Relationships between the Raindrop Size Distribution and Properties of the Environment and Clouds Inferred from TRMM

S. Joseph Munchak *

Earth System Science Interdisciplinary Center, University of Maryland, College Park, Maryland

Christian D. Kummerow, and Gregory Elsaesser

Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado

* *Corresponding author address:* Dr. S. Joseph Munchak, ESSIC/UMCP, Code 613.1, NASA/GSFC, Greenbelt, MD 20771.

E-mail: s.j.munchak@nasa.gov

ABSTRACT

Variability in the raindrop sized distribution (DSD) has long been recognized as a source of uncertainty in relationships between radar reflectivity Z and rain rate R. In this study, we analyze DSD retrievals from two years of data gathered by the Tropical Rainfall Measuring Mission (TRMM) satellite and processed with a combined radar-radiometer retrieval algorithm over the global oceans equatorward of 35°. Numerous variables describing properties of each reflectivity profile, large-scale organization, and the background environment are examined for relationships to the reflectivity-normalized median drop diameter, ϵ_{DSD}. In general, we find that higher freezing levels and relative humidities are associated with smaller ϵ_{DSD}. Within a given environment, the mesoscale organization of precipitation and the vertical profile of reflectivity are associated with DSD characteristics. In the tropics, the smallest ϵ_{DSD} values are found in large but shallow convective systems, where warm rain formation processes are thought to be predominant, whereas larger sizes are found in the stratiform regions of organized deep convection. In the extratropics, the largest ϵ_{DSD} values are found in the scattered convection that occurs when cold, dry continental air moves over the much warmer ocean after the passage of a cold front. The geographical distribution of the retrieved DSDs is consistent with many of the observed regional $Z-R$ relationships found in the literature as well as discrepancies between the TRMM radar-only and radiometer-only precipitation products. In particular, mid-latitude and tropical regions near land tend to have larger drops for a given reflectivity, whereas the smallest drops are found in the eastern Pacific Intertropical Convergence Zone.

1. Introduction

The raindrop size distribution (DSD) is a fundamental quantity in radar meteorology and other remote sensing applications and has been the subject of numerous of studies including measurements via disdrometer (*e.g.*, Marshall and Palmer (1948), Waldvogel (1974), Tokay and Short (1996)) and radars (*e.g.*, Williams et al. (1995), Bringi et al. (2003)), parameterizations (*e.g.*, Ulbrich (1983), Haddad et al. (1996), Sempere-Torres et al. (1998), Testud et al. (2001)), and numerical simulations (*e.g.*, List et al. (1987), Brown (1989), Hu and Srivastava (1995), Prat and Barros (2007)). Various moments of the DSD describe physical quantities, such as the liquid water content W, rain rate R and median volume diameter D_0, as well as quantities important for microwave remote sensing such as radar reflectivity Z and specific attenuation k. Relationships between the remotely-sensed and physical quantities are often sought after, particularly the reflectivity-rain rate ($Z - R$) relationship, which is frequently parameterized as the power law $Z = aR^b$. It has been known since the early days of radar meteorology (Atlas and Chmela 1957) that a single unique $Z - R$ relationship does not exist and instead, local relationships were often derived over long periods of time in order to provide radar rainfall estimates that were reasonable on seasonal and yearly scales at a given location (Battan 1973).

The variability of reported $Z - R$ relationships, both between different locations and at the same location at different times, provides some information about the microphysical processes that shape the DSD, although it is difficult to separate effects of drop concentration and drop size on the coefficients of the $Z - R$ relationship (Steiner et al. 2004). Rosenfeld and Ulbrich (2003) classified DSDs by dynamics (convective vs. stratiform) and microphysics

(continental vs. maritime). Stratiform and continental DSDs are characterized by large D_0 for a given W, whereas convective and maritime DSDs of the same W have lower D_0 (and thus, lower Z). Although the names "continental" and "maritime" suggest that the proximity to the ocean is associated with DSD type, these designations do not reveal the mechanism(s) behind the differences between the two ends of the continuum. In fact, maritime DSDs have been observed over land (e.g., Fujiwara and Yanase (1968), Carey et al. (2001), Bringi et al. (2003)) and continental DSDs have been measured in tropical oceanic locations such as the Florida Keys (Tokay et al. 2003). Therefore, it is useful to review the processes that affect the DSD to understand why observed DSD characteristics are often, but not always, found in the expected locations.

The formation of rain is typically classified microphysically as either a warm or cold process. Warm rain formation involves the growth of cloud droplets via collision to a critical size where fall speed is enhanced, allowing the rapid collection of additional drops as the fall speed of the growing raindrop increases with its mass. Eventually, the largest drops break up due to hydrodynamic instability. Various models (List et al. (1987), Hu and Srivastava (1995)) have shown the collision-coalescence and breakup processes to result in an equilibrium shape to the DSD regardless of overall concentration which acts as a scaling factor. This has been observed in tropical convection ((Atlas and Ulbrich 2000),(Uijlenhoet et al. 2003)), which has the requisite rainfall rates and above-freezing column depth to achieve equilibrium. Cold rain formation occurs with the melting of frozen hydrometeors such as snow, graupel, or hail. These frozen particles are larger than the cloud droplets out of which warm rain forms and melt into correspondingly larger rain drops. As these fall, they too are subject to breakup which will reduce their size, although the extent to which

this occurs depends on the depth of the above-freezing layer and the initial DSD.

Cloud dynamics influences the relative importance of warm and cold processes via updraft strength and vertical structure. Convective rain can contain a mixture of warm and cold microphysics; cold microphysics becomes more important with stronger updrafts and cloud tops that reach above the freezing level. Stratiform rain can occur due to large-scale ascent or in convective outflow anvils. In either case, updrafts are weaker and limited to a shallower layer than in convection, and stratiform rain usually forms via cold processes. Besides formation and internal processes, external processes such as evaporation and size sorting can also influence the DSD. Evaporation preferentially acts on small drops, thereby increasing D_0 when rain falls into a subsaturated layer. The influence of size sorting by wind shear and turbulence on the DSD depends on the particular situation and may act to increase or decrease the median drop size.

Considering all of the above processes, one would expect DSDs with smaller drops for a given Z to fall from clouds where warm rain processes are predominant and in environments with deep, humid above-freezing layers. Meanwhile, larger drops would be expected in drier locations with a preference for deeper convection and/or more stratiform rain. Although these expectations qualitatively match observed DSDs, the relative influence of environmental and dynamical effects is not well known. Understanding their role could aid in understanding the effects of aerosol loading on precipitation. Studies have suggested both suppression (Rosenfeld 2000) and enhancement (van den Heever et al. 2006) of rainfall with increasing aerosol burden, depending on the aerosol properties and interaction between cloud microphysics and dynamics (Givati and Rosenfeld 2005). These are also expected to affect the DSD via changing the relative importance of warm and cold rain formation processes.

Improved understanding of the relative importance of environmental, dynamical, and microphysical effects on the rain DSD can also benefit global satellite-based estimates of rainfall, which all rely on DSD assumptions in retrieval algorithms. Microwave radiometer-derived estimates, available on a number of satellite platforms, are physically tied to the emission signal (over oceans), which is roughly proportional to column-integrated W. The relationship between W and R is not as variable as the $Z-R$ relationship (R is approximately proportional to the 3.67th moment of the DSD, whereas Z is to the 6th and W is to the 3rd), but uncertainties in this relationship can still cause errors of as much as 10% (Wilheit et al. 2007) in R. Spaceborne radar-based estimates from the TRMM (Kummerow et al. 1998) precipitation radar (PR) rely on a set of default $Z - R$ relationships (Iguchi et al. 2000) that are modified to match the attenuation inferred by the apparent decrease in the surface reflection in heavy rain (Meneghini et al. 2000). Given the noise inherent in rain-free estimates of the surface cross section, this method is only reliable in rain rates exceeding approximately 10 mm hr $^{-1}$, and, in lighter rain, the default Z-R relationship must be assumed. Rain estimates from CloudSat (Stephens et al. 2002), which uses a higher frequency (94 GHz) that is subject to far greater attenuation than the TRMM PR, use the surface reference technique exclusively, disregarding the reflectivity information (Haynes et al. 2009), although a DSD is still implied in the $k - R$ relationship.

In order to improve understanding of DSD formation processes, their geographic distribution, and how they may affect global satellite rainfall estimates, a combined radar-radiometer algorithm, previously developed by the author (Munchak and Kummerow (2011); hereafter MK11), is utilized. A brief description of the algorithm and its sensitivity to underlying assumptions is examined in section 2. While a satellite retrieval cannot provide as detailed

and precise DSD information as in-situ data from field campaigns, they can be used to put the data from these campaigns into the global context. To achieve this objective, we analyze the output of this algorithm as applied to two years of Tropical Rainfall Measuring Mission (TRMM) data. In section 3, we describe a database containing the retrieval results as well as ancillary variables that represent the rainfall formation processes described previously. Their influence upon the DSD is analyzed in section 4. In section 5, the geographical patterns of all factors that are associated with the rain DSD are examined and it is shown that these patterns are largely consistent with the TMI/PR bias patterns in Berg et al. (2006) and the DSD map of Kozu et al. (2009). Conclusions are presented in section 6.

