Source of Acquisition
NASA Marshall Space Flight Center

The Tropical Convective Spectrum. Part I: Archetypal Vertical Structures

DENNIS J. BOCCIPPIO

Earth Science Department, NASA Marshall Space Flight Center, Huntsville, Alabama

WALTER A. PETERSEN AND DANIEL J. CECIL

University of Alabama in Huntsville, Huntsville, Alabama

(Manuscript received 2 December 2003, in final form 10 August 2004)

ABSTRACT

A taxonomy of tropical convective and stratiform vertical structures is constructed through cluster analysis of 3 yr of Tropical Rainfall Measuring Mission (TRMM) "warm-season" (surface temperature greater than 10°C) precipitation radar (PR) vertical profiles, their surface rainfall, and associated radar-based classifiers (convective/stratiform and brightband existence). Twenty-five archetypal profile types are identified, including nine convective types, eight stratiform types, two mixed types, and six anvil/fragment types (nonprecipitating anvils and sheared deep convective profiles). These profile types are then hierarchically clustered into 10 similar families, which can be further combined, providing an objective and physical reduction of the highly multivariate PR data space that retains vertical structure information. The taxonomy allows for description of any storm or local convective spectrum by the profile types or families. The analysis provides a quasi-independent corroboration of the TRMM 2A23 convective/stratiform classification. The global frequency of occurrence and contribution to rainfall for the profile types are presented, demonstrating primary rainfall contribution by midlevel glaciated convection (27%) and similar depth decaying/stratiform stages (28%–31%). Profiles of these types exhibit similar 37- and 85-GHz passive microwave brightness temperatures but differ greatly in their frequency of occurrence and mean rain rates, underscoring the importance to passive microwave rain retrieval of convective/stratiform discrimination by other means, such as polarization or texture techniques, or incorporation of lightning observations. Close correspondence is found between deep convective profile frequency and annualized lightning production, and pixel-level lightning occurrence likelihood directly tracks the estimated mean ice water path within profile types.

1. Introduction

It is widely held that identifiable convective regimes exist in nature, although precise definitions of these are elusive. Examples span spatial and temporal scales, and include land–ocean distinctions (Zipser 1994), break/monsoon behavior (Rutledge et al. 1992; Williams et al. 1992), seasonal differences in the Amazon (Williams et al. 2002), and phases of the Madden–Julian oscillation (DeMott and Rutledge 1998b; Anyamba et al. 2000) and of tropical easterly waves (Petersen et al. 2003). These regimes are often described by differences in the realized local convective spectra and measured by various metrics of convective intensity, depth, areal coverage, and rainfall. These traditional metrics, when derived from radar reflectivity, typically involve reduced, scalar quantities thought to be important to convective dynamics: maximum radar reflectivity, cloud-top height, 30–35-dBZ echo top height, rain rate, etc. Individually, these metrics may be deficient, as their interpretation is often nonunique (the same metric value may signify different physics in different storm realizations or stages of storm evolution). Such metrics also fail to capture the coherence and interrelationships available in full vertical structure information available from volumetric radar datasets, although nonparametric analysis of their distributions may still be informative (Petersen and Rutledge 2001). Alternative means of data reduction that seek to preserve the information

Corresponding author address: Dr. Dennis J. Boccippio, National Space Science and Technology Center, NASA Marshall Space Flight Center SD-60, Marshall Space Flight Center, Huntsville, AL 35812.
E-mail: Dennis.Boccippio@nasa.gov

content of hydrometeor vertical structure may thus be warranted.

One alternative approach is the discovery of natural partitions of vertical structure in a globally representative dataset, or "archetypal" reflectivity profiles. A simple profile separation was performed by Liu and Fu (2001) and Fu and Liu (2001) using empirical orthogonal functions (EOFs), and L'Ecuyer et al. (2004) have studied rainfall relationship to vertical structure using a classification technique. In this paper, such a discovery is instead accomplished through cluster analysis of a very large sample [$O(10^6)$] of Tropical Rainfall Measuring Mission (TRMM) precipitation radar (PR) reflectivity columns. The rain-conditional and unconditional distributions of archetypal profile type frequencies at a given location and/or season provide a description of the local convective spectrum that retains vertical structure information. Such a taxonomy of profile types also allows evaluation of which vertical structures are most important to global rainfall and provides a possible link between empirical convective observations and related latent heating profiles [since synthetic radar reflectivity profiles are a feasible output of many cloud-resolving models (CRMs)]. This link may be critical in objective approaches to data assimilation of convective observations into forecast models. Such a taxonomy also allows evaluation of the nonuniqueness inherent in convective observables that only implicitly contain, or integrate over, vertical structure information, such as passive microwave brightness temperature or lightning flash rate. This nonuniqueness problem may be quite important, as global volumetric radar observations are rare [the TRMM PR, and possibly a future National Aeronautics and Space Administration (NASA) Global Precipitation Mission], while operational passive microwave observations are available from a constellation of platforms, and lightning observations are available from a variety of ground-based and, potentially, space-based platforms. The TRMM PR volumetric data thus forms a natural nexus through which to couple convective theory, models, and observations, with ultimate application to more common or cost-effective (albeit potentially lower information content) satellite-based observations.

This first paper in a series establishes, through cluster analysis, an objective, ordinally ranked and hierarchically clustered classification of radar vertical profiles. The classification is used to illustrate regional differences in annualized rain-conditional and unconditional convective spectra. The classification is then used to examine the profile types' frequency of occurrence and contribution to tropical rainfall, as well as their passive microwave and lightning properties. Subsequent studies will address regional and seasonal variability in these profile spectra, passive microwave vertical structure diagnosis, and rainfall estimation errors as a function of vertical structure, storm feature decomposition into cells using the profile classification, and objective classification of local convective regimes.

2. Scope and methodology

a. Philosophy

A nonhierarchical cluster analysis simply answers the question: Given a set n of p-parameter descriptions of individual cases or instances, find k natural clusters, or partitions, of these cases in the p-space The number of clusters requested, k, is arbitrary and must be prescribed, although iterative examination of analyses while varying k can often reveal when too few or too many clusters are sought. In the case of clustering reflectivity profiles, a reasonable goal is separation into clusters that appear to indicate different convective or microphysical states. In this study, cluster analysis is used as an empirical, quasi-objective means of multivariate data reduction (TRMM PR data are highly multivariate, including reflectivity at 80 vertical levels, surface rainfall estimates, convective/stratiform and bright band classifiers, etc.). From field experience, we know that the radar reflectivity vertical profile spectrum (in a highly multidimensional data space) does not consist of cleanly separated, spherical clusters (e.g., as a biological species data space might). Rather, it consists of a continuous sequence of deeper and more intense profiles, with possible frequency "bumps" (i.e., weak modes) corresponding to physically distinct warm rain profiles, glaciated midlevel profiles, and deep convective profiles, more or less monotonically declining in frequency, and a parallel branch of decaying stratiform types. Traditionally, rule-based (and often univariate) techniques are used to partition (reduce) convective data spaces for subsequent analysis (e.g., radar echo tops deeper than altitude X or brightness temperature colder than temperature Y). We posit that a more robust and physically meaningful data space reduction can be accomplished through cluster analysis of highly multivariate data.

With no a priori expectation of spherical data clusters, the application of cluster analysis to vertical radar data is thus philosophically a search for a *useful* data space reduction. There is thus also no objectively "optimal" cluster analysis design (such as choice and weighting of input parameters, selection of distance metrics, specification of the number of clusters sought, etc.)—or indeed any objective means of assessing "optimality." As with other techniques used for data re-

duction or modeling (e.g., specification of windowing parameters in spectral analysis, or basis functions in nonlinear regression), the model design is ultimately subjective (although guided by expert opinion), and success is qualitatively gauged by the coherence and usefulness of the results. However, we suggest that scientific analysis of convective observations using a model derived from highly multivariate data is significantly *less* sensitive to subjective expert opinion than, for example, analysis using univariate rules. In this study, we further mitigate subjectivity in the model design by seeking *more* radar profile types (clusters) than would be likely used in analysis and objectively grouping these into related families of types (using a secondary, hierarchical cluster analysis). Ultimately, the success of the technique must be (and is) gauged subjectively by the consistency of the classification in actual storm scenes and the consistency of regional convective spectrum variations with our a priori expectations.

The methodology employed here (profile-level analysis, discretized by TRMM PR 4-km diameter vertical columns) also places the results and conclusions within a specific context. Part of this context is driven by our interest in and focus on improving pixel-scale *retrieval* of physical quantities from orbital sensors (radar, passive microwave, and lightning). Inferences from these results related to convective physics must be interpreted within this context. Results from global analyses at "storm" (precipitation feature) scale (Nesbitt et al. 2000; Toracinta et al. 2002; Cecil et al. 2005) provide a complementary view of the tropical convective spectrum. Successful statistical analysis of global radar data at scales between the pixel and storm scales (i.e., the "cell" scale) has yet to be achieved, although we note that objective classification of pixel-level profile types may be a useful tool in eventual objective cell-scale decomposition.

b. Data sources

In addition to the precipitation radar, the TRMM platform hosts a passive microwave (PM) sensor [the TRMM Microwave Imager (TMI)], observing at frequencies including 10, 19, 21, 37, and 85 GHz, and an optically based total (intracloud and cloud-to-ground) lightning imager [the Lightning Imaging Sensor (LIS)]. The PR pixel resolution at nadir is 4.3 km; TMI pixel resolution at 85 and 37 GHz is approximately 7 km × 5 km and 16 km × 9 km, respectively (although there are gaps between adjacent TMI scans), and LIS resolution varies from 4 to 10 km across the PR swath portion of its field of view [the TMI/PR pixel sampling can be seen in Fig. 11 of Hong et al. (1999)].[1] The LIS has a diurnally varying detection efficiency of 73%–93% (Boccippio et al. 2002) and median dwell time of 83 s, corresponding to an ability to detect lightning occurrence if flash rates exceed approximately 1 flash (fl) min^{-1}. The TMI and LIS swaths both encompass the much smaller PR swath.

The PR and TMI pixel data are taken from the TRMM 1Z99 dataset [1Z99 is a new TRMM product, also referred to as the "University of Utah TRMM Precipitation Feature Dataset" (Nesbitt et al. 2000)]. The TMI pixel closest to each PR column is identified. Lightning data from the LIS v4.1 product are used; lightning flash and "area" (loosely, thunderstorm cell) radiance-weighted optical centroids are also paired to PR pixels. Closest prior 6-hourly National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis (Kalnay et al. 1996) data are also paired to each pixel.

c. Methodology overview

Identification of radar vertical profile types, and families of types, is accomplished in three stages. First, in the *training phase*, a "flat" (nonhierarchical) cluster analysis is performed on a subset of TRMM PR vertical profiles, using 47 observations for each profile, and seeking 25 clusters. This analysis identifies the 47-dimensional centroids of the distinct clusters. A similar cluster analysis approach was recently applied to tropical sounding data to discover archetypal sounding types (Lucas and Zipser 2000). Next, in the *full classification* phase, 3 yr of TRMM PR profiles are classified based on their proximity to these centroids, and the mean radar, PM, and lightning properties of profiles in each of the 25 clusters are recorded. Finally, in the *taxonomy construction* phase, these means (which include the original centroids) are used as the descriptors in a further hierarchical cluster analysis of the 25 types into similar families. Thus, radar-only measurements are used for primary profile *classification*, while radar, PM, and lightning measurements are used for *interpretation* (i.e., the taxonomy construction tasks of naming, ranking, and further aggregation of the profiles). The PM and lightning data were not used in the primary classification itself as we desire a scheme that can operate on *any* volumetric radar dataset (orbital, suborbital, or ground based), regardless of the availability of concurrent PM and lightning observations.

[1] Readers are cautioned to consider these issues of sensor resolution, undersampling, viewing angle, and parallax correction when interpreting conclusions drawn from "concurrent" multisensor measurements presented below.

d. Data filtering and preprocessing

Some filtering and preprocessing of TRMM PR profiles is performed prior to inclusion in the primary cluster analysis. First, only "warm-season" radar reflectivity profiles are considered, specifically those that occur when the near-surface atmospheric temperature $T_a >$ 10°C. NCEP reanalysis surface data are used to estimate this criterion for each PR column. While arbitrary, the warm-season restriction 1) helps exclude heavily sheared reflectivity profiles from strongly baroclinic winter systems, since this analysis is primarily locally vertical, and 2) helps preserve meaningfulness and consistency of vertical profile reflectivity data at low levels in the atmosphere, since the analysis is performed using temperature as a vertical coordinate (again using NCEP reanalysis data) rather than the native PR vertical coordinate altitude. During winter seasons, a significant number of PR columns in midlatitudes thus remain unclassified after the cluster analysis (22.3% of the total sample).

Second, only PR columns that participate in TRMM 1Z99 precipitation features (Nesbitt et al. 2000) are considered. 1Z99 precipitation features are collections of contiguous PR columns that are either precipitating or that exhibit TMI passive microwave ice-scattering signatures (hence adjacent high-reflectivity, nonprecipitating anvil overhang may be included). A noise filter is applied to columns suspected of containing spurious high reflectivity values aloft.[2] Note that no feature minimum size criterion such as that used by Toracinta et al. (2002) has been applied to the data in this study. Since the 47 inputs (described below) span different dynamic ranges and are in different (or no) units, they are all standardized to have zero mean and unit variance.

e. Training phase

During the training phase, a subset of all PR columns from December 1997 to November 2000 (2.5% randomly subsampled; 3.2 million columns) are used to identify the p-dimensional centroids of the profile clusters. Only columns passing the T_a criterion and near nadir (i.e., from PR rays 10–39) are used for cluster centroid identification, thus guarding against contamination by resolution issues, off-nadir viewing near the corners of the PR swath, and nonstationarity across the swath in some of the input parameters.

[2] Specifically, reflectivities greater than 25 dBZ in the original 80-level PR data, which also differ by more than 5 dBZ from both adjacent levels, are considered noise, rejected, and replaced by the average of the adjacent levels.

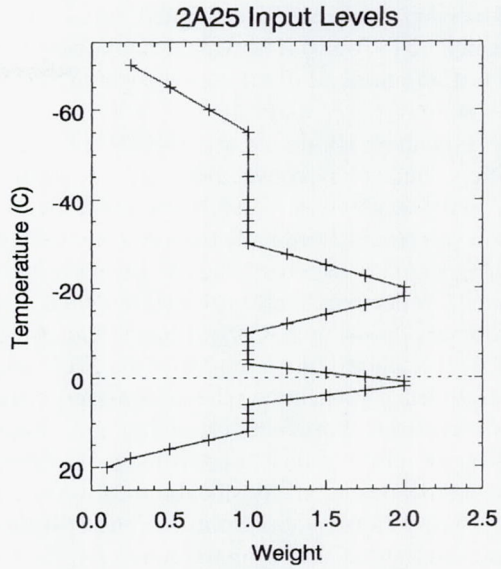

FIG. 1. Relative weights assigned to reflectivity at each input temperature level in the cluster analysis.

