-simulated
759 precipitation statistics together with cloud errors. One surprising result is that the CAM
760 has as much or greater amounts of congestus as observations. This contrasts with all
761 previous studies using ISCCP data which had concluded that large-scale models lack
762 congestus. A possible reconciliation of our results with the previous studies is that we
763 primarily use precipitation profiles in this study to detect congestus whereas the other
764 studies using ISCCP data rely on identification of congestus through the visible and
765 infrared cloud properties. Although the results from the CAM may not apply to other
766 climate models, it may be that models do produce congestus (middle level topped
767 precipitating convection) but that the cloud properties of the congestus regime are
768 seriously biased. Indeed, a preliminary comparison of ISCCP simulator results when the
769 CAM simulates congestus clouds (as identified by CloudSat) suggests that the model
770 cloud properties for the congestus regime are indeed biased when compared to MODIS
771 observations of visible optical thickness and highest cloud top pressure.

772 This paper provides a possible methodology to use the merged dataset from the radar
773 and lidar observations to evaluate model performance. In the future, we will exploit the
774 synergy of the A-Train to deliver complementary measurements of the same
775 environmental phenomena and the collocated large-scale variables along the CloudSat
776 flight track, to further understand model deficiencies. For example, Clouds and the
777 Earth's Radiant Energy System (CERES; Wielicki et al., 1996) radiative fluxes will be
778 used to describe the radiative characteristics of the individual regimes and address the
779 impact of the cloud regimes on the cloud radiative forcing at the top of atmosphere. With
780 the rapid evolution of the physical parameterizations in CAM, our future evaluation
781 efforts will focus on the next officially released version, CAM4. In order to explore the
782 physical reasons for the differences between model and observation, we also plan to
783 perform more sensitivity experiments on specific aspects of the cloud parameterizations
784 to identify future model improvements.
785
786
787 **Acknowledgements:**
788 Yuying Zhang was funded through a grant from the NASA Modeling and Analysis
789 and Prediction Program, Don Anderson, program manager. Support for the work of Jim
790 Boyle and Stephen Klein at LLNL was provided by the Atmospheric Radiation
791 Measurement and Climate Change Prediction Programs which are directed from the
792 Office of Science at the U.S. Department of Energy. This research was performed under
793 the auspices of the U.S. Department of Energy by Lawrence Livermore National
794 Laboratory under contract DE-AC52-07NA27344. Support for this work (GM) at the
795 University of Utah was provided by NASA through a contract issued by the Jet
796 Propulsion Laboratory, California Institute of Technology under a contract with NASA
797 and NASA Grant NNX07AT45G. We greatly appreciate the DIME website from which
798 we obtained the cluster program. We also thank Richard Neale for assistance in
799 constructing the simulation of CAM3.5 with undilute plumes. The comments of the three
800 anonymous reviewers are appreciated.
801

802 **Reference:**
803
804 Anderberg, M. R. (1973), Cluster Analysis for Applications, 359 pp., Academic, New
805 York.
806
807 Bodas-Salcedo A., M. J. Webb, M. E. Brooks, M. A. Ringer, K. D. William, S. F. Milton,
808 D. R. Wilson, 2008: Evaluating cloud systems in the Met Office global forecast
809 model using simulated CloudSat radar reflectivities. J. Geophys. Res., 113, D00A13,
810 doi:10.1029/2007JD009620.
811
812 Bony, S., et al. (2004), On dynamic and thermodynamic components of cloud changes,
813 Clim. Dyn., 22, 71– 86.
814
815 Bony, S., R. Colman, V. M., Kattsov, R. P. Allan, C. S. Bretherton, J-L. Dufresne, A.
816 Hall, S. Hallegatte,M. M. Holland, W. Ingram, D. A. Randall, B. J. Soden, G.
817 Tselioudis, and M. J. Webb, 2006: How well do we understand and evaluate climate
818 change feedback processes? J. Climate, 19, 3445-2461.
819
820 Boccippio, D. J., W. A. Petersen, and D. J. Cecil, 2005: The tropical convective
821 spectrum. Part I: Archetypal vertical structures. *J. Clim.*, **18**, 2744-2769.
822
823 Boyle, J., S. Klein, G. Zhang, S. Xie, and X. Wei, 2008: Climate model forecast
824 experiments for toga coare, Mon. Wea. Rev, 136, 808832.
825
826 Cess, R. D., and co-authors, 1990: Intercomparison and interpretation of climate feedback
827 processes in 19 atmospheric general circulation models. J. Geophys. Res., 95, D10,
828 16,601-16,615.
829
830 Chen, Y., and A. D. Del Genio, 2008: Evaluation of tropical cloud regimes in
831 observations and a general circulation model. Clim. Dyn., 32, 355-369.
832
833 Chevallier, F., and P. Bauer, 2003: Model rain and clouds over oceans: Comparison with
834 SSM/I observations. Mon. Wea. Rev., 131, 1240-1255.
835
836 Chiriaco, M., et al., 2006: The ability of MM5 to simulate ice clouds: systematic
837 comparison between simulated and measured fluxes and Lidar/Radar profiles at the
838 SIRTA atmospheric observatory. Mon Weather Rev, 134, 897-918.
839
840 Chepfer, H., M. Chiriaco, R. Vautard, and J. Spinhirne, 2007: Evaluation of MM5
841 Optically Thin Clouds over Europe in Fall Using ICESat Lidar Spaceborne
842 Observations. Mon Weather Rev, 135, 2737-2753.
843
844 Chepfer, H., S. Bony, D. Winker, M. Chiriaco, J.-L. Dufresne, G. Sèze, 2008: Use of
845 CALIPSO lidar observations to evaluate the cloudiness simulated by a climate model.
846 Geophys. Res. Lett., 35, L15704, doi:10.1029/2008GL034207.
847
848 Collins, W. D., et al., 2006: The formulation and atmospheric simulation of the
849 Community Atmospheric Model: CAM3. J. Clim., 19, 2144-2161.

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, Aiguo, 2006: Precipitation characteristics in eighteen coupled climate models. J. Clim., 19, 4605-4630.

Del Genio, A. D., and W. Kovari, 2002: Climatic properties of Tropical precipitating convection under varying environmental conditions. J. Climate, 15, 2597-2615.

Frisch, A. S., C. W. Fairall, and J. B. Snider, 1995: Measurement of stratus cloud and drizzle parameters in ASTEX with Ka-band Doppler radar and microwave radiometer. J. Atmos. Sci., 52, 2788-2799.

Haynes, J. M., G. L. Stephens, 2007: Tropical oceanic cloudiness and the incidence of precipitation: Early results from Cloudsat. Geophys. Res. Lett., L09811 DOI 10.1029/2007GL029335.

Gordon, N.D., et al., 2005: Cluster analysis of cloud regimes and characteristic dynamics of midlatitude synoptic systems in observations and a model. *J. Geophys. Res.*, 110, D15S17.

Houghton, J.T., and et al., 2001: Climate change: the scientific basis. Contribution of working group I to the third assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge, pp 525-582.

Im, E., S. L. Durden, and C. Wu, 2006: Cloud Profiling Radar for the CloudSat mission. IEEE Aerosp. Electron. Syst. Mag., 20, pp. 15-18.

Jakob, C., and G. Tselioudis 2003, Objective identification of cloud regimes in the Tropical Western Pacific, Geophys. Res. Lett., 30(21), 2082, doi: 10.1029/2003GL018367.

