Tropical cloud and precipitation regimes as seen from near-simultaneous TRMM, CloudSat, and CALIPSO observations and comparison with ISCCP

Zhengzhao Johnny Luo1,2, Ricardo C. Anderson1, William B. Rossow2, and Hanii Takahashi3,4

1Department of Earth and Atmospheric Sciences, City College of City University of New York, New York, New York, USA, 2NOAA-CREST Institute, City College of City University of New York, New York, New York, USA, 3Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, California, USA, 4Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA

Abstract Although Tropical Rainfall Measuring Mission (TRMM) and CloudSat/CALIPSO fly in different orbits, they frequently cross each other so that for the period between 2006 and 2010, a total of 15,986 intersect lines occurred within 20 min of each other from 30°S to 30°N, providing a rare opportunity to study tropical cloud and precipitation regimes and their internal vertical structure from near-simultaneous measurements by these active sensors. A k-means cluster analysis of TRMM and CloudSat matchups identifies three tropical cloud and precipitation regimes: the first two regimes correspond to, respectively, organized deep convection with heavy rain and cirrus anvils with moderate rain; the third regime is a convectively suppressed regime that can be further divided into three subregimes, which correspond to, respectively, stratocumulus clouds with drizzle, cirrus overlying low clouds, and nonprecipitating cumulus. Inclusion of CALIPSO data adds to the dynamic range of cloud properties and identifies one more cluster; subcluster analysis further identifies a thin, midlevel cloud regime associated with tropical mountain ranges. The radar-lidar cloud regimes are compared with the International Satellite Cloud Climatology Project (ISCCP) weather states (WSs) for the extended tropics. Focus is placed on the four convectively active WSs, namely, WS1–WS4. ISCCP WS1 and WS2 are found to be counterparts of Regime 1 and Regime 2 in radar-lidar observations, respectively. ISCCP WS3 and WS4, which are mainly isolated convection and broken, detached cirrus, do not have a strong association with any individual radar and lidar regimes, a likely effect of the different sampling strategies between ISCCP and active sensors and patchy cloudiness of these WSs.

1. Introduction

Clouds are produced by atmospheric circulation and are manifestations of weather systems. At the same time, clouds feed back to the circulation that generates them through radiative and latent heating effects. Clouds also participate in the hydrological cycle by generating precipitation. To understand these different roles of clouds, it is important to map out the three-dimensional distributions of cloud properties and to relate them to the atmospheric conditions that produce them. In this study, we focus on the vertical dimension and elucidate the vertical structures of cloud and precipitation.

While various satellite observations have been used since 1970s to characterize the horizontal variations of cloudiness and cloud optical and microphysical properties [e.g., Rossow and Schiffer, 1991, 1999; Plantnick et al., 2003], it is not until after late 1990s and early 2000s that we have established the capability to directly survey the cloud vertical structures from space [Stephens et al., 2002; Stephens and Kummerow, 2008]. Observing vertically resolved cloud and precipitation properties usually requires active remote sensing systems such as radar and lidar. In the past two decades we have witnessed a number of such systems being developed and deployed, the most notable examples being TRMM (Tropical Rainfall Measuring Mission) [Kummerow et al., 2000], CloudSat [Stephens et al., 2002, 2008], and CALIPSO (Cloud-Aerosol Lidar and Infrared Pathfinder Observation) [Winker et al., 2010].

Clouds are made up of a wide range of hydrometeors from nonprecipitating droplets of about 10 μm in radius, through a range of ice crystal sizes from 10 to over 100 μm, to large raindrops and graupel of several millimeters. Droplet number density and cloud water content also vary from one cloud to another by orders of magnitude. Consequently, it is difficult for a single radar or lidar system to observe the entire cloud due to factors such as limited sensitivity of the instrument and attenuation of signals by hydrometeors. Each
instrument is usually most sensitive to a certain part of the cloud. For example, Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) on board CALIPSO is most sensitive to small cloud droplets and ice crystals and thus can easily detect tenuous clouds such as thin cirrus or the topmost part of deep convective clouds. However, lidar signals are quickly attenuated and are unable to penetrate deep into thick clouds [Winker et al., 2010].

W-band cloud-profiling radar (CPR) on board CloudSat operates at 94 GHz and is sensitive to large cloud droplets and precipitation-size particles, but it misses tenuous clouds such as thin cirrus and small cumulus. Also, the CPR echoes are severely attenuated by heavy precipitation [Stephens et al., 2008]. The Ku-band (13.8 GHz) precipitation radar (PR) on board TRMM is capable of capturing moderate to heavy precipitation but misses light rain, drizzle, and nonprecipitating clouds entirely [Kummerow et al., 1998].

Numerous studies have analyzed TRMM and CloudSat/CALIPSO data individually to characterize the vertical structures of tropical cloud and precipitation systems (Masunaga et al. [2005], Liu and Zipser [2005], Hoynes and Stephens [2007], and Zhang et al. [2007], to name a few). CloudSat and CALIPSO data have sometimes been studied together [e.g., Mace et al., 2009] because the two satellites fly in close formation as part of the A-Train constellation [Stephens et al., 2002; L’Ecuyer and Jiang, 2010]. However, no previous study has brought all three active measurements together to take full advantage of their differing but complementary sensitivities to clouds and precipitation. Although TRMM and CloudSat/ CALIPSO fly in different orbits, they frequently cross each other. TRMM follows a highly inclined orbit covering the tropics, which allows it to frequently cut across the Sun-synchronous polar orbiting path of CloudSat and CALIPSO: over the period from August 2006 to December 2010, a total of 15,986 intersects between CloudSat/CALIPSO-TRMM occurred within 20 min of each other, with each intersect being a line of matchup that is ~400 km long. Figure 1 shows an example of the line of matchup, and Figure 2 shows the corresponding TRMM, CloudSat, and CALIPSO observations. One of the main purposes of this study is to exploit the synergy of nearly simultaneous observations by these different active systems in an effort to elucidate the vertical structure of tropical cloud and precipitation systems.