Radar reflectivity values (in dBZ units) at various temperature levels make up 40 of the $p = 47$ multiparameter descriptors for each column, taken from the TRMM 2A25 product (Iguchi et al. 2000). The chosen temperature levels, and relative weights γ assigned to them, are shown in Fig. 1. The distribution of temperature levels selected as inputs, and their weights, reflects our emphasis on radar measurements most likely to reveal significant kinematic or microphysical differences, specifically, near the melting level and in the lower mixed phase levels of storms. An artificial value of 0 dBZ is assigned to levels without radar echo, or with reflectivities less than the nominal 17-dBZ PR threshold. This value, while arbitrary, is sufficiently removed from the 17-dBZ threshold to implicitly strongly weight occurrence versus nonoccurrence at a given temperature level, while still allowing the dynamic range of reflectivity when echoes do occur (approximately 17–65 dBZ) to contribute information. In cases where any specified input temperature level is warmer than the local T_a, the lowest altitude reflectivity value is simply "copied down" the temperature profile, ensuring consistent behavior across a wide dynamic range of possible surface temperature and altitude. Since inclusion of nonphysical data has obvious drawbacks, we significantly underweight the warmest temperature level inputs in the cluster analysis; this also mitigates systematic bias that might be incurred by use of path integrated attenuation (PIA)-corrected reflectivity profiles from the 2A25 dataset (Iguchi et al. 2000; Meneghini et al. 2000). To offset this low-level underweighting, we include an additional input, the TRMM

2A25 version 4 algorithm estimated surface rainfall, overweighted ($\gamma = 2.0$). The wide dynamic range of this input further increases its effective weight in the cluster analysis.

The remaining six inputs are TRMM 2A23 rainfall classifiers: binary "is convective" and "is stratiform" flags, both weighted $\gamma = 0.5$, their respective confidence levels (ranked ordinally from 0–9 and weighted $\gamma = 0.025$), a binary "is other" flag (mostly corresponding to "anvil," reflectivity-aloft profiles), weighted $\gamma = 0.5$, and a binary "possibly has bright band" flag, weighted $\gamma = 1.5$. It is important to note that the 2A23 convective/stratiform classification scheme considers both vertical structure *and* horizontal variability; thus some spatial variance information is used to nudge the otherwise purely vertical column classification. However, since the 2A23 classifiers comprise only 5 of the 47 inputs and are underweighted, the primary convective/stratiform separation in the analysis derives directly from the vertical profiles themselves.

For the flat (nonhierarchical) cluster centroid identification, the Interactive Data Language (IDL) CLUSTER_WTS routine is employed; this identifies centroids in the very large dataset via a (user transparent) neural-network-based optimization engine. This routine assumes spherical clusters in the standardized data space (although our weighting of some parameters provides some ellipticity) and uses Euclidean distance as a distance metric. We consider the assumption of sphericity to be irrelevant as the underlying data themselves are at best weakly modal rather than true clusters (see discussion above). The technique simply provides a means of identifying data-driven multidimensional boundaries within the data space.

Analyses were attempted seeking $k = 2, 3, \ldots 40$ clusters; $k = 25$ yielded a reasonable trade-off between identification of "primary" profile types without an overabundance of clusters consisting of very rare/anomalous profiles.[3] Since anomalous profiles tend to be very far removed in the $p = 47$ space, it is important to give the analysis the freedom to cluster them during the training phase, thus allowing physically important and distinct but multidimensionally more proximate profile types to still be isolated. While the final selection of $k = 25$ was mostly subjective, this subjectivity is mitigated by our eventual aggregation of these clusters into related families during the taxonomy construction phase; that is, more clusters are identified than will actually be used for scientific analysis.

There are thus three subjective factors in this analysis: the parameters selected as inputs, the weights assigned to those parameters, and the final number of clusters sought. During the exploratory phase of the study, these were varied in a trial-and-error fashion using the probability distributions of the resultant reflectivity profile clusters for assessment. No *objective* optimization criteria (other than convergence of the iterative cluster analysis algorithm itself) were applied to the experimental design; the goal was to create a versatile and useful tool for partitioning the highly multivariate input parameter space, rather than one optimized for a specific application, for example, surface rainfall retrieval.

f. Full classification phase

During the full classification phase, the 47-dimensional centroids are used to classify a full 3-yr (December 1997–November 2000) dataset (212.5 million columns). During this phase, the outermost PR rays are included; we thus implicitly tolerate some off-nadir classification error in the final analyses, so long as the cluster centroids themselves are robust. The IDL CLUSTER routine is used; it simply classifies each profile based on its Euclidean distance to the nearest of the 25 centroids found during the training phase. The primary clusters found during the training phase and their properties computed during the full classification phase are presented in section 3a.

g. Taxonomy construction phase

Assignment into 1 of 25 arbitrarily numbered profile clusters is arguably not a significant improvement over 47 input parameters, or approximately 90 raw observations and derived parameters. To be useful as a means of data reduction, the clusters need to be qualitatively interpretable, carry meaningful names, exist in some sort of grouping or ordinal ranking relative to each other, and be capable of being consolidated into fewer, broader (but still physically meaningful) categories. As will be seen below, many of the clusters exhibit readily recognizable vertical structures, such as a series of increasingly deeper convective types, stratiform types, and anvil types. However, several are ambiguous, and it is not readily apparent where they exist in relation to the others in the 47-parameter data space.

To construct a cluster-naming scheme, we perform a second cluster analysis on the mean properties of pro-

[3] We are bounded at the high end of k by the size of the training sample; for k larger than 25, "outlier" clusters contained too few cases to yield coherent mean properties. For k lower than 25, several clusters clearly consisted of "composites" of clusters with visually distinct vertical structure in the $k = 25$ run. There is nothing intrinsically meaningful about $k = 25$, it simply "works" for the selected set of input parameters and training sample size.

files in the clusters themselves. These mean properties include the original 47-parameter centroids *and* observations not included in the training phase. These latter include the median 17-dBZ echo top temperature, the mean area and weighted area (defined below) of precipitation features in which profiles of each type occur, the mean TRMM TMI 85-GHz polarization-corrected temperature (PCT), mean PM 37-GHz PCT, mean 37–85-GHz depression, the likelihood that profiles of this type occur anywhere in a thunderstorm complex, and the likelihood that a thunderstorm center ("area") or lightning flash observed by the TRMM LIS occurs within 5, 10, and 15 km of profiles of each type.

With only 25 data objects to be clustered, a hierarchical clustering is feasible. Agglomerative hierarchical clustering was performed in the R statistical programming language. A number of different distance metrics (Euclidean and Manhattan) and linkage schemes (single, average, complete, weighted, and Ward) were tested, all yielding essentially the same results (assessed by inspecting standard dendrograms); the design selected for use was Euclidean distance with complete linkage. As will be shown in section 3b, the lowest-level pairs in this clustering are visually and physically similar profile types that aggregate in physically meaningful ways. A 3–4 character naming scheme was constructed in which the first letter denotes the primary profile class [convective (C), stratiform (S), mixed (M), or anvil (A)], and the second numeral denotes relative depth of the profile [1 for warm (tops near or warmer than 0°C), 2 for midlevel (tops in the mixed-phase region), 3 for deep (tops above the mixed-phase region), 4 for deep/near-tropopause depth]. The third letter and fourth numeral denote subtypes of this family within the dendrogram. Thus, for example, S_{1_b} denotes warm (1) stratiform (S), of subtype b, while $C_{3_{a_2}}$ denotes deep (3) convective (C) of subtype a_2. A color coding scheme was constructed to visually represent the clusters, in which convective profiles carry warm colors [varying with depth from purple (warm) to yellow (midlevel) to orange (deep) to red (very deep)], stratiform and mixed profiles take on cool colors [from blue (warm) to cyan (mid) to green (deep) to yellow-green (deep/mixed). Low-precipitation S and M variants have darker shades of the same hue, and anvil profiles have gray shades.[4]

[4] A byte-coding scheme was also devised to encode profile type for storage based on its location in the hierarchical clustering dendrogram, with most to least significant bits corresponding to successive branch "decisions" down the cluster tree.

3. Cluster analysis results

a. Primary profile types

The results of the full classification phase are shown in Figs. 2–4. Each cluster's plot shows the conditional frequency distribution of reflectivity at each temperature level. Characteristic values of three of the seven nonreflectivity descriptors are included (the percentage in each cluster classified by 2A23 as stratiform and convective, and as "bright band possible or certain"). In 14 of the 17 C and S series profiles, agreement with the TRMM 2A23 convective/stratiform classification exceeds 90%. Since the 2A23 classifiers make up only a few, strongly underweighted inputs to the cluster analysis, this can be seen as a quasi-independent corroboration of the 2A23 classification algorithm (section 2).

The rationale for the names given to each type will be presented below, although simple visual inspection of the reflectivity distributions conveys much. Within the convective C profiles, the C_1 family corresponds to warm convective profiles (C_{1_a}) and warm convective profiles with 17-dBZ echo tops colder than 0°C but little evidence of mixed-phase growth (C_{1_b}) (or at least, little contribution from mixed-phase growth profile rainfall). The C_2 "midlevel" family of convective profiles exhibits mixed-phase growth with tops from approximately −5° to −25°C (C_{2_a}) and −20° to −40°C (C_{2_b}). The C_3 family is clearly "deep" and will be loosely termed "garden variety" deep convective. The C_4 family consists of two unique types; C_{4_a} is not the deepest profile type (indeed, it is slightly shallower than C_{3_b}) but has the highest surface reflectivity and very high reflectivity in the mixed-phase region, with much less of an abrupt reflectivity decline at temperatures colder than 0°C than the other profiles (it lacks the "inflection point" at −10° to −15°C seen in other deep C profiles). We will loosely refer to this as the "wet growth" deep profile, an interpretation that will be corroborated by its prevalence in midlatitude and mesoscale convective system (MCS)-prone regions (section 5, below). Profile C_{4_b} exhibits very cold tops, approximately corresponding to tropical tropopause temperatures. This profile is thus loosely termed "deep tropical."

An interesting feature of the convective profile clusters is that they strongly suggest that significant regions of the data space are statistically "off limits" in nature. For example, convective profiles with echo tops near −40°C *imply* a fairly narrow range of "allowable" reflectivities lower in the profiles. This can be seen as "good news" for estimation of reflectivity values from CRMs, cross-comparison between CRMs and observa-

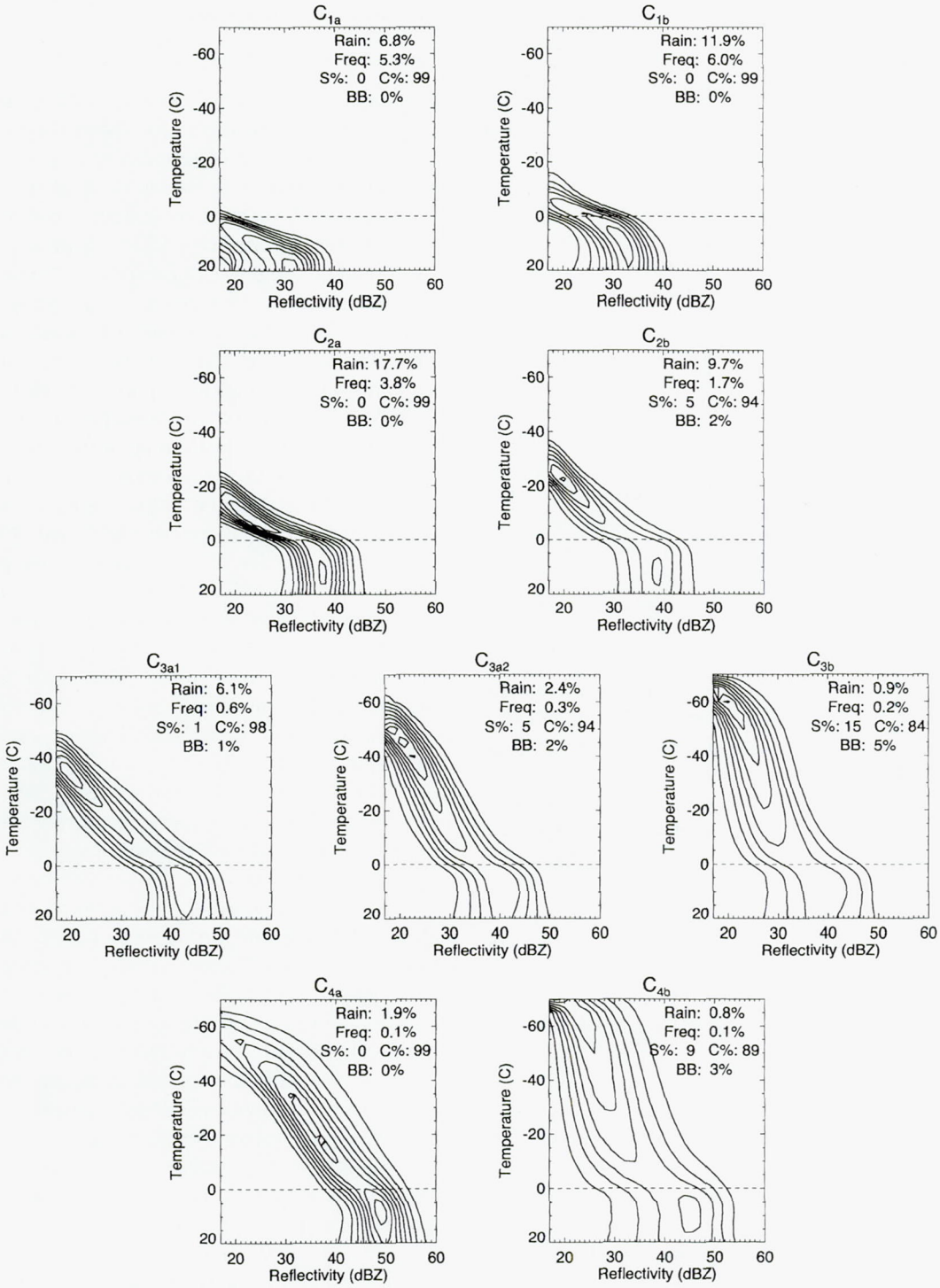

FIG. 2. Reflectivity vs temperature frequency distributions of the *C* series clusters. Contours are eight uniformly spaced levels, scaled to the peak frequency in each plot. Insets show each cluster's contribution to warm-season rainfall, frequency in the sample, and fraction classified by TRMM 2A23 as *S*, *C*, and possibly having bright band (BB).

