Bayesian Retrievals of Vertically Resolved Cloud Particle Size Distribution Properties

DEREK J. POSSELT

University of Michigan, Ann Arbor, Michigan

JAMES KESSLER

NOAA/Great Lakes Environmental Research Laboratory, Ann Arbor, Michigan

GERALD G. MACE

University of Utah, Salt Lake City, Utah

(Manuscript received 16 August 2016, in final form 30 November 2016)

ABSTRACT

Retrievals of liquid cloud properties from remote sensing observations by necessity assume sufficient information is contained in the measurements, and in the prior knowledge of the cloudy state, to uniquely determine a solution. Bayesian algorithms produce a retrieval that consists of the joint probability distribution function (PDF) of cloud properties given the measurements and prior knowledge. The Bayesian posterior PDF provides the maximum likelihood estimate, the information content in specific measurements, the effect of observation and forward model uncertainties, and quantitative error estimates. It also provides a test of whether, and in which contexts, a set of observations is able to provide a unique solution. In this work, a Bayesian Markov chain Monte Carlo (MCMC) algorithm is used to sample the joint posterior PDF for retrieved cloud properties in shallow liquid clouds over the remote Southern Ocean. Combined active and passive observations from spaceborne W-band cloud radar and visible and near-infrared reflectance are used to retrieve the parameters of a gamma particle size distribution (PSD) for cloud droplets and drizzle. Combined active and passive measurements are able to distinguish between clouds with and without precipitation; however, unique retrieval of PSD properties requires specification of a scene-appropriate prior estimate. While much of the uncertainty in an unconstrained retrieval can be mitigated by use of information from 94-GHz passive brightness temperature measurements, simply increasing measurement accuracy does not render a unique solution. The results demonstrate the robustness of a Bayesian retrieval methodology and highlight the importance of an appropriately scene-consistent prior constraint in underdetermined remote sensing retrievals.

1. Introduction

Clouds are a critical component of Earth's climate system, acting to influence the radiation budget, the hydrologic cycle, and the vertical distribution of water and energy (Stephens 2005). The details of cloud radiative effect, as well as precipitation rate and distribution, are sensitive to the vertical and horizontal distribution of hydrometeor size and phase within clouds. In particular, the distribution of cloud particle sizes [the cloud particle size distribution (PSD)] plays a role in determining the vertical transport of water mass, latent and radiative heating, and precipitation rate. Estimates of cloud particle sizes at and near the tops of clouds have been available for some time from passive spaceborne measurements [e.g., the Moderate Resolution Imaging Spectroradiometer (MODIS); Platnick et al. 2003]. Spaceborne radar has provided information on the vertical distribution of cloud-sized particles (*CloudSat*; Stephens et al. 2008) and precipitation-sized particles [e.g., from the Tropical Rainfall Measuring Mission (TRMM; Kummerow et al. 2000) and Global Precipitation Measurement mission (GPM; Hou et al. 2014)]. These measurements, along with a growing volume of observations from field campaigns, are providing crucial global data on cloud properties.

Corresponding author e-mail: Derek J. Posselt, derek.posselt@jpl.nasa.gov

DOI: 10.1175/JAMC-D-16-0276.1

While the observational volume continues to grow, unambiguous retrieval of cloud particle size distributions remains a significant challenge. In a recent paper (Posselt and Mace 2014, hereafter PM14), we explored the information content of upward-looking active and passive measurements of precipitating orographic clouds. We found that, in the absence of a sophisticated cloud structure algorithm and a priori estimate, a unique cloud PSD retrieval was not possible. The cloud observed in PM14 was mixed phase, and the difficulties retrieving ice PSD properties are well known (Wood et al. 2014). In this study, we follow the work of Lebsock et al. (2011), Mace et al. (2016), and Leinonen et al. (2016), and examine data from A-Train measurements of liquid-only clouds over the subtropical Pacific Ocean. The physical scenario is far simpler than in our previous experiments, as these clouds contain only one condensed phase. They may also be expected to contain two particle modes: a small (nonprecipitating) cloud mode and a larger (drizzle) mode. The droplet mode is directly related to the ambient concentrations of cloud condensation nuclei (CCN; Hegg et al. 2012). The evolution of the precipitation mode is connected to within-cloud microphysics, dynamics, and radiative heating, and is fundamentally a cloud-lifetime-limiting process (Wood et al. 2009). Restricting our observations to liquid clouds (with and without drizzle) also removes the influence of the particle shape on measurements, reducing the sources of uncertainty in the cloud PSD estimate.

Recent work has demonstrated that a combination of a single radar frequency, with a passive (microwave and/or visible/infrared) constraint on integrated cloud mass, has promise for retrieving warm rain precipitation (Lebsock et al. 2011; Mace et al. 2016; Leinonen et al. 2016) and mixed phase (PM14) cloud properties. However, a detailed, and in some cases complex, algorithm is needed to mitigate the cloud particle size/number ambiguity associated with use of a single radar frequency (Lebsock et al. 2011; Mace et al. 2016; Leinonen et al. 2016). Proper incorporation of prior knowledge may allow for a unique retrieval of cloud properties, even in the absence of sufficient observational constraint. In many cases, such detailed prior constraints are necessary to provide convergence to a solution in an iterative algorithm. In this paper, we utilize a Bayesian Markov chain Monte Carlo (MCMC) algorithm that explicitly samples the probability distribution of cloud properties and does not specifically require a prior estimate or cloud structure model. We explore the information contained in observations available from both passive and active remote sensing observations of cloud properties, and also explore the extent to which prior information may help to constrain the retrieval. Specifically, we utilize the flexibility and generality of MCMC to address the following questions:

- What is the structure of the retrieval space for estimates of low cloud properties from *CloudSat*? How does the information space change with addition of passive visible, shortwave infrared (SWIR), and microwave measurements?
- Is there sufficient information in combined active–passive measurements from A-Train to place a unique constraint on the PSD properties of liquid-only clouds?
- Previous studies indicate *CloudSat* is capable of distinguishing precipitating from nonprecipitating clouds (Haynes et al. 2009). How clear is this distinction in the full retrieval space returned by the MCMC algorithm?
- What additional information is contained in 94-GHz passive microwave observations obtained by the *CloudSat* Cloud Profiling Radar (CPR)?