While tropical clouds and precipitation exhibit a variety of forms, previous studies have shown that they tend to cluster around a small number of dominant regimes [e.g., Jakob and Tselioudis, 2003; Ros sow et al., 2005]. Each cloud regime has a distinctive distribution pattern of cloud properties that appear repetitively throughout the observations and can thus be objectively identified through standard cluster analysis such as the k-means clustering method [Anderberg, 1973]. Based on different combinations of cloud properties, a number of studies have been conducted to characterize tropical and global cloud and precipitation regimes [e.g., Ros sow et al., 2005; Zhang et al., 2007; Elsaesser et al., 2010; Tselioudis et al., 2013]. These studies suggested that compared to traditional cloud types as in the International Cloud Atlas, the cloud regimes consisting of a mix of different cloud types can better represent the “building blocks” of tropical cloud systems, as they show good correspondence to distinctive characteristics in tropical wave dynamics [Tromeur and Rossow, 2010; Mekonnen and Ros sow, 2011], cloud radiative effects [Oreopoulos and Ros sow, 2011], latent heating [Jakob and Schumacher, 2008], and vertical layering [Tselioudis et al., 2013]. It is for this reason that cloud regimes are also called weather states to reflect the fact that they are indicative of mesoscale weather patterns [Ros sow et al., 2005].
In this study, we follow the cloud regime approach and use a cluster analysis method to examine tropical cloud and precipitation vertical structures as seen by the nearly coincident TRMM, CloudSat, and CALIPSO observations. To our knowledge, this is the first cluster analysis of the combined TRMM-CloudSat-CALIPSO measurements. Each of the measurements is most sensitive to a certain part of the cloud. Put together, they paint a more complete picture of the whole cloud and precipitation systems. Moreover, we compare the TRMM-CloudSat-CALIPSO cloud regimes with spatially and temporally collocated weather state data from the International Satellite Cloud Climatology Project (ISCCP) that are based on cluster analysis of visible and IR retrievals of cloud top height and optical depth [Rossow et al., 2005] to understand the similarities and differences between passive and active remote sensing of clouds and between horizontal and vertical patterns of clouds.

Section 2 briefly introduces the data used and describes the analysis method. Section 3 presents the cloud and precipitation regime analysis results based on joint analysis of TRMM, CloudSat, and CALIPSO observations. Emphasis is placed on the vertical structure of the tropical cloud systems. Section 4 compares cloud/precipitation regimes with ISCCP weather states. Finally, section 5 summarizes the study and discusses a few new findings.

2. Data and Analysis Method

2.1. TRMM, CloudSat, and CALIPSO

TRMM satellite was launched in November 1997. It follows a 35°-inclined orbit that covers the low latitudes at an altitude of 350 km (402 km after August 2001). The TRMM precipitation radar or PR, which is used in this study, is a 13.8 GHz (Ku-band) radar having horizontal and vertical resolutions of 4.3 km and 0.25 km, respectively, at nadir. PR scans between ±17° and the result is a swath of 215 km. For a detailed description of TRMM sensor package, see Kummerow et al. [1998].

Figure 2. A convective cloud system off the coast of tropical Africa (near 3°S, 5°W) observed near-simultaneously by (a) CloudSat CPR and CALIOP and (b) TRMM PR along the matchup line (see Figure 1). Radar observations (CPR and PR) are shown in radar reflectivities. CALIOP-only information (i.e., clouds that are missed by radar but detected by lidar) is included in gray color in Figure 2a. (c and d) The normalized H-dBZ histograms. Following Zhang et al. [2010], we group the CALIOP information as a separate dBZ bin below the CloudSat CPR detection limit.
CloudSat and CALIPSO fly in close formation to maintain the desired spatial overlap of footprints [Stephens et al., 2002]. CloudSat carries with it a 94 GHz (W-band) cloud-profiling radar (CPR) with the minimum detectable signal about ~30 dBZ. The CPR footprint size is 2.5 km along track and 1.4 km across track, and the effective vertical resolution is 480 m, oversampled at 240 m resolution [Stephens et al., 2008]. CloudSat is most sensitive to thick clouds and precipitating particles. The primary instrument carried on CALIPSO is a nadir viewing two-wavelength (532 nm and 1064 nm) polarization-sensitive lidar called CALIOP [Winker et al., 2010]. The CALIOP has higher horizontal and vertical resolutions than the CPR and can detect optically thin clouds with optical thickness of 0.01 or less [Winker et al., 2010], but its signals are quickly attenuated beyond optical thickness of about 3 [Mace et al., 2009]. CloudSat and CALIPSO measurements complement each other in characterizing the vertical structure of clouds.

Three CloudSat products, which incorporate TRMM and CALIPSO data, are used in this study: (1) 2B-GEOPROF [Mace et al., 2009], which provides cloud mask and radar reflectivity; (2) 2B-GEOPROF-LIDAR [Mace et al., 2009], which is similar to 2B-GEOPROF but incorporates cloud information from the CALIPSO level 2 vertical feature mask product; and (3) 2D-CLOUDSAT-TRMM, containing CloudSat CPR and TRMM PR intersects that occur within a defined time limit (up to 50 min, but we only use the intersects within 20 min in this study). More details on CPR and PR intersect and matchup will be presented in section 2.3 using an example. The study period is from August 2006 to December 2010. All data are downloaded from CloudSat Data Processing Center at www.cloudsat.cira.colostate.edu.