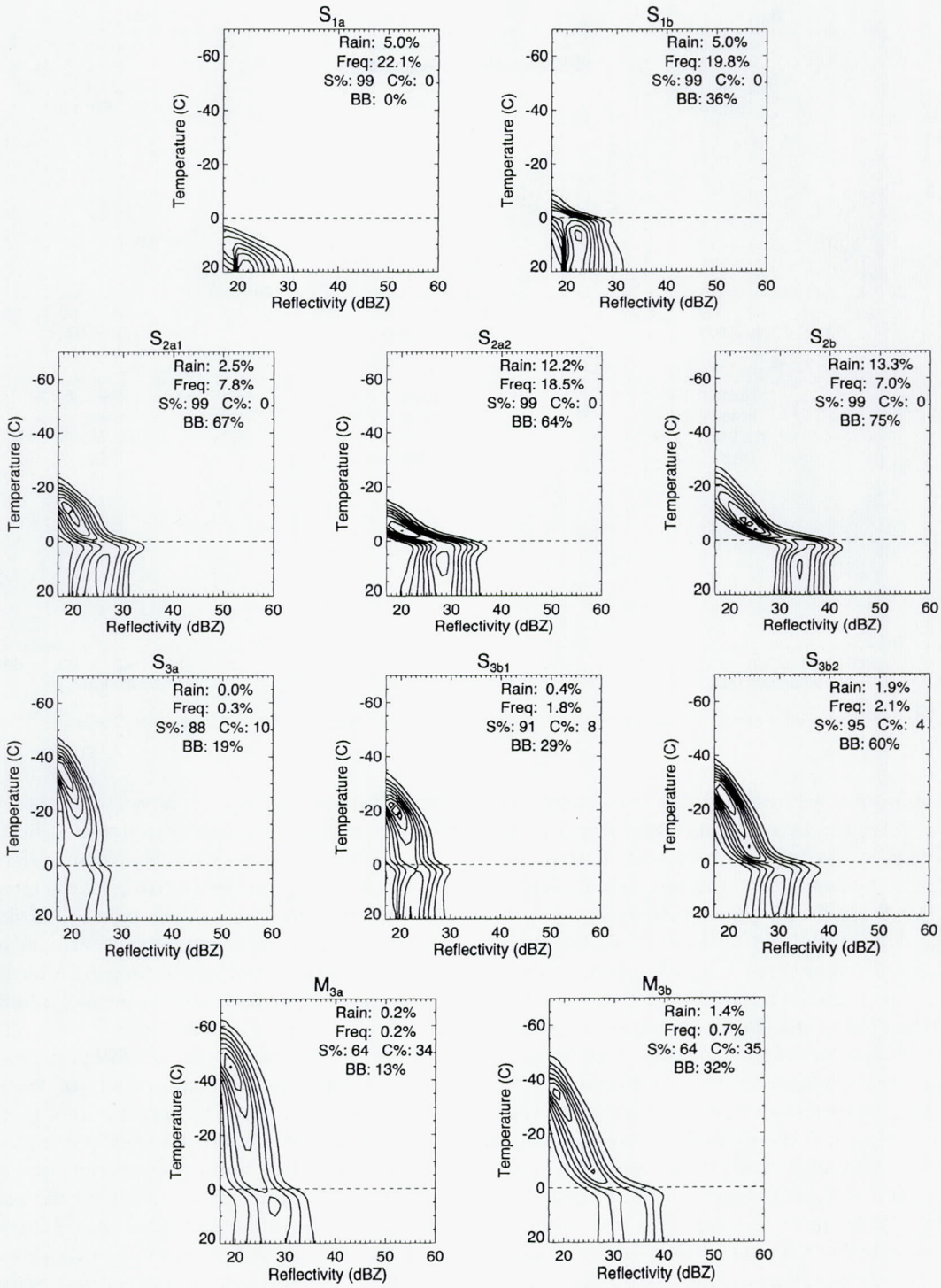

FIG. 3. Temperature–reflectivity frequency distributions: S and M series profile clusters.

tions, validation of CRM physical fidelity, or estimation of precipitation ice and precipitation water content from observed reflectivities using CRMs as guidance.

The stratiform S and mixed M profiles are more difficult to ordinally rank based only on visual inspection. The S_1 family includes what appears to be shallow stratocumulus (S_{1_a}) and stratiform rain with tops at 0°C but infrequent bright band S_{1_b}. The S_2 family includes

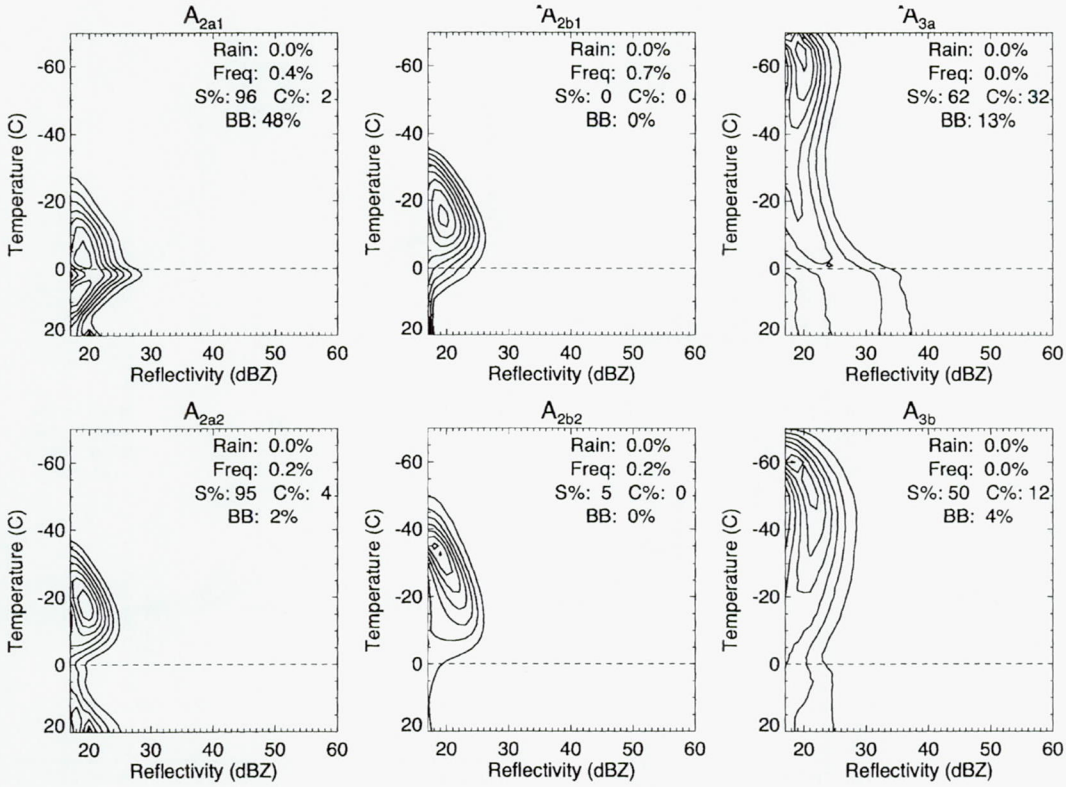

FIG. 4. Temperature–reflectivity frequency distributions: anvil/fragment (A) series profile clusters.

midlevel stratiform with bright band, with depths suggesting occurrence as decaying stages of isolated midlevel profiles, as components of weak, loosely organized oceanic systems, or as the rearmost-trailing regions of MCS stratiform precipitation; we term this family "garden variety cold stratiform." The S_3 family, alternatively, includes deep, low precipitation variants (S_{3_a}, $S_{3_{b_1}}$), resembling MCS transition zones, and a deep, high precipitation profile ($S_{3_{b_2}}$), similar to the leading regions of MCS stratiform precipitation. The M_3 mixed profiles are more ambiguous; both are more often classified by 2A23 as stratiform than as convective but have relatively low incidence of bright band. As will be shown below, they share many similarities with S_3 profiles, and together S_3 and M_3 are loosely termed "MCS stratiform." While the *interpretation* of M_3 profiles is unclear, we note that they nonetheless exhibit a narrow range of reflectivities along their profiles and do appear to comprise distinct classes of profiles.

The six anvil A types are comparatively rare and include purely aloft anvils, overhanging anvils, and fragments of sheared convective or stratiform profiles along the edges of precipitation features.

Additional radar, passive microwave, and lightning characteristics are shown in Tables 1 and 2 (characteristics of rare A types are not reported, for brevity). The additional radar parameters are shown in bubble plots in Fig. 5. In Fig. 5a, each type's median surface reflectivity is plotted against its median echo top temperature (bubble size here denotes the type's contribution to total rainfall). The ordering of convective, mixed, deep stratiform and midlevel stratiform pairs is loosely reminiscent of a "growth/decay" sequence, adding confidence to the naming convention.

In Fig. 5b, the mean area of 1Z99 precipitation features containing at least one profile of the specified type (each feature area is counted once if a profile occurs anywhere in it) and the weighted mean area of 1Z99 precipitation features containing the specified profiles (each feature area is counted the number of times the profile is found within it) are plotted (bubble size here denotes each profile's total frequency of occurrence). As expected, the deeper and more intense profiles tend to occur in larger features. The weighted mean area metric provides additional information: it peaks for midlevel profiles, both convective and stratiform, suggesting that when features containing these profile types occur, the types comprise significant fractions of the features themselves. This observation is important, as midlevel convective and stratiform pro-

TABLE 1. Characteristics of the convective (C) family profile clusters. Echo tops and surface reflectivity are medians; areas and brightness temperatures are means.

Parameter	C_{1_a}	C_{1_b}	C_{2_a}	C_{2_b}	$C_{3_{a_1}}$	$C_{3_{a_2}}$	C_{3_b}	C_{4_a}	C_{5_b}
2A23 convective (%)	100.0	100.0	99.5	94.9	98.0	94.8	84.1	99.0	89.9
2A23 stratiform (%)	0.0	0.0	0.5	5.1	2.0	5.2	15.9	1.0	9.8
2A23 other (%)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3
2A23 bright band (%)	0.0	0.2	0.5	2.4	1.8	2.8	5.3	1.0	3.6
17-dBZ echo top (°C)	3	−6	−16	−27.5	−37.5	−50	−60	−55	−60
Surface reflectivity (dBZ)	30	32	37	38	43	40	38	49	42
Feature area (km^2)	37	65	133	231	332	360	417	446	481
Feature area, weighted (km^2)	223	555	1056	1132	1017	844	756	953	766
85-GHz PCT (K)	280.0	270.3	256.3	238.1	219.6	200.0	180.7	177.9	156.4
37-GHz PCT (K)	283.0	279.0	273.8	269.7	264.4	261.1	256.6	244.3	245.0
37–85 (GHz) depression (IWP; K)	3.0	8.7	17.4	31.5	44.8	61.1	75.9	66.3	88.6
In thunderstorm complex (%)	7.4	16.5	30.9	52.3	72.3	72.6	75.5	96.7	84.9
Thunderstorm <5 km (%)	0.2	0.5	1.3	4.1	10.8	11.1	11.7	24.0	15.9
Thunderstorm <10 km (%)	0.7	1.7	4.6	12.9	30.6	32.7	35.2	62.1	45.0
Thunderstorm <15 km (%)	1.5	3.6	8.3	21.4	45.9	50.3	54.9	83.3	66.9
Flash <5 km (%)	0.3	0.9	2.7	8.8	25.6	28.2	33.5	67.8	48.9
Flash <10 km (%)	1.1	2.7	6.9	20.0	46.9	51.2	56.2	89.1	69.9
Flash <15 km (%)	1.9	4.7	10.6	26.9	55.1	59.4	64.4	92.8	76.7

files contribute a dominant portion of warm-season rainfall section (Fig. 5a; sections 3b and 6 below).

b. *Profile families*

As discussed in section 2g, the 25 primary profile types are hierarchically clustered, using their 47-parameter radar centroids and the additional radar, PM, and lightning parameters shown in Tables 1 and 2 as descriptors (all values are standardized prior to analysis). The results (using agglomerative clustering, a Euclidean distance metric, and complete linkage scheme) are shown in Fig. 6 as an "enhanced" dendrogram. Branch cuts on the dendrogram occurring at larger height values correspond to greater separation of the agglomerated clusters in the input parameter space (thus, the greatest multidimensional separation occurs between the "deep convective" branch and the "everything else" branch, while the most similar components are the $S_{2_{a_1}}$ and $S_{2_{a_2}}$ individual profiles). Obviously the results of the analysis depend upon the selected set of descriptors, although that

TABLE 2. Characteristics of the stratiform (S) and mixed (M) family profile clusters.

Parameter	S_{1_a}	S_{1_b}	$S_{2_{a_1}}$	$S_{2_{a_2}}$	S_{2_b}	S_{3_a}	$S_{3_{b1}}$	$S_{3_{b_2}}$	M_1	M_2
2A23 convective (%)	0.2	0.1	0.7	0.0	0.0	10.6	8.4	4.3	34.6	35.5
2A23 stratiform (%)	99.7	99.9	99.3	100.0	100.0	88.0	91.2	95.7	65.0	64.5
2A23 other (%)	0.2	0.0	0.0	0.0	0.0	1.4	0.3	0.0	0.4	0.0
2A23 bright band (%)	0.5	36.0	67.4	64.6	75.4	19.5	29.4	60.6	13.8	32.8
17-dBZ echo top (°C)	8	−1	−8	−14	−18	−10	−16	−27.5	−30	−35
Surface reflectivity (dBZ)	23	23	25	28	34	21	23	29	27	34
Feature area (km^2)	14	32	157	77	293	476	286	403	415	399
Feature area, weighted (km^2)	259	990	1603	1634	2100	1041	1180	1520	824	1404
85-GHz PCT (K)	280.6	270.1	256.7	262.2	245.2	233.0	248.1	236.9	211.1	222.3
37-GHz PCT (K)	282.7	277.4	273.0	274.0	268.1	271.0	272.7	269.0	266.7	266.4
37–85-GHz depression (IWP) (K)	2.0	7.3	16.2	11.8	22.9	38.0	24.6	32.1	55.5	44.1
In thunderstorm complex (%)	3.7	9.5	21.8	15.8	30.2	67.0	42.1	54.8	62.0	61.2
Thunderstorm <5 km (%)	0.0	0.1	0.2	0.1	0.2	1.8	0.7	1.0	4.2	3.1
Thunderstorm <10 km (%)	0.1	0.3	0.7	0.3	0.9	7.7	3.2	4.1	16.2	11.0
Thunderstorm <15 km (%)	0.4	0.7	1.7	0.8	2.0	17.8	7.3	8.9	32.2	20.9
Flash <5 km (%)	0.1	0.1	0.4	0.2	0.5	5.5	1.8	2.6	13.5	8.0
Flash <10 km (%)	0.2	0.5	1.3	0.6	1.6	17.0	5.9	7.9	32.2	20.3
Flash <15 km (%)	0.5	1.0	2.5	1.3	2.9	28.8	10.8	13.4	44.2	29.1

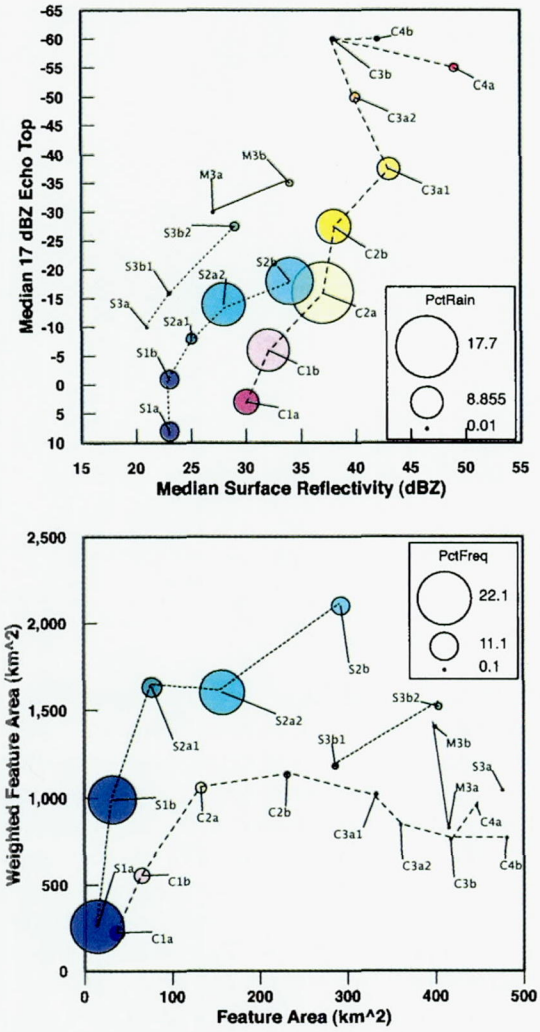

FIG. 5. Radar characteristics of the profile clusters. (a) Median surface reflectivity and echo top temperature; bubble size denotes profile contribution to net rainfall. (b) Median area of features containing each profile type; unweighted (each feature area counted once per profile occurrence) and weighted (each feature area counted the number of times the profiles occur within it). Bubble size denotes relative frequency of each profile type in the entire dataset.

set is fairly comprehensive and includes some nonlocal information such as the containing-feature area.