Jakob, C., and C. Schumacher, 2008: Precipitation and latent heating characteristics of the major tropical western pacific cloud regimes. J. Clim., 21, 4348-4364.

Jensen, M. P., and A. D. Del Genio, 2006: Factors limiting convective Cloud-Top height at the ARM Nauru Island climate research facility. J. Climate, 19, 2105-2117.

Klein, S. A., and D. L. Hartmann, 1993: The seasonal cycle of low stratiform clouds. J. Climate, 6, 1587-1606.

Klein, S. A., X. Jiang, J. Boyle, S. Malyshev, S. Xie, 2006: Diagnosis of the summertime warm and dry bias over the U.S. Southern Great Plains in the GFDL climate model using a weather forecasting approach, Geophys. Res. Lett., 33, L18805, doi:10.1029/2006GL027567.

897 Klein, S. A, and C. Jakob, 1999: Validation and sensitivities of frontal clouds simulated
898 by the ECMWF model. Mon Weather Rev, 127(10):2514-2531.
899
900 Khairoutdinov, M. F., and D. A. Randall, 2003: Cloud resolving modeling of the ARM
901 Summer 1997 IOP: Model formulation, results, uncertainties, and sensitivities, J.
902 Atmos. Sci., 60, 607-625.
903
904 Liu, C. and E. J. Zipser, 2005: Global distribution of convection penetrating the tropical
905 tropopause. *J. Geophys. Res.*, **110**, D23104.
906
907 Liu, C. and E. J. Zipser, 2008: Diurnal cycles of precipitation, clouds, and lightning in the
908 tropics from 9 years of TRMM observations. Geophys. Res. Lett., 35, L04819,
909 doi:10.1029/2007GL032437.
910
911 Mace, G. G. (2004), Level 2 GEOPROF product process description and interface control
912 document, Coop. Inst. for Res. in the Atmos., Fort Collins, Colo.
913
914 Mace, G.G., M. Deng, B. Soden, E. Zipser, 2006: Association of tropical cirrus in the 10-
915 15-km layer with deep convective sources: an observational study combining
916 Millimeter Radar data and satellite-derived trajectories. J. Atmos. Sci., 63, 480-503.
917
918 Mace, G. G., R. Marchand, Q. Zhang, and G. Stephens (2007), Global hydrometeor
919 occurrence as observed by CloudSat: Initial observations from summer 2006,
920 Geophys. Res. Lett., 34, L09808, doi:10.1029/2006GL029017.
921
922 Mace, G. G., Q. Zhang, M. Vaughn, R. Marchand, G. Stephens, C. Trepte, D. Winker,
923 2009: A Description of Hydrometeor Layer Occurrence Statistics Derived from the
924 First Year of Merged Cloudsat and CALIPSO Data. J. Geophys. Res., 114, D00A26.
925
926 Marchand, R. T., Mace. G.G., and Ackerman T.P. 2008: "Hydrometeor Detection using
927 CloudSat - an Earth Orbiting 94 GHz Cloud Radar", J. of Atmos. Oceanic Technol.,
928 25, 519-533, doi:10.1175/2007JTECHA1006.1.
929
930 Marchand, R., J. Haynes, G. G. Mace, T. Ackerman, G. Stephens (2009), A comparison
931 of simulated cloud radar output from the multiscale modeling framework global
932 climate model with CloudSat cloud radar observations, J. Geophys. Res., 114,
933 D00A20, doi:10.1029/2008JD009790.
934
935 Neale, R. B., J. H. Richter, and M. Jochum (2008), The impact of convection on enso:
936 From a delayed oscillator to a series of events, Journal of Climate, preprint (2008),
937 0000–0000
938
939 Norris, J. R., and C. P. Weaver (2001), Improved techniques for evaluating GCM
940 cloudiness applied to the NCAR CCM3, J. Clim., 14, 2540–2550.
941
942 O'Connor, E. J., R. J. Hogan, A. J. Illingworth, and C. D. Westbrook, 2009: How do
943 model parameterizations of drizzle compare to radar and lidar observations? J.
944 Climate., submitted.

O'Dell, C. W., P. Bauer, R. Bennartz, 2007: A Fast Cloud Overlap Parameterization for Microwave Radiance Assimilation. J. Atmos. Sci., 64, 3896-3909.

Phillips, T. J., et al., 2004: Evaluating parameterizations in GCMs: Climate simulation meets weather prediction. Bull. Am. Meteorol. Soc., 85,1903-1915.

Pincus, Robert, R. Hemler, and S. Klein, 2006: Using stochastically generated subcolumns to represent cloud structure in a Large-scale model. Monthly Weather Rev., 134, 3644-3656.

Randall, D. A., and et al., 2007: Climate models and their evaluation. In Climate Chanfe 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H. L. Miller(eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Ringer, M., and R. Allan, 2004: Evaluating climate model simulations of tropical cloud. Tellus A 56:308-327.

Richter, J. H., and P. J. Rasch (2008), Effects of convective momentum transport on the atmospheric circulation in the community atmopshere model, version 3 (cam3), J. Climate.

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

Rossow, W. B., and R. A. Schiffer: Advances in understanding clouds from ISCCP. *Bull. Amer. Meteor. Soc.,* 80, 2261-2288, 1999.

Salomonson, V. V., and D. L. Toll, 1991: Execution phase (C/D) spectral band characteristics of the EOS Moderate Resolution Imaging Spectrometer-Nadir (MODIS-N) facility instrument. Adv. Space Res., 11, 231-236.

Sassen, K., and G. G. Mace, 2002: Cirrus. Ground-Based Remote Sensing of Cirrus Clouds. D. K. Lynch et al., Eds., Oxford University Press, 168-196.

Stephens, G. L., et al. (2002), The CloudSat mission and the A-Train, Bull. Am. Meteorol. Soc., 83, 1771–1790.

Stephens, G. L., and N. B. Wood, 2007: Properties of tropical convection observed by Millimeter-Wave Radar Systems. Mon. Wea. Rev., 135, 821-842.

Soden, B. J., A. J. Broccoli, R. S. Hemler, 2004: On the use of cloud forcing to estimate cloud feedback. J. Climate, 17(19): 3661-3665.