### 2.2. ISCCP Weather States

ISCCP collects and calibrates all available visible and infrared (IR) radiance measurements made from operational weather satellites (both geostationary and polar-orbiting satellites) and retrieves cloud properties such as cloud top pressure (CTP) and optical depth (Tau) based on the calibrated radiances [Ros sow and Schiffer, 1991; Ros sow and Schiffer, 1999]. ISCCP data include CTP-Tau joint histograms. Cluster analysis of the ISCCP D1 data (280 km × 280 km grid boxes at each three-hourly sunlit intervals) defines distinctive cloud regimes or weather states (WSs) [Ros sow et al., 2005]. A total of eight WSs are identified for the extended tropics (35°S–35°N): WS1–4 correspond to convectively perturbed regime with a predominance of high-level clouds, while WS5–8 are associated with convectively suppressed regimes. Figure A1 in the Appendix shows the centroids of the eight ISCCP WSs, as documented in a number of previous publications [e.g., Mekonnen and Ros sow, 2011]. The data period is from July 1983 to June 2008. The overlapping period with TRMM-CloudSat-CALIPSO is August 2006 to June 2008. The ISCCP WS data are downloaded from ISCCP website: http://isccp.giss.nasa.gov/tcluster.html.

### 2.3. H-dBZ Histograms and k-Means Clustering Method

For the analysis period (August 2006 to December 2010), a total of 15,986 CloudSat/CALIPSO-TRMM intersects occurred within 20 min of each other over 30°S–30°N. Each intersect covers a line of matchup of ~400 km, as shown in Figure 1. We excluded data poleward of 30°N and 30°S because the sample size of the matchups becomes disproportionately large between 30°N–35°N and 35°S–30°N, while they are more constantly distributed as a function of latitude within 30°S and 30°N. Since TRMM PR footprint (~5 km resolution) is larger than that of CloudSat CPR (2.5 km along track and 1.4 km across track), each matchup within the intersect line contains one PR observation and multiple CPR observations. To quantify the vertical structure of the observed cloud and precipitation features, we first construct a joint histogram of height and radar reflectivity for each 400 km matchup line, which we call H-dBZ histogram after Zhang et al. [2007]. The H-dBZ histogram is similar to the contoured frequency by altitude diagram as introduced in Yuter and Houze [1995], although we normalize the histogram in a slightly different way, such that the sum of the histogram at each height level corresponds to the percentage of cloud cover of that level, following Zhang et al. [2010]. The H-dBZ histogram is made for CloudSat and TRMM separately. Eight dBZ bins are chosen from ~30 dBZ to 50 dBZ (with 10 dBZ increment) covering the radar reflectivity ranges of both CloudSat CPR and TRMM PR. Although PR reflectivities do not go below 15 dBZ and CPR reflectivities rarely exceed 25 dBZ (due to differences in sensitivities and attenuation), we choose the common bins (~30 dBZ and 50 dBZ) for both H-dBZ histograms to emphasize the complementary nature of CPR and PR (which is clearly shown in Figures 2c and 2d). To facilitate comparison, height levels of the H-dBZ histogram are set to match the seven pressure levels used in ISCCP, i.e., surface −800 hPa (due to surface cluttering and contamination, radar data for the first 1 km are not included in the analysis), 800–680 hPa, 680–560 hPa, 560–440 hPa, 440–310 hPa,
310–180 hPa, and 180–50 hPa. So each H-dBZ histogram is a 7 × 8 matrix. The H-dBZ histogram constructed this way succinctly summarizes the mesoscale variation of cloud and precipitation vertical structures observed by the radars.

Both TRMM and CloudSat observations can be easily transformed into the H-dBZ histogram. To absorb CALIPSO lidar measurements, we follow the approach by Zhang et al. [2010]; that is, an additional dBZ bin is added at the left end of the H-dBZ diagram of CloudSat to represent CALIPSO-observed clouds that are missed by CloudSat. This additional dBZ bin can be thought of as representing clouds with radar reflectivity less than the minimum detection signal of the CPR (~30 dBZ). The CPR and PR H-dBZ histograms are normalized separately. When less than 5% of the CloudSat/CALIPSO combined profiles have clouds, the matchup is deemed as clear and is not included in the k-means cluster analysis. Figure 2 shows a deep convective cloud off the coast of tropical Africa (near 3°S, 5°W) observed near-simultaneously by CloudSat and CALIPSO (Figure 2a) and TRMM (Figure 2b). Also shown are the corresponding H-dBZ histograms (Figures 2c and 2d).

The standard k-means cluster algorithm [e.g., Anderberg, 1973] is applied to the H-dBZ histograms to search for a predefined number (K) of cluster centroids. The analysis procedure is as follows: for each intersect line (~400 km), two H-dBZ histograms are constructed, one for CloudSat and the other for TRMM each having the dimension of 7 × 8, i.e., 7 in height dimension and 8 in radar reflectivity dimension. An example is given in Figures 2c and 2d. Then the two H-dBZ histograms are combined into a single array (one appended to the other to create an array of 7 × 8 + 7 × 8 = 7 × 16), and this combined array is passed to the k-means analysis. A total of 15,986 combined arrays are compiled that go into the cluster analysis. Once the k-means cluster analysis program generates the cluster centroids (7 × 16 arrays), they are returned back to the original two separate H-dBZ histograms representing the CloudSat and TRMM components of the cluster centroids. When CALIPSO data are included, they are grouped into H-dBZ histogram in a separate dBZ bin so the reflectivity dimension increases from 8 to 9. Several different K values are tried in the analysis. To determine the optimal number of clusters, we follow the criteria as suggested in Rossow et al. [2005]; that is, the correlations between the resulting centroids should be low enough so that they are significantly different from each other, and the geographical distributions of these centroids should be significantly different, too. These centroids are representative of distinct cloud and precipitation regimes.