The enhanced dendrogram encodes two additional pieces of information. The *width* of each branch or subbranch corresponds linearly to the total rainfall contribution from profiles within that branch. The *color* of each branch corresponds to the frequency of occurrence of profiles within that branch (varying on a cool-to-warm color rainbow color scale). Thus, for example, within the midlevel branch, C_2 and S_3/M_3 profiles occur with comparable frequency, but C_2's contribute significantly more rainfall to the total. For emphasis (as it will be used repeatedly), the color table used to code profile types is included at left and in the profile labels.

After the deep convective–everything else split, the A_3 and A_{2_b} anvil types are broken off, consistent with their somewhat anomalous structure. The next major split is then "warm" (C_1 and S_1) versus midlevel (C_2, S_3, M_3, and S_2). Within the warm category, C_{1_a} and C_{1_b} pair closely together (which is why C_{1_b} was named as it was, rather than being included in a midlevel family, despite having echo tops colder than 0°C; a similar rationale holds for S_{1_b}). Within the midlevel category, midlevel convective C_2 is more similar to the MCS stratiform S_3 and mixed M_3 families than to the S_2 garden variety stratiform with bright band family. Similarly, S_3 and M_3 pair together in their own family. The precise ordering of profile types and substructure of branches within the S_3/M_3 family is the *only* dendrogram feature that varied when different distance metrics or linkage schemes were tested, suggesting some ambiguity in their interpretation. Fortunately, net rainfall contributions from these profiles are small, although $S_{3_{b_2}}$ and M_{3_b} profiles do occur with some frequency.

As an example of the importance of this objective, hierarchical clustering approach, our initial inclination (prior to this analysis) was to include the $C_{3_{a_1}}$ profile with the midlevel profiles, as its echo tops occur near the top of the mixed-phase region (−30°C to −50°C). However, within the PR/TMI/LIS data space considered, this profile type clearly has far more in common with other deep convective profiles than it does with midlevel convective profiles. This distinction is important, as $C_{3_{a_1}}$ contributes the most rainfall of the deep convective profiles, and misclassifying it as midlevel would nontrivially skew the "warm/midlevel/deep" rainfall decompositions to be presented below.

The hierarchical clustering is performed on the mean or median radar, passive microwave, and lightning characteristics of profile types. It is also instructive to view the identified profile types and hierarchically clustered families within the original data space. Figure 7a renders the profile type distribution within the first three principal components of the (unweighted) 47-parameter PR profile training dataset. In this plot, the profiles' color table entries [red/green/blue (RGB) values] are "blended" at each principle component (PC) pair weighted by their frequency of occurrence at that pair. Each profile type cluster is well bounded in either the PC1, 2 or PC1, 3 pairs. Additionally, profile families are proximate in ways consistent with the hierarchical clustering (i.e., similar-shaded C_1, C_2, C_3, C_4, S_1, S_2, S_3, M_3, and A regions are adjacent).

While physical interpretation of principal components is imprecise, it appears that PC1 loosely corre-

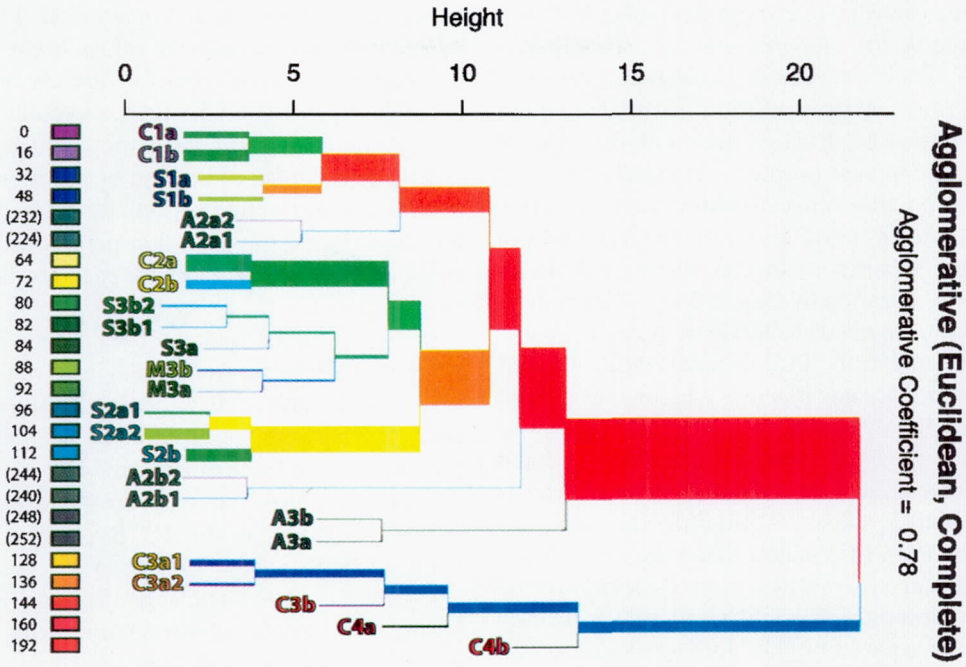

FIG. 6. Agglomerative hierarchical clustering of the 25 profile types into families. The level ("height") of each branch cut denotes distance in the multiparameter descriptor space. Branch width denotes the relative contribution to total rainfall from profiles in each branch, while branch color denotes relative frequency of profiles in each branch in the overall sample. The color table used to render profile types is shown at left and in the profile labels. A most-to-least significant bit byte coding for profile types is included at far left for reference.

sponds to a "depth/rainfall/intensity" mode of variability, PC2 loosely to a midlevel mode (C_{1_b}, C_2, S_{1_b}, and S_2 exhibit positive PC2 values; warmer and colder profiles are negative in PC2), and PC3 loosely to a "convective/stratiform" separation. Among the warmer profiles, PC3 is required to distinguish between C_1 and S_1 (the families occupy the same domain in the PC1, PC2 decomposition), consistent with the occurrence of their branch cut at a relatively low height value on the hierarchical clustering dendrogram and the known difficulties in separating warm stratiform and convective profiles in the TRMM dataset (Schumacher and Houze 2003a,b).

The actual frequency distribution of radar observa-

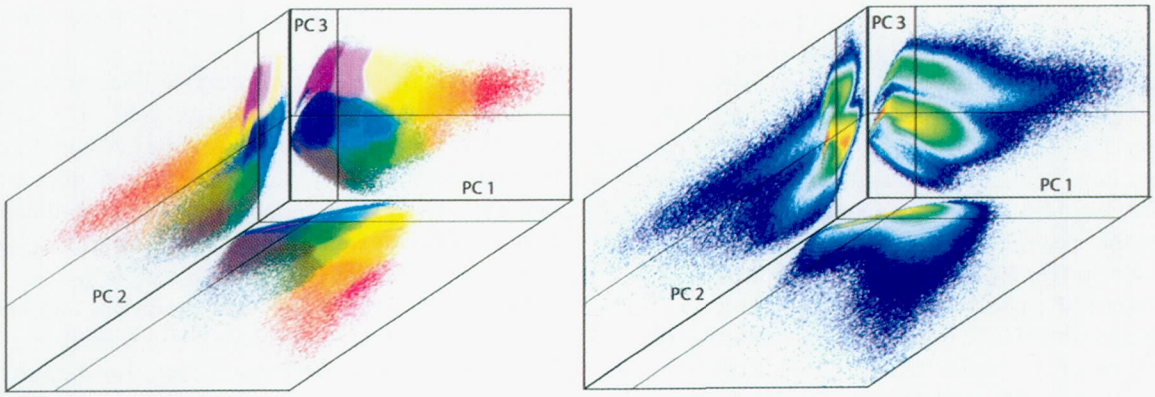

FIG. 7. (a) The PR training data, projected onto its first three PCs and shaded based on the columns' assigned profile clusters. Each point on the grids is shaded with a blend of the profile color table entries for profiles occurring at that locus, with colors weighted by the frequency of occurrence of each profile type. (b) Same PC projection, with logarithm of observation frequency contoured (irrespective of profile type).

tions within the (unweighted) input data space is shown in Fig. 7b (note that the *logarithm* of frequency is contoured). To a certain extent, this visualization corroborates our a priori conceptualization of the data space presented in section 2a, that is, a weakly modal series of increasingly deep/intense profiles with rapidly declining frequency and parallel convective/stratiform branches. The frequency distribution also corroborates some of the key profile separations in the cluster analysis. In particular, the separation of C_3 and C_4 profiles is justified by the "kink" in the distribution in the PC1, 3 panel and in the long tail in the PC2, 3 panel. While difficult to discern given the color schemes and angular projection, warm C_1 family and S_1 family profiles comprise single modes in the PC1, 2 space, but in the PC1, 3 and PC2, 3 spaces weak, separate modes exist for C_{1_a} versus C_{1_b} and S_{1_a} versus S_{1_b}. There is also little evidence (particularly in the PC1, PC3 panel) that a different choice of primary cluster analysis design or algorithm (using these input parameters) would yield a significantly "better" profile separation; the types *appear* meaningfully separated in physically interpretable ways. We recall the original premise that subjectivity in the classification scheme design is tolerable so long as the results are *useful*. Further evidence of the utility of the scheme is presented next.

4. Example storm complexes

Six examples of PR column classification are shown in Fig. 8. The scenes include four insets: the profile classification, shaded using the color table from Fig. 6 (left insets); the 17-dBZ echo top temperature (center left); the near-surface reflectivity (center right); and an RGB rendering of 37-GHz PCT (red), 85–37-GHz PCT depression, which loosely corresponds to total ice water path (blue) and total lightning optical pulse count (green) (right). In the RGB panels, warm rain appears as red areas, cold rain as magenta, and weak rain with cold tops (or anvil) as blue, while lightning "whitens" regions by adding in green. Approximately 13 000 such scenes have been rendered and examined, centered on the largest TRMM 1Z99 precipitation feature in each orbit; the scenes shown in this figure and the next were selected to represent "interesting" features.

Overall, the classification results show good agreement and consistency with traditional radar metrics and strong coherence within scenes. Crudely, the classification can be considered an "echo tops plus" product; greater coherence is demonstrated with radar echo tops than with surface reflectivity. This, of course, is by design, as the cluster analysis was designed with vertical structure separation as a primary goal. The fact that a range of surface reflectivity values occurs in the plots for individual profile types (as already shown in Figs. 2–4) is not problematic; rather, a hypothetical application of the classification scheme could be to improve radar-based rainfall estimation by partitioning observations [i.e., constructing reflectivity–rainfall (Z–R) relationships] based on vertical structure, in which case a range of surface reflectivity for each profile type would be *required*.

In Figs. 8a,b, the context of M_{3_b} family profiles (light green in the classification insets) relative to parent convection is illustrated; these surround active convective cores and smoothly transition to deep stratiform S_{3_b} (green) and midlevel stratiform S_2 (cyan). Anvil A series profiles surround the cores and correspond to "bluer" regions in the RGB passive microwave/lightning insets.

Figure 8c illustrates a continuous gradation of profile types. Scattered warm convective profiles occur in advance of an organized convective line, which itself smoothly transitions from warm to midlevel to deep convective profiles. Behind the convective line, an $S_{3_{b_2}}$ to $S_{3_{b_1}}$ to A transition occurs on the right flank; on the rear flank, an $M_{3_b}/S_{3_{b_2}}/S_{2_b}/S_{2_a}/S_{1_b}$ gradation occurs.

Figure 8d illustrates a case where weak lightning occurs in the trailing stratiform region composed of M_3 and S_3 profiles, but not in the regions dominated by S_2 profiles (despite the fact that these latter exhibit greater surface reflectivity). A number of other scenes were observed in which S_3 or even scattered C_2 occurrence in trailing stratiform regions corresponds to low flash rate lightning occurrence. In the RGB plots, the light green/light blue shading is characteristic of these cases (modest rain, infrequent lightning, and some ice), suggesting a TMI + LIS signature for deep trailing stratiform regions.

Figure 8e illustrates a convective line comprised of smaller, more isolated convective cells. The column classification here seems to reveal the cellular structure more comprehensively than either the echo tops or surface reflectivity fields alone. Weak embedded convection is also suggested in the core of the trailing stratiform region.