993 Soden, B. J., and I. M. Held, 2006: An assessment of climate feedbacks in coupled ocean
994 atmosphere models. J. Climate, 19, 3354-3360.
995
996 Wielicki, B. A., B. R. Barkstrom, E. F. Harrison, R. B. Lee III, G. L. Smith, and J. E.
997 Cooper, 1996: Clouds and the Earth's Radiant Energy System (CERES): An Earh
998 observing system experiment. Bull. Am. Meteorol. Soc., 77, 853-868.
999
1000 Webb, M., C. Senior, S. Bony, J. J. Morcrette, 2001: Combining ERBE and ISCCP data
1001 to assess clouds in the Hadley Centre, ECMWF and LMD atmospheric climate
1002 models. Clim Dyn, 17:905-922.
1003
1004 Williams, K. D., C. A. Senior, J. Mitchell, 2001: Transient climate change in the Hadley
1005 Centre models: the role of physical processes. J. Climate, 14, 12:2659-2674.
1006
1007 Williams, K. D., M. A. Ringer, C. A. Senior, 2003: Evaluating the cloud response to
1008 climate change and current climate variability. Clim. Dyn, 20, 705-721.
1009
1010 Williams, K. D., C.A. Senior, A. Slingo and J.F.B. Mitchell, 2005: Towards evaluating
1011 cloud response to climate change using clustering technique identification of cloud
1012 regimes. Clim. Dyn. 24, 701-719 doi:10.1007/s00382-004-0512-z.
1013
1014 Williams, K. D., and et al., 2006: Evaluation of a component of the cloud response to
1015 climate change in an intercomparison of climate models. Clim. Dyn., 26, 145-165.
1016 Doi 10.1007/s00382-005-0067-7.
1017
1018 Williams, K. D., and G. Tselioudis, 2007: GCM intercomparison of global cloud regimes:
1019 present-day evaluation and climate change response. Climate Dynamics, 29, 231-240.
1020
1021 Williams, K. D., and M. Webb, 2008**:** A quantitative performance assessment of cloud
1022 regimes in climate models Clim. Dyn. In press.
1023
1024 Winker, D. M., B. H. Hunt, and M. J. McGill (2007), "Initial performance assessment of
1025 CALIOP", Geophys. Res. Lett., 34, L19803, doi:10.1029/2007GL030135
1026
1027 Xie S., M. Zhang, J. S. Boyle, R. T. Cederwall, G. L. Potter, W. Lin, 2004: Impact of a
1028 revised convective triggering mechanism on Community Atmosphere Model, Version
1029 2, simulations: Results from short-range weather forecasts, J. Geophys. Res., 109,
1030 D14102, doi:10.1029/2004JD004692.
1031
1032 Xie, S. C., J. Boyle, S. Klein, X. Liu, and S. Ghan, 2008: Simulations of Arctic mixed-
1033 phase clouds in forecasts with CAM3 and AM2 for M-PACE. J. Geophys. Res., 113,
1034 D04211, doi: 10.1029/2007JD009225.
1035
1036 Zipser, E. J., et al. (2006), Where are the most intense thunderstorms on Earth?, Bull. Am.
1037 Meteorol. Soc., 87, 1057–1071.
1038

1039 Zhang, G. J., and N. A. McFarlane 1995:Sensitivity of climate simulations to the
1040 parameterization of cumulus convection in the Canadian Climate Center general
1041 circulation model, Atmos. Ocean, 33, 407-446.
1042
1043 Zhang, M., W. Lin, C. S. Bretherton, J. J. Hack, and P. J. Rasch, 2003: A modified
1044 formulation of fractional stratuform condensation rate in the NCAR Community
1045 Atmospheric Model (CAM2). J. Geophys. Res., 108, 4035,
1046 doi:10.1029/2002JD002523.
1047
1048 Zhang, M. H., and co-authors, (2005), Comparing Clouds And Their Seasonal Variations
1049 in 10 Atmospheric General Circulation Models With Satellite Measurements, *J.*
1050 *Geophys. Res., 110,* D15s03, doi:10.1029/2004JD005021.
1051
1052 Zhang, Y., and G. G. Mace, 2006: Retrieval of Cirrus Microphysical Properties with a
1053 Suite of Algorithms for Airborne and Spaceborne Lidar, Radar, and Radiometer Data.
1054 J. Appli. Meteorol. And Climatol., 45, 1665-1689.
1055
1056 Zhang Y., S. Klein, G. G. Mace, J. Boyle (2007), Cluster analysis of tropical clouds using
1057 CloudSat data, Geophys. Res. Lett., 34, L12813, doi:10.1029/2007GL029336.
1058
1059 Zhang, Y., S. A. Klein, C. Liu, B. Tian, R. T. Marchand, J. M. Haynes, R. B. McCoy, Y.
1060 Zhang, and T. P. Ackerman, 2008: On the diurnal cycle of deep convection, high-
1061 level cloud, and upper troposphere water vapor in the Multiscale Modeling
1062 Framework, J. Geophys. Res., 113, D16105, doi:10.1029/2008JD009905.

	low clouds with less precip		low clouds with precip		thin cirrus		congestus		cirrus anvils		deep convection with heavy precip		Clear
	RFO	TCC	RFO	TCC	RFO	TCC	RFO	TCC	RFO	TCC	RFO	TCC	RFO
Observation	35%	0.63	18%	0.69	9%	0.84	9%	0.84	8%	0.90	13%	0.93	8%
CAM 3.1	25%	0.46	18%	0.45	5%	0.77	6%	0.60	1%	0.88	33%	0.91	12%
CAM 3.5 undilute	23%	0.43	19%	0.42	5%	0.70	6%	0.57	1%	0.84	33%	0.90	13%
CAM 3.5	22%	0.36	21%	0.32	3%	0.58	13%	0.44	1%	0.74	27%	0.85	13%

Table 1. The data distributions for observations, simulations from CAM3.1, CAM3.5 with undilute plume, and CAM3.5 in the six cloud clusters and clear-sky condition with TCC lower than 5%. The data listed are the relative frequency of occurrence (RFO, left column), and the total cloud coverage (TCC, right column). The numbers of elements are 54,828 and 913,536 for observations and model simulations, respectively.