2. Algorithm Description

Although the full details of the combined algorithm used to retrieve the DSD properties are given by MK11, a brief summary of the relevant output parameters and their sensitivity to internal assumptions is provided here. The core of the algorithm is a radar profiling algorithm that operates similarly to the standard TRMM rain profiling algorithm (2A25; Iguchi et al. (2000, 2009)). A gamma distribution is assumed for the rain DSD: $N(D) = N_0 D^{\mu} e^{-\Lambda D}$, with an intercept parameter (N_0), shape parameter (μ), and slope parameter (Λ), which is related to the median volume diameter D_0 via relation $\Lambda = (3.67 + \mu)/D_0$ (Ulbrich 1983). This formulation implies a power-law relationship between Z and D_0 of the form $D_0 = aZ^b$. In MK11, initial values for a and b are set by rain type indicated by the TRMM rain-classification algorithm (2A23), which identifies profiles as stratiform, convective, or other based on bright band detection, horizontal homogeneity, and maximum reflectivity

(Awaka et al. 2007). The coefficient a is modified by a multiplicative factor ϵ_{DSD} in order to match estimates of the path-integrated attenuation (PIA) provided by the surface reference technique (SRT; Meneghini et al. (2000)), as well as the microwave brightness temperatures ($\mathrm{T_b}$s) at 10, 19, and 37 GHz. Values of ϵ_{DSD} less (greater) than 1 represent mean drop sizes that are smaller (larger) than the default relationship, containing more (less) liquid water at the same reflectivity. Table 1 provides $Z - R$ coefficients for selected values of ϵ_{DSD} to aid in the interpretation and application of results presented in this study.

In addition to adjusting the rain DSD, the combined algorithm also adjusts the ice particle size distribution (PSD) with an analogous factor ϵ_{ICE} in order to match the scattering signal observed at 85 GHz. Values of ϵ_{ICE} less (greater) than 1 imply more (less) scattering than the default $Z - PSD$ relationship implies. However, the physical interpretation of ϵ_{ICE} is somehat ambiguous since this change in scattering could be a result of changes in ice density, morphology, or supercooled cloud water amount as well as changes in the PSD.

A cloud water adjustment is also made in order to match the microwave $\mathrm{T_b}$s while being consistent with rain rates estimated by ground-based polarimetric radars. A default cloud water profile containing approximately 3% (stratiform) or 7% (convective) of the rain water content is assumed, and the integrated cloud liquid water path (cLWP) is modified by the multiplicative factor ϵ_{CLW}. Since cloud water and rain water have similar radiometric signatures, the relative sensitivity of adjustments to ϵ_{CLW} and ϵ_{DSD} was constrained with ground validation data in MK11, but nevertheless remains a source of uncertainty in the combined algorithm.

The retrieval itself is done in the optimal estimation framework, minimizing a cost function (1) consisting of the departure of the modeled PIA and brightness temperatures $f(\mathbf{x})$

from their observed values $\mathbf{y}$, normalized by their covariances $\mathbf{S}_{\mathrm{y}}$, and the departure of the state vector $\mathbf{x}$ consisting of ϵ_{DSD}, ϵ_{ICE}, and ϵ_{CLW} from their default values $\mathbf{x}_{\mathrm{a}}$, normalized by their covariances $\mathbf{S}_{\mathrm{x}}$. This process is carried out over large scenes consisting of as many as a thousand radar pixels (more computational details are given in MK11).

$$\Phi = (\mathbf{y} - f(\mathbf{x}))^{\mathrm{T}}\mathbf{S}_{\mathrm{y}}^{-1}(\mathbf{y} - f(\mathbf{x})) + (\mathbf{x} - \mathbf{x}_{\mathrm{a}})^{\mathrm{T}}\mathbf{S}_{\mathrm{a}}^{-1}(\mathbf{x} - \mathbf{x}_{\mathrm{a}}) \tag{1}$$

In this work, a slight departure is made from the default coefficients a and b and cloud water profiles given by MK11. In that work, different default vales of these coefficients for stratiform and convective rain were selected to replicate the Z-R coefficients used by the 2A25 algorithm. Here, no *a priori* convective/stratiform separation is made because one of the goals of this work is to determine the extent to which DSD is correlated with observables related to these dynamics. Since the optimal estimation method used by the combined algorithm retains some of the *a priori* relationships, depending on the information content in the SRT PIA and $\mathrm{T_b}$s, meaningful comparisons between convective and stratiform DSDs can not be made. Thus, a single weighted average (85% stratiform, 15% convective, which represents their proportion in the version 6 TRMM products) of the coefficients and cloud water profiles is used as the default for this study.

To test the sensitivity of the retrieved value of D_0 to the default assumptions, one month (January 2001) of data was processed twice assuming stratiform and convective $D_0 - Z$ coefficients and cloud/ice profiles. The root-mean-square (rms) difference between the retrieved D_0 is compared to the rms difference between the retrievals and default values as a function of two information content metrics, the $\mathbf{A}$ and $\mathbf{S}_{\mathrm{x}}$ diagonal values (Rodgers 2000) in Figures

1a and 1b, respectively. Assuming linearity of the Jacobian $\mathbf{K}$ and no error in the forward model used in the retrieval, $\mathbf{A}$ represents the fractional weight of the observations in the retrieved value of D_0 (the remainder coming from the *a priori* assumption):

$$\mathbf{A} = \mathbf{S}_{\mathrm{x}}\mathbf{K}^{\mathrm{T}}\mathbf{S}_{\mathrm{y}}^{-1}\mathbf{K}. \tag{2}$$

Likewise, the retrieval covariance matrix $\mathbf{S}_{\mathrm{x}}$, defined by

$$\mathbf{S}_{\mathbf{x}} = (\mathbf{K}^{\mathrm{T}}\mathbf{S}_{\mathrm{y}}\mathbf{K} + \mathbf{S}_{\mathrm{a}}^{-1})^{-1}, \tag{3}$$

can be compared to the *a priori* covariance matrix $\mathbf{S}_{\mathrm{a}}$ (defined in MK11) to evaluate the information content of the observations. L'Ecuyer et al. (2006) note that $\mathbf{S}_{\mathrm{x}}$ and $\mathbf{S}_{\mathrm{a}}$ both define areas in the retrieval parameter space. The amount by which the observations reduce the space represented by $\mathbf{S}_{\mathrm{x}}$ from that represented by $\mathbf{S}_{\mathrm{a}}$ is another measure of the information present in the retrieval.

For both metrics, as the information content increases, the rms difference between the retrieved values of D_0 under different DSD assumptions decreases. At the same time, the rms difference between the retrieved and default (*a priori*) values of D_0 increases. Where these values cross each other can be thought of as the point where the observations and default assumptions equally contribute to the retrieved value of D_0. This occurs near an $\mathbf{A}$ diagonal value of 0.007 and $\mathbf{S}_{\mathrm{a}}\mathbf{S}_{\mathrm{x}}^{-1}$ value of 0.015. Under the definitions of these statistics, these thresholds may seem rather low, but because of the two-dimensional, multi-parameter nature of the retrieval, the off-diagonal elements of $\mathbf{A}$ and $\mathbf{S}_{\mathrm{x}}$, which represent covariances with other parameters (particularly ϵ_{CLW}) and spatial covariances (due to the large radiometer fields-of-view relative to the radar footprint), are large. Thus, the retrieved D_0 values in the

absence of high-resolution radar path-integration-attenuation estimates can only strictly be considered representative over the radiometer FOV, which is 18 by 30 km at 19 GHz, the channel most sensitive to rain, and under the cloud water-rain water partitioning described in MK11.

For the analyses in sections 4 and 5, we choose $\mathbf{A}$ as the information content metric to determine thresholds subsets of data where the retrieved DSD can be considered robust. This is not to discard $\mathbf{S}_{\mathrm{x}}$, but simply recognizes their redundancy which is clear in Figure 1 and in their definitions (2 and 3). At the point of crossover with respect to $\mathbf{A}$, the rms uncertainty in the retrieved value of D_0 is about 0.15 mm, and 60% of the retrieved profiles exceed this threshold. This further decreases asymptotically to around 0.05 mm at an $\mathbf{A}$ diagonal value of 0.07, but only 20% of profiles obtain this higher threshold. These asymptotic values appear to represent the upper limit to which D_0 can be retrieved using the method of MK11.

3. Profile Database

Two years of TRMM data were processed with the MK11 algorithm, one representing the pre-orbit-boost period (August 1999-July 2000) and one representing the post-boost period (January-December 2006). In order to speed computations and avoid biases associated with ground clutter (Shimizu et al. 2009), only the central 25 PR angle bins were processed. Due to uncertainties in surface emissivities (a necessary component of the combined algorithm) over land, only over-ocean retrievals were considered in this analysis. These two years provided 65,782,705 precipitation profiles geographically distributed as shown in Figure 2a. The distribution of profiles in the database is a function of both the frequency of occurrence

of rain and TRMM's orbital geometry. The latter enhances the number of profiles in the mid-latitudes, which the central PR swath samples more often than the equator due to more frequent orbit overlaps.

The fraction of profiles within each 1° grid cell that exceed the $\mathbf{A} > 0.007$ and $\mathbf{A} > 0.07$ thresholds established in section 2 are shown in Figures 2b and 2c, respectively. The profiles exceeding each information content threshold are not evenly distributed, with relatively few of these profiles in the already sparsely-precipitating subsidence regions west of the subtropical continents. Since the method of MK11 relies upon the 10, 19, and 37 GHz channels on TMI along with the radar PIA to adjust ϵ_{DSD}, unequal distribution of profiles with high information content reflects unequal distribution of the ability of the algorithm to make use of these measurements. The TMI observations are only used when rain coverage within the radiometer FOV exceeds 50%; thus isolated profiles are not adjusted. The PIA is only used when it exceeds the natural variability (noise) in the surface reflectivity cross-section from which it is derived; this variability is usually 2-3 dB (Meneghini et al. 2000). The PIA is strongly related to the rain liquid water path (LWP), thus shallow and light rain DSDs cannot be retrieved with it, and in fact this is one of the primary weaknesses of single-frequency radar rain profiling algorithms such as 2A25. To illustrate the differences between the general population of profiles and those that exceed each information content threshold, the distribution of each population is shown as a function of precipitation feature size and PIA in Figures 3a and 3b, respectively. These differences are an important caveat to be kept in mind in the ensuing analyses.