3. Tropical Cloud and Precipitation Regimes and Vertical Structure as Seen by Coincident TRMM, CloudSat, and CALIPSO Measurements

3.1. Combining TRMM PR and CloudSat CPR—Synergy Between Ku-Band and W-Band Radars

We first analyze the coincident measurements from TRMM PR and CloudSat CPR focusing on the synergy between the two radar systems in characterizing tropical clouds and precipitation systems, since such a study has not been done in the past. CALIPSO observations will be added later (section 3.2). k-means cluster analysis of coincident TRMM and CloudSat data suggests that there are three cloud and precipitation regimes in the tropics. Experimenting with larger cluster numbers produces new clusters that are highly correlated with existing ones.

Figure 3 shows the H-dBZ histograms for the three cloud (Figures 3a–3c) and precipitation (Figures 3d–3f) regimes, and Figure 4 shows their geographical distributions in terms of relative frequency of occurrence (RFO) of each regime. The first regime (Cluster 1; Figures 3a and 3d) is primarily made up of deep convective clouds: large CPR reflectivities (10 dBZ and higher) characteristic of deep convection extend all the way to the upper troposphere; severe attenuation of the CPR echoes by heavy rains occurs below the melting level (~590 hPa for the tropical region). The TRMM PR also shows heavy precipitation near the surface. The “bulge” in the H-dBZ histogram near the melting level corresponds to the bright band often observed by precipitation radars in the stratiform rain regions. Cluster 1 accounts for only 4.6% of the total cases and is concentrated over the tropical regions where deep convective clouds are frequently observed by previous studies, namely, the Intertropical Convergence Zone, South Pacific Convergence Zone, the Warm Pool, Tropical Africa, and Amazon. The patchy nature of the geographical distribution (Figure 4) is attributed to relatively small sample size of the coincident TRMM and CloudSat cases: of the total 8484 cloudy matchup lines, only 389 belong to Cluster 1 (note that 7502 of the 15,986 matched cases are considered clear scenes and do not enter the cluster analysis). The total cloud cover (TCC), defined as the fraction of CloudSat profiles that contain cloud signals, is 0.97 on average for Cluster 1. Similarly, we calculate the total precipitation cover...
Figure 3. The three cloud regimes or clusters based on joint analysis of (a–c) CloudSat CPR and (d–f) TRMM PR. Relative frequency of occurrence (RFO) and mean total cloud cover (TCC) are shown on top of each H-dBZ histogram. Similarly, the mean total precipitation cover (TPC) is shown on top of each H-dBZ histogram.

Figure 4. Geographical distributions (within each 10° × 10° grid box) of the relative frequency of occurrence (RFO) for each of the three cloud and precipitation regimes as shown in Figure 3.
(TPC), defined the same way as TCC except using TRMM PR observations to search for precipitation signal near the surface. The mean TPC for Cluster 1 is 0.50. The large coverage fraction of both clouds (most of them being high-level clouds) and significant near-surface precipitation suggests that Cluster 1 consists of not only ordinary cumulonimbus plumes but also highly organized convection, most likely mesoscale convective systems (MCSs) covering almost the entire 400 km TRMM-CloudSat matchup line. Figure 2 shows a typical example.

It is worth noting that CPR and PR measurements provide highly complementary information for depicting deep convective clouds. Berg et al. [2010] discussed the CPR-PR synergy in estimating the distribution of rainfall over tropical and subtropical oceans, as the PR provides the best information on the total rain volume, while the CPR provides more sensitive rain detection and is thus better for estimating drizzle and light rain. Here we emphasize the PR-CPR (or Ku-W band) synergy in observing the vertical structure of tropical deep convection. Such synergy can be appreciated by dividing the Cluster 1 H-dBZ histograms (Figures 3a and 3d) into three height ranges that correspond to different cloud phases: (1) ice phase (above 310 hPa): CPR with high sensitivity at $\sim$30 dBZ is capable of observing most of the ice crystals inside thick cirrus anvils or the topmost part of convective cores. PR, however, misses most of the non-graupel ice particles due to low sensitivity (about 18 dBZ), as is evident in Figure 3. (2) Mixed phase (310 hPa–590 hPa): over this height range, both PR and CPR are capable of observing cloud and precipitation particles, and therefore, their H-dBZ histograms show some overlap, although differences in radar frequencies and the Mie scattering effect make the CPR effective reflectivity somewhat smaller than that of the PR. (3) Liquid phase (below the 590 hPa): below the melting level, heavy rain severely attenuates the CPR signal, while the PR echoes are much less affected and continue to increase toward the surface. Therefore, the CPR-PR synergy in observing deep convection works in such a way that the CPR is best positioned for observing the ice-phase part of the cloud system (above 310 hPa), while the PR works the best below the melting level ($\sim$590 hPa) under heavy rain; for the mixed-phase range, both CPR and PR measurements contain useful information about cloud vertical structure and microphysics. The Ku- and W-band synergy should help in the design of future spaceborne missions that focus on studying convective clouds and storms.

The second regime (Cluster 2; Figures 3b and 3e) is composed of mostly thick cirrus anvils, as the maximum occurrence frequency of the CPR H-dBZ histogram is concentrated in the upper troposphere from 180 to 440 hPa with radar reflectivity of $\sim$30 to 0 dBZ. Within the cirrus anvils, radar reflectivity tends to increase with decreasing height, suggesting that ice crystals grow as they sediment, consistent with the finding by Yuan et al. [2011]. The CPR H-dBZ histogram for Cluster 2 also shows signals of deep convection (i.e., radar echo greater than 10–15 dBZ extending from near the melting level to the upper troposphere), but the occurrence frequency is relatively low compared to that of cirrus anvils. The PR H-dBZ histogram shows that Cluster 2 contains a significant amount of surface precipitation, although not as heavy or frequent as Cluster 1. Cluster 2 accounts for 15% of the TRMM-CloudSat intersect cases (i.e., 3–4 times as frequent as Cluster 1). Their geographical distribution is similar to that of Cluster 1, with a wider spread (Figure 4b). The average TCC of Cluster 2 is 0.88 and average TPC is 0.12, both smaller than those of Cluster 1 (TCC and TPC for Cluster 1 are 0.97 and 0.50, respectively). Hence, Cluster 2 consists of scenes that are made up of deep convection and anvils, but they are less organized than Cluster 1 and contain less precipitation. The geographical association between Cluster 1 and Cluster 2 and the characteristics of their H-dBZ histograms suggest that Cluster 1 likely represents MCS at the mature stage, while Cluster 2 represents the dissipating stage, as discussed in Takahashi and Luo [2014] [see also Lee et al., 2013]. We will return to this point when comparing results with the ISCCP weather states.