In Fig. 8f, the organized convection has detrained a significant anvil shield. In the echo tops inset, the anvil is shown to exhibit the coldest echo tops (color table "saturation" at black). Apparently, this anvil extends well ahead of the convective cluster at reflectivities near or below the radar threshold: in the shallow line ahead of the complex, half the line "incorrectly" exhib-

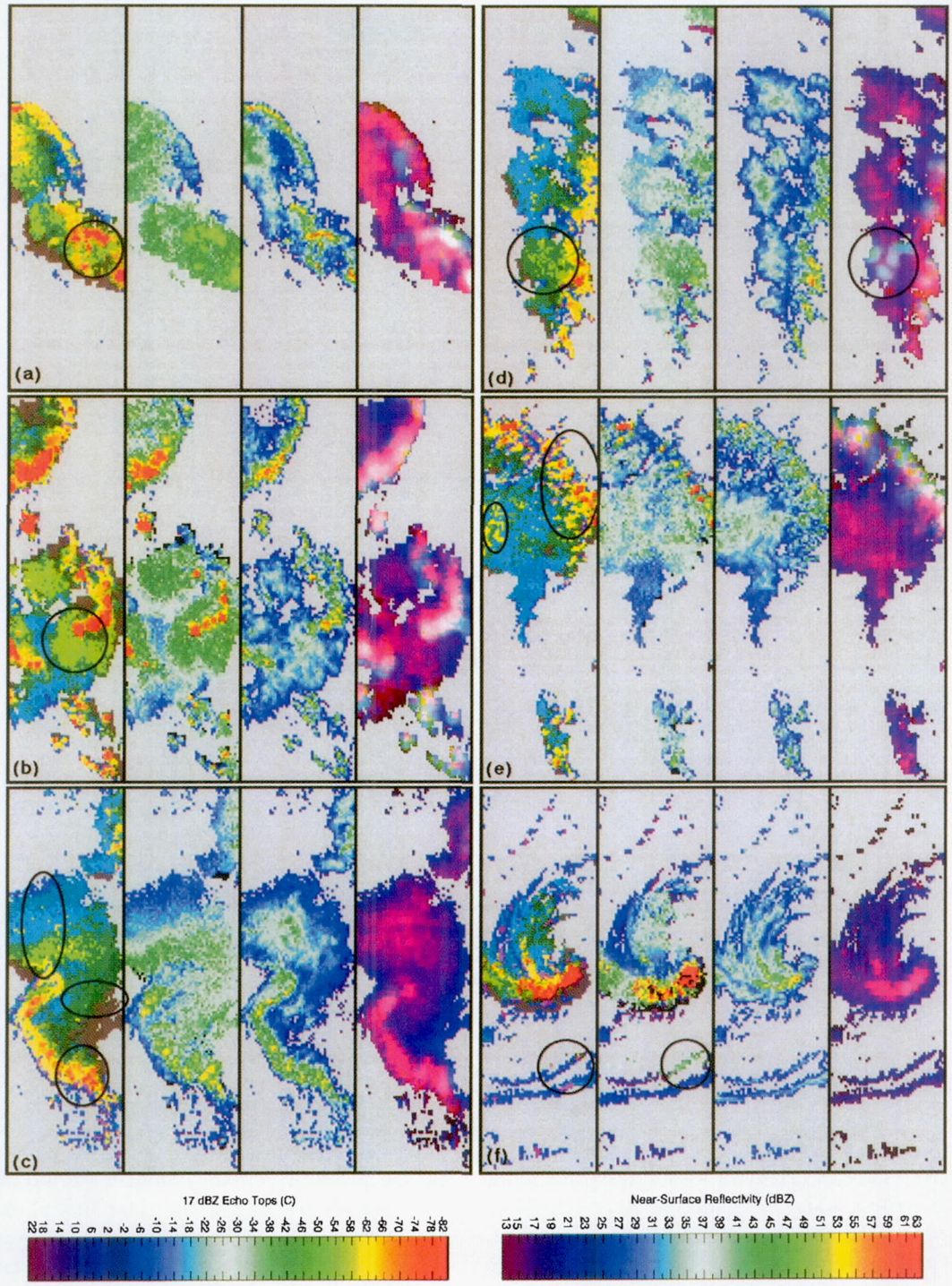

FIG. 8. Sample scenes of large precipitation features. In each scene, the insets correspond to (left) profile type classification, using the color scheme shown above; (left center) 17-dBZ radar echo top temperature; (right center) near-surface reflectivity; and (right) passive microwave/lightning RGB composite, with the red channel mapping 37-GHz PCT, the blue channel mapping the 37–85-GHz PCT difference (loosely, IWP), and the green channel mapping lightning optical pulse frequency. Features discussed in the text are highlighted by enclosing circles.

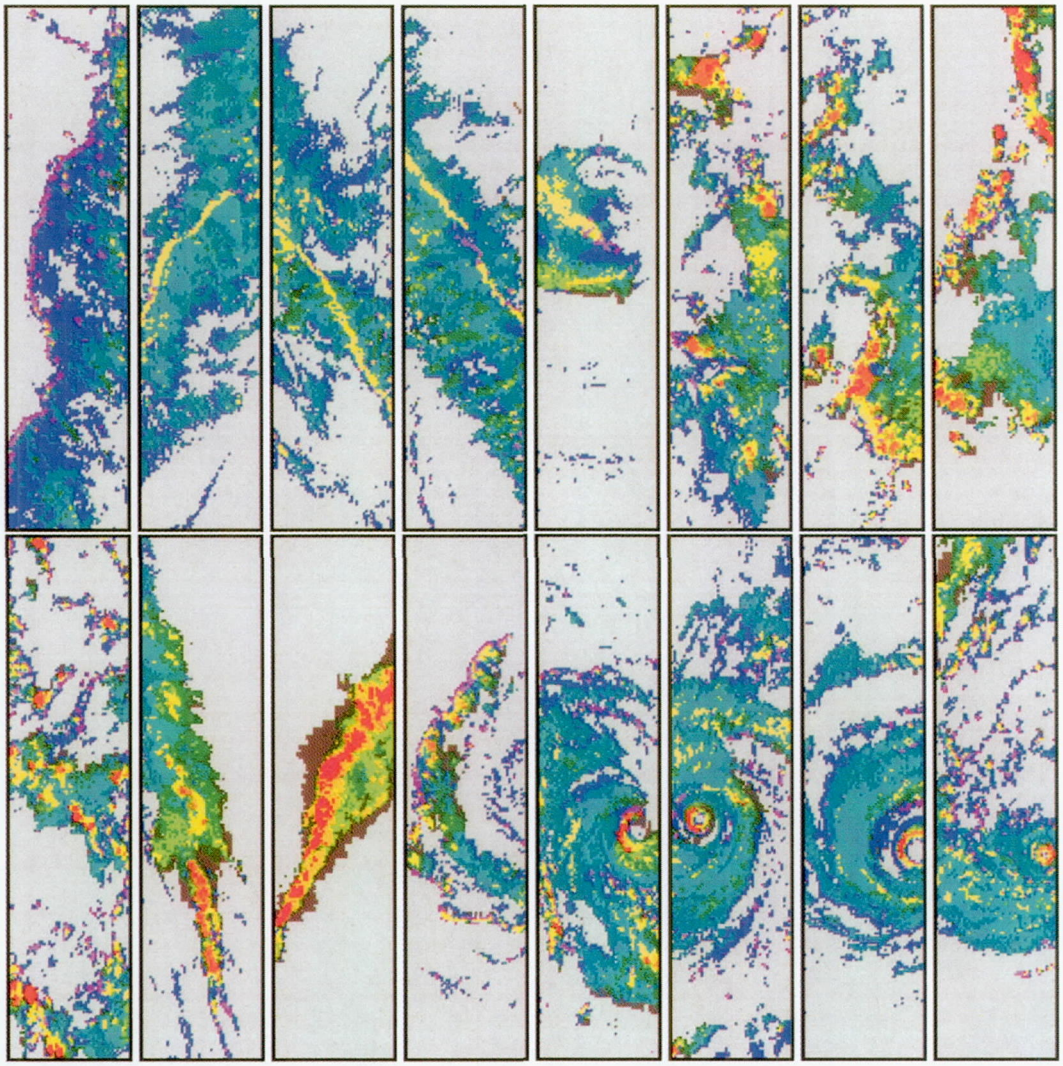

FIG. 9. Additional sample scenes (profile classification only).

its fairly cold echo tops. Interestingly, the classification scheme nonetheless correctly identifies these profiles as warm convective and stratiform rain.

Figure 9 shows 16 additional large precipitation feature scenes (profile classifications only), including five cases with a warm or midlevel convective line embedded in large stratiform regions, four cases of disorganized or weakly organized deep convective outbreaks, three additional deep convective line cases, and four cases of tropical cyclonic features with deep convection occurring on one side of the eyewall. Again, the coherence of profile classification in these scenes, and its visual consistency with known patterns of convective mesoscale organization, is presented as purely subjective "validation" of the usefulness of the technique.

5. Geographic distributions

a. Conditional frequency and rainfall spectra

The rain-conditional frequencies of occurrence of selected profile types and families, on an annualized basis, are shown in Fig. 10. The first three panels show a warm (C_1 + S_1)/midlevel (C_2 + S_2 + S_3 + M_3)/deep (C_3 + C_4) decomposition, corresponding to a vertical cut at a height of about 11 on the dendrogram in Fig. 6 (ignoring the A types and "overriding" the separation of C_{4_b} into its own category at this level). The remaining panels show a decomposition corresponding to a vertical cut at heights between 4 and 6 on the dendrogram.

The warm–midlevel–deep decomposition shows the known dominance of warm profiles over oceans and

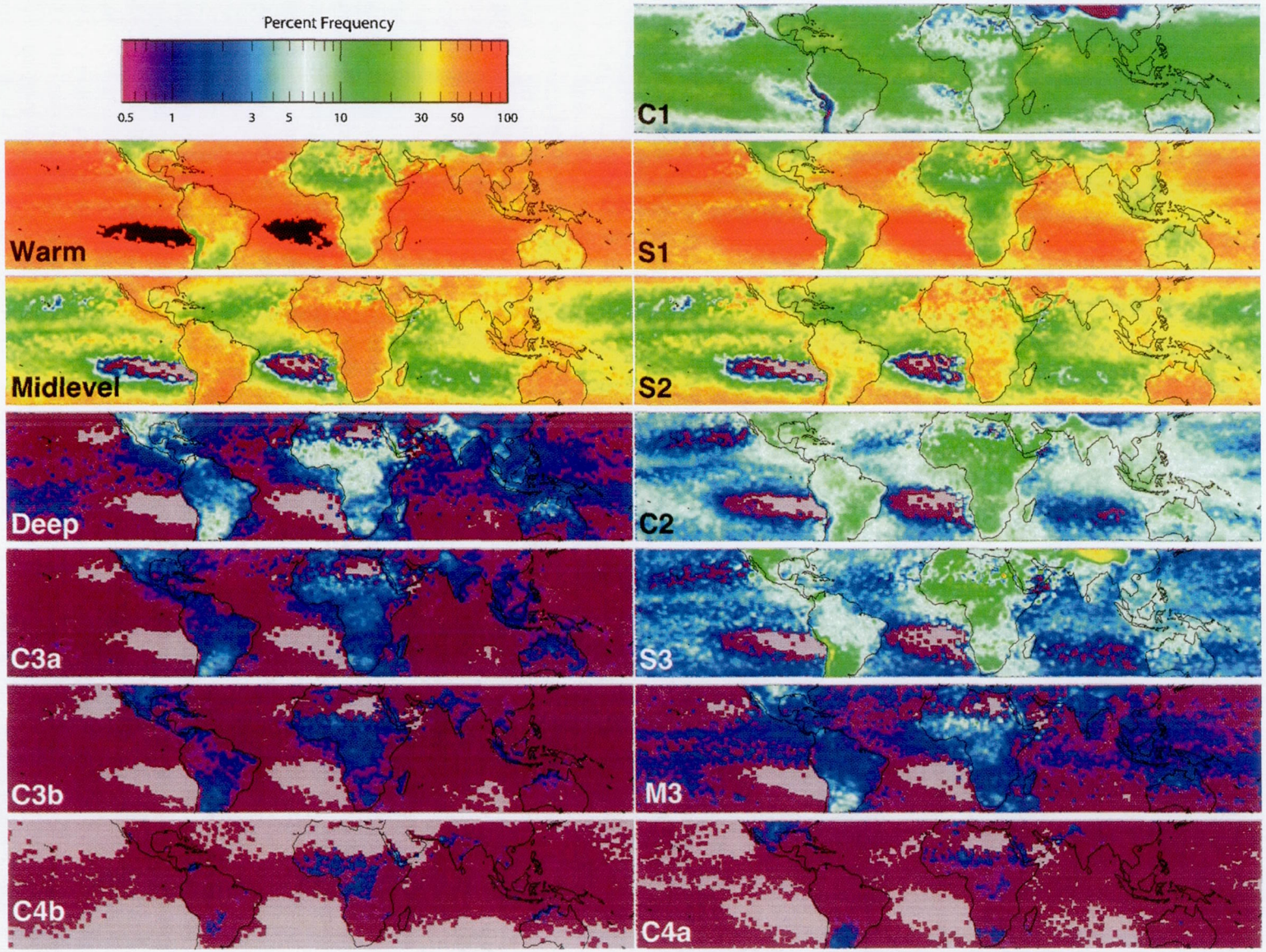

FIG. 10. Annualized rain-conditional frequencies of selected warm-season profile types and families.

their relative depletion in the conditional spectra over continents (most severe at midlatitudes and over Africa). This depletion is less over the Amazon, India, southeast Asia, and the Maritime Continent, which appear "intermediate" between continental and ocean frequencies. The separation into C_1 and S_1 warm families shows that the warm depletion over Africa occurs for both convective and stratiform types, while for the Amazon, India, and the Maritime Continent, it occurs only for stratiform types, and over southeast Asia, it is barely evident (indeed, southeast Asia seems unique among continental regions by having near-oceanic relative frequencies of S_1 profiles). The S_1 profiles also dominate the oceanic warm profile spectrum. Notable in all three panels is the "intrusion" of oceanic warm profile frequencies into eastern coastal areas of South America and Africa. Also notable is the sharp contrast in warm rain frequencies between the Red Sea and adjoining land areas.

The warm rain continental depletion is balanced primarily by increases in midlevel profile frequency (and to a lesser extent, deep profile frequency). The land/ocean distinction is sharpest for C_2, S_3, and M_3 profile families. Again, the Amazon, India, southeast Asia, and the Maritime Continent appear intermediate between land and ocean spectra. The S_3 and M_3 families—which we have previously termed MCS deep stratiform—comprise relatively high portions of the profile spectra in midlatitudes and in the sub-Saharan Africa region, consistent with interpreting them as characteristic of MCSs.

The relative frequency of deep convective profiles is elevated, as expected, over tropical continents, consistent with global lightning and convection studies (Orville and Henderson 1986; Boccippio et al. 2000, Nesbitt et al. 2000; Toracinta and Zipser 2001; Petersen and Rutledge 2001; Christian et al. 2003) and particularly in the same midlatitude and African MCS regions. This is also expected as the frequencies are rain conditional; in these regions, profile occurrence is less frequent (more "suppressed," or lower duty cycle) but more explosive (greater relative frequency of deep convective types). Deep convective profiles over the Amazon are relatively depleted, consistent with some interpretations of the region as a "green ocean" (Petersen and Rutledge 2001; Williams et al. 2002).

Figure 11 shows the same profile type/family decomposition, but with relative contribution to rainfall (rather than relative frequency of occurrence) contoured. Among the warm profiles, the relative contribution to rainfall is more "balanced" between C_1 and S_1 types than their relative frequencies. Over oceans, the extremely high frequency of S_1 profiles offsets their very low rain rates and makes these profiles "competitive" with the C_1 types. Over Africa and in midlatitudes, S_1 profiles contribute negligibly to total rainfall. Interestingly, over the Amazon, S_1 profiles contribute as much rainfall as the deepest convective types (C_{3_a}, or C_{3_b} and C_4 combined). A similar situation exists over India for S_1 and $C_{3_b} + C_4$ profiles. This is likely due to enhanced S_1 occurrence during these regions' monsoonal seasons (December–February and June–August, respectively).