In examining the influence of prior information, we consider two limiting cases. The *narrowly defined prior* (which we refer to as ''scene consistent'') assumes the geophysical context and the cloud types are known, so that the ranges of possible cloud particle size distribution properties are tightly constrained. In our specific application, this means that we assume the PSD properties are restricted to those that are consistent with shallow drizzling liquid clouds over the Southern Ocean. The *broadly defined prior* (which we refer to as ''climatological'') assumes no knowledge of the specific scene, but instead sets broad constraints on the PSD properties consistent with the expected range of conditions observable on Earth.

In addition to exploring the characteristics of the solution space in a remote sensing retrieval, we test a new version of our MCMC algorithm (three-stage delayed-reject MCMC) suitable for higher-dimensional retrieval problems. We illustrate how this algorithm can be applied to retrieval of vertical profiles of cloud properties, maintaining the flexibility of the Bayesian framework and simultaneously accounting for covariance in cloud properties in the vertical. We wish to emphasize at the outset that we do not expect radar-only (or even radar + passive microwave and visible/shortwave IR reflectance) observations to provide sufficient constraint on a bimodal cloud particle size distribution. Rather, our aim is to determine how much information the observations contain, both if one assumes scene-dependent constraints and if one relaxes the prior constraints. In addition, we do not seek to describe a new operational retrieval algorithm, and there are certainly important sources of variability that we will not be able to consider (e.g., radar beamfilling, attenuation). Our experiments lay the foundation for future research that will examine several related questions:

TABLE 1. List of observations and the assumed uncertainty. Values in parentheses are used in sensitivity experiments (see text for details).

Observation	Uncertainty
0.55-μm reflectance (%)	25 (5)
2.10-μm reflectance (%)	25 (5)
Cloudy − clear Tb94 (K)	5.0
Radar reflectivity (dBZ)	1.0

- What is the effect of adding measurements in the form of additional active or passive microwave frequencies, Doppler spectrum observations, infrared frequencies, or multiangle imaging? Such observations have been collected in recent field campaigns [e.g., the joint Radar Definition Experiment–Integrated Precipitation and Hydrology Experiment (RADEX–IPHEX)], and forthcoming experiments will examine this question.
- What is the utility of adding information on the internal cloud dynamics and thermodynamic state? Is it possible to utilize a numerical model in the retrieval to enforce a scene-consistent prior? Such questions are the purview of data assimilation, a topic that is beyond the scope of this paper.
- What is the effect of imposing more restrictive structure functions (e.g., allowing cloud particles to have approximately the same number and size at and near the cloud top, but very different at cloud base)? We leave this as a suggestion for future work, and plan to present results of experiments using different prior assumptions in a subsequent paper.

The remainder of this paper is structured as follows. We describe the observations, forward models, and retrieval algorithm in section 2. Section 3 contains a presentation of the retrieval results, followed by a discussion of the limitations of our work and suggestions for future work in section 4. We offer a summary and conclusions in section 5, followed by three appendices that describe the details of the Bayesian MCMC algorithm.

2. Observations, forward models, and retrieval algorithm

a. A-Train observations and cloud scene description

Observations used in the retrieval consist of MODIS 0.55- and 2.1-μm reflectance and *CloudSat* radar reflectivity and 94-GHz brightness temperature. While our previous work has illustrated the value of lower-frequency brightness temperatures (e.g., 23 and 31 GHz), the effective footprint sizes of these observations in the A-Train are much larger than those of *CloudSat* and MODIS. Incorporation of these measurements into the retrieval would have required additional treatment of the effects of partial cloudiness in these low broken clouds, and therefore we have decided not to include them here. MODIS data were obtained from collection 6 downloaded from the Level-1 and Atmosphere Archive and Distribution System (LAADS) site and extracted from the granule observed at 1955 UTC 6 January 2007 over the eastern South Pacific Ocean (http://dx.doi.org/10.5067/MODIS/MYD01.006). *CloudSat* radar reflectivity was obtained from the *CloudSat* data processing center from granule 3691 from version 4 of the 2B-GEOPROF product. The 94-GHz brightness temperature is an experimental product produced by using the *CloudSat* radar antenna as a passive microwave observing system between observations of the active radar return. Because there is no onboard calibration, rather than using the absolute brightness temperatures, we utilize the cloudy–clear brightness temperature difference. This quantity is computed as the difference between the Tb94 observed for the cloudy footprint and an along-track average of the Tb94 in clear air on either side of the cloud. We ensure the profiles are cloud free by using the cloud mask flag in the 2B-RL-GEOPROF product (Mace and Zhang 2014) that combines the radar cloud mask (Marchand et al. 2008) and the cloud mask produced by the *CALIPSO* lidar. We note that recent work has reported on retrievals of cloud properties using *CloudSat* Tb94 calibrated against AMSR-E and AMSR-2 (Lebsock and Suzuki 2016).