The third regime (Cluster 3; Figures 3c and 3f) is the most prevalent regime, accounting for 80% of the cases. Judging by the H-dBZ histograms, Cluster 3 is a convectively suppressed regime with a mixture of different cloud types. A large fraction of clouds in Cluster 3 are low-level clouds below 680 hPa. They do not contain significant amount of rainfall as indicated by lack of TRMM PR echoes; nevertheless, there are some occurrences of light rain from shallow convection, as suggested by CPR reflectivities in the range of $\sim$10 to 0 dBZ near the surface, consistent with finding by Stephens and Wood [2007] that near-surface W-band radar reflectivities of $\sim$10 to 0 dBZ correspond to light rain. Another important cloud type within Cluster 3 is cirrus at 440–180 hPa with CPR reflectivity of $\sim$30 to $\sim$10 dBZ. They are optically thinner than the cirrus anvils in Cluster 2. The geographical distribution of Cluster 3 appears to be widely spread over the whole tropics.
There are a few local concentrations near the west coast of major continents that seem to correspond to stratocumulus and trade cumulus that have been identified in a number of previous cluster analysis studies [Rossow et al., 2005; Zhang et al., 2007; Tan et al., 2013; Tselioudis et al., 2013]. Thin cirrus clouds could be distributed throughout the tropics. Some of them are associated with detrainment from deep convection and others form in situ, as found in Luo and Rossow [2004]. The mean TCC and TPC for Cluster 3 are, respectively, 0.37 and 0.02.

Because Cluster 3 shows a mixture of high-level and low-level clouds, we made an attempt to separate them by increasing the cluster number from 3 to 4. However, results are unsatisfactory: cirrus and low-level clouds are still largely combined in a single cluster; the new cluster generated is a mixture of cirrus and some deep convective clouds. A probable reason for not being able to obtain a clean separation of these cloud types is likely due to the small sample size (8484 cases that contain clouds) and limited dynamic range of cloud properties that can be observed by radars (it will be shown in section 3.2 that adding lidar information helps the separation to some extent). However, we can get around the problem by conducting a separate k-means cluster analysis for the Cluster 3 cases only, that is, a nested cluster analysis. Figure 5 shows the resulting

![Figure 5](image)

**Figure 5.** The same as Figures 3a–3c, but for the three subregimes associated with Cluster 3 in Figure 3c. See text for details.

![Figure 6](image)

**Figure 6.** The same as Figure 4, but for the three convectively suppressed subregimes as shown in Figure 5.
subregime CloudSat H-dBZ histograms ($K = 3$ is the optimal cluster number), and Figure 6 shows the corresponding geographical distribution. Based on the H-dBZ histograms and the geographical distribution maps, it can be concluded that the three subregimes consist of, respectively, stratocumulus clouds with drizzle or light precipitation (Figure 5a, RFO: 20%; TCC: 0.63), cirrus overlying lower level clouds (Figure 5b, RFO: 11%; TCC: 0.66), and nonprecipitating cumulus (Figure 5c, RFO: 68%; TCC: 0.24). TPC is not counted because there is virtually no precipitation signal from the TRMM PR. Among these three subregimes, cumulus is the most prevalent (RFO = 68%), but their TCC is the lowest (0.24). In other words, the cumulus subregime is most widespread across the whole tropics (as shown in Figure 6c), but the cumulus scenes consist of broken clouds covering only a small fraction of the scenes. In contrast, the stratocumulus subregime covers a smaller fraction of the tropics (RFO = 20%), but those scenes contain more continuous cloud decks (TCC = 0.63).

3.2. A Joint TRMM-CloudSat-CALIPSO View

CALIPSO brings new information to the TRMM-CloudSat view of tropical clouds. When CALIPSO observations are incorporated in the H-dBZ histogram as an additional dBZ category, $k$-means cluster analysis gives $K = 4$ as the optimal number for clusters. Figure 7 shows the centroids of these four regimes in terms of H-dBZ histograms, and Figure 8 shows the geographical distributions of their RFO.

The first two regimes (Clusters 1 and 2) in Figures 7a, 7b, 7e, and 7f and 8a and 8b are very similar to those of the TRMM-CloudSat analysis (see section 3.1 and Figures 3a, 3b, 3d, and 3e and 4a and 4b). They are associated with deep convective clouds and cirrus anvils, respectively. Cluster 1 accounts for 5.8% of the cases, and Cluster 2 accounts for 19%. The mean TCCs for Clusters 1 and 2 are, respectively, 0.99 and 0.98. Note that Cluster 2 (Figure 7b) has much higher TCC (0.98) compared to the counterpart in the TRMM-CloudSat analysis (0.88; Figure 3b) because CALIPSO can detect more tenuous cirrus clouds. The mean TPCs for Clusters 1 and 2 are, respectively, 0.41 and 0.06.