In order to determine the effect of variables related to the background environment, storm structure, and microphysics on the retrieved DSD, each profile was associated with

the variables listed in Table 2. Many of these variables come from products derived from various instruments on board the TRMM satellite, ensuring coincidence in time and space. The combined algorithm, in addition to providing the retrieval parameters (ϵ_{DSD},ϵ_{ICE}, and ϵ_{CLW}) and their associated information content metrics, calculates the attenuation-corrected reflectivity profile. Vertical reflectivity structure has been related to the DSD in a number of studies (L'Ecuyer et al. 2004). For example, the difference in reflectivity above and below the freezing level has been related to updraft strength and the relative importance of cold and warm rain formation (Shige et al. 2008), and Xu et al. (2008) identified a warm rain signature where reflectivity increases towards the surface below the melting level[1]. Thus, reflectivities at levels relevent to these relationships are included in the database to test them with respect the the MK11-derived DSD. The strength of the bright band is used to determine the density of the melting particles as described in MK11 and Zawadzki et al. (2005).

A number of variables are derived from PR products 2A23 (rain characteristics) and 2A25 (rain profile). These include the storm echo top, precipitation feature size (number of contiguous raining pixels), local time, and local standard deviation (within 25 km) of near-surface rain rate and reflectivity. In order to classify the dynamic environment, several parameters used by Elsaesser et al. (2010) to classify tropical convection are also included in the database. These are the number of profiles with echo tops less than 5 km, between 5 km and 9 km, and greater than 9 km within a 1° box[2] surrounding each profile. The same

[1]In our database, this is defined as the near-surface reflectivity minus the lowest valid reflectivity within 1 km below the melting level.

[2]A 25×25 PR pixel box, approximately 100 km on each side.

echo top classes are again used for convective profiles only. The 1° average convective rain rate and convective rain fraction are also used in this classification scheme.

Background parameters total precipitable water (TPW) and sea surface temperature (SST) were derived from TMI data using the methods of Elsaesser and Kummerow (2008) and Gentemann et al. (2004), respectively. Note that these represent the nearest value outside of the raining area. Column relative humidity was calculated by dividing the retrieved TPW by the saturated TPW derived from a temperature profile consistent with the SST and freezing level. Additional meteorological parameters augmenting those available from TRMM observations were taken from the Modern Era Retrospective-Analysis For Research And Applications (MERRA; Bosilovich (2008)) in order to further identify meteorological regimes that might be associated with the DSD. These include temperatures and geopotential heights at selected pressure levels (850mb, 700mb, and 500mb), the 850-500mb and 850-300mb wind shear magnitude, the surface-850mb lapse rate, 700mb vertical velocity, boundary layer height[3] and relative humidity below the freezing level and in the boundary layer. As with any reanalysis data, these variables should be considered representative of the synoptic environment, and moisture/vertical velocity values in particular may be in error near convective rain.

A number of variables related to cloud microphysics are included. The 12 μm channel on the TRMM Visible and Infrared Scanner (VIRS) instrument (Kummerow et al. 1998) was used to determine the cloud top temperature. The cloud top effective radius (R_e) is retrieved from the VIRS data using the method of Nakajima and King (1990). The slope of

[3]Defined as the height at which potential temperature exceeds the surface value by more than 3K; output was insensitive to a range from 2-5K

effective radius with respect to cloud top temperature and the depth of the column where R_e exceeds 15 μm over a 1° grid cell are also included to indicate the presence of warm rain processes as suggested by Rosenfeld and Lensky (1998). Since the visible-infrared retrieval technique only works during the daytime, daily and monthly composites of these variables were constructed and used where coincident data were unavailable. The lightning flash rate comes from TRMM's Lightning Imaging Sensor (Boccippio et al. 2002). The SPRINTARS (Takemura et al. 2000) aerosol optical depth (AOD) reanalysis was included as an additional microphysics variable.

Table 2 lists all of these variables, their distribution shape, and their correlation to ϵ_{DSD} at both thresholds established in section 2. For those variables distributed lognormally, the correlation coefficient was derived in log space. Since ϵ_{DSD} itself is distributed lognormally, all correlations here and elsewhere in this study are actually in relation to $\ln(\epsilon_{\mathrm{DSD}})$. Many of the observed and theoretical relationships in section 1 are confirmed with this data. For example, ϵ_{DSD} decreases with increasing melt density (weaker bright bands) and increasing spatial variability of reflectivity, both of which are commonly used to identify convective rain (Awaka et al. 2004). Microphysics within the profile are also important; large amounts of ice, lightning activity, and an absence of the warm rain signature in the slope of the reflectivity profile below the melting level are also associated with high values of ϵ_{DSD}. However, background environment microphysics (cloud R_e and AOD) are uncorrelated with ϵ_{DSD}. There also appears to be an environmental relationship, with warmer, more humid environments favoring smaller ϵ_{DSD}. Although many of these relationships make sense from a physical point of view, many of these variables are correlated with each other. Thus we will examine the relationship between ϵ_{DSD} and multiple variables in section 4 to identify those which

have significant predictive ability.

4. Sources of DSD variability

The purpose of this section is to more clearly identify the variables in Table 2 with the physical mechanisms described in section 1, simultaneously describing as much of the variability in ϵ_{DSD} as possible given the limitations of the retrieval itself, described by MK11 and in section 2. Because many of the variables in Table 2 only take on physically meaningful values in cold rain (*e.g.*, melt density, IWP), we first separate the database into warm and cold rain using a simple test of whether or not a valid echo exists within 500m of the freezing level as determined by the top of the interpolated bright band height. Within the warm and cold rain subsets, we performed a principle component (PC) analysis of those variables most strongly correlated with ϵ_{DSD}. This analysis creates new proxy variables (the PCs) that represent correlated behavior amongst the original variables. These PCs are also by definition uncorrelated with each each other. The empirical orthogonal functions (EOFs) which come out of this analysis are a regression of the original (standardized) variables onto the PCs. An important consideration in this type of analysis is assessing the significance of each mode. For the purposes of this section, we consider a mode significant if a similar mode, explaining a similar fraction of variance in the database and having a similar correlation with ϵ_{DSD}, is present in subsets of the data (central pixels only and single pre/post-boost years), and that mode explains more variance than a single independent variable (*i.e.*, for a subset of n variables, the variance explained must be greater than $1/n$).

In warm rain, the individual variables most strongly correlated with ϵ_{DSD} are the echotop,

the total number of echo tops under 5km within the 1° surrounding each radar pixel, the boundary layer relative humidity, lapse rate, and freezing level[4]. Cloud top temperature was also included, since cold cloud tops may indicate the influence of cold rain processes even if the detected echo top is below the freezing level. The first three PCs (Table 3) of these five variables are significant under the criteria established previously. The first mode consists primarily environmental variables: high boundary layer relative humidity, high freezing levels, and small lapse rates together are negatively correlated with ϵ_{DSD}. The second mode and third modes represent the organization of precipitation in terms of low cloud concentration, cloud top temperature, and echo top height.

The behavior of ϵ_{DSD} with respect to these three modes at the $\mathbf{A} > 0.007$ threshold is illustrated in Figure 4 (Similar behavior occurs at the $\mathbf{A} > 0.07$ level). The smallest values of ϵ_{DSD} are noted when PC1, PC2, and PC3 are all positive; this represents warm-topped, shallow precipitation in tropical environments with numerous low clouds, indicative of large areas of weak convection (Elsaesser et al. 2010). The largest values, meanwhile, occur when PC1 and PC2 are negative and PC3 is positive, representing colder-topped clouds in extratropical environments with numerous deep clouds. The presence of colder clouds tops in this mode may be an indicator of cold rain processes even though the echo top does not extend above the freezing level. In these profiles, there may be errors in the interpolated freezing height and/or there may be undetected cold processes due to extension of cloud top above the 17-dBZ echo top or influence of neighboring pixels (Liu and Zipser 2009). Additionally, since these are occurring in extratropical environments the underlying forcing

[4]Although temperatures at various levels have higher correlations than some of these, they are largely redundant with lapse rate and freezing level.

may be different (we will examine these relationships in different meteorological regimes in section 5).

Aside from the possible intrusion of cold rain processes, the primary mechanisms affecting the DSD in warm rain are sub-cloud-base humidity and echo top height. The effect of humidity is consistent with theory; smaller values of ϵ_{DSD} are retrieved in more humid environments where the effect of evaporation on DSDs below cloud base is minimized. Echo top increases towards negative values of PC2 and PC3 (the lower right of the PC2-PC3 plane in Figure 4); and a corresponding increase of ϵ_{DSD} is consistent with the longer path for drop growth via collision.

In cold rain, additional variables not available in warm rain are included in the PC analysis. These additional variables are the density of melting particles (a proxy for bright band strength), the difference in maximum reflectivity above and below the melting layer, and the slope of reflectivity below the melting layer. Cloud top temperature and echo top height have little correlation with the DSD in cold rain once the reflectivity structure is accounted for, so they were removed. As with warm rain, three significant modes of variability are present among these variables. The first mode primarily represents environments with high freezing level heights, high relative humidity in the boundary layer, and low concentrations of shallow clouds and vice-versa. The warmer, more humid environments in this mode tend towards smaller values of ϵ_{DSD}. The second mode represents the coordinated variation in the properties of the vertical reflectivity structure. Profiles with low reflectivity above the melting layer relative to below, weak bright bands, and an increase in reflectivity towards the surface within the rain layer tend to have smaller values of ϵ_{DSD}. The third mode represents a different combination of environment and organization from the first mode; this time, stable

lapse rates and high humidity are positively correlated with numerous low clouds.