Uncertainties in the observations are specified consistent with optimal estimation retrievals (Mace et al. 2016) and are listed in Table 1. Examination of the retrieval PDFs is conducted for three different cloud profiles over the remote subtropical ocean. The spatial context may be seen in the MODIS scene depicted in Fig. 1. The cloud located farthest to the south (profile 1) has relatively small visible reflectance and near-infrared reflectance (~0.33 and ~0.2, respectively; Table 2; Fig. 2c), indicating this cloud likely contained primarily drizzle drops. The reflectance at near-infrared wavelengths decreases with increasing cloud droplet size, and reflectance at 2.1-μm wavelength derives from near the cloud top (Platnick 2000). In contrast, the profiles determined to have both precipitation and cloud droplets (profile 2) and cloud droplets only (profile 3) both exhibit larger visible reflectance (0.69 and 0.77, respectively) and SWIR reflectance (0.30 and 0.31, respectively; Table 2; Fig. 2c). Evidence for differences in cloud internal structure can be seen in the *CloudSat* radar reflectivity (Fig. 2a). Cloud 1 exhibits relatively large radar reflectivity (maximum of −4 dBZ), but with a cloud top that is lower than clouds 2 and 3 by approximately one *CloudSat* range bin. Cloud 2 almost certainly contains precipitation-sized

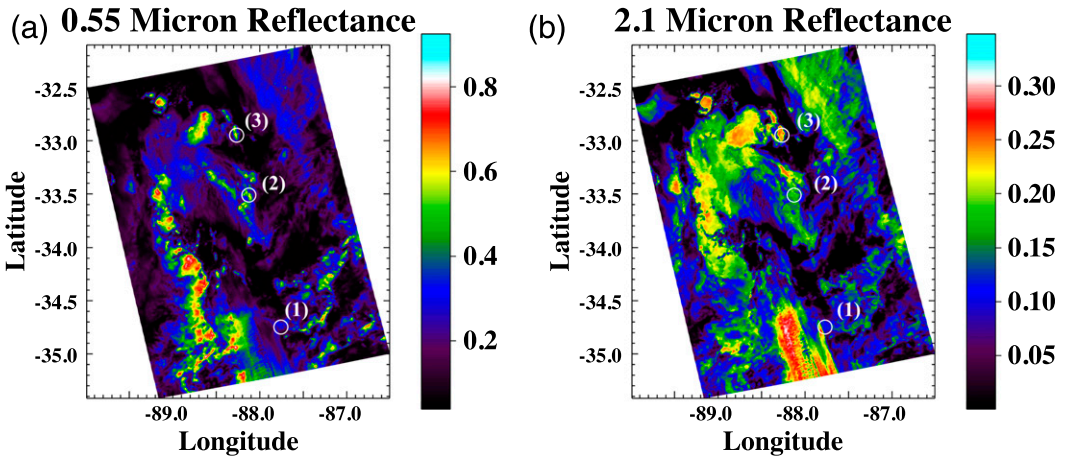

FIG. 1. MODIS (a) 0.55- and (b) 2.10-μm reflectance over the scene of interest. Cloud profiles analyzed in the text are circled and labeled consistent with the annotations in Fig. 2, below.

droplets, as it exhibits reflectivity values greater than zero in the lower levels of the cloud. Reflectivity in cloud 3 does not exceed −21 dBZ and therefore likely contains only small cloud particles, albeit enough of them to lead to relatively large visible and SWIR reflectance values. The fact that the cloud-top properties are very similar for clouds 2 and 3, while the interior structure is clearly very different, highlights the importance of a multivariate approach to retrieval of cloud properties (Mace et al. 2016). We do not plot *CloudSat* reflectivity lower than 1000 m above the surface, as the signal-to-noise ratio is low in this region because of the presence of ground clutter.

b. Forward models

All retrievals of cloud properties from remote sensing measurements rely on some form of model to connect the retrieved quantities to the observations. In our case, the measurements consist of 94-GHz radar reflectivity and brightness temperature, as well as 0.55- and 2.1-μm reflectances. Consistent with observations of drizzling or lightly precipitating cumulus clouds over the ocean (Rauber et al. 2007; Mace et al. 2016), we assume there are two populations of condensed phase particles present: a small (cloud) mode and a larger (drizzle) mode. We further assume the PSD can be described by a modified version of a gamma probability density function of the form

$$n_i(D_i) = N_{0_i}\left(\frac{D_i}{D_{0_i}}\right)^{\alpha_i}\exp\left(-\frac{D_i}{D_{0_i}}\right), \qquad (1)$$

where the subscript i refers to either the small mode (subscript s) or the large particle mode (subscript l). The retrieval returns estimates of the parameters of the gamma function: the reference number N_{0_i} and diameter D_{0_i}. Except where noted, we also retrieve the width parameter α_i. Note that the distribution defined in (1) has a single mode that occurs at a diameter of $D_{\text{mode}_i} = \alpha_i D_{0_i}$ and with a number of particles per unit diameter (the number density) at the modal diameter (D_{mode_i}) $N(D_{\text{mode}_i}) = N_{0_i}\alpha_i^{\alpha_i}\exp(-\alpha_i)$. Since the parameters N_{0i} and D_{0_i} set the number of particles and diameter at the mode of the distribution for species i for a given width parameter α_i, we will refer to them henceforth as the *modal* number and diameter, respectively.

The radar forward model is documented in PM14 and computes radar reflectivity using Mie theory for liquid (Bohren and Huffman 1983). For the shallow cumulus clouds observed in this study, we neglect multiple scattering. Microwave brightness temperatures are computed using the Eddington approximation (Kummerow 1993), with radiative properties consistent with Mie theory and surface emissivity set to 0.99. Visible and near-infrared reflectances are computed using the Radiant eigenmatrix solver (Christi and Gabriel 2003), again with radiative properties consistent with Mie theory.