For the two convectively active regimes in Figure 7 (i.e., Clusters 1 and 2; Figures 7a, 7b, 7e, and 7f), additional information from CALIPSO is mostly thin cirrus associated with deep convective clouds that are missed by CloudSat. For example, the H-dBZ histogram for Cluster 2 shows a local maximum occurrence frequency at the 310–50 hPa height level with signal strength below the detection threshold of the CloudSat CPR (i.e., only CALIPSO can see these clouds). They are obviously very thin cirrus, including tropopause subvisual cirrus [Winker and Trepte, 1998]. Some are found in the lower stratosphere above the tropopause (the height range of 180–50 hPa). These lower stratospheric clouds are possibly generated by overshooting deep convection, as discussed by Iwasaki et al. [2015]. It is interesting to note that the lower stratospheric clouds in Cluster 1 are optically much thicker than the counterpart...
in Cluster 2, as the former can reach up to CPR reflectivities greater than 5 dBZ, while the latter stops at -20 dBZ and mostly concentrated in the lidar-only category (Figures 7a–7d). This contrast is also seen in Figures 3a–3c for CloudSat-only H-dBZ histograms. We speculate that the lower stratospheric clouds with large CPR radar echoes (as in Cluster 1) are the uppermost part of active overshooting plumes, whereas those with marginal radar echo or with lidar-only signals (as in Cluster 2) are likely lower stratospheric clouds indirectly induced by overshooting such as “jumping cirrus” [e.g., Wang, 2004].

Clusters 3 and 4 (Figures 7c, 7d, 7g, and 7h) are convectively suppressed regimes. Cluster 3 is composed of mostly lower level clouds. A large fraction of these low clouds are below the detection limit of CloudSat CPR (i.e., only “visible” to CALIPSO). Nevertheless, some of the low clouds are thick enough to produce drizzle or light rain as indicated by CPR reflectivity greater than -10 dBZ near the surface. A small subset of them even produces rain rates large enough for TRMM PR to detect. The geographical distribution of Cluster 3 (Figure 8c) is well aligned with locations of tropical and subtropical stratocumulus [e.g., Klein and Hartmann, 1993]. In addition to lower level clouds, CALIPSO also adds some thin cirrus at the height range of 310–50 hPa. Cluster 3 accounts for 25% of the cases, and the corresponding mean TCC and TPC are, respectively, 0.85 and 0.03.

Cluster 4 is a mixture of thin, low-level clouds (e.g., trade cumulus) and thin cirrus, as suggested by both the CPR H-dBZ histogram (Figures 7d and 7h) and the corresponding geographical distribution (Figure 8d). They contain virtually no precipitation at the surface, according to both the PR and CPR observations. This regime accounts for 49% of the cases. The mean TCC and TPC are, respectively, 0.62 and 0.01.

Figure 8. The same as Figure 4, but for the joint radar-lidar observations (i.e., TRMM, CloudSat, and CALIPSO) cluster analysis.
The joint TRMM-CloudSat-CALIPSO analysis produced one more cluster or regime than the analysis that only involves TRMM and CloudSat by separating the convectively suppressed regime of the latter (Cluster 3 in Figures 3c and 3f) into two regimes (Clusters 3 and 4 in Figures 7c, 7d, 7g, and 7h). This suggests that inclusion of CALIPSO data adds to the dynamic range of cloud properties that can be observed by TRMM and CloudSat, better separating more continuous low clouds from low, broken, and thin cirrus clouds.

Following the nested clustering approach in section 3.1, we conducted a similar subregime analysis on Cluster 4 (Figures 7d and 7h) because it appears to consist of a mixture of different types of clouds. (We applied the same subregime analysis on Cluster 3, but it does not break into any meaningful subclusters, which probably makes sense because the geographical distribution of Cluster 3 in Figure 8c suggests that this regime is mostly marine stratocumulus clouds.) Figure 9 shows the subregimes associated with Cluster 4 ($K = 3$ is the optical cluster number). Examining the H-dBZ histograms and the corresponding geographical distributions of these subregimes suggests that (1) subcluster 1 (Figures 9a and 9d) consists of broken cirrus and cumulus clouds (we inferred that they are broken clouds because TCC is only 0.54) that are spread out over the whole tropics. Subcluster 1 is the most prevalent subcluster and accounts for 76.5% of the cases. (2) Subcluster 2 (Figures 9b and 9e; 11.4% of the case) contains mostly very thin cirrus near the tropopause that are only observed by CALIPSO. Judging by the geographical locations (Figure 9e), they are most likely tropopause cirrus connected directly or indirectly to deep convection. TCC for Subcluster 2 is as high as 0.94, consistent with the large-scale coverage of the tropopause cirrus [Winker and Trepte, 1998]. (3) Subcluster 3 (Figures 9c and 9f; 12.1% of the cases) is most interesting: the H-dBZ histogram shows that they are spread out in the height and reflectivity range, but most of them are midlevel clouds observed only by CALIPSO (or in the lowest CPR reflectivity bin), so they are optically very thin. Their geolocations suggest that these clouds are mostly associated with major mountain ranges in the lower latitudes such as Himalayas, Andes, and the Ethiopian Highlands. It is remarkable that the nested $k$-means cluster analysis can single
out these mountain clouds! The origin of these clouds is beyond the scope of this paper but should be a very interesting topic to follow-up.

4. Comparison With ISCCP Weather States

It is of interest to compare tropical cloud regimes derived from the active sensors with ISCCP weather states (WSs), as the former are based on characteristics of cloud vertical structures, while the latter rely on horizontal distribution of cloud properties (optical thickness and cloud top pressure). A recent study by Tsiloudis et al. [2013] composited the CloudSat-CALIPSO cloud vertical structure with respect to each ISCCP WS and showed that the horizontal distributions of cloud properties as seen by ISCCP are associated with distinct patterns of the vertical layering of the clouds based on radar and lidar measurements. Here we adopt a somewhat different approach by directly cross-comparing the two versions of cloud regimes. Since cloud regimes describe recurring patterns in the tropics, it is reasonable to expect that regimes from radar and lidar observations may have counterparts in the ISCCP WSs.