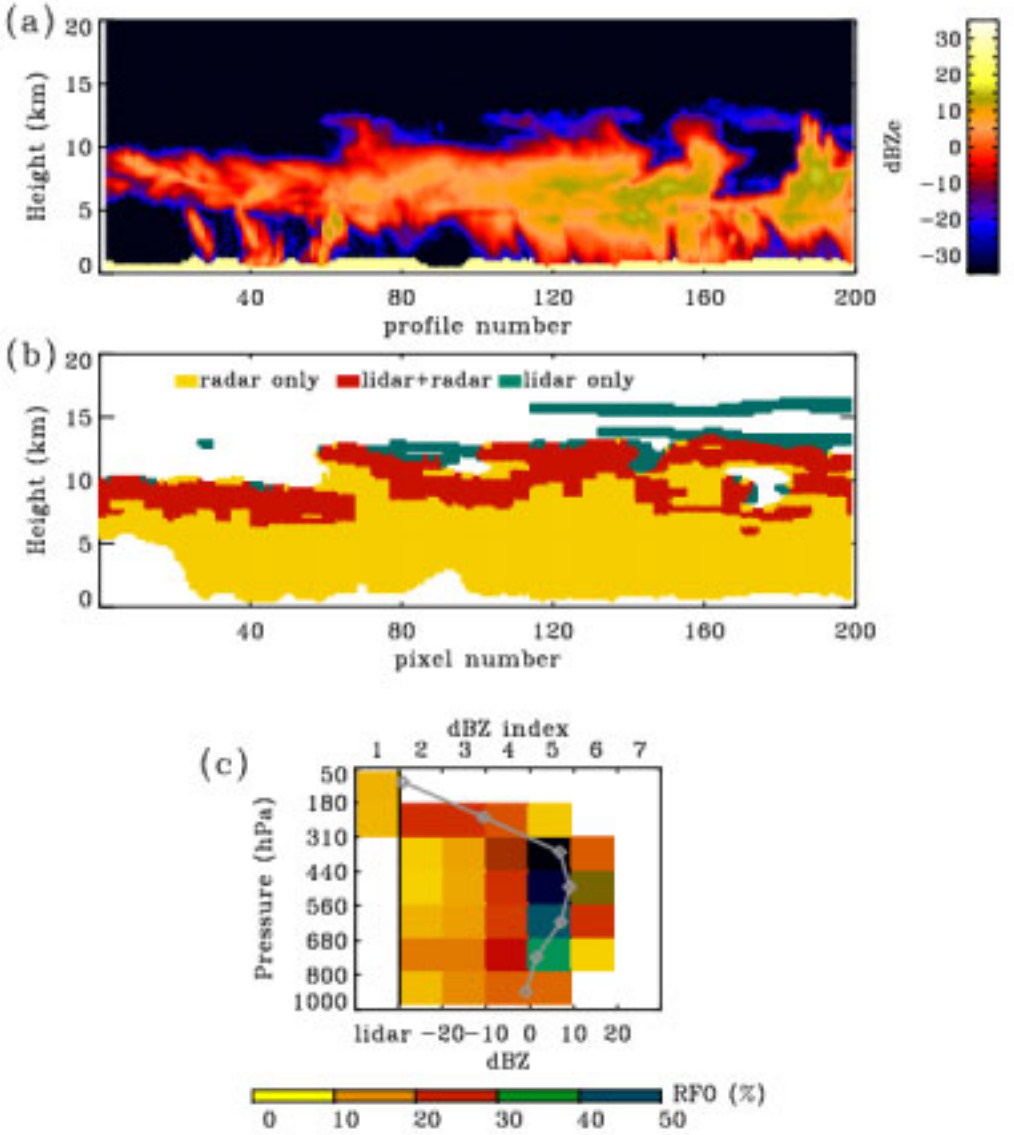

Figure 1. A case from Tropical Western Pacific Ocean (2°N 140°E) on Jul 15, 2006 that illustrates the creation of the joint histogram of atmospheric pressure and signal strength for a subgrid-scale cloud pattern. (a) Radar reflectivity from CloudSat observations of 200 adjacent profiles; (b) Hydrometeor mask by combining radar and lidar data; (c) The joint histogram of atmospheric pressure and signal strength for this sample. The shading indicates the relative frequency of occurrence (RFO) of clouds or precipitation at each bin of atmospheric pressure; The left column depicts the cloud fraction detected by lidar but missed by radar ('lidar only' clouds in panel b); The line with diamonds depicts the vertical profile of the normalized mean dBZ index for this histogram.

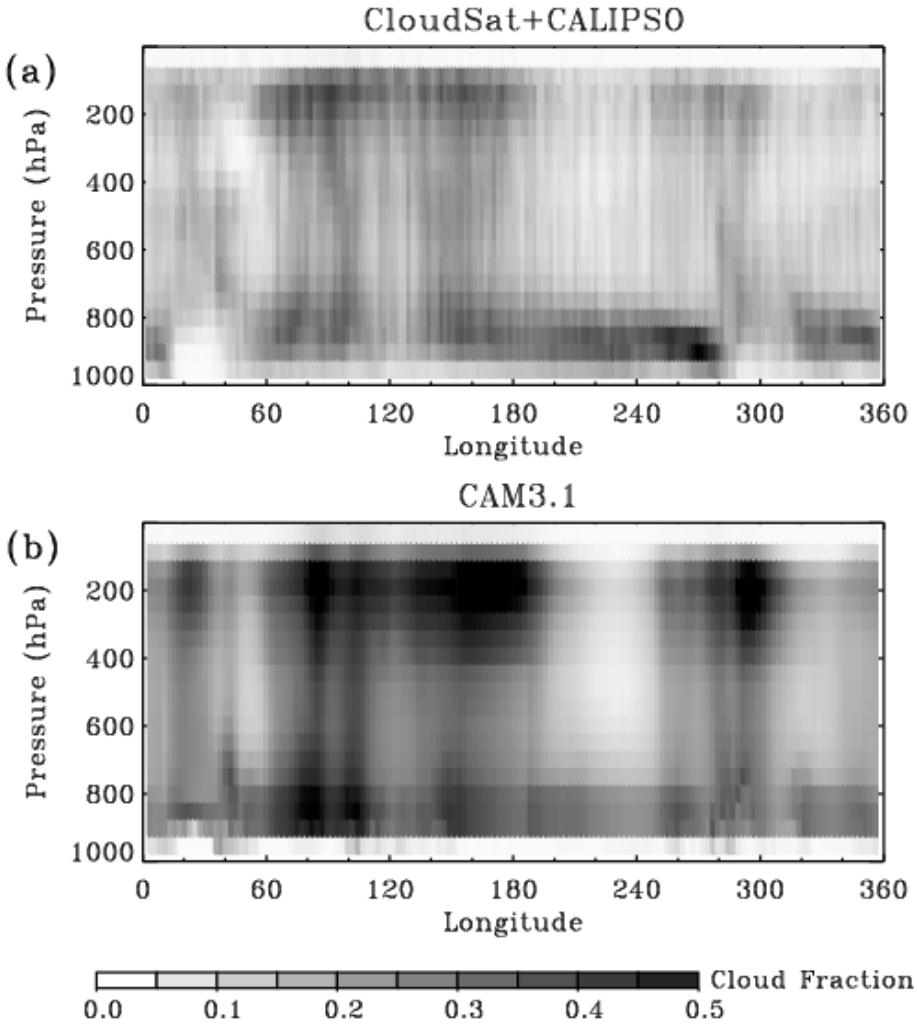

Figure 2. Comparison of the meridional-mean cloud occurrence frequency for the tropical region (23.5˚S-23.5˚N) during June-September 2006: (a) Observations from CloudSat and CALIPSO (b) Simulator output of the cloud simulations from CAM 3.1 day-2 forecasts.