Outside of cold ocean gyre regions, rainfall contribution is dominated by midlevel profiles. Indeed, the midlevel profile contribution exhibits remarkably little variability between land and ocean (notable given the significant land/ocean differences in relative frequency of occurrence). The contributions are again balanced between convective (C_2) and stratiform (S_2) types, with very little contribution from MCS stratiform S_3 and M_3 types (while comparable in relative frequency to C_2 profiles, their lower rain rates significantly diminish their net rain contribution).

As expected, deep convective profiles contribute significantly more to net rainfall over continents than over oceans, with the greatest relative contributions occurring in midlatitudes and over the Congo and sub-Saharan Africa. However, a decomposition of the deep spectrum is instructive. The "MCS region" deep convective dominance is least pronounced in the garden variety C_3 and deep tropical C_{4_b} types, and most pronounced in the wet growth C_{4_a} type. The frequency and rainfall maps of S_3, M_3, and C_{4_a} all support the interpretation of these families as preferentially occurring in MCSs, and in a "disproportionate" contribution to rainfall frequency and amount in these regions by mesoscale systems. An interesting contrast is observed in Africa north of the equator: just north of the MCS region latitudes, extending into the more arid sub-Saharan and Saharan regions, C_2 profiles contribute a large portion of the net rainfall. This represents a transition from organized systems to more isolated, midlevel convective outbreaks in the more arid regions.

b. *Unconditional frequency spectra*

The conditional spectra discussed above mask significant regional variations in absolute (unconditional) rainfall occurrence. The unconditional profile variability is illustrated by decomposing the (annualized) local spectra into a series of three-parameter descriptors, and rendering these in three-channel RGB composites. The selection of profile families for these decompositions is again determined by vertical "cuts" on the dendrogram in Fig. 6.

Four RGB composites are shown in Fig. 12. In each

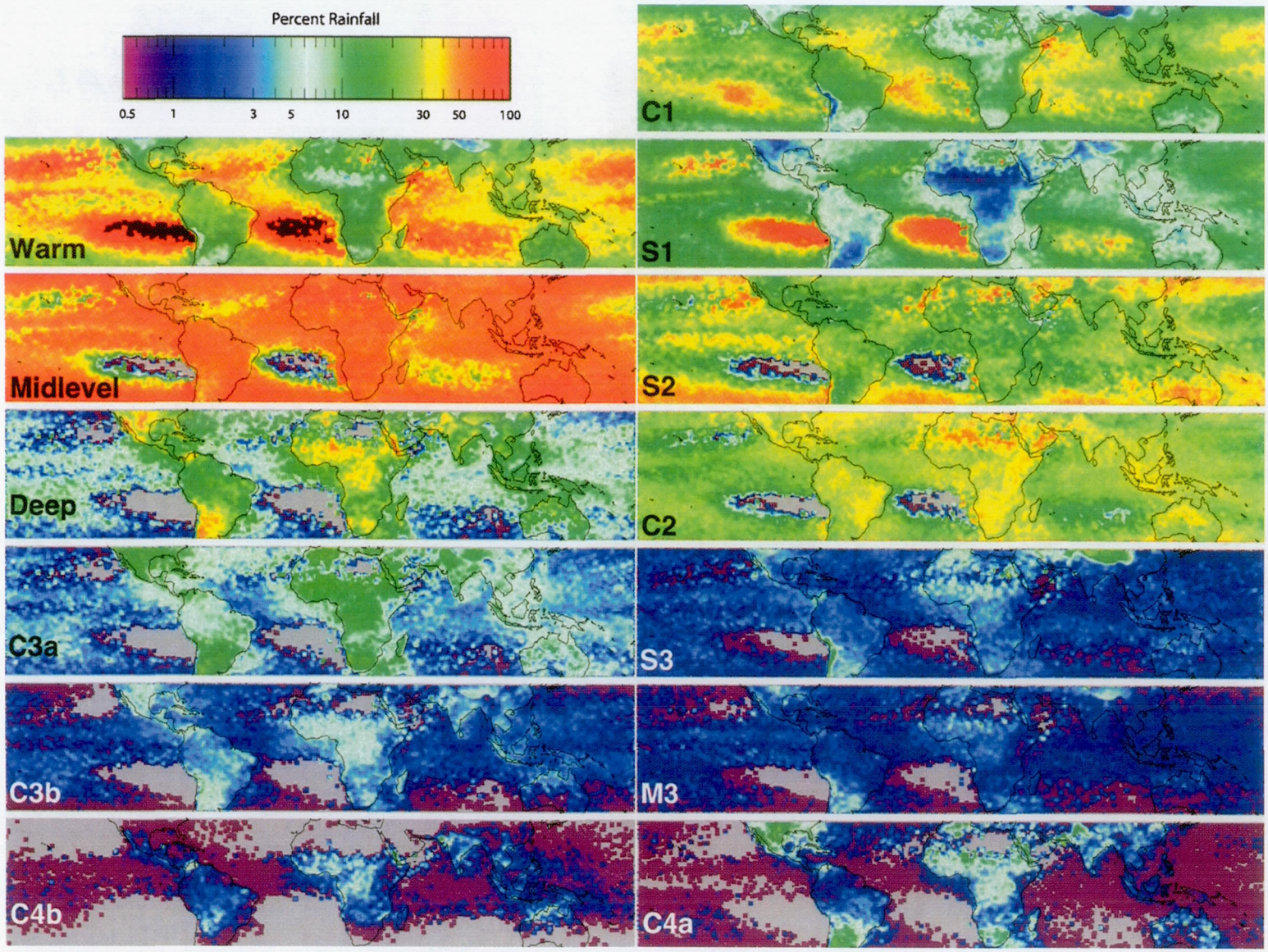

FIG. 11. Annualized rain-conditional percent rainfall contribution from selected profile types and families. Panel sequence is the same as in Fig. 10.

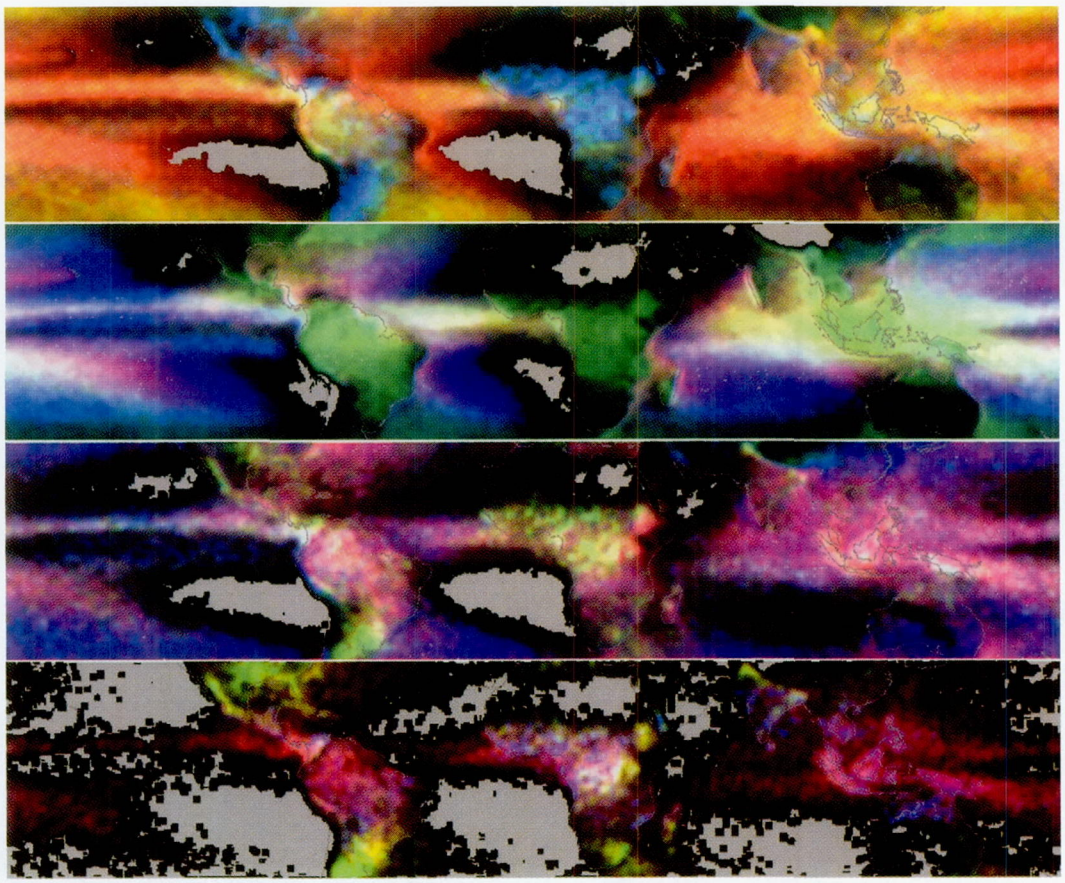

FIG. 12. Three-parameter RGB decompositions of the unconditional (a) overall, (b) warm, (c) midlevel, and (d) deep profile spectra. RGB mappings are (a) $C_1 + S_1/C_2 + S_2 + S_3 + M_3/C_3 + C_4$ (warm/midlevel/deep), (b) $C_{1a}/C_{1b}/S_{1a}$, (c) $C_2/S_3 + M_3/S_2$, and (d) $C_3/C_{4a}/C_{4b}$. Each channel is normalized to its own minimum/maximum range, rather than the range of all channels in each plot, to enhance contrast.

panel, each channel maps the unconditional frequency of occurrence of a set of profile types or families. Since these span very different dynamic ranges, each channel in each map is normalized to its own minimum and maximum frequency of occurrence (otherwise the maps would be dominated by single colors). Thus, for example, in these maps, regions with a "blue dominance" do not necessarily have *more* blue profiles than red profiles; the blue frequency is simply greater in these regions than elsewhere in the maps. The maps are not intended to convey quantitative information; rather, they are intended to illustrate simple, three-parameter descriptions of climatological "convective regimes" (unique colors in the maps correspond to different regimes).

Figure 12a shows the basic warm/midlevel/deep (red/green/blue) profile family decomposition. The map shows much of the same variability evident in the rain-conditional maps. It also reveals very interesting warm rain enhancements in a number of narrow, offshore coastal bands, including off the west coasts of South America, Madagascar, and to a lesser extent, Australia, and off the east coasts of India, southeast Asia, and Indonesia. Figure 12b decomposes the warm portion of the profile spectrum into C_{1_a} (red), C_{1_b} (green), and S_{1_a} (blue) types (S_{1_b} is neglected as three parameters are visual overload enough!). The offshore/coastal enhancements are apparently dominated by warm convective C_{1_a} profiles. The green dominance over Africa illustrates the "promotion" of C_{1_a} profiles there to C_{1_b} types and absence of S_1 profiles. The continuous gradation westward from the southeast Pacific cold ocean gyre illustrates an evolution from exclusively shallow stratocumulus profiles to include increasingly deeper warm convective types as the South Pacific convergence zone (SPCZ) is approached. The warm spectrum in the central/eastern Pacific intertropical convergence zone (ITCZ) contains discernibly fewer C_{1_b} profiles than, for example, the west and east Pacific warm pools and the Atlantic ITCZ. The warm spectrum over the

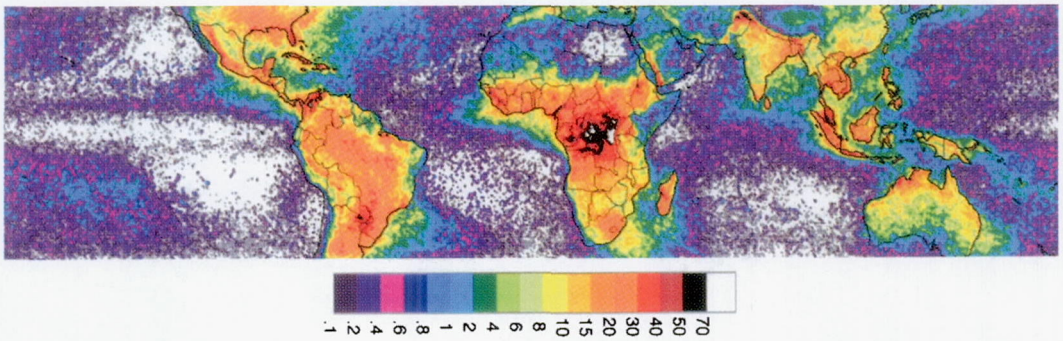

FIG. 13. Annualized total (intracloud and cloud to ground) lightning flash rate density (fl km^{-2} yr^{-1}), derived from 5 yr (1995–2000) of Optical Transient Detector and 5 yr (1998–2002) of TRMM Lightning Imager Sensor data.

Maritime Continent is very similar over both island and ocean regions, and distinct from the Indian Ocean to the west and the Pacific warm pool to the east (indeed, the interisland ocean regions there have a nearly unique oceanic warm spectrum, more comparable to the Amazon basin than to other oceanic regions). A final feature of interest is the upstream/downstream difference in warm profile composition around the Hawaiian island chain, demonstrating a clear shift from stratiform to convective profile types.

Figure 12c decomposes the midlevel spectrum into C_2 (red), S_3/M_3 (green), and S_2 (blue) families. The primary features of interest here are the (blue) enhancement of S_2 profiles over midlatitude oceans and the MCS stratiform (green) enhancements discussed previously in midlatitude continents and sub-Saharan Africa. The Pacific ITCZ is also significantly "narrower" at midlevels. Both features are also evident in the deep convective spectrum decomposition in Fig. 12d into C_3 (red), C_{4_a} (green), and C_{4_b} (blue). The enhancement of wet growth C_{4_a} profiles in midlatitudes is particularly striking (the absence of blue in these regions is due to the fact that C_{4_b} profiles can only occur with tropical-depth tropopauses). Figure 12d underscores the fact that the deep convective spectrum itself is neither monolithic nor described completely by cloud depth (recall that $C_{3_{a_2}}$, C_{3_b}, and C_{4_a} profiles are of comparable depth).

There is close correspondence between the spatial variability of midlevel and deep convective profile frequency in these maps and annualized global lightning production, as observed by the Optical Transient Detector (Boccippio et al. 2000) and TRMM Lightning Imaging Sensor. Figure 13 shows total lightning production, computed following the methodology of Christian et al. (2003) but also including five years of cross-calibrated TRMM LIS data. The agreement extends to very local scales, including features in the Congo basin, Madagascar, Colombia, Papua New Guinea, Borneo, and the Central American west coast. While not surprising, the spatial agreement is in stark contrast to many other remotely observed convective properties, such as outgoing longwave radiation (OLR) or net rainfall.