TABLE 2. Observations from each profile used in this study.

	Rain	Cloud + rain	Cloud
0.55-μm reflectance	0.33	0.69	0.77
2.10-μm reflectance	0.20	0.30	0.31
Cloudy − clear Tb94 (K)	10.74	16.20	16.23
dBZ, 2520 m	—	−29.62	−28.29
dBZ, 2280 m	−30.57	−13.81	−27.14
dBZ, 2040 m	−12.88	−4.40	−21.48
dBZ, 1800 m	−6.96	−0.19	−22.84
dBZ, 1560 m	−4.85	0.42	−35.00
dBZ, 1320 m	−4.07	−0.52	—
dBZ, 1080 m	−5.69	−0.37	—

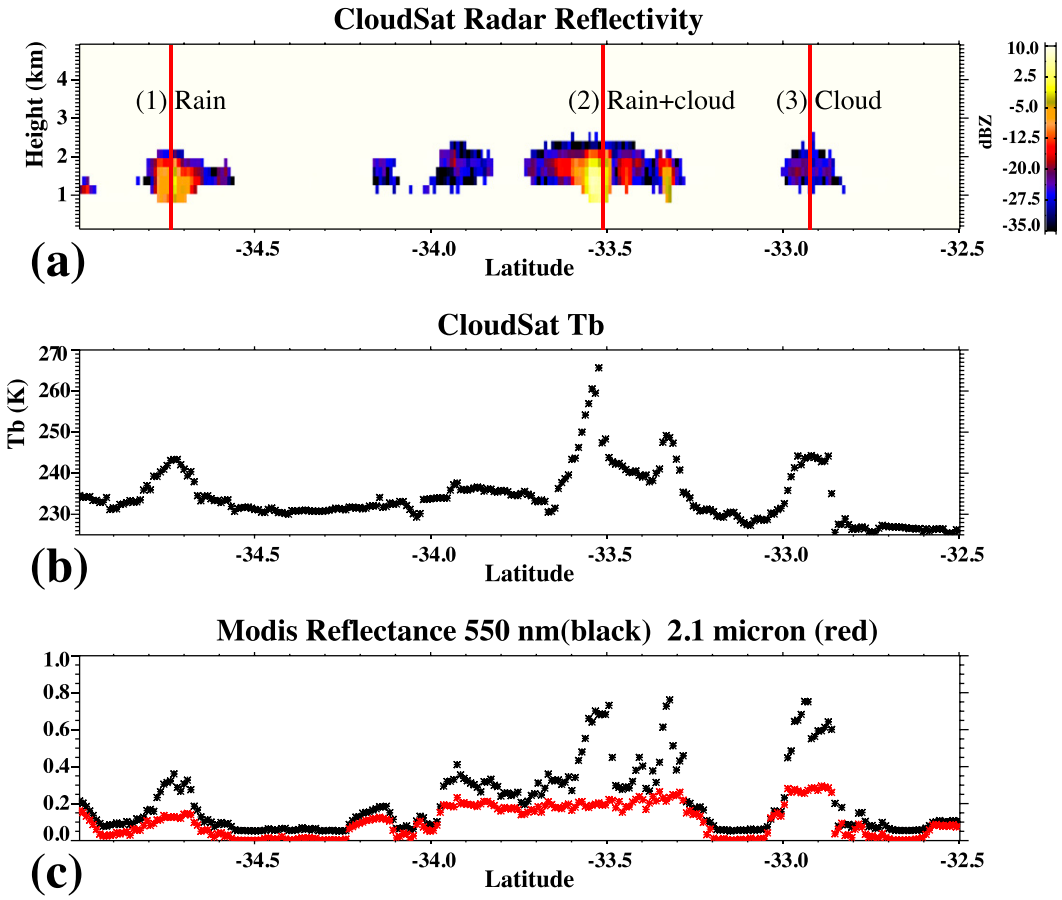

FIG. 2. (a) *CloudSat* reflectivity (dBZ; color shading) along with the locations of the three cloud profiles examined in the text, (b) *CloudSat* 94-GHz brightness temperature (K), and (c) MODIS 0.55- (black) and 2.10-μm (red) reflectance.

c. *Markov chain Monte Carlo algorithm*

The solution to a retrieval problem may be formulated probabilistically (e.g., Rodgers 2000) by computing the conditional probability of the retrieved variables, given prior knowledge and the set of available observations:

$$P(\mathbf{x}\,|\,\mathbf{y}) = \frac{P(\mathbf{x})P(\mathbf{y}\,|\,\mathbf{x})}{P(\mathbf{y})}. \tag{2}$$

In this statement of Bayes's relationship between conditional probabilities, $P(\mathbf{x})$ represents prior knowledge of the retrieved variables, $P(\mathbf{y})$ represents the probability space containing all possible values of the observations, and the likelihood $P(\mathbf{y}\,|\,\mathbf{x})$ represents the uncertainty in the observations and in the forward models (cf. Posselt et al. 2015). While many algorithms assume the probabilities in (2) are Gaussian so that the posterior distribution can be obtained via least squares error variance minimization, a more complete description of the retrieval solution may be obtained if we assume no specific form for the posterior distribution. MCMC algorithms generate a Markov chain that consists of samples from the posterior distribution produced by the combination of the prior and likelihood. A description of the theory that underlies MCMC can be found in Posselt (2013) and PM14, as well as references therein.