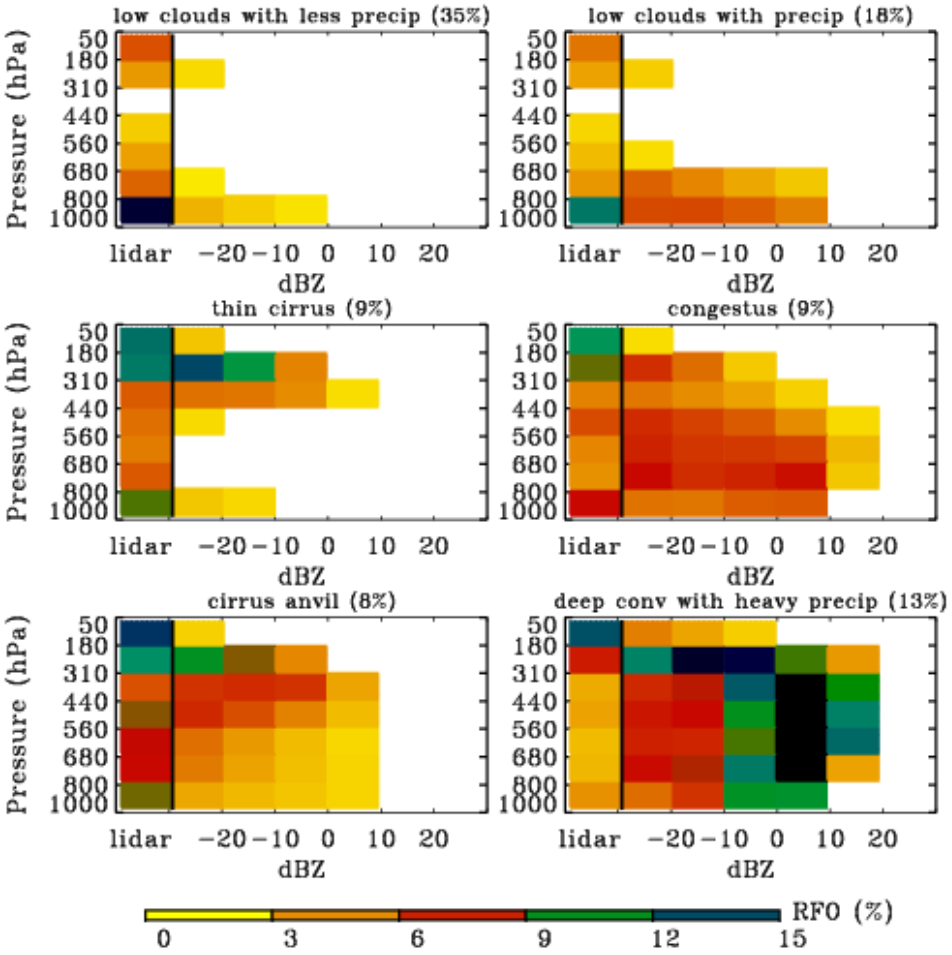

Figure 3. Joint histograms of atmospheric pressure and signal strength for the centroids of the six tropical clusters from the CloudSat and CALIPSO observations collected in June-September 2006. These clusters are named by the primary cloud morphology. The RFO for each cluster is shown in brackets.

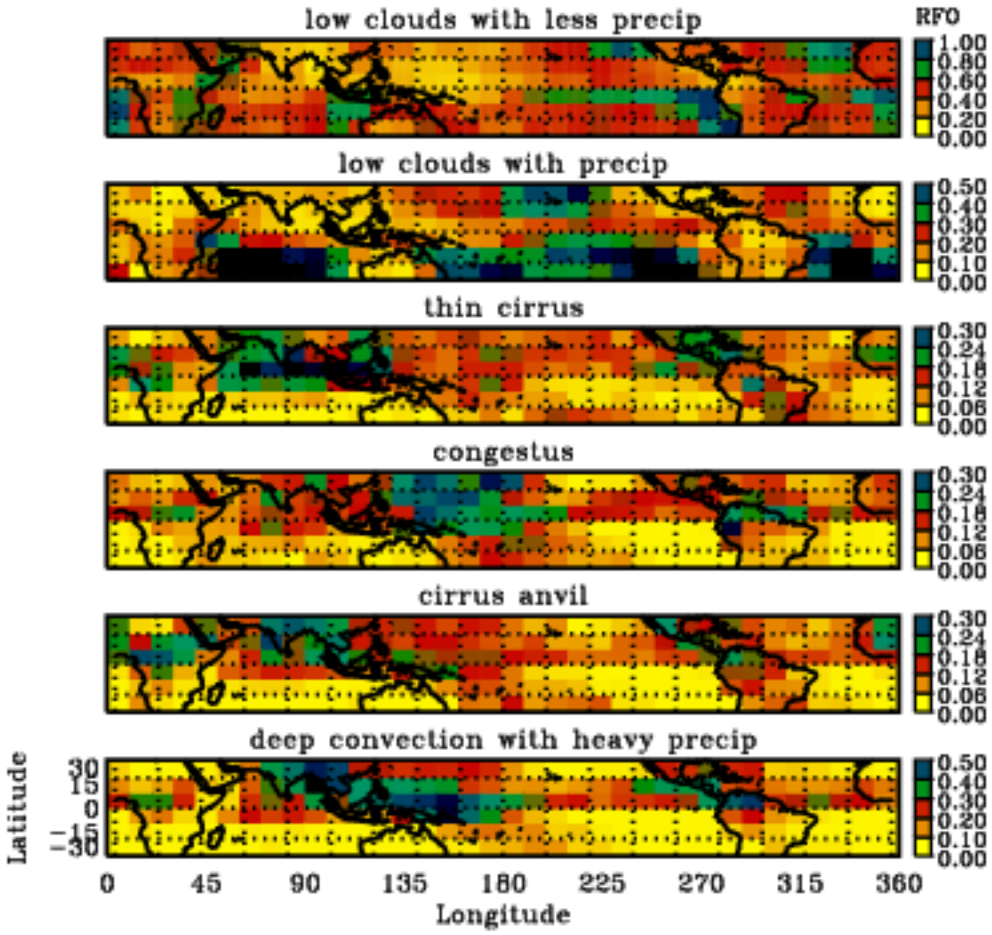

Figure 4. The time-averaged occurrence fraction of each CloudSat-CALIPSO cluster. The sum of the frequencies across clusters represents the frequency of cloudy patterns in a 10˚-by-10˚ box.

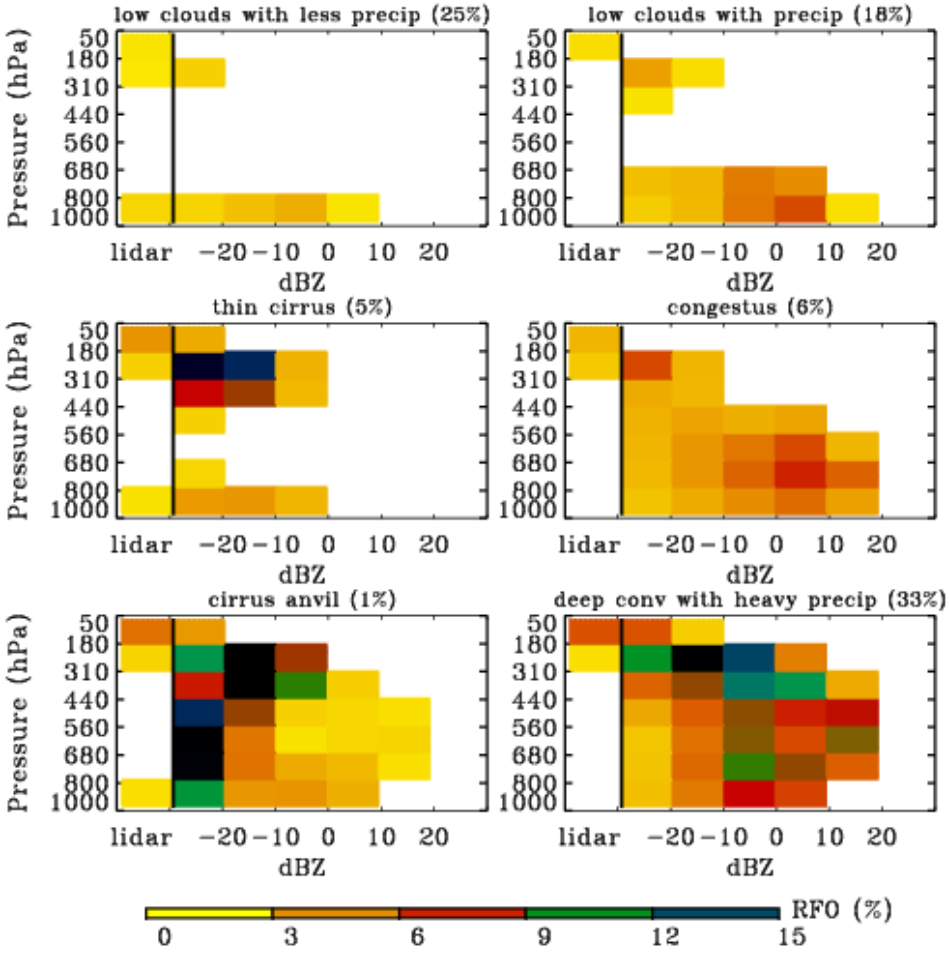

Figure 5. Joint histograms of cluster centroids from CAM3.1 by assigning cloud simulations into observational clusters based on the minimum Euclidean distance.