The mean value of ϵ_{DSD} as a function of the first three PCs for cold rain is illustrated in Figure 5. The smallest values of ϵ_{DSD} are found in tropical environments with numerous shallow precipitating clouds and all of the profile characteristics of warm rain: weak bright bands, high reflectivities below the melting layer than above, and an increase in reflectivity towads the surface indicating an active coalescence process. Large values of ϵ_{DSD} are found in dry extratropical environments with steep lapse rates. Interestingly, the trend in ϵ_{DSD} with respect to the profile shape is different in the extratropics than in the tropics, with an increase in ϵ_{DSD} in profiles with weaker bright bands and high reflectivities below the melting layer than above. Steiner and Smith (1998) find that the dense particles in weak bright bands may be composed of either small, heavily rimed ice particles or larger graupel or hail, with the latter being preferred in stronger updrafts. In extratropical environments, convective updrafts can be stronger than in the tropics due to larger thermal buoyancy and stronger dynamic forcing (Xu and Randall 2001). The increase in drop size with weaker bright bands in these colder environments is consistent with both of these tendencies. In addition, the distribution of profiles in the PC1-PC3 plane implies that many of these colder environments are also dry. Thus, these profiles may be more representative of graupel-containing convection (consistent with the weak bright band) and with evaporation offsetting any warm rain processes in the shallow sub-melting layer.

In order to determine the total variance in ϵ_{DSD} explained by the first three principle components of the warm and cold rain database variables, three-dimensional look-up tables were created (the two-dimensional means of this tables are shown in Figures 4 and 5) with 100 indices in each dimension. The mean value of ϵ_{DSD} for each threshold of information

content was then taken at each index. The value predicted from this table was then compared to the actual retrieved value. By this method, the database principle components explain 23% of the variance in ϵ_{DSD} at the $\mathbf{A} > 0.007$ threshold and 20% at the $\mathbf{A} > 0.07$ threshold.

5. Distribution of DSD variability by geographic region and meteorological regime

Global maps of the mean and PC-predicted values of ϵ_{DSD} are presented in Figure 6. Many of the observed global patterns are reproduced by the PC-predicted values, including the maximum over the Mediterranean Sea and other mid-latitude locations, along with the minima over the eastern Pacific and southern Indian oceans. The increase in ϵ_{DSD} from the eastern to western Pacific is also predicted, but underestimated in magnitude. Also, high values of ϵ_{DSD} in the Caribbean, Gulf of Mexico, and south-central Pacific are underestimated by the PC-based prediction. Increasing the information content threshold to $\mathbf{A} > 0.07$ does not eliminate these residual biases, so they are likely not an artifact of limited information content biasing the mean ϵ_{DSD} in some regions more than others.

In order to determine if the relationships derived in section 4 are equally valid under different meteorological conditions, a meteorological regime classification was performed using a k-means clustering technique (Anderberg 1973) on selected parameters in Table 2. First, the background environment was classified into three regimes (tropical, subtropical, and extratropical) by TPW and 850mb temperature. Within the tropical regime, precipitation was classified as belonging to shallow, mid-level, or deep regimes as defined by Elsaesser

et al. (2010). These clusters represent different modes of organization in convection fields (both in a horizontal spatial extent and vertical extent). The subtropical and extratropical regimes were both broken into two categories by precipitation area, cloud top temperature, and convective fraction. In both environments, a cluster representing organized frontal precipitation, with large precipitation areas, cold cloud tops, and low convective fractions and a cluster representing isolated, shallow convective precipitation were identified. In subtropical environments the former category can be thought of as precipitation associated with "atmospheric rivers" (Zhu and Newell 1998), long but narrow plumes of moisture extending from the tropics to mid-latitudes. In extratropical environments this same category may be found as part of the warm and cold conveyors of extratropical cyclones (Browning 1986). The shallow isolated cluster in the subtropics of exists often under a subsidence inversion, whereas its extratropical counterpart is often triggered when cold continental air is brought over the warm ocean surface after a frontal passage and the resulting instability forces shallow convection in an otherwise subsident environment.

The mean retrieved and predicted value of ϵ_{DSD} in each meteorological regime and information content threshold is given in Table 5. The mean of most clusters closely matches the predicted value, although the tropical mid-level and subtropical isolated shower means are overestimated and both extratropical classifications are underestimated. An examination of maps of the residual error for each cluster (not shown) produces no regional patterns for the extratropical clusters, but the subtropical and tropical clusters do produce patterns that constribute to the overall biases. In the subtropical clusters, ϵ_{DSD} is under-predicted near land areas and over-predicted in the mid-latitude oceans far from land, whereas in the tropical clusters, ϵ_{DSD} is under-predicted near land areas and over-predicted over the eastern

Pacific and southern Indian oceans. These regional patterns suggest that the relationships identified in section 4, while generally valid, do not fully account for all of the processes that affect ϵ_{DSD}. Differences in ϵ_{DSD} from one cluster to another and the difference between the cluster mean and PC predicted may not be the result of differences in observable background parameters, but instead may be related to cloud system scale parameters that influence organization of convection that are largely unobservable from satellite or realized in re-analysis datasets. One possibility is that convective updraft strength, which modulates the warm rain formation process by controlling the rate at which cloud droplets grow (Rosenfeld and Ulbrich 2003), is higher near land due to the origination of systems over land with higher convective available potential energy (CAPE) (Zipser 1994), while the opposite is true over the eastern Pacific (Shige et al. 2008). Therefore, caution should be exercised when applying the relationships derived here to systems over land. In addition, the eastern Pacific contains more "pure" warm rain profiles that are not part of a larger system that extends above the freezing level (Liu and Zipser 2009), and these are not fully accounted for by the variables that define the first three warm PCs in section 4.

6. Summary and Conclusions

In this study we have used the combined radar-radiometer retrieval technique of MK11 to analyze two years of rain DSD retrievals from the TRMM satellite, focusing on the factors that influence the reflectivity-normalized median drop size ($< \epsilon_{\mathrm{DSD}} >$) and how these are related to properties of clouds and their environment. Previous studies, summarized by (Rosenfeld and Ulbrich 2003), have pointed to a variety of sources of variability in the rain

DSD and its expression in the coefficients of $Z - R$ power laws. We have found that:

i. Smaller median drop sizes (both in absolute and reflectivity-normalized values) are found in warm rain than cold rain, as defined by the presence of a radar echo within 500m of the freezing level;

ii. Within the warm rain subset, the smallest drops are found in organized but shallow convective systems in humid tropical environments;

iii. Within the warm rain subset, drop size increases with echo top height which is consistent with the longer path through which drop growth via collision takes place;

iv. Within the cold rain subset, smaller drops are found in more tropical environments where there is also evidence of warm rain processes in the vertical profile of reflectivity (weak bright band and an increase of reflectivity below the melting level);

v. In cold environments, bright band strength does not correlate with $< \epsilon_{\mathrm{DSD}} >$ as strongly as in tropical environments. This is consistent with stronger convective updrafts in the extratropics, which form larger graupel and hail particles than weaker updrafts in tropical convection which form heavily rimed small ice particles.

Together, these environment and cloud properties explain about 23% of the variability in retrieved values of $< \epsilon_{\mathrm{DSD}} >$, which is sufficient to reproduce much of the observed regional variation in reflectivity-normalized drop size. The remaining variability might be related to factors unobservable by the TRMM instruments and inadequately represented in the MERRA reanalysis, such as updraft strength. Inadequate resolution of the low-frequency microwave footprints used to adjust the DSD or temporal variability within a given set of

environmental, microphysical, and dynamical factors could also be a sources of the large amount of variability unexplained in this analysis.

Despite the large amount of unexplained variability at the individual pixel level, the regional patterns of DSD are captured quite well by the priciple components identified in section 4. These patterns, which have been produced for both stratiform and convective rain, are generally similar to those presented by Kozu et al. (2009) for convective rain, although absolute values of the Z-R coefficients differ due to the inclusion of stratiform rain in this study. These regional patterns of DSD can be largely explained by patterns in the dynamical, environmental, and microphysical factors that shape DSD. Much of the bias between PR and TMI rain estimates appears to be related to these DSD assumptions via two pathways: 1) Insufficient adjustments to the default DSD by the PR 2A25 algorithm, especially in light and moderate rain where surface reference estimates of the path-integrated attenuation do not exceed the noise level, and 2) Incorrect assumption of DSD and/or vertical distribution of rain water in the database of profiles used by the Goddard Profiling Algorithm (GPROF) algorithm for TMI, which affects the liquid water content-rain rate conversion. The former issue could be addressed by including a "warm" vs. "cold" rain identification process and default DSDs in addition to the stratiform vs. convective identification in future versions of the PR 2A25 algorithm. Biases introuduced by the latter issue should be reduced substantially when a database of radiometer-adjusted PR precipitation profiles, with T_bs that are consistent with Z and R, are used in place of cloud resolving model-derived profiles in upcoming versions of passive radiometer rain retrieval algorithms (Kummerow et al. 2011); but this remains to be seen.

Much work remains to be done to verify these relationships, and in particular to identify biases in the combined radar-radiometer algorithm that may create spurious relationships between the DSD adjustment and unrelated factors. Nevertheless, the relationships we have found are consistent with what is known about the processes that shape the rain DSD and may be used to create time-varying Z-R relationships for ground-based radars or to enhance over-land TRMM PR retrievals, where radiometer-enhanced retrievals are complicated by the unknown factors related to surface emissivity and radar-only retrievals must rely on the surface reference estimate of attenuation, which is noisier over land than water. However, it should be emphasized that caution must be used in extending these relationships over land, as some regimes (*e.g.*, orographic precipitation) may be unsampled over the ocean. The upcoming Global Precipitation Measurement (GPM) mission, scheduled to launch in 2013, will carry a dual-frequency radar with the ability to retrieve two parameters of the DSD at each range gate (Kuo et al. 2004), reducing much of the ambiguity in DSD retrievals over land and ocean. At that time it will be worthwhile to revisit the relationships noted in this work.

Acknowledgments.

This work was supported by NASA Headquarters under the NASA Earth and Space Science Fellowship Program. Additional funds for publication were provided by grant NNX10AG75G.

REFERENCES

Anderberg, M. R., 1973: *Cluster Analysis for Applications*. Academic Press, 359 pp.