6. Frequency of occurrence; contributions to rainfall and lightning

The relevance of each profile type in terms of total frequency of occurrence, contribution to total rainfall, and total lightning production (associating each flash's radiance-weighted centroid with the nearest PR profile) is shown in the first three columns of Table 3. Note that these are percentages only of those PR columns classified by this study, that is, those that were not rejected as being cold season, noise, or isolated from a 1Z99 precipitation feature. Convective profiles occur 18% of the time but are associated with 58% of rainfall and 77% of lightning; stratiform and mixed profiles occur 80% of the time and are associated with 42% of rainfall and 20% of lightning.

The dominant rainfall-producing profile is C_{2_a} (17.7%), followed by the similar-depth stratiform S_{2_a} family and S_{2_b} profile (14.7%, 13.3%), then by the next warmer and next colder convective profiles, C_{1_b} and C_{2_b} (11.9% and 9.7%). Taken together, midlevel profiles contribute 55% ($C_2 + S_2$) to 59% ($C_2 + S_2 + S_3 + M_3$) of all rainfall and are most proximate to 23%–35% of all lightning. Warm profiles contribute 29% of all rainfall, while deep convective profiles contribute only 12% (with half of this coming from the shallowest $C_{3_{a_1}}$ profile and two-thirds from the C_{3_a} family; Figs. 2, 6); their low overall frequency of occurrence is not sufficiently offset by higher rainfall rates to cause them to dominate. This is consistent with a regional study of the Tropical Ocean Global Atmosphere Coupled Ocean–Atmosphere Response Experiment (TOGA COARE) west Pacific warm pool by DeMott and Rutledge

TABLE 3. Summary of profile major type frequency, rainfall contribution, and lightning contribution.

Profile p	% Frequency $N = \frac{n_p}{n}$	% Rain $R = \frac{r_p}{r}$	% Lightning $L = \frac{l_p}{l}$	$\frac{R}{N} = \frac{r_p}{n_p}\frac{n}{r}$	$\frac{L}{N} = \frac{l_p}{n_p}\frac{n}{l}$	$\frac{L}{R} = \frac{l_p}{r_p}\frac{r}{l}$
C	**18.1**	**58.2**	**76.8**	**3.2**	**4.2**	**1.3**
C_1	11.3	18.7	5.7	1.7	0.5	0.3
C_{1_a}	5.3	6.8	1.6	1.3	0.3	0.2
C_{1_b}	6.0	11.9	4.1	2.0	0.7	0.3
C_2	5.5	27.4	17.8	5.0	3.2	0.7
C_{2_a}	3.8	17.7	6.9	4.7	1.8	0.4
C_{2_b}	1.7	9.7	10.9	5.7	6.4	1.1
C_3	1.1	9.4	30.4	8.5	28	3.2
C_{3_a}	0.9	8.5	23.1	9.4	26	2.7
C_{3_b}	0.2	0.9	7.3	4.5	37	8.1
C_4	0.2	2.7	28.0	14	140	10
C_{4_a}	0.1	1.9	13.8	19	140	7.3
C_{4_b}	0.1	0.8	9.2	8.0	92	12
S, M	**80.3**	**41.9**	**20.3**	**0.5**	**0.3**	**0.5**
S_1	41.9	10.0	2.8	0.2	0.07	0.3
S_{1_a}	22.1	5.0	1.1	0.2	0.05	0.2
S_{1_b}	19.8	5.0	1.7	0.3	0.09	0.3
S_2	33.3	28.0	5.2	0.8	0.2	0.2
S_{2_a}	26.3	14.7	3.4	0.6	0.1	0.2
S_{2_b}	7.0	13.3	1.8	1.9	0.3	0.1
S_3	4.2	2.3	6.2	0.6	1.5	2.7
S_{3_a}	0.3	0.0	1.2	0.0	4.0	—
S_{3_b}	3.9	2.3	5.0	0.6	1.3	2.2
M_3	0.9	1.6	6.2	1.8	6.9	3.9
M_{3_a}	0.2	0.2	2.4	1.0	12	—
M_{3_b}	0.7	1.4	3.8	2.0	5.4	2.7
A	**1.6**	**<0.1**	**2.9**	**<0.1**	**1.8**	**>29**

(1998a), who found midlevel profiles dominant contributors to total rainfall.

The midlevel *profile* dominance in rainfall contribution must be reconciled with prior findings that large and often deep mesoscale systems contribute a disproportionately large amount of rainfall (Mohr et al. 1999; Nesbitt et al. 2000; Toracinta et al. 2002). Figure 5b shows that the median area of precipitation features containing C_2 profiles is fairly small, corresponding to 13–17-km diameter storms (assuming circular features). However, the median *weighted* area of precipitation features (counting feature areas the number of times each profile type occurs within them) *peaks* for midlevel profiles, confirming that when large features occur, midlevel profiles make up a large portion of their total area. Midlevel profile importance to total rainfall thus stems from both a large number of "isolated" convective cells *and* a much smaller number of large features, in which many of the midlevel profiles only exist within (and because of) larger-scale convective organization and deep convective elements. The present results simply confirm that surface rain rates below the deepest, highest reflectivity profiles *within* these large systems are not the dominant contributors. Regardless of the parent storm structure of C_2 and S_2 profiles, this information may be important, as it guides emphasis in improvement of active and passive microwave remote sensing correction algorithms and physical retrievals away from high-reflectivity, deep convective elements and toward the middle of the spectrum, where even convective/stratiform separation can be challenging (section 7). This guidance of emphasis also applies to field campaign/ground validation studies, where there is a natural tendency to focus on the deepest, most intense convective features.

The relevance of warm rain processes can also be summarized: 19% of rainfall comes from convective warm rain profiles C_1, and 36% of rainfall comes from $C_1 + C_{2_a}$ (profile types that arguably do not contain significant contributions to rainfall from mixed-phase growth); 5% (S_{1_a}) to 10% (S_1) of rain comes from stratiform warm profiles, with the understanding that some shallow convective warm rain may be classified as stratiform given TRMM PR limitations (Schumacher

and Houze 2003a; Schumacher and Houze 2003b), which could affect not only 2A23 convective/stratiform classifiers but also this cluster analysis. Hence 24%–46% of all rain comes from profiles without indication (lower bound), or with questionable indication (upper bound) of concurrent or prior precipitation-sized mixed-phase growth. Conversely, 54%–76% of all rain is associated with glaciated column types. All decompositions reported here are obviously subject to significant regional variability within the TRMM domain (Fig. 11) and again, reflect only warm-season rainfall (section 2d).

The final three columns of Table 3 provide further insight into the process physics associated with each profile type p. The variables n_p, r_p, and l_p denote the total number of pixels, amount of rainfall, and number of lightning flashes associated with each profile type, and n, r, and l the respective totals within the warm-season dataset; $N = n_p/n$, $R = r_p/r$, $L = l_p/l$ thus represent the relative frequency, fraction of total rainfall, and fraction of total lighting associated with each type, and R/N, L/N, and L/R thus vary as the mean rain rate of each profile type, mean number of flashes associated with each profile type, and number of flashes per unit rainfall associated with each profile type (each scaled by an invariant global constant). The value of R/N confirms the importance of moderate rain-rate profiles to total rainfall, given the apportionment of R shown in the second column.

Here L/N demonstrates that midlevel convective profiles with comparatively low flash rates make up a disproportionate percentage of global lightning production (the C_3: C_2 ratio of lightning contribution L is 1.7:1, while the ratio of flash rates L/N is 8.5:1); this is consistent with results by Boccippio et al. (2000) demonstrating that the flash rate spectrum is heavily dominated by low flash rate storm cells and that regional variability in total lightning production is dominated by storm cell *occurrence* rather than variability in instantaneous flash *rates*. This inference may have relevance to parameterization of lightning production in regional or global chemistry models for purposes of NO_x estimation, elevating the importance of resolving the threshold process by which a storm cell becomes electrified and produces lightning, versus resolving the dependency of instantaneous flash rate on storm cell kinematic and microphysical state.

The value of L/R implies two conclusions: 1) lightning is most relevant for rainfall estimation for deep convective profiles (ignoring M and A profiles, which contribute little to the global total), and 2) the profile-level "rain yield per flash" $\sim R/L$ depends on vertical structure. This latter observation has relevance to results by Petersen and Rutledge (1998), who found that daily lightning rain yield relationships varied with convective regime. The current results help explain that finding: the convective regimes of Petersen and Rutledge (1998) exhibit different profile spectra (using the current classification), and their large-scale rain yields thus reflect a regime-dependent weighted average of the profile-level, instantaneous rain yields presented here.

7. Microwave and lightning properties

As described in section 2b, concurrent passive microwave and lightning data are paired to each radar column; these can thus be binned by the 25 archetypal profile types. For the lightning statistics, search distances of 5, 10, and 15 km from each column were examined. Additionally, the percentage of time in which a precipitation feature containing each column type has lightning *anywhere* within its bound is computed (i.e., likelihood that a column type occurs in a thunderstorm complex).

Tables 1 and 2 have shown, for each profile type, the mean TMI 85- and 37-GHz PCTs (Toracinta et al. 2002) and the mean 37–85-GHz PCT depression [a metric loosely related to total Ice Water Path (IWP; Vivekanandan et al. 1991)] within the profile. Note that because of resolution differences, particularly at 37 GHz, these means may contain a signal from significant sub-TMI–pixel variability in the actual column types.

The complete distribution of profile types within the 37–85-GHz PCT space is rendered in Fig. 14. The same "frequency-weighted color blending" approach is used as in Fig. 7; at each PCT pair, the profiles' color table RGB values are composited with a weighting given by each profile's frequency of occurrence at that pair. In contrast to the radar principal component decomposition in Fig. 7, the profile distributions in Figs. 14a,d are significantly "muddier," revealing the (known; Grecu and Anagnostou 2001) fact that the 37–85-GHz PCT space does not map uniquely to vertical structure. The ambiguity is worst for warm and midlevel profiles, where far more common stratiform type profiles "mask" less common convective profiles in the overall distribution. The C and S profiles are separated in Figs. 14b,c,e,f (i.e., positing the existence of a "perfect" C/S separation algorithm operating on high frequency brightness temperatures).

Within these plots, the suspected wet growth deep convective profile C_{4_a} (darkest red) exhibits an unusu-

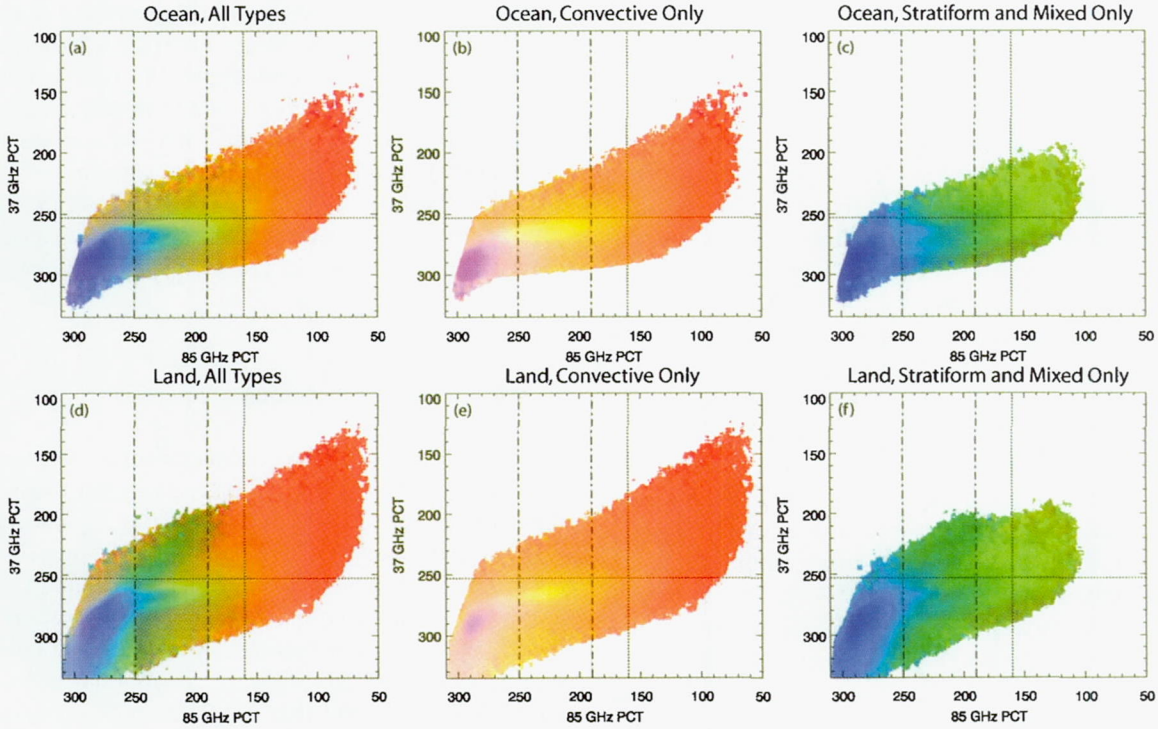

FIG. 14. (a)–(f) Profile type distribution within the (85 GHz, 37 GHz) PCT space, over (a)–(c) ocean and (d)–(f) land. Colors denote a frequency-weighted blending of profile color table entries as in Fig. 7.

ally warm 85-GHz mean PCT. This may reflect the emission effects of high supercooled liquid water contents aloft (Vivekanandan et al. 1991; Toracinta et al. 2002).

Overlaid on the panels in Fig. 14 (dashes–dots) is the 250-K 85-GHz PCT thershold used by, for example, Mohr and Zipser (1996) and Nesbitt et al. (2000) as an ice-scattering signature and the 190-K 85-GHz threshold used by Toracinta and Zipser (2001) as a strong ice-scattering threshold. Also overlaid (dots) are 85- (160 K) and 37-GHz (253 K) thresholds for the upper 1% of the cumulative distribution function (CDF) of minimum PCTs in precipitation features, found by Cecil et al. (2005) ("CAT-2" in their study). The 190-K 85-GHz strong ice-scattering threshold appears reasonable; the most common convective profiles colder than this threshold are C_3 and C_4 (orange and red), while the most common stratiform types are S_3 and M_3 (green and olive green). The 250-K ice scattering threshold is less well defined; significant excursions of C_{1_b} (which is "barely cold" and does not exhibit strong evidence of mixed phase growth) below this threshold occur, as well as some excursions of the deeper C_{2_a} above it.