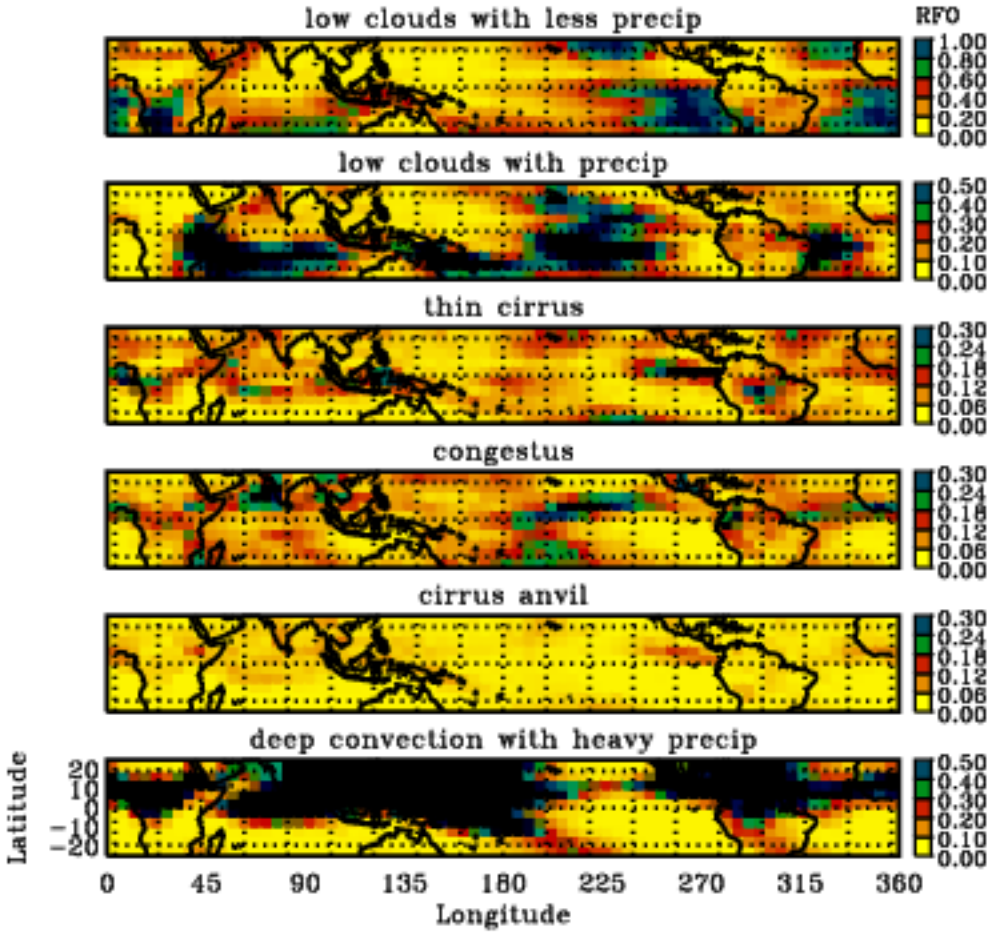

Figure 6. The temporal-averaged occurrence fraction of each cluster from cloud simulations in CAM3.1 forecasts.

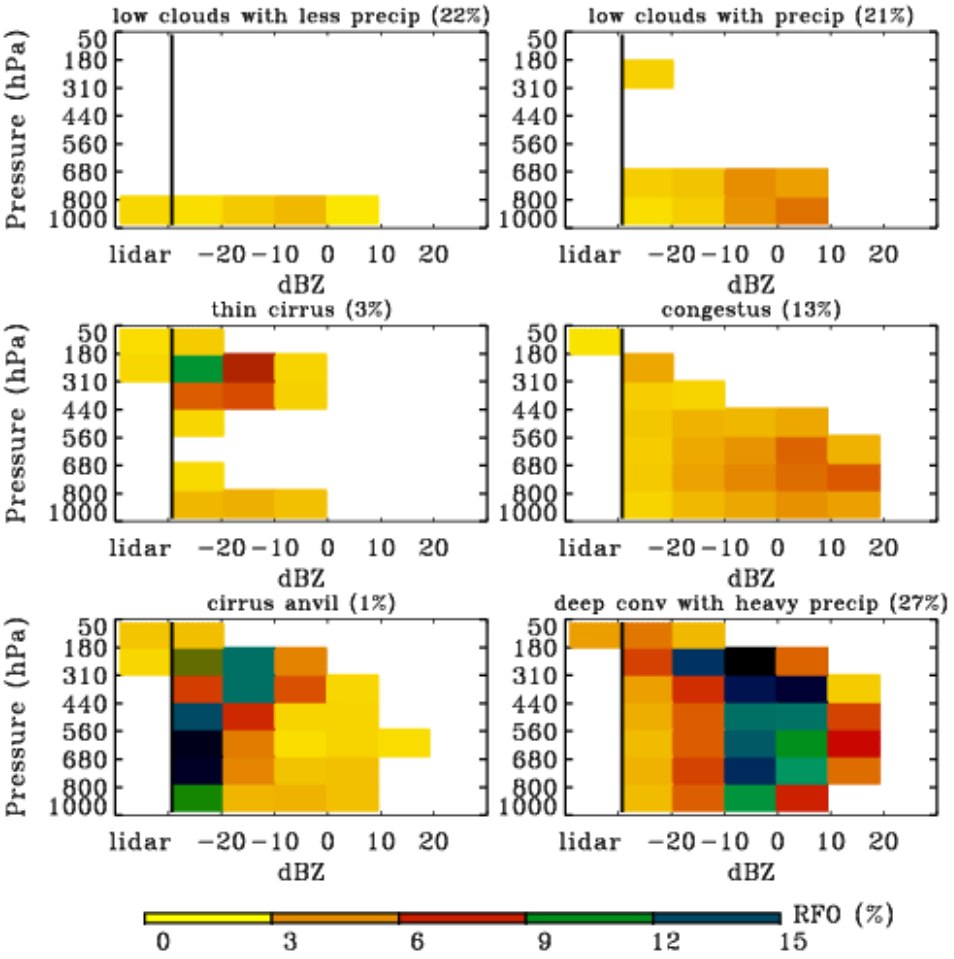

Figure 7. As in Fig. 5 but from CAM3.5.