In this study, the vector of retrieved variables $\mathbf{x}$ contains the large and small cloud modal diameter and number and the gamma width parameter alpha, in each *CloudSat* range bin. As such, the retrieval vector is 6 × (number of radar range bins) long. In each experiment we assume no prior knowledge on $\mathbf{x}$, except that each retrieved variable lies within a realistic range of values (Table 3). We set ranges on the various components of $\mathbf{x}$ depending upon whether a scene-consistent or climatological prior is assumed. The observation vector $\mathbf{y}$ consists of the two reflectance values, the 94-GHz brightness temperature difference, and the 94-GHz reflectivity in each *CloudSat* range bin. As such, the observation

TABLE 3. List of retrieved state variables, along with their prior ranges for scene-consistent and climatological experiments.

	Scene		Climate	
	Min	Max	Min	Max
D_0, cloud droplets (μm)	0.1	15.0	0.1	30.0
N_0, cloud droplets (cm^{-4})	0.001	1000.0	0.001	50 000.0
α, cloud droplets	1.0	2.0	1.0	5.0
D_0, precipitation (μm)	0.1	100.0	0.1	1000.0
N_0, precipitation (cm^{-4})	0.0001	5.0	0.0001	500.0
α, precipitation	1.0	2.0	1.0	3.0

vector is 3 + (number of radar range bins) in length. In the MCMC algorithm, each step in the Markov chain is generated via random perturbation of the six state variables in a single layer. Acknowledging that the cloud properties will be correlated layer to layer, we spread this perturbation in the vertical using a Gaussian decorrelation function (see appendix A for details).

To simultaneously allow for rapid exploration of a large range of cloud properties and detailed analysis of local features in the solution space, we implemented a three-stage delayed-reject algorithm based on the work of Green and Mira (2001). A detailed derivation of the three-stage delayed-reject algorithm is presented in appendix B. An uncorrelated Gaussian proposal distribution was used to generate perturbations, and proposal variance was tuned to an acceptance rate of greater than 30% in each experiment. Delayed-reject algorithms allow for a generally larger accept rate than standard Metropolis–Hastings algorithms, and an accept rate between 30% and 60% was found in our case to strike a balance between thorough and efficient sampling. We utilize 24 simultaneous MCMC chains to more efficiently explore the retrieval space, and each chain is run to 100 000 iterations so that each MCMC experiment generates 2.4×10^6 forward model solutions per profile. Chains are initiated from values of the PSD parameters that are distributed evenly in the prior space, but we found that the results were completely insensitive to the choice of chain starting point. We assess convergence of the algorithm by computing the R statistic (Gelman et al. 2004), which compares variance between Markov chains with variance within each Markov chain (see appendix C for details). Each experiment converged to an R value of less than 1.1, indicating sufficient mixing was achieved.

3. Retrieval results

a. *Comparison of three cloud profiles with scene-consistent prior*

We first examine results from the scene-consistent prior, and compare the multidimensional retrieved probability density function for each of the three clouds of interest. As mentioned above, MCMC produces a sample of the full multidimensional probability distribution of the retrieved variables. In our experiments, the dimension of the retrieval space (the number of retrieved parameters) is equal to the number of cloudy layers multiplied by the number of retrieved variables. As the default is to allow variability in all three PSD parameters (modal length and number and PSD width) for cloud and precipitation droplet modes, each layer has a six-dimensional retrieval space. Examination of the output revealed that there was near complete nonuniqueness in the determination of the width parameter in the scene-consistent case, and, for this reason, we show only a depiction of the four-dimensional space consisting of modal diameter and number for cloud and precipitation modes. Examination of this space indicated that plots of the two-dimensional marginal distributions of the droplet and drizzle number parameters N_0 were also flat; nearly any combination of the two was capable of producing forward-modeled values consistent with the observations. There are interesting features in the 2D droplet modal diameter–drizzle modal diameter marginal and in the 2D marginal of the modal number parameter and modal diameter for each cloud species. Consequently, we show this subset of three two-dimensional marginals in Fig. 3. To illustrate the differences between the clouds, and the changes in retrieval space with depth, we plot two-dimensional marginal PDFs for the cloud and precipitation modal diameter and number for the cloud top, cloud top minus 480 m (third *CloudSat* range bin from the top) and cloud top minus 960 m (fifth *CloudSat* range bin from the top).

Comparison of the retrieval space at the cloud top in all three clouds (Fig. 3, top row) reveals little difference in the solution among the three profiles. The precipitation-mode diameter is very well constrained, and ranges from approximately 15–30 μm for all three cases. Cloud-mode diameter probability density is maximized at approximately 8–10 μm, with smaller droplets also highly likely. There is little change in probability density with change in precipitation-mode number, indicating changes in the precipitation-mode number have a small effect on the forward-modeled reflectivity, brightness temperature, or reflectance. Larger cloud droplet diameters are feasible in profiles that contain precipitation (profiles 1 and 2), but where the cloud droplet diameter is largest, the number parameter for cloud droplets is restricted to be small ($<\sim$200 cm^{-4}). In contrast, the cloud-only profile (profile 3) centers the cloud droplet modal diameter probability density maximum at approximately 8 μm. Differences in the retrieval space begin to emerge 500 m