Atlas, D. and A. C. Chmela, 1957: Physical-synoptic variations of drop-size parameters. *Proc. 6th Weather Radar Conf.*, Boston, MA, Amer. Meteor. Soc., 21–29.

Atlas, D. and C. W. Ulbrich, 2000: An observationally based conceptual model of warm oceanic convective rain in the tropics. *J. Appl. Meteor.*, **39**, 2165–2181.

Awaka, J., T. Iguchi, and K. Okamoto, 2004: On rain type classification algorithm TRMM PR 2A23 V6. *2nd TRMM International Science Conference*, Tokyo, Japan.

Awaka, J., T. Iguchi, and K. Okamoto, 2007: *Measuring Precipitation from Space - EURAINSAT and the Future*. Springer, New York, 213-224 pp.

Battan, L. J., 1973: *Radar Observation of the Atmosphere*. University of Chicago Press, Chicago, 324 pp.

Berg, W., T. L'Ecuyer, and C. Kummerow, 2006: Rainfall climate regimes: the relationship of regional TRMM rainfall biases to the environment. *J. Appl. Meteor. and Climatol.*, **45**, 434–454.

Boccippio, D. J., W. J. Koshak, and R. J. Blakeslee, 2002: Performance assessment of the optical transient detector and lightning imaging sensor. part i: Predicted diurnal

variability. *J. Atmos. Oceanic Technol.*, **19 (9)**, 1318–1332, doi:10.1175/1520-0426(2002)019⟨1318:PAOTOT⟩2.0.CO;2.

Bosilovich, M., 2008: NASA's modern era retrospective-analysis for research and applications: Integrating earth observations. *Earthzine*, URL `http://www.earthzine.org/2008/09/26/nasas-modern-era-retrospective-analysis/`.

Bringi, V. N., V. Chandrasekhar, J. Hubbert, E. Gorgucci, W. L. Randeu, and M. Schoenhuber, 2003: Raindrop size distribution in different climatic regimes from disdrometer and dual-polarized radar analysis. *J. Atmos. Sci.*, **60**, 354–365.

Brown, P. S., 1989: Coalescence and breakup-induced oscillations in the evolution of the raindrop size distribution. *J. Atmos. Sci.*, **46**, 1186–1192.

Browning, K. A., 1986: Conceptual models of precipitation systems. *Wea. Forecasting*, **1 (1)**, 23–41.

Carey, L. D., R. Cifelli, W. A. Petersen, S. A. Rutledge, and M. A. F. S. Dias, 2001: Characteristics of Amazonian rain measured during TRMM-LBA. *30th Conf. on Radar Meteorology*, Munich, Germany, Amer. Meteor. Soc.

Elsaesser, G. and C. Kummerow, 2008: Towards a fully parametric retrieval of the non-raining parameters over the global oceans. *J. Appl. Meteor.*, **47**, 1599–1618.

Elsaesser, G. S., C. D. Kummerow, T. S. L?Ecuyer, Y. N. Takayabu, and S. Shige, 2010: Observed self-similarity of precipitation regimes over the tropical oceans. *J. Climate*, **23 (10)**, 2686–2698, doi:10.1175/2010JCLI3330.1, URL `http://journals.ametsoc.org/doi/abs/10.1175/2010JCLI3330.1`.

Fujiwara, M. and T. Yanase, 1968: Raindrop z-r relationships in different altitudes. *Preprints 13th Conf. on Radar Meteorology*, Boston, MA, Amer. Meteor. Soc., 380–383.

Gentemann, C. L., F. Wentz, C. Mears, and D. Smith., 2004: In-situ validation of TRMM microwave sea surface temperatures. *J. Geophys. Res.*, **109**, C04 021.

Givati, A. and D. Rosenfeld, 2005: Separation between cloud-seeding and air-pollution effects. *J. Appl. Meteor.*, **44**, 1298–1314.

Haddad, Z. S., S. L. Durden, and E. Im, 1996: Parameterizing the raindrop size distribution. *J. Appl. Meteor.*, **35**, 3–13.

Haynes, J. M., T. S. L'Ecuyer, G. L. Stephens, S. D. Miller, C. Mitrescu, N. B. Wood, and S. Tanelli, 2009: Rainfall retrieval over the ocean with spaceborne W-band radar. *J. Geophys. Res.*, **114**, D00A22, doi:10.1029/2008JD009 973.

Hu, Z. and R. Srivastava, 1995: Evolution of the raindrop size distribution by coalescence, breakup, and evaporation: Theory and observations. *J. Atmos. Sci.*, **52**, 1761–1783.

Iguchi, T., T. Kozu, J. Kwiatkowski, R. Meneghini, J. Awaka, and K. Okamoto, 2009: Uncertainties in the rain profiling algorithm for the TRMM precipitation radar. *J. Met. Soc. Japan*, **87A**, 1–30.

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.

Kozu, T., T. Iguchi, T. Kubota, N. Yoshida, S. Seto, J. Kwiatkowski, and Y. N. Takayabu,

2009: Feasibility of raindrop size distribution parameter estimation with TRMM precipitation radar. *J. Met. Soc. Japan*, **87A**, 53–66.

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

Kummerow, C. D., S. Ringerud, J. Crook, D. Randel, and W. Berg, 2011: An observationally generated a-priori database for microwave rainfall retrievals. *J. Atmos. Oceanic Technol.*, **28**, 113–130.

Kuo, K.-S., E. A. Smith, Z. Haddad, E. Im, T. Iguchi, and A. Mugnai, 2004: Mathematical-Physical Framework for Retrieval of Rain DSD Properties from Dual-Frequency KuKa-Band Satellite Radar. *J. Atmos. Sci.*, **61 (19)**, 2349–2369, doi:10.1175/1520-0469(2004)061⟨2349:MFFROR⟩2.0.CO;2.

L'Ecuyer, T., H. Masunaga, and C. Kummerow, 2006: Variability in the characteristics of precipitation systems in the tropical Pacific. Part II: implications for atmospheric heating. *J. Climate*, **19**, 1388–1406.

L'Ecuyer, T. S., 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 (5)**, 831–849, doi:10.1175/1525-7541(2004)005⟨0831:TAGMOR⟩2.0.CO;2, URL `http://journals.ametsoc.org/doi/abs/10.1175/1525-7541%282004%29005%3C0831%3ATAGMOR%3E2.0.CO%3B2`.

List, R., T. B. Low, N. Donaldson, E. Freire, and J. R. Gillespie, 1987: Temporal evolution

of drop spectra to collisional equilibrium in steady and pulsating rain. *J. Atmos. Sci.*, **44**, 362–372.

Liu, C. and E. J. Zipser, 2009: Warm Rain in the Tropics: Seasonal and Regional Distributions Based on 9 yr of TRMM Data. *J. Climate*, **22 (3)**, 767–779.

Marshall, J. and W. M. Palmer, 1948: The distribution of raindrops with size. *J. Meteor.*, **5**, 165–166.

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

Munchak, S. J. and C. D. Kummerow, 2011: A modular optimal estimation method for combined radar-radiometer precipitation profiling. *J. Appl. Meteor. Climatol.*, **50**, 433–448.

Nakajima, T. and M. D. King, 1990: Determination of the optical thickness and effective particle radius of clouds from reflected solar radiation measurements. Part I: Theory. *J. Atmos. Sci.*, **47**, 1878–1893.

Prat, O. P. and A. P. Barros, 2007: Exploring the use of a column model for the characterization of microphysical processes in warm rain: results from a homogeneous rainshaft model. *Adv. Geosci.*, **10**, 145–152.

Rodgers, C. D., 2000: *Inverse Methods for Atmospheric Sounding: Theory and Practice.* World Scientific Publishing Co., Singapore, 238 pp.

Rosenfeld, D., 2000: Suppression of rain and snow by urban and industrial air pollution. *Science*, **287**, 1793–1796.

Rosenfeld, D. and I. M. Lensky, 1998: Satellite-based insights into precipitation formation processes in continental and maritime convective clouds. *Bull. Amer. Meteor. Soc.*, **79**, 2457–2476.

Rosenfeld, D. and C. W. Ulbrich, 2003: Cloud microphysical properties, processes and rainfall estimation opportunities. *Radar and Atmospheric Science: A Collection of Essays in Honor of David Atlas, Meteor. Monogr. No. 52*, Amer. Meteor. Soc., 237–258.

Sempere-Torres, D., J. M. Porr, and J.-D. Creutin, 1998: Experimental evidence of a general description for raindrop size distribution properties. *J. Geophys. Res.*, **103 (D2)**, 1785–1797.

Shige, S., T. Watanabe, H. Sasaki, T. Kubota, S. Kida, and K. Okamoto, 2008: Validation of western and eastern Pacific rainfall estimates from the TRMM PR using a radiative transfer model. *J. Geophys. Res.*, **113**, D15 116.

Shimizu, S., R. Oki, T. Tagawa, T. Iguchi, and M. Hirose, 2009: Effects of the orbit boost of the TRMM satellite on PR rain estimates. *J. Met. Soc. Japan*, **87A**, 83–92.

Steiner, M. and J. A. Smith, 1998: Convective versus stratiform rainfall: An ice-microphysical and kinematic conceptual model. *Atmos. Res.*, **47-48**, 317–326.

Steiner, M., J. A. Smith, and R. Uijlenhoet, 2004: A microphysical interpretation of radar reflectivityrain rate relationships. *J. Atmos. Sci.*, **61 (10)**, 1114–1131.

Stephens, G. L., D. G. Vane, and coauthors, 2002: The Cloudsat mission and the A-train. *Bull. Amer. Meteor. Soc.*, **83**, 1771–1790.

Takemura, T., H. Okamoto, Y. Maruyama, A. Numaguti, A. Higurashi, and T. Nakajima, 2000: Global three-dimensional simulation of aerosol optical thickness distribution of various origins. *J. Geophys. Res.*, **105**, 17 853–17 873.