The mean PCT pairs for each profile type are plotted in Fig. 15a. The severe ambiguity between midlevel convective and stratiform types (e.g., C_{1_b} and S_{1_b}, C_{2_a} and $S_{2_{a_2}}$, C_{2_b} and $S_{3_{b_2}}$, $C_{3_{a_1}}$ and M_{3_b}) is illustrated. These exhibit nearly identical mean PCT pairs. This ambiguity can be quite important; within the (C_{2_b}, $S_{3_{b_2}}$) pair, the convective profile contributes 7 times more rainfall than the stratiform profile, but the stratiform profile occurs twice as often. The severity of the problem is underscored by the overall dominance in tropical rainfall by midlevel profile contributions (Figs. 6, 11; Table 3).[5]

The difficulty encountered in inferring vertical structure from high-frequency PCT observations alone [e.g., the MLR or P_{GSCAT} vertically integrated hydrometeor algorithms of Grecu and Anagnostou (2001)] is problematic; as an extreme example, the overland implementation of the TRMM 2A12 rainfall product relies solely on 85-GHz brightness temperature. Passive microwave rainfall retrievals using

[5] Also in Fig. 15a, many of the shallow and midlevel convective profiles exhibit 37-GHz mean PCT warmer than similar-depth stratiform profiles, perhaps reflecting subpixel variability issues with convective profiles. This may represent disagreement with studies that find a greater increase in 37-GHz brightness temperatures than at 85 GHz associated with melting layers (Olson et al. 2001a).

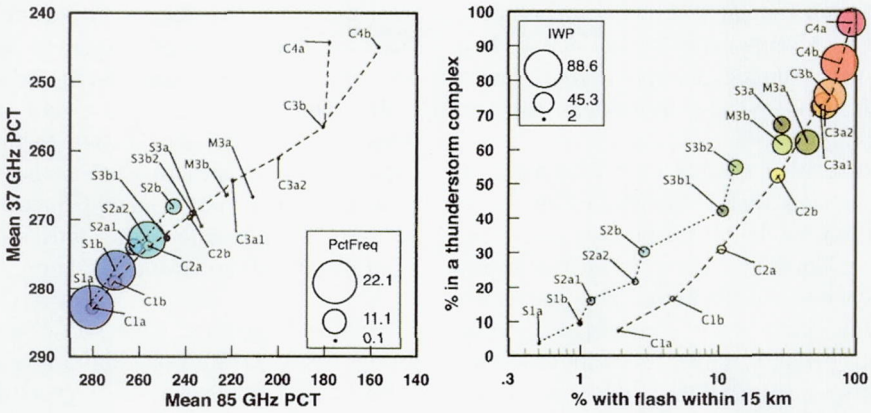

FIG. 15. High-frequency passive microwave and lightning characteristics of the profile clusters. (a) Mean 37- and 85-GHz PCT (bubble size denotes profile frequency in the entire dataset), illustrating significant convective/stratiform ambiguity for important midlevel profiles. (b) Probability that a profile of each type has an LIS-observed lightning flash centroid within 15 km and probability that a profile of each type occurs anywhere in a thunderstorm complex (1Z99 precipitation feature containing lightning). Bubble size denotes IWP, proxied by the 37–85-GHz PCT difference.

blended physical/statistical retrievals (Kummerow et al. 2001) could plausibly be significantly improved by a priori estimation of the hydrometeor profile vertical structure itself. Passive microwave-based profile type classification could be used as "branch points" in rainfall retrieval algorithms to constrain the space of physically plausible scattering/emission paths. The results shown in Fig. 14 show that *at minimum*, convective/stratiform separation, perhaps using horizontal texture or polarization information (Hong et al. 1999; Olson et al. 2001b), must be performed. For optimal prediction of vertical structure in absence of radar observations, inclusion of low-frequency passive microwave, lightning, or other ancillary observations (Miller and Emery 1997; Del Genio and Kovari 2002) as inputs to a multivariate model would likely be required.

As an illustration of the potential for lightning observations to assist in convective/stratiform separation or profile typing, Tables 1 and 2 show, for each profile type, the likelihood that a column of that type occurs in a thunderstorm complex, and the probability that a lightning area (thunderstorm cell center) or flash radiance–weighted centroid falls within 5, 10, or 15 km of the column. The participation-in-a-thunderstorm complex and flash-within-15-km metrics are plotted and joined in Fig. 15b, with mean PCT depression (loosely, ice water path) shown by relative symbol size. Notable in these plots are the separation of "ambiguous" midlevel convective/stratiform cluster pairs in their lightning probabilities. This is a demonstration of how lightning information might statistically (and expectedly) help remove convective/stratiform ambiguity in passive microwave observations.[6]

8. Conclusions

This study demonstrates that vertical structure information in radar reflectivity profiles is sufficiently unique to organize the profiles into physically distinct clusters. Evidence that this organization is meaningful and useful is provided by the clusters' coherence in actual storm scenes, distinct passive microwave and lightning characteristics, physically intuitive (and sometimes locally very high contrast) geographic distributions, and significant variability in their frequency of occurrence and fractional rainfall contribution. The clusters can be aggregated into more general families of profiles using joint passive microwave and lightning observations, as well as radar characteristics not originally used as profile descriptors (e.g., the mean area of containing storms). A meaningful, reduced multiparameter description of local convective spectra can thus be constructed from the conditional and unconditional frequency of occurrence of these families; the hierarchical

[6] We posit a nonlinear, multiparameter-input classification model such as neural network within which the presence or absence of lightning does not provide deterministic classification, but simply helps to "nudge" the decision surface between classes in the right direction. Figure 15b implies that lightning inputs should benefit such a model, although the quantitative skill gains could only be verified through full model training and validation.

clustering dendrogram indicates levels of decomposition (family specificity) supported by the actual data. Such convective spectrum descriptions *retain vertical structure information* in ways that traditional scalar radar metrics do not.

Analysis of the relative rainfall contributions from profile families confirms that 55%–59% of all warm-season rainfall within the TRMM orbit bounds is associated with midlevel profiles. These are often associated with large systems that contain deep convective elements, although those deep convective columns themselves are not primary rainfall contributors. Glaciated column types contribute 54%–76% of all rainfall (the lower number includes only profiles with evidence of concurrent or prior mixed-phase growth; the latter includes all profiles with echo tops colder than 0°C). Convective warm rain contributes 19%–36%. Convective-type profiles make up 18% of the sample, contribute 58% of rainfall, and are most proximate to 77% of all lightning flash centers. Stratiform and mixed profiles make up 80% of the sample, contribute 42% of rainfall, and are most proximate to 20% of lightning (the latter dominated by MCS stratiform and mixed stratiform/convective profiles).

Several profile types are identified in which the same passive microwave 85–/37-GHz PCT pair corresponds to significantly different radar vertical structure and surface reflectivity; the worst ambiguity occurs for midlevel profiles highly important to warm-season rainfall within the TRMM domain. Supplementary lightning observations provide at least one means to help resolve this nonuniqueness problem. Use of the profile typing results can provide a useful tool to subset PM observations and examine whether, and how, supplementary observations such as brightness temperature gradients or variance, polarized brightness temperatures or lower frequency observations can further resolve the nonuniqueness problem.

This classification scheme can be used to study seasonal variability in local convective spectra, as well as forcing/response behavior relative to the local environment. It can also be used to identify convectively "similar" locations and as one component of an objective convective regime definition. From the standpoint of ground validation, profile classification may provide a useful tool for analysis of systematic biases and errors in retrieval algorithms [e.g., binning cross-algorithm retrieval discrepancies by profile types derived from ground validation (GV) data (L'Ecuyer et al. 2004)]. As a data reduction technique, it may have application in other convective problems of interest (e.g., predicting lightning occurrence based on radar-only observations). From a data assimilation standpoint, a vertical-structure-based convective spectrum description provides an objective target for ensemble CRM runs and prescription of associated latent heating profiles. Finally, joint passive microwave and lightning observations could potentially be used to predict a "virtual" radar reflectivity structure in instances where volumetric radar data are not available. Such a prediction could be useful in physically based retrieval of rainfall rates from passive microwave observations themselves.

Acknowledgments. This study was funded under Grant NRA 99-OES-03, under the direction of Dr. Ramesh Kakar. The authors gratefully acknowledge helpful discussions with E. Zipser, S. Nesbitt, T. L'Ecuyer, and S. Heckman, and early provision of the TRMM 1Z99 dataset by S. Nesbitt and E. Zipser.

REFERENCES

Anyamba, E., E. R. Williams, J. Susskind, A. C. Fraser-Smith, and M. Füllekrug, 2000: The manifestation of the Madden–Julian oscillation in global deep convection and in the Schumann resonance intensity. *J. Atmos. Sci.,* **57,** 1029–1044.

Boccippio, D. J., and Coauthors, 2000: The Optical Transient Detector (OTD): Instrument characteristics and cross-sensor validation. *J. Atmos. Oceanic Technol.,* **17,** 441–458.

——, W. Koshak, and R. Blakeslee, 2002: Performance assessment of the Optical Transient Detector and Lightning Imaging Sensor. Part I: Predicted diurnal variability. *J. Atmos. Oceanic Technol.,* **19,** 1318–1332.

Cecil, D., S. Goodman, D. Boccippio, E. Zipser, and S. Nesbitt, 2005: Three years of TRMM precipitation features. Part I: Radar, radiometric, and lightning characteristics. *Mon. Wea. Rev.,* **133,** 543–566.

Christian, H., and Coauthors, 2003: Global frequency and distribution of lightning as observed from space by the Optical Transient Detector. *J. Geophys. Res.,* **108,** 4005, doi:10.1029/2002JD002347.

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

DeMott, C., and S. Rutledge, 1998a: The vertical structure of TOGA COARE convection. Part I: Radar echo distributions. *J. Atmos. Sci.,* **55,** 2730–2747.

——, and ——, 1998b: The vertical structure of TOGA COARE convection. Part II: Modulating influences and implications for diabatic heating. *J. Atmos. Sci.,* **55,** 2730–2747.

Fu, Y., and G. Liu, 2001: The variability of tropical precipitation profiles and its impact on microwave brightness temperatures as inferred from TRMM data. *J. Appl. Meteor.,* **40,** 2130–2143.

Grecu, M., and E. N. Anagnostou, 2001: Overland precipitation estimation from TRMM passive microwave observations. *J. Appl. Meteor.,* **40,** 1367–1380.

Hong, Y., C. D. Kummerow, and W. S. Olson, 1999: Separation of convective and stratiform precipitation using microwave brightness temperature. *J. Appl. Meteor.,* **38,** 1195–1213.

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

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Kummerow, C., and Coauthors, 2001: The evolution of the Goddard Profiling Algorithm (GPROF) for rainfall estimation from passive microwave sensors. *J. Appl. Meteor.,* **40,** 1801–1820.

L'Ecuyer, T., C. Kummerow, and W. Berg, 2004: Toward a global map of raindrop size distributions. Part I: Rain-type classification and its implications for validating global rainfall products. *J. Hydrometeor.,* **5,** 831–849.

Liu, G., and Y. Fu, 2001: The characteristics of tropical precipitation profiles as inferred from satellite radar measurements. *J. Meteor. Soc. Japan,* **79,** 131–143.

Lucas, C., and E. J. Zipser, 2000: Environmental variability during TOGA COARE. *J. Atmos. Sci.,* **57,** 2333–2350.

Meneghini, R., T. Iguchi, T. Kozu, L. Liao, K. Okamoto, J. Jones, and J. Kwiatkowski, 2000: Use of the surface reference technique for path attenuation estimates from the TRMM precipitation radar. *J. Appl. Meteor.,* **39,** 2053–2070.

Miller, S. W., and W. J. Emery, 1997: An automated neural network cloud classifier for use over land and ocean surfaces. *J. Appl. Meteor.,* **36,** 1346–1362.

Mohr, K. I., and E. J. Zipser, 1996: Mesoscale convective systems defined by their 85-GHz ice scattering signature: Size and intensity comparison over tropical oceans and continents. *Mon. Wea. Rev.,* **124,** 2417–2437.

——, J. S. Famiglietti, and E. J. Zipser, 1999: The contribution to tropical rainfall with respect to convective system type, size, and intensity estimated from the 85-GHz ice-scattering signature. *J. Appl. Meteor.,* **38,** 596–606.

Nesbitt, S. W., E. J. Zipser, and D. J. Cecil, 2000: A census of precipitation features in the Tropics using TRMM: Radar, ice scattering, and lightning observations. *J. Climate,* **13,** 4087–4106.

Olson, W. S., P. Bauer, C. D. Kummerow, Y. Hong, and W.-K. Tao, 2001a: A melting-layer model for passive/active microwave remote sensing applications. Part II: Simulation of TRMM observations. *J. Appl. Meteor.,* **40,** 1164–1179.

——, Y. Hong, C. D. Kummerow, and J. Turk, 2001b: A texture-polarization method for estimating convective–stratiform precipitation area coverage from passive microwave radiometer data. *J. Appl. Meteor.,* **40,** 1577–1591.

Orville, R. E., and W. Henderson, 1986: Global distribution of midnight lightning: September 1977 to August 1978. *Mon. Wea. Rev.,* **11,** 2640–2653.

Petersen, W. A., and S. A. Rutledge, 1998: On the relationship between cloud-to-ground lightning and convective rainfall. *J. Geophys. Res.,* **103,** 14 025–14 040.

——, and ——, 2001: Regional variability in tropical convection: Observations from TRMM. *J. Climate,* **14,** 3566–3586.

——, R. Cifelli, D. Boccippio, S. Rutledge, and C. Fairall, 2003: Convection and easterly wave structures observed in the eastern Pacific warm pool during EPIC-2001. *J. Atmos. Sci.,* **60,** 1754–1773.

Rutledge, S. A., E. R. Williams, and T. D. Keenan, 1992: The Down Under Doppler and Electricity Experiment (DUNDEE): Overview and preliminary results. *Bull. Amer. Meteor. Soc.,* **73,** 3–16.

Schumacher, C., and R. A. Houze Jr., 2003a: Stratiform rain in the Tropics as seen by the TRMM Precipitation Radar. *J. Climate,* **16,** 1739–1756.

——, and ——, 2003b: The TRMM precipitation radar's view of shallow, isolated rain. *J. Appl. Meteor.,* **42,** 1519–1524.

Toracinta, E. R., and E. J. Zipser, 2001: Lightning and SSM/I-ice-scattering mesoscale convective systems in the global Tropics. *J. Appl. Meteor.,* **40,** 983–1002.

——, D. Cecil, E. Zipser, and S. Nesbitt, 2002: Radar, passive microwave, and lightning characteristics of precipitating systems in the Tropics. *Mon. Wea. Rev.,* **130,** 802–824.

Vivekanandan, J., J. Turk, and V. Bringi, 1991: Ice water path estimation and characterization using passive microwave radiometry. *J. Appl. Meteor.,* **30,** 1407–1421.

Williams, E. R., S. A. Rutledge, S. G. Geotis, N. O. Rennó, E. N. Rasmussen, and T. M. Rickenbach, 1992: A radar and electrical study of tropical "hot towers." *J. Atmos. Sci.,* **49,** 1386–1395.

——, and Coauthors, 2002. Contrasting convective regimes over the Amazon: Implications for cloud electrification. *J. Geophys. Res.,* **107,** 8082, doi:10.1029/2001JD000380.

Zipser, E. J., 1994: Deep cumulonimbus cloud systems in the Tropics with and without lightning. *Mon. Wea. Rev.,* **122,** 1837–1851.