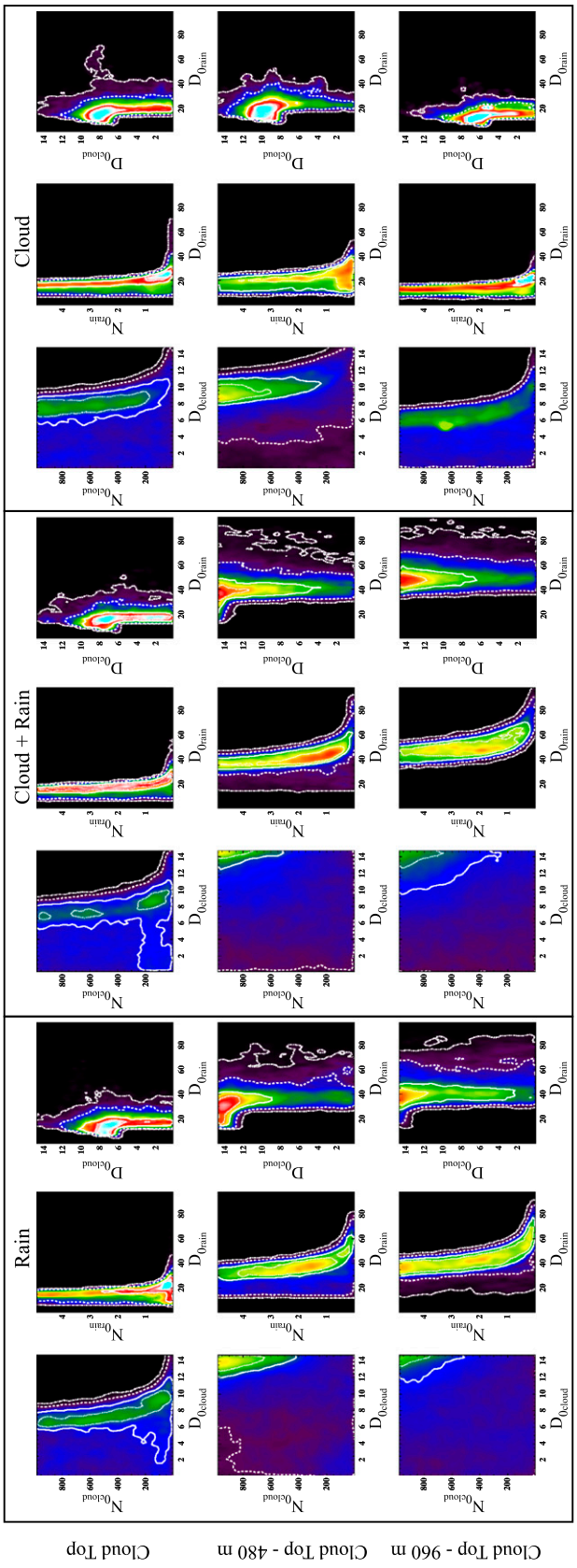

FIG. 3. Two-dimensional marginal probability density plots (color contours). White dot–dashed, dashed, solid, and dotted lines correspond to the 99.7%, 95%, 68%, and 38% probability contours, respectively. Color filled contours correspond to the probability density normalized to 1.0 over 51 bins in each parameter. For each cloud case, plots show the joint marginals of cloud droplet PSD parameters (in the left column), precipitation PSD parameters (in the center column), and cloud droplet and precipitation modal diameter (in the right column).

into the cloud (Fig. 3, center row). The profiles containing precipitation (profilesm1 and 2) have similar retrieval space characteristics, with modal precipitation drop diameters of approximately 40 μm and a maximum cloud droplet modal diameter of 15 μm (the maximum value allowed in the retrieval). In contrast, the profile that likely contains primarily cloud droplets (profile 3) exhibits a maximum in the cloud droplet modal diameter probability at 10 μm, and a maximum precipitation modal diameter probability at approximately 25 μm. Examination of the layers nearer the bottom of the cloud (as detected by the radar; bottom row, Fig. 3) reveals that the clouds containing precipitation continue to exhibit increases in precipitation-mode size with increasing depth below cloud top, while the cloud droplet-only profile (profile 3) exhibits smaller cloud and precipitation diameters than the layer immediately above. The small cloud and precipitation diameters in the lowest layer of profile 3 are likely due to the fact that this layer is at the bottom of the cloud.

In profiles 1 and 2, there is a small increase in precipitation-mode diameter with decreasing precipitation-mode number. While the sensitivity of radar reflectivity to changes in number is relatively small, in this case the increase in diameter compensates for the decrease in number to maintain the same radar reflectivity (and vice versa). In addition, all three profiles exhibit a decrease in precipitation-mode diameter as the cloud-droplet-mode diameter increases. In fact, the increase in curvature in the mode of the joint distribution of raindrop modal number and diameter space is due specifically to the constraint that the small mode diameter must always be smaller than the large mode diameter. As the small mode diameter increases, the large mode diameter must decrease to maintain the same radar reflectivity.

The features observed in the two-dimensional marginal distributions in Fig. 3 may be understood more generally in terms of the functional dependence of the observations on the PSD parameters. Specifically, note that the radar reflectivity may be very coarsely approximated as $Z \sim N_t D^6$. If we further assume that the reflectivity is additive so that it consists of the sum of contributions from droplets and rain, then we may see that 1) the major contribution to reflectivity sensitivity derives from the dependence on drop size and 2) there is nonuniqueness in the system in that a change in droplet number may be compensated by a change in precipitation number, etc.