Testud, J., S. Oury, R. Black, P. Amayenc, and X. Dou, 2001: The concept of "normalized" distribution to describe raindrop spectra: A tool for cloud physics and remote sensing. *J. Appl. Meteor.*, **40**, 1118–1140.

Tokay, A. and D. A. Short, 1996: Evidence from tropical raindrop spectra of the origin of rain from stratiform versus convective clouds. *J. Appl. Meteor.*, **35**, 1671–1687.

Tokay, A., D. B. Wolff, K. R. Wolff, and P. Bashor, 2003: Rain gauge and disdrometer measurements during the Keys Area Microphysics Project (KAMP). *J. Atmos. Oceanic Technol.*, **20**, 1460–1477.

Uijlenhoet, R., J. A. Smith, and M. Steiner, 2003: The microphysical structure of extreme precipitation as inferred from ground-based raindrop spectra. *J. Atmos. Sci.*, **60 (10)**, 1220–1238.

Ulbrich, C. W., 1983: Natural variations in the analytical form of the drop size distribution. *J. Climate Appl. Meteor.*, **22**, 1764–1775.

van den Heever, S. C., G. G. Carri, W. R. Cotton, P. J. Demott, and A. J. Prenni, 2006: Impacts of nucleating aerosol on florida storms. Part I: Mesoscale simulations. *J. Atmos. Sci.*, **63**, 1752–1775.

Waldvogel, A., 1974: The n0 jump of raindrop spectra. *J. Atmos. Sci.*, **31**, 1067–1078.

Wilheit, T. T., V. Chandrasekhar, and W. Li, 2007: Impact of uncertainty in the drop size distribution on oceanic rainfall retrievals from passive microwave observations. *IEEE Trans. Geosci. Remote Sensing*, **45**, 3160–3164.

Williams, C. R., W. L. Ecklund, and K. S. Gage, 1995: Classification of precipitating clouds in the tropics using 915-MHz wind profilers. *J. Atmos. Oceanic Technol.*, **12**, 996–1012.

Xu, K.-M. and D. A. Randall, 2001: Updraft and downdraft statistics of simulated tropical and midlatitude cumulus convection. *J. Atmos. Sci.*, **58 (13)**, 1630–1649.

Xu, X., K. Howard, and J. Zhang, 2008: An automated radar technique for the identification of tropical precipitation. *J. Hydrometeor.*, **9 (5)**, 885–902, doi:10.1175/2007JHM954.1.

Zawadzki, I., W. Szyrmer, C. Bell, and F. Fabry, 2005: Modeling of the melting layer. Part III: The density effect. *J. Atmos. Sci.*, **62**, 3705–3722.

Zhu, Y. and R. E. Newell, 1998: A proposed algorithm for moisture fluxes from atmospheric rivers. *Mon. Wea. Rev.*, **126 (3)**, 725–735.

Zipser, E. J., 1994: Deep cumulonimbus cloud systems in the tropics with and without lightning. *Mon. Wea. Rev.*, **122 (8)**, 1837–1851.

List of Tables

1 Coefficients of the relationship $Z = AR^B$ and $R = \alpha Z^\beta$ for selected values of ϵ_{DSD} in the relationship $D_0 = \epsilon_{\mathrm{DSD}} a Z^b$, where $a = 0.5794$, $b = 0.1094$, and Z is in units of mm^6 m^{-3}, R is in mm hr^{-1}, and D_0 is in mm, and a gamma DSD with shape parameter $\mu = 3$ is assumed. The values for a and b were selected to represent an 85% stratiform-weighted average of the $Z - R$ coefficients given by Iguchi et al. (2000). 35

2 List of profile database variables with their source and distribution shape. The correlation coefficient of $\ln(\epsilon_{\mathrm{DSD}})$ with each variable for profiles exceeding the **A** threshold of 0.007 (0.07) is given by r_1 (r_2). 36

3 Significant EOFs of warm rain variables in order of variance explained (VE). The correlation of each PC with the number of echo tops under 5km within 1° (N5), echo top height (ETH), boundary layer relative humidity (BLRH), lapse rate (LR), freezing level height (FLH), Cloud top temperature (CT), and ϵ_{DSD} (r_1 and r_2 have the same meaning as in Table 2) is given in the table. Correlations above 0.5 are bolded to highlight the variables most strongly represented by each mode. 37

4 Significant EOFs of cold rain variables in order of variance explained (VE). In addition to the variables for warm rain in Table 3 this table includes melting particle density (RHOM), maximum reflectivity above the melting layer minus maximum reflectivity below melting layer (ZDIFF), and the slope of reflectivity below the melting layer (ZS). Correlations above 0.5 are bolded to highlight the variables most strongly represented by each mode. 38

5 Mean and predicted (P) values of $< \epsilon_{\mathrm{DSD}} >$ by meteorological regime and information content threshold. 39

TABLE 1. Coefficients of the relationship $Z = AR^B$ and $R = \alpha Z^\beta$ for selected values of ϵ_{DSD} in the relationship $D_0 = \epsilon_{\mathrm{DSD}} a Z^b$, where $a = 0.5794$, $b = 0.1094$, and Z is in units of mm^6 m^{-3}, R is in mm hr^{-1}, and D_0 is in mm, and a gamma DSD with shape parameter $\mu = 3$ is assumed. The values for a and b were selected to represent an 85% stratiform-weighted average of the $Z - R$ coefficients given by Iguchi et al. (2000).

ϵ_{DSD}	A	B	α	β
0.50	40	1.29	0.0576	0.775
0.75	114	1.32	0.0274	0.760
1.00	258	1.34	0.0156	0.748
1.25	510	1.35	0.0097	0.743
1.50	902	1.34	0.0063	0.745
1.75	1440	1.32	0.0041	0.757
2.00	2085	1.29	0.0027	0.775

Table 2. List of profile database variables with their source and distribution shape. The correlation coefficient of $\ln(\epsilon_{\mathrm{DSD}})$ with each variable for profiles exceeding the **A** threshold of 0.007 (0.07) is given by r_1 (r_2).

Variable	Source	Distribution	r_1	r_2
Melt density	PR	Normal	-.21	-.19
Total precipitable water (TPW)	TMI	Normal	-.09	-.13
Ice water path (IWP)	TMI+PR	Lognormal	.19	.19
Sea surface temperature (SST)	TMI	Normal	-.02	-.07
Near-surface dBZ	PR	Normal	-.07	.02
Maximum dBZ in rain layer	PR	Normal	-.02	.09
Maximum dBZ in melting layer	PR	Normal	.15	.24
Maximum dBZ in ice layer	PR	Normal	.13	.10
Reflectivity slope in rain layer	PR	Normal	-.26	-.25
Cloud Top Temperature	VIRS	Multimodal	-.10	-.08
Mean cloud effective radius	VIRS	Normal	.00	.00
Cloud effective radius slope	VIRS	Normal	.00	.00
Warm rain depth	VIRS	Normal	-.02	-.04
Lightning flash rate	LIS	Lognormal	.11	.14
Aerosol Optical Depth	SPRINTARS	Lognormal	-.02	-.01
Echo Top Height	PR	Multimodal	.11	.07
Precipitation feature size	PR	Lognormal	-.04	-.01
Profiles with echo top $<$ 5km within 1°	PR	Lognormal	-.20	-.17
Convective “”	PR	Lognormal	-.10	-.10
Profiles with 5km $<$ echo top $<$ 9km within 1°	PR	Lognormal	.02	.05
Profiles with echo top $>$ 9km within 1°	PR	Lognormal	.08	.05
Convective “”	PR	Lognormal	.01	-.01
1° average convective rain rate	PR	Lognormal	-.06	-.05
1° average convective rain fraction	PR	Normal	-.08	-.12
25-km reflectivity standard deviation	PR	Normal	-.14	-.20
Surface-850mb lapse rate	MERRA	Normal	.14	.10
850mb temperature	MERRA	Normal	-.13	-.14
700mb temperature	MERRA	Normal	-.12	-.15
500mb temperature	MERRA	Normal	-.16	-.17
850mb height	MERRA	Normal	-.07	-.06
700mb height	MERRA	Normal	-.12	-.12
500mb height	MERRA	Normal	-.14	-.16
Freezing level	MERRA	Normal	-.16	-.17
Mean relative humidity below freezing level	MERRA	Normal	-.13	-.09
850-500mb shear	MERRA	Normal	.06	.07
850-300mb shear	MERRA	Normal	.07	.08
700mb vertical velocity	MERRA	Normal	.04	.00
Boundary layer height	MERRA	Normal	.09	.07
Boundary layer relative humidity	MERRA	Normal	-.18	-.14

Table 3. Significant EOFs of warm rain variables in order of variance explained (VE). The correlation of each PC with the number of echo tops under 5km within 1° (N5), echo top height (ETH), boundary layer relative humidity (BLRH), lapse rate (LR), freezing level height (FLH), Cloud top temperature (CT), and ϵ_{DSD} (r_1 and r_2 have the same meaning as in Table 2) is given in the table. Correlations above 0.5 are bolded to highlight the variables most strongly represented by each mode.

Mode	VE (%)	FLH	LR	BLRH	N5	CT	ETH	r_1	r_2
1	28.2	**.55**	**-.70**	**.79**	.13	-.32	.35	-.26	-.22
2	23.5	.28	-.35	.17	**-.51**	**.78**	**-.55**	-.01	.01
3	19.2	-.39	-.08	.34	**.71**	.06	**-.60**	-.22	-.24

Table 4. Significant EOFs of cold rain variables in order of variance explained (VE). In addition to the variables for warm rain in Table 3 this table includes melting particle density (RHOM), maximum reflectivity above the melting layer minus maximum reflectivity below melting layer (ZDIFF), and the slope of reflectivity below the melting layer (ZS). Correlations above 0.5 are bolded to highlight the variables most strongly represented by each mode.