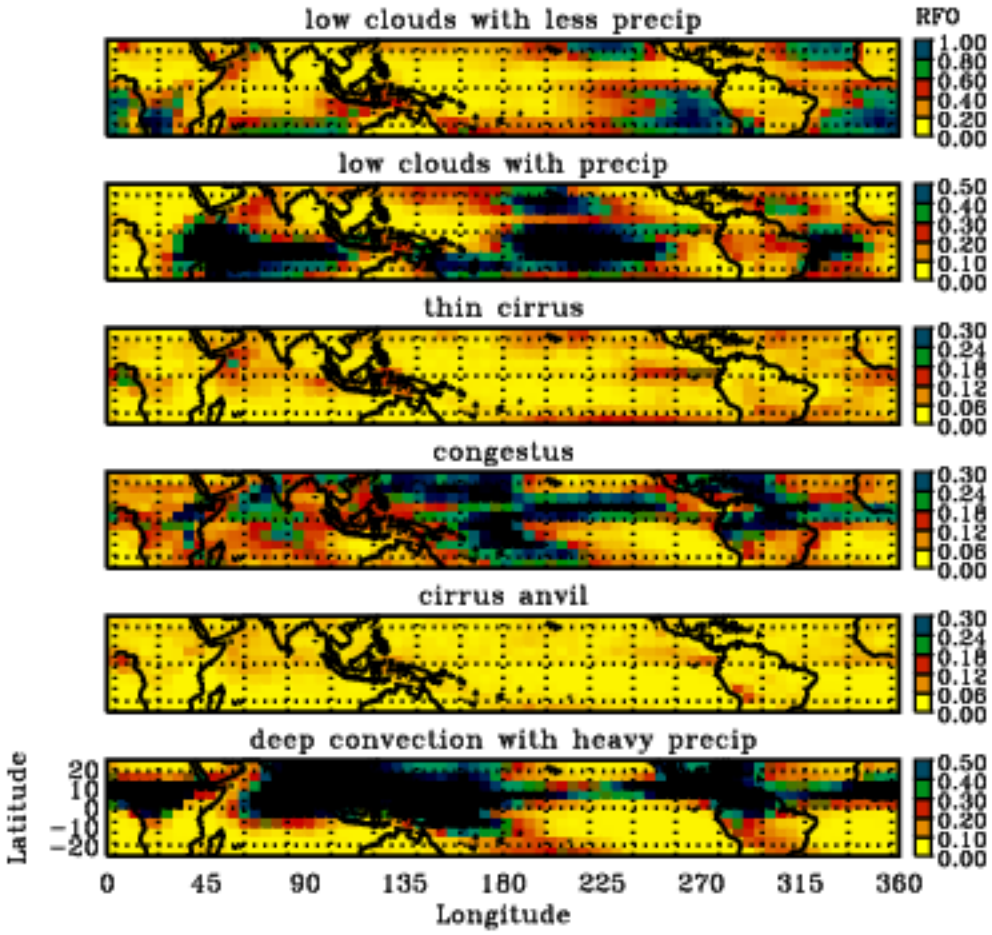

Figure 8. As in Fig. 6 but from CAM3.5.

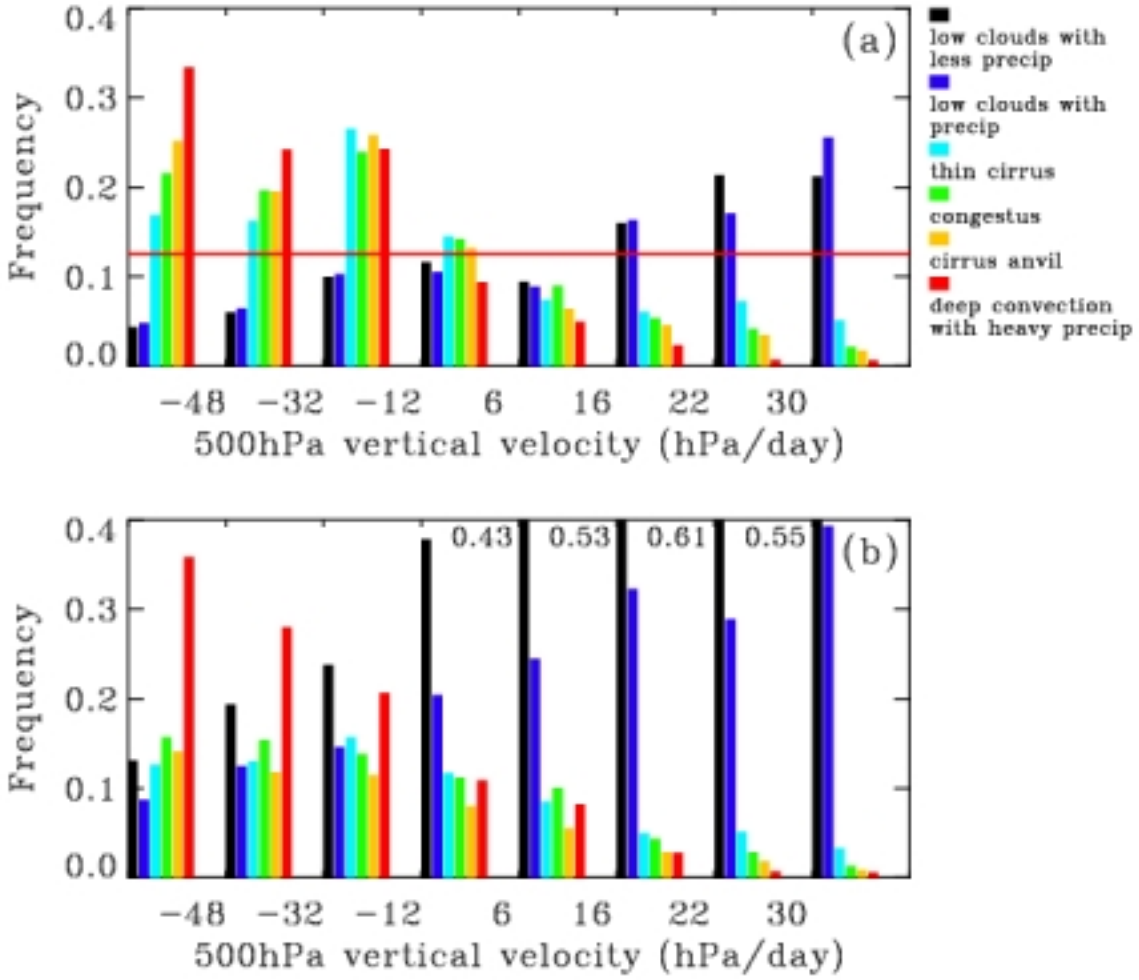

Figure 9. The frequency of the occurrence for each cluster from CloudSat-CALIPSO observation as a function of large-scale dynamics defined by the monthly mean vertical velocity at 500hPa calculated using NCEP analysis data. The boundaries for each omega bin are determined such that each bin represents the equivalent occurrence frequency of vertical velocities. (a) The fraction of elements of a given cluster which occur in the given vertical velocity bin. For this measure, the sum of the frequencies in the eight vertical velocity bins for each cluster is 1. The red line indicates the occurrence frequency if there were no association of cloud clusters with the 500hPa vertical velocity. (b) The fraction of cloudy elements of a given vertical velocity bin which belong to a given cluster. For this measure, the sum of the frequencies of the six clusters in each vertical velocity bin is 1. The frequency is labeled beside the bars hitting the top limit.