Examination of the retrieval in the retrieved variable space provides the most direct information on the solution characteristics, but it is also instructive to examine the retrieval solution in terms of cloud mass and number. The liquid water content (g m^{-3}) can be computed from the modal diameter, number, and distribution width according to

$$q_i = a_{m,i} N_{0_i} D_{0_i}^{(b_{m,i}+1)} \Gamma(\alpha_i + b_{m,i} + 1). \quad (3)$$

Here, i corresponds to either cloud or rain, $a_{m,i}$ and $b_{m,i}$ are the coefficient and exponent (respectively) in the mass–dimensional power-law relationship, D_{0_i} and N_{0_i} are the modal diameter and number (respectively), and α_i is the width of the gamma distribution. In contrast to clouds that may contain ice particles, the assumption that liquid particles are spheres with constant liquid water density (ρ_L) prescribes the values of $a_{m,i}$ and $b_{m,i}$ to $\rho_L(\pi/6)$ and 3.0, respectively. The total number of particles per unit volume (cm^{-3}) is

$$N_{t,i} = N_{0_i} D_{0_i} \Gamma(\alpha_i + 1). \quad (4)$$

Plots of the retrieved vertical profiles of cloud and rainwater content and number are shown in Fig. 4. Because MCMC returns a sample of the multivariate posterior probability density function, it is possible to compute all manner of statistics of the retrieval. In these plots, we depict the center of mass of the distribution (mean and median) as well as the 5%, 25%, 75%, and 95% quantiles and the interquartile range (IQR). Viewed this way, we can determine 1) whether the mean is a robust measure of the center of mass, 2) whether distributions are skewed or symmetric, and 3) the distribution dispersion (commonly taken to be the single measure of retrieval uncertainty). Rather than plotting the standard deviation of the solution, as is typically done, we utilize the interquartile range as it is a more robust measure of dispersion in the distribution, and is resistant to outliers (Wilks 2011).

Examination of the retrievals across all three clouds of interest reveals the posterior distribution is strongly skewed in cloud number and cloud water content for the profiles that contain nonnegligible drizzle content (profiles 1 and 2). The mean and median differ, and the 95% quantile extends to relatively large values. In these cases, the mean is not an appropriate measure of the true cloud water content or number, and the standard deviation is not a representative measure of the error. In contrast, the distributions of rainwater content for profiles 1 and 2 are very symmetric, and the mean and median are nearly identical. Precipitation number is more symmetric than cloud number, but the fact that the numbers are very low makes the distributions skewed, as the values are hard bounded at zero. Uncertainties in cloud droplet content and number (as measured by the interquartile range) are largest a few hundred meters below cloud top. Uncertainty is smallest at cloud top in