Mode	VE (%)	FLH	LR	BLRH	N5	ZDIFF	RHOM	ZS	r_1	r_2
1	27.1	**.83**	-.44	**.52**	**-.61**	-.44	.23	.15	-.18	-.17
2	24.9	-.19	.38	-.19	.29	**-.73**	**.77**	**.54**	-.21	-.20
3	17.4	-.30	**-.57**	**.60**	**.64**	.01	-.05	.24	-.26	-.21

TABLE 5. Mean and predicted (P) values of $< \epsilon_{\mathrm{DSD}} >$ by meteorological regime and information content threshold.

Environment	Cluster	$\mathbf{A} > 0.007$		$\mathbf{A} > 0.07$	
		$< \epsilon_{\mathrm{DSD}} >$	$< \epsilon_{\mathrm{DSD}} P >$	$< \epsilon_{\mathrm{DSD}} >$	$< \epsilon_{\mathrm{DSD}} P >$
Tropical	Shallow	0.92	0.93	0.94	0.93
	Mid-level	0.91	0.94	0.92	0.94
	Deep	0.95	0.96	0.95	0.96
Subtropical	Organized Frontal	0.91	0.91	0.94	0.94
	Isolated Shallow	0.88	0.91	0.91	0.92
Extratropical	Organized Frontal	0.96	0.95	0.98	0.96
	Isolated Shallow	1.04	1.02	1.04	1.03

List of Figures

1 a) Root-mean-square difference between retrieved D_0 under stratiform assumptions and retrieved D_0 under convective assumptions (black), retrieved and default D_0 under stratiform assumptions, and retrieved and default D_0 under convective assumptions (red) as a function of $\mathbf{A}$ diagonal value. b) same as a), except as a function of $\mathbf{S}_{\mathrm{a}}$ diagonal value divided by $\mathbf{S}_{\mathrm{x}}$ diagonal value. In both panels, the fraction of profiles exceeding the information content value on the x-axis is indicated by the dashed line and tick marks on the right y-axis. 41

2 Top panel: number of profiles in $1{\times}1^\circ$ grid boxes. Lower panels: fraction of profiles in each grid box that exceed the threshold of information content indicated. 42

3 a) Histogram of profiles by precipitation area for different information content thresholds. b) same as a), except as a function of surface reference path-integrated attenuation. 43

4 Mean value of ϵ_{DSD} in the PC1-PC2, PC2-PC3, and PC1-PC3 planes for warm rain. 44

5 Mean value of ϵ_{DSD} in the PC1-PC2, PC2-PC3, and PC1-PC3 planes for cold rain. 45

6 Mean and predicted values of ϵ_{DSD} at the $\mathbf{A} > 0.007$ threshold gridded at 1° resolution. 46

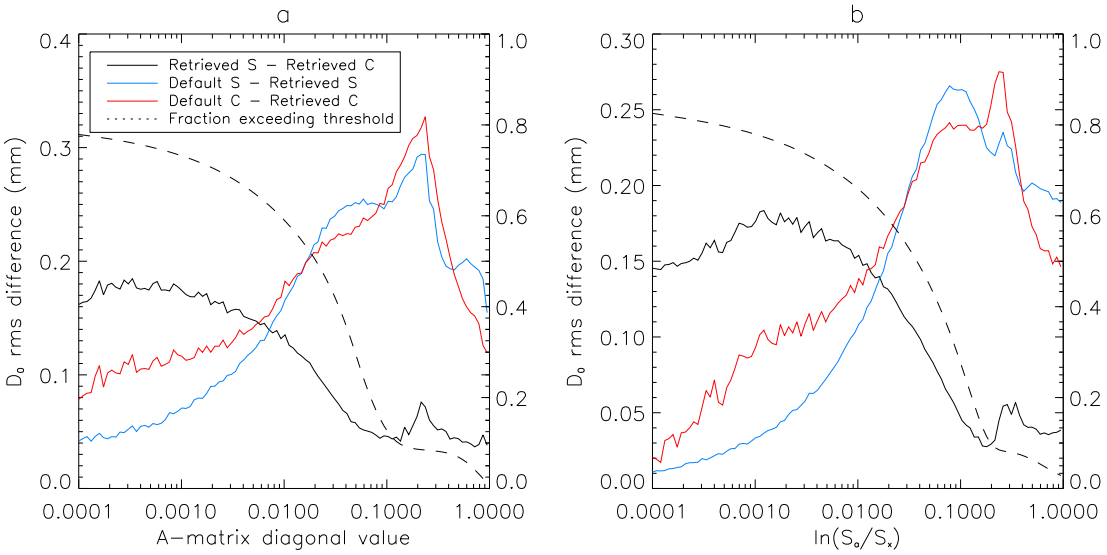

FIG. 1. a) Root-mean-square difference between retrieved D_0 under stratiform assumptions and retrieved D_0 under convective assumptions (black), retrieved and default D_0 under stratiform assumptions, and retrieved and default D_0 under convective assumptions (red) as a function of $\mathbf{A}$ diagonal value. b) same as a), except as a function of $\mathbf{S}_\mathrm{a}$ diagonal value divided by $\mathbf{S}_\mathrm{x}$ diagonal value. In both panels, the fraction of profiles exceeding the information content value on the x-axis is indicated by the dashed line and tick marks on the right y-axis.

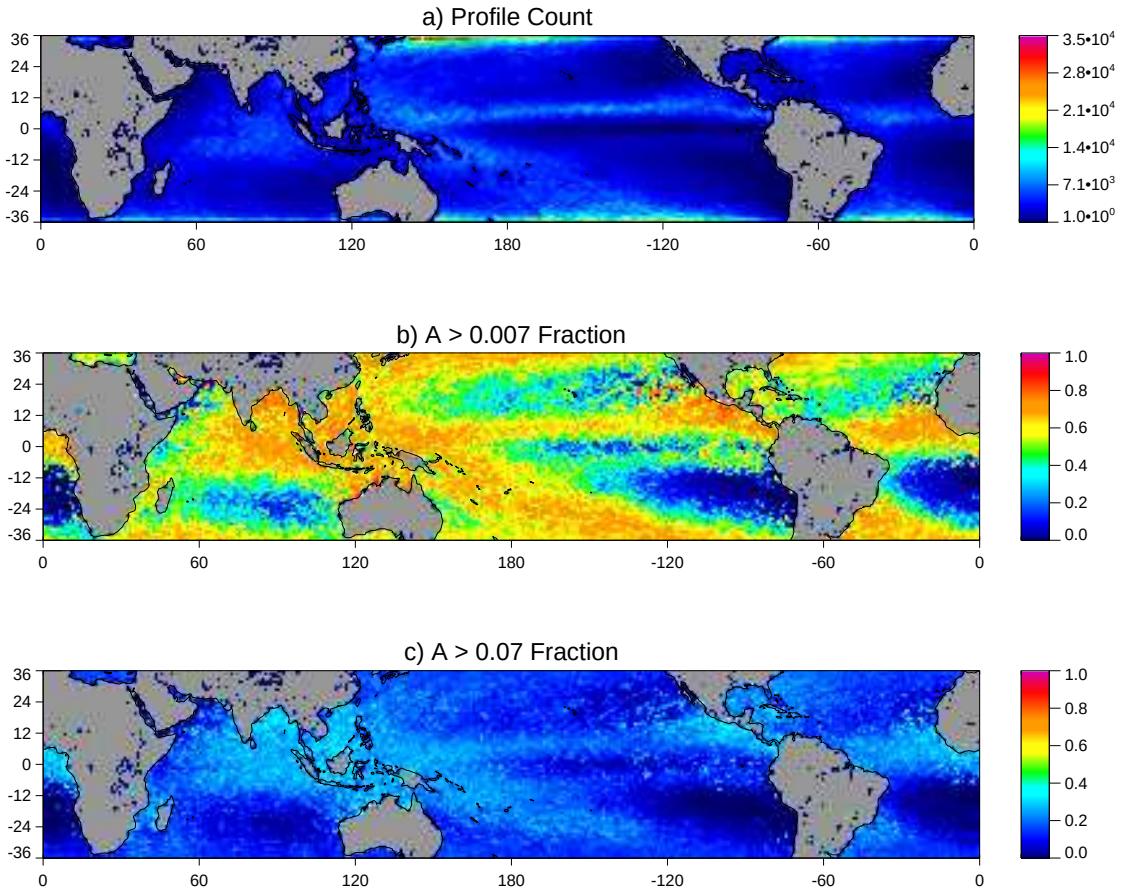

FIG. 2. Top panel: number of profiles in $1\times1^\circ$ grid boxes. Lower panels: fraction of profiles in each grid box that exceed the threshold of information content indicated.

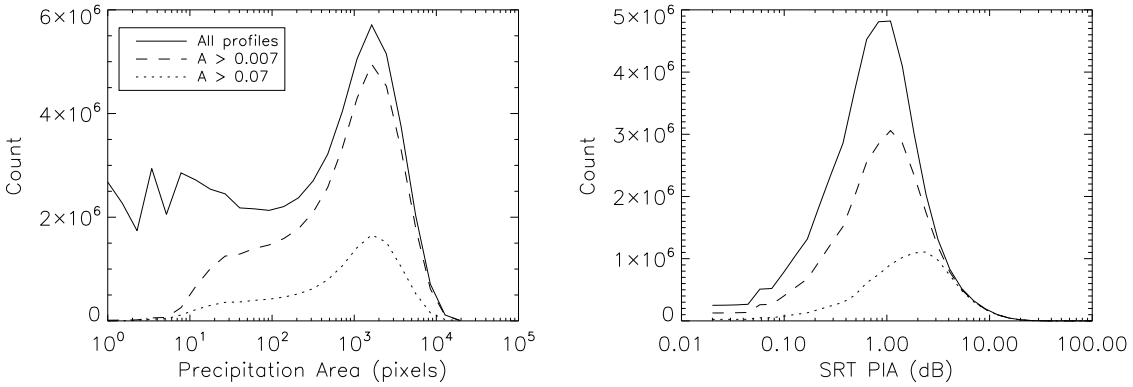

FIG. 3. a) Histogram of profiles by precipitation area for different information content thresholds. b) same as a), except as a function of surface reference path-integrated attenuation.