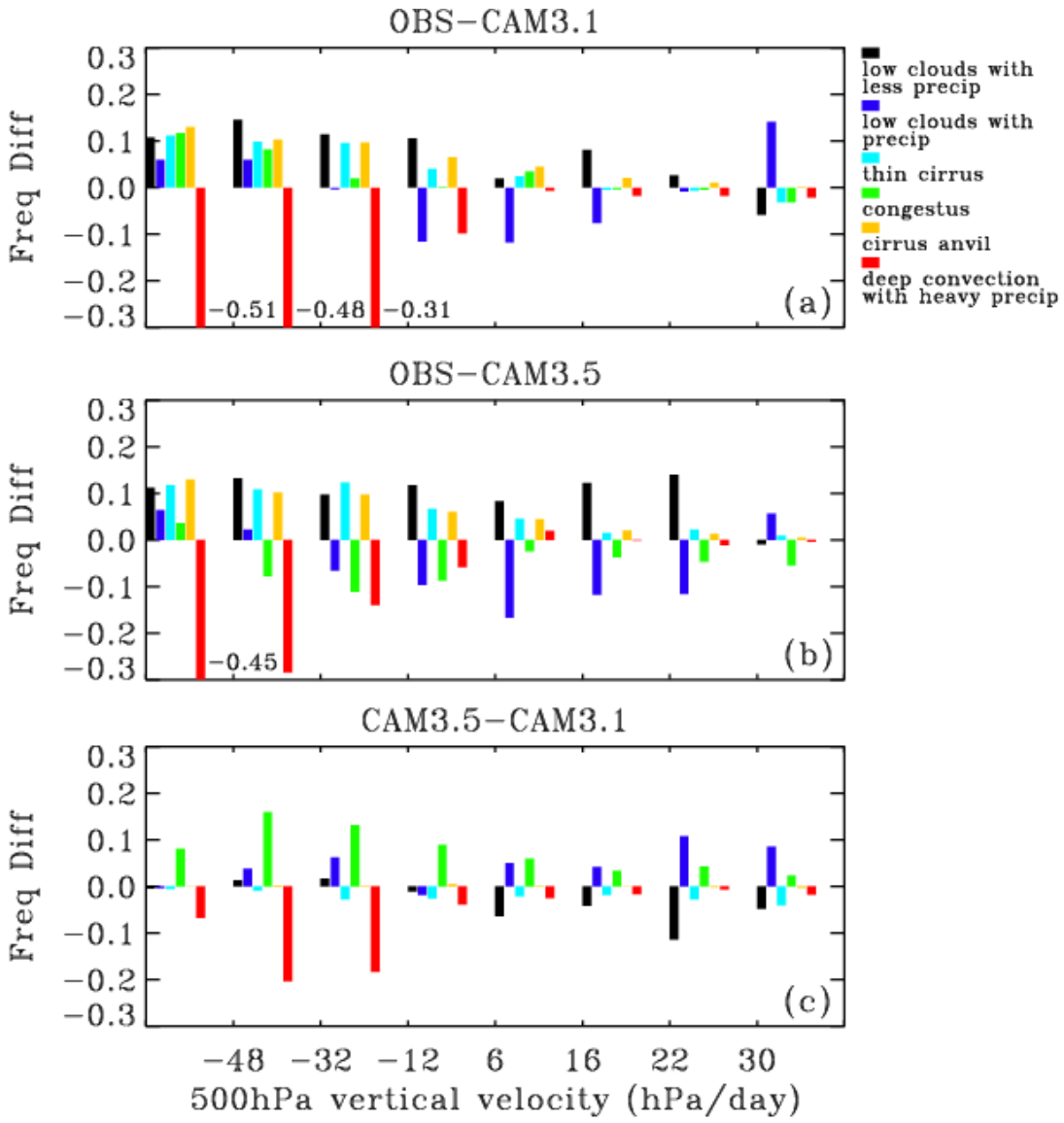

Figure 10. The frequency difference of the six clusters in each vertical velocity bin for the fraction of elements of a given vertical velocity bin which belong to a given cluster. This measure is the same as was displayed in Figure 9(b): (a) The difference between observations and CAM3.1forecasts, (b) the difference between observations and CAM3.5, and (c) the difference between CAM3.5 and CAM3.1. The frequency difference is labeled beside the bars hitting the bottom limit.

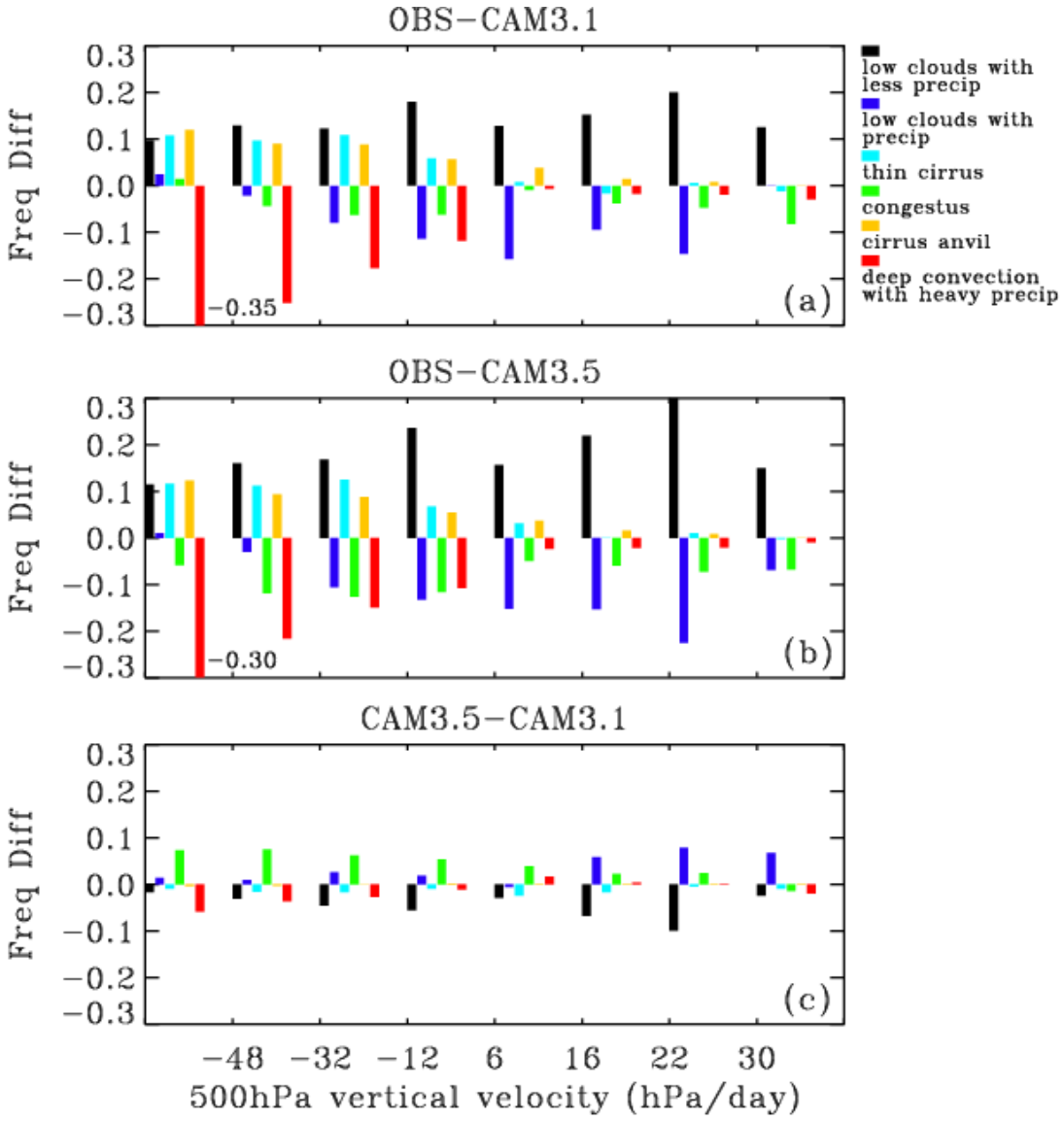

Figure 11. As in Fig. 10 but for CAM3 climate integrations.