We match up TRMM-CloudSat/CALIPSO Cloud Regimes with ISCCP WS by dropping each PR-CPR intersect into the corresponding ISCCP 2.5° × 2.5° grid based on the central latitude and longitude of the intersect line. Figure 10 shows the RFO of the ISCCP WSs sorted by TRMM-CloudSat Regimes (i.e., radar-only regimes; Figure 10a) and TRMM-CloudSat-CALIPSO Regimes (i.e., radar-lidar regimes; Figure 10b) as described in section 3. To reduce clutter and to place more emphasis on convective clouds, we combine the four ISCCP WSs representing convectively suppressed conditions (WS5–8) into a single set. The nature of the four convectively active ISCCP WSs (i.e., WS1–4) is as follows [Mekonnen and Rossow, 2011; Oreopoulos and Rossow, 2011; Tan et al., 2013]: WS1 represents organized deep convective clouds with almost the entire 2.5° × 2.5° (280 km × 280 km) area covered by optically thick, high top clouds. WS2 is mostly thick cirrus anvils and is usually found in the vicinity of WS1. As explained in Tan et al. [2013], WS1 and WS2 likely represent different life stages of MCSs with WS1 being the mature stage of the MCSs and WS2 being the dissipating stage. WS3 is a mixture of different types of tropical convective clouds such as cumulus congestus, deep convection, and cirrus, and their cloud coverage is significantly less than that of WS1 so they are most likely disorganized or isolated convection in the tropics. WS4 is mostly cirrus not attached to deep convection, but they are likely connected to tropical deep convection through some indirect ways [e.g., Luo and Rossow, 2004], so we include WS4 in the convectively active group.

Figure 10 shows that ISCCP WS1 and WS2 are mostly associated with Cluster 1 and Cluster 2 of both the radar-only and radar-lidar observations. In particular, WS1 is the most prominent weather state in Cluster 1 (both radar-only and radar-lidar versions) and WS2 is most often found in Cluster 2. So broadly speaking, ISCCP
WS1 is the counterpart or equivalent of the radar-lidar Cluster 1, and ISCCP WS2 is the equivalent of the radar-lidar Cluster 2. Hence, the recurring patterns of tropical convective clouds have distinct horizontal (ISCCP) and vertical structures (radar and lidar) that well correspond well to each another, consistent with Tselioudis et al. [2013]. Inclusion of collocated TRMM PR profiles adds new insight into the precipitation character of the tropical convective regimes: Figures 3d and 3e and 7e and 7f show that Cluster 1 not only has heavier near-surface precipitation than Cluster 2 [cf., Lee et al., 2013; Rossow et al., 2013; Tan et al., 2015] but also contains more precipitation-size hydrometeors throughout the whole cloud volume.

ISCCP WS3 has similar RFO among all regimes in radar and lidar observations, although there is a slight preference for Cluster 2. ISCCP WS4 behaves somewhat similar to WS3. The fact that ISCCP WS3 and WS4 are nearly equally distributed among all the radar-lidar regimes can be attributed to different sampling strategies and viewing geometries between these observational systems, as ISCCP WS is based on horizontal view of cloud properties within a larger 280 km × 280 km area, whereas TRMM-CloudSat-CALIPSO matchup line is just an ~400 km vertical “curtain.” Rossow et al. [1993] illustrate the sampling effects when comparing point to area measurements, and Rossow and Zhang [2010] explicitly test the effects of different spatial scales on the amount of apparent disagreement between CloudSat-CALIPSO and ISCCP. WS3 and WS4 are geographically associated with WS1 and WS2 (recall that all WS1–4 are convectively active regimes) but are much more frequent [Rossow et al., 2005], so that they are usually near the convective systems. Lee et al. [2013] illustrates the association in time of WS3 and WS4 with WS1 and WS2.

The convectively suppressed WSs (WS5–8) are mostly concentrated in the corresponding nonconvective regimes in radar and lidar regimes, namely, Cluster 3 in TRMM-CloudSat observations and Clusters 3 and 4 in TRMM-CloudSat-CALIPSO observations. Only a small fraction of ISCCP WS5–8 falls into the convectively active regimes by radar and lidar. The reason is probably similar to that responsible for ISCCP WS3–4 being mixed up with convectively suppressed regimes as explained in the previous paragraph, namely, patchy cloudiness and different sampling geometries.

5. Summary and Discussion

We analyzed over 4 years of coincident measurements by TRMM, CloudSat, and CALIPSO missions to elucidate tropical cloud and precipitation regimes and their internal vertical structures. Although TRMM and CloudSat/CALIPSO fly in different orbits, they cross each other frequently so that for the period from August 2006 through December 2010, a total of 15,986 intersect lines (~400 km long) occurred within 20 min of each other, which provides a rare opportunity for investigating cloud and precipitation vertical structures using near-simultaneous active sensors that highly complement one another by covering a very wide range of hydrometers from tiny and sparse ice crystals to large raindrops. To identify the distinct patterns of tropical cloud and precipitation, we adopted the well-documented k-means cluster analysis: radar and lidar observations from each of the intersect line were first converted into the H-dBZ histogram, and then a total of 15,986 H-dBZ histograms were passed through the cluster analysis. The main findings of the study are summarized as follows:

1. We first analyzed the coincident measurements from TRMM PR and CloudSat CPR focusing on the synergy between the two radars with different wavelengths and sensitivities to hydrometeors. k-means cluster analysis identified three cloud and precipitation regimes in the joint TRMM-CloudSat H-dBZ histogram. The first two clusters correspond to, respectively, organized deep convection with heavy rain (RFO: 5%) and thick anvil clouds with moderate rain (RFO: 15%). The third regime is a convectively suppressed regime (RFO: 80%), and it can be further divided into three subregimes through a separate k-means cluster analysis, which correspond to stratocumulus clouds with some drizzle, cirrus overlying low clouds, and nonprecipitating cumulus, respectively.