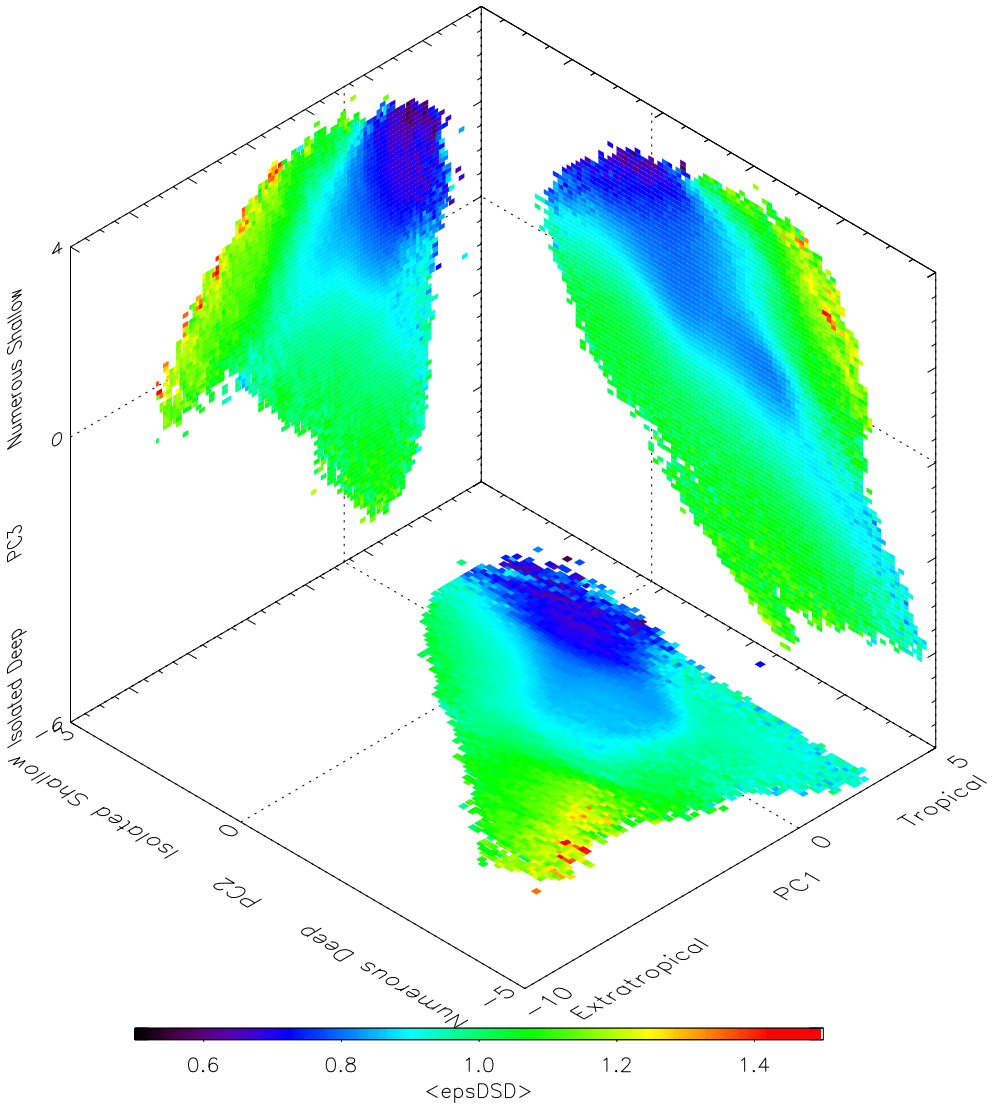

FIG. 4. Mean value of ϵ_{DSD} in the PC1-PC2, PC2-PC3, and PC1-PC3 planes for warm rain.

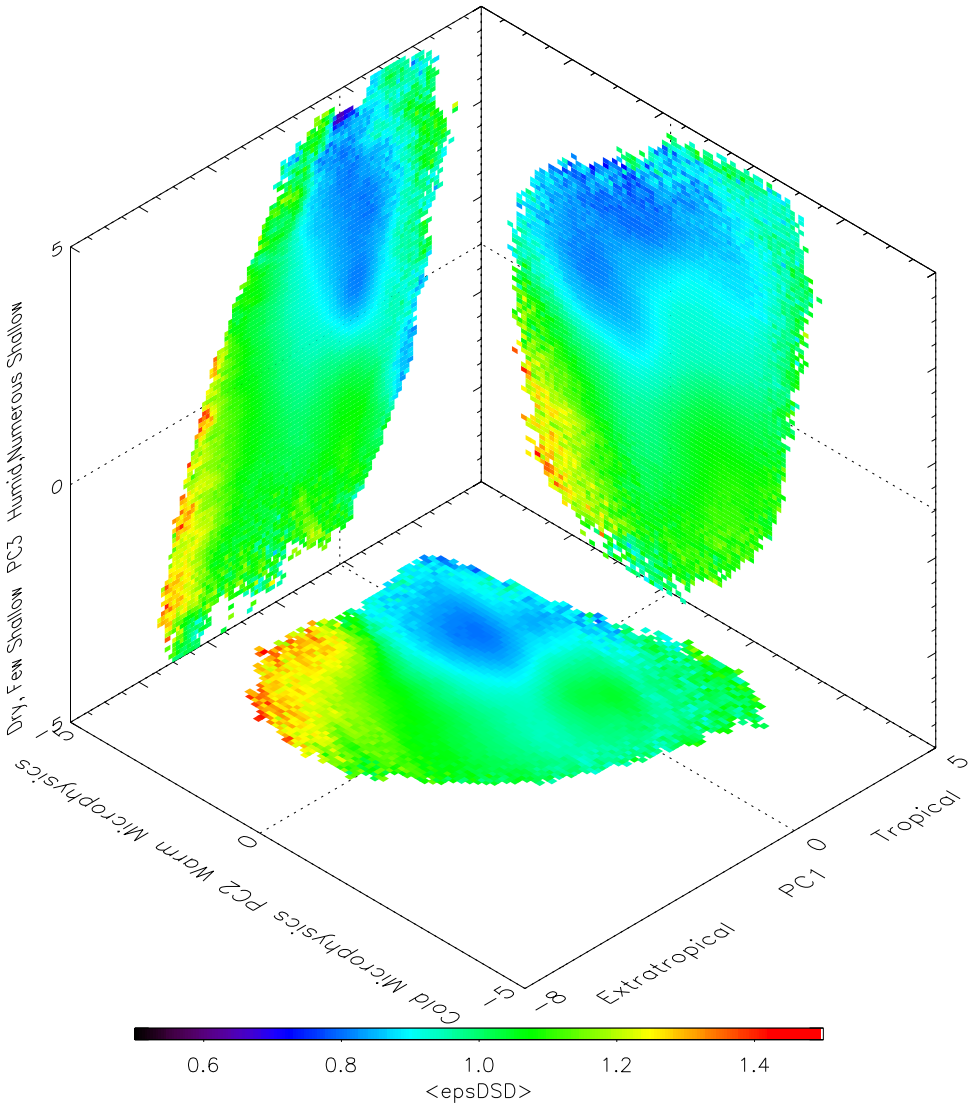

FIG. 5. Mean value of ϵ_{DSD} in the PC1-PC2, PC2-PC3, and PC1-PC3 planes for cold rain.

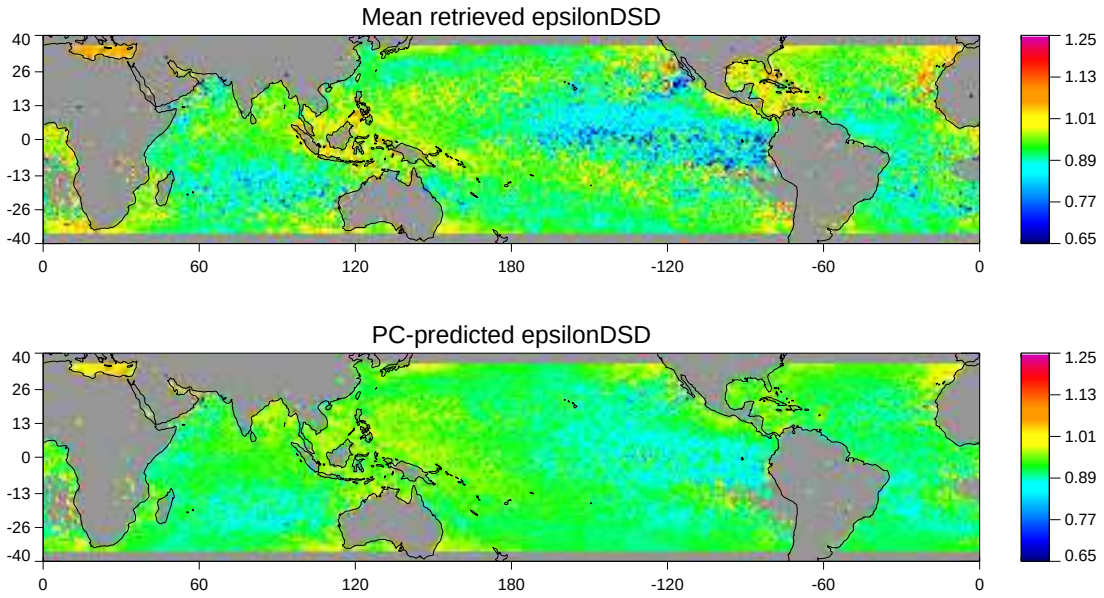

FIG. 6. Mean and predicted values of ϵ_{DSD} at the $\mathbf{A} > 0.007$ threshold gridded at 1° resolution.