2. Inclusion of CALIPSO data adds to the dynamic range of cloud properties that can be detected and picks up very thin cirrus and shallow cumulus clouds that are missed by radars and identifies one more cloud regime. Interestingly, subregime cluster analysis of the dominant convectively suppressed regime successfully singles out some very thin midlevel clouds over the tropical mountain ranges that are mostly observed by CALIPSO only. The origin of these clouds remains to be investigated in a future study.

3. The radar-lidar cloud regimes are compared with the ISCCP weather states (WSs). ISCCP WS1 and WS2 are found to be counterparts of Cluster 1 and Cluster 2 in radar-lidar observations, respectively. Evidence from
Both this and previous studies suggest that radar-lidar Cluster 1 (or ISCCP WS1) represents organized deep convection or MCSs most likely in a mature stage, whereas radar-lidar Cluster 2 (or ISCCP WS2) consists of extensive coverage of thick anvils with some precipitation usually found in the vicinity of Cluster 1 (or WS1) that are either the anvil part of larger systems or MCSs in a dissipating stage. ISCCP WS3 (mixture of isolated convection) and WS4 (cirrus) do not have a strong association with any individual radar and lidar regimes probably due to different sampling strategies and the patchy nature of these clouds in the sampling area.

Findings (1) and (2) can be compared with a previous work by Zhang et al. [2010] who conducted cluster analysis of CloudSat and CALIPSO data in a similar region. Zhang et al. [2010] identified six clusters for the tropics. Our study found four clusters initially; however, the fourth cluster can be further broken into three subclusters so the total number of clusters is comparable between the two studies. One small difference is that we see less low cloud regime than Zhang et al. [2010] but pick up a different cluster, namely, the midlevel thin clouds over the major tropical mountain ranges. One possible reason could be attributable to the differences in data period and sampling method. Zhang et al. [2010] only analyzed 3–4 months of data, but their sample is complete and unconditional, whereas we analyzed more than 4 years of data but are confined to observations within 20 min of TRMM overpass in order to bring together near-simultaneous PR, CPR, and CALIOP measurements.

Some of the results in this paper corroborate previous works concerning tropical cloud and precipitation regimes [e.g., Rossow et al., 2005; Zhang et al., 2010; Lee et al., 2013; Rossow et al., 2013; Tan et al., 2013], but a few new lessons are learned that are worth discussing. First, the synergy between TRMM PR (Ku band) and CloudSat CPR (W-band) in depicting tropical convective cloud structures is revealed through the clustering analysis of the joint H-dBZ histograms. From a cloud microphysics perspective, a deep convective cloud column can be divided into three stacked layers: ice-phase hydrometeors on top (−38°C), mixed-phase particles in between (0°C to −38°C), and liquid-phase droplets (>0°C) at the bottom. Our results show that CPR with high sensitivity is most suitable for characterizing the ice-phase component of a deep convective cloud. PR is best positioned to retrieve rain information below the melting level (because of smaller attenuation compared to CPR), and both CPR and PR contain useful information content for the mixed-phase part. Hence, a dual PR-CPR or Ku-W band radar package will deliver better performance in studying the vertical structure and internal processes of deep convective clouds than a single-band radar or a Ku-Ka dual band radar with low sensitivity (such as the Global Precipitation Measurement dual-frequency precipitation radar).

Adding a lidar will further benefit study of cirrus resulting from convection that is usually too tenuous for radars to detect. Elucidation of these synergies may become a useful consideration for designing future spaceborne missions or constellation that aim at studying convective storm systems.

A second lesson learned from this study concerns the k-means cluster analysis of cloud properties. Recalling the cluster analysis of the joint TRMM-CloudSat observations produces two convectively active regimes (Cluster 1 and Cluster 2) and one single convective suppressed regime (Cluster 3). The RFOs of the former two are 5% and 15% respectively, while the RFO of the latter is 80%. Based on examination of the H-dBZ histogram and geographical distribution, Cluster 3 appears to be a mixture of different types of clouds and can be further divided into some subregimes. However, increasing K from 3 was not able to break Cluster 3 in a clean way. We get around the problem by running a separate k-means cluster analysis on the scenes that belong to Cluster 3 only, that is, a nested cluster or subcluster analysis. The nested k-means analysis successfully divides Cluster 3 into three subregimes that are separated geographically and in cloud properties, which correspond to, respectively, stratocumulus off the west coasts of major continent, cirrus over lower level clouds near convectively active region, and shallow cumulus throughout the tropics and subtropics. Another successful example of the nested cluster approach is the analysis of the joint TRMM-CloudSat-CALIPSO observations, in which the original Cluster 4 is divided into three subregimes; an interesting subcluster emerges that corresponds to midlevel clouds near major mountain ranges of the lower latitudes that are mostly observed by CALIPSO only (very thin).

Previous studies have also explored the nested cluster analysis approach. For example, Mason et al. [2014] identified unique subregimes within the general midtopped cloud regime over the Southern Ocean. They were not able to obtain a clean separation of these subregimes but succeeded only through the nested cluster approach. A similar example is found in Oreopoulos et al. [2016], in which they started from a small number.
of baseline or "core" clusters and then conduct subregime analysis for each of the "cores" until an optimal set of clusters is achieved. Hence, for researchers who are interested in using the cluster technique, we strongly recommend these different options.

Finally, we note that adding lidar information to the two radar observations helps separate the regimes to some extent (although not as effectively as the nested cluster method), suggesting that increasing the dynamic range of observation will generate more clusters.

Appendix A

For the sake of completeness, we present here the centroids of the eight ISCCP weather states (WSs) for 35°S–35°N shown in ISCCP CTP-TAU histograms (Figure A1). This figure is adopted from Mekonnen and Rossow [2011].

**References**