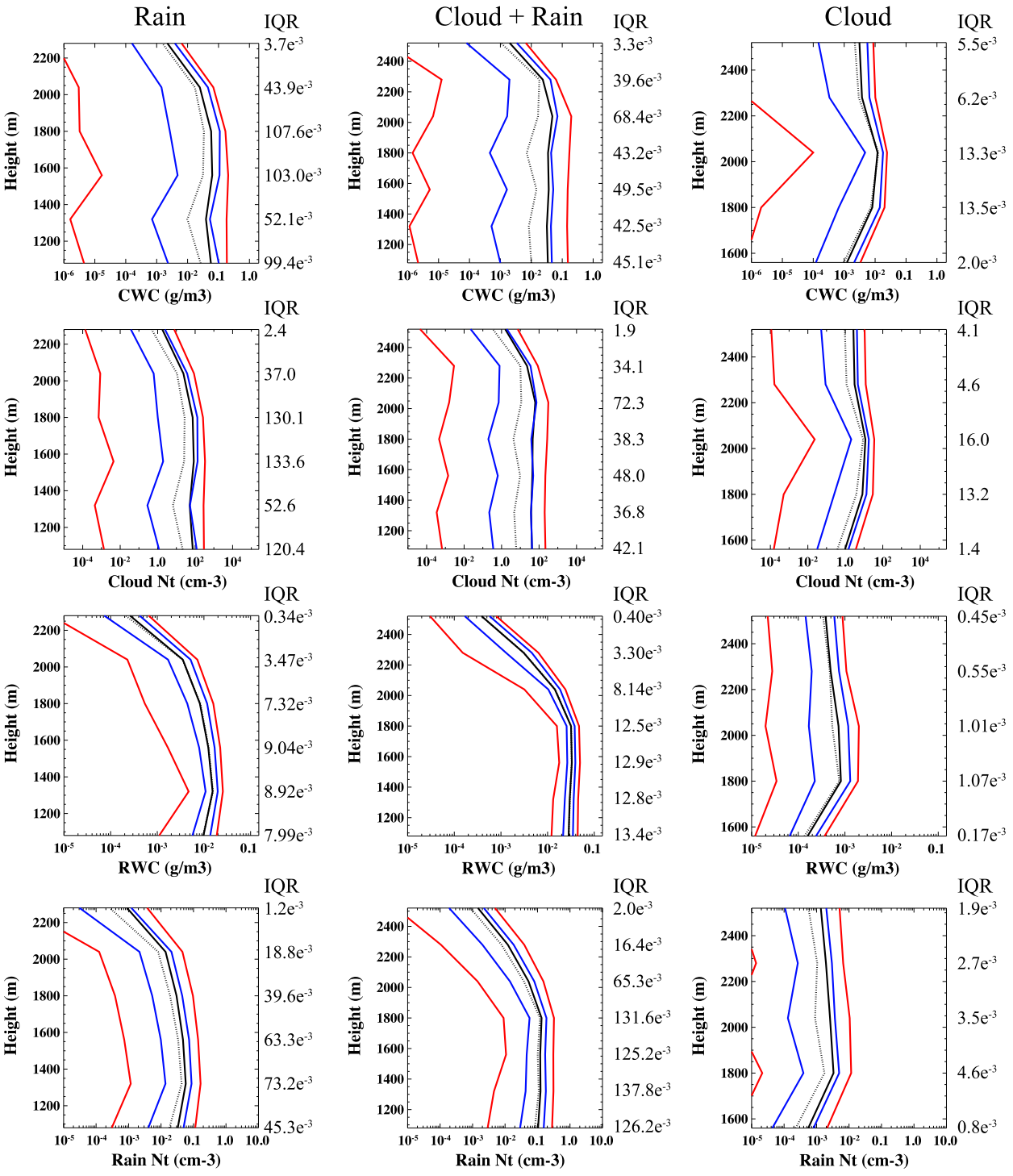

FIG. 4. Plots of the statistics of the retrieved (top) cloud droplet liquid water content (g m^3), (top middle) droplet number concentration (cm^{-3}), (bottom middle) precipitation water content (g m^{-3}), and (bottom) precipitation drop number concentration (cm^{-3}) for the profile containing (left) primarily precipitation, (center) a combination of cloud droplets and precipitation, and (right) primarily cloud droplets. Each plot depicts the mean (solid black line), median (dashed black line), 25th and 75th percentiles (blue lines), and 5th and 95th percentiles (red lines) of the retrieved probability distribution. Note that all variables are plotted on a log scale for ease of interpretation. The IQR is reported as a column of values along the right-hand side of each plot.

the profiles that contain precipitation, with IQR smallest at the cloud base in the nonprecipitating profile. Uncertainty in precipitation content and number is maximized in the middle and lower portions of each cloud, and is again smallest at the cloud top. The very small retrieval uncertainty at and near cloud top results from the assumption of a constant 1 dB error in the radar reflectivity. Small changes in modal number and diameter will lead to large changes in reflectivity when the reflectivity is low. Consequently, the water content and number are very tightly constrained for low radar reflectivity but less well constrained at higher reflectivity. In fact, examination of the radar reflectivity values for each of the three profiles (Table 2) indicates the uncertainty scales with the reflectivity value itself.

A brief discussion of these results is warranted before we proceed to exploration of the specific influence of the various observations. First, a unique retrieval estimate implies that a single most likely set of values of the PSD properties can be obtained. It is clear from the probability contour plots (Fig. 3) that the precipitation-mode diameter is relatively well constrained, and that the observations place an upper bound on the cloud droplet size at and near the cloud top. However, there is very little information in the observations to constrain the droplet number. In each profile and cloud layer, any value of cloud droplet and precipitation number concentration that lies in the range between the prior minimum and maximum values would produce very similar forward-modeled observations. It should be noted that we used a very conservative estimate for the error on the visible reflectance to account for cloud 3D effects across this very spatially inhomogeneous scene. We shall explore the influence of increased visible and SWIR reflectance accuracy presently.

In addition, we utilized a restrictive range on the retrieval space, but we applied no further information on the relationship between cloud and precipitation-mode diameter and number. Knowledge of the geographic region, the cloud-top height, and the thermodynamic and dynamic environment could be used to place additional constraints on the relationship between cloud modes that may very well lead to a unique retrieval. It may also be possible to apply constraints based on spatial correlation between retrieval parameters. We shall discuss this further in section 4.

b. *Observation-specific influence on the retrieval*

We now focus on the profile containing cloud and precipitation (profile 2), and examine the influence of each type of observation on the retrieval. We analyze the retrieval space variables (modal diameter and number), as these are the quantities directly estimated in the MCMC algorithm. To examine the influence of each observation, we conduct several different experiments in which we utilize various combinations of reflectance, reflectivity, and 94-GHz brightness temperature. Each involves a separate run of the MCMC algorithm, so that each has an estimate of the full joint posterior distribution. We compare results for the distribution center of mass (mean or median) and dispersion (standard deviation or interquartile range). Recall that the prior distribution for all parameters is bounded uniform, and the minimum is set to nearly zero. The median and interquartile range are both trivial to compute for the uniform distribution, and are equal to $a + 0.5(b - a)$ and $0.5(b - a)$, respectively (Fig. 5; dashed black line), where a and b represent the minimum and maximum values used to bound the uniform distribution (Table 3).

The control experiment (Fig. 5; black solid line) utilizes all available observations, and assigns to them the default values of observation uncertainty (Table 1). In a retrieval that uses only radar reflectivity as a constraint (Fig. 5; purple line), the median and IQR are approximately equal to the prior for the cloud modal diameter and number, reflecting the fact that the reflectivity contains very little information on the small cloud mode for this case. In contrast, the drizzle diameter and number are constrained nearly as well with reflectivity as with all observations, indicating the bulk of the information on the drizzle mode is derived from the radar. When only reflectance is used as an observation (Fig. 5; blue line), there is near zero information on any of the retrieved parameters, even the cloud droplet mode. This is due to the fact that the errors in reflectance are assumed to be 25%, as we shall see shortly.

An experiment run without 94-GHz brightness temperature (Fig. 5; cyan line) yields results nearly identical to the experiment that includes Tb94 as an observation. Tb94 has been shown to provide an important early constraint in optimal estimation retrievals (Mace et al. 2016), but in our case, the radar provides the bulk of the information, and little is added by the passive microwave observations. It is possible that use of different types of observations, with smaller uncertainty, may change the balance of information (Lebsock and Suzuki 2016), and we leave such an analysis for future study. The effect of the assumption of 25% error in the MODIS reflectance can be assessed by rerunning the reflectance-only and radar reflectivity + reflectance cases with an assumed reflectance observation uncertainty of 5% (Fig. 5; green and red lines). In the case in which reflectance and reflectivity observations are used, there is little difference in the drizzle-mode solution. However, the median cloud droplet modal diameter and number in layers below the cloud top increase to near their maximum allowed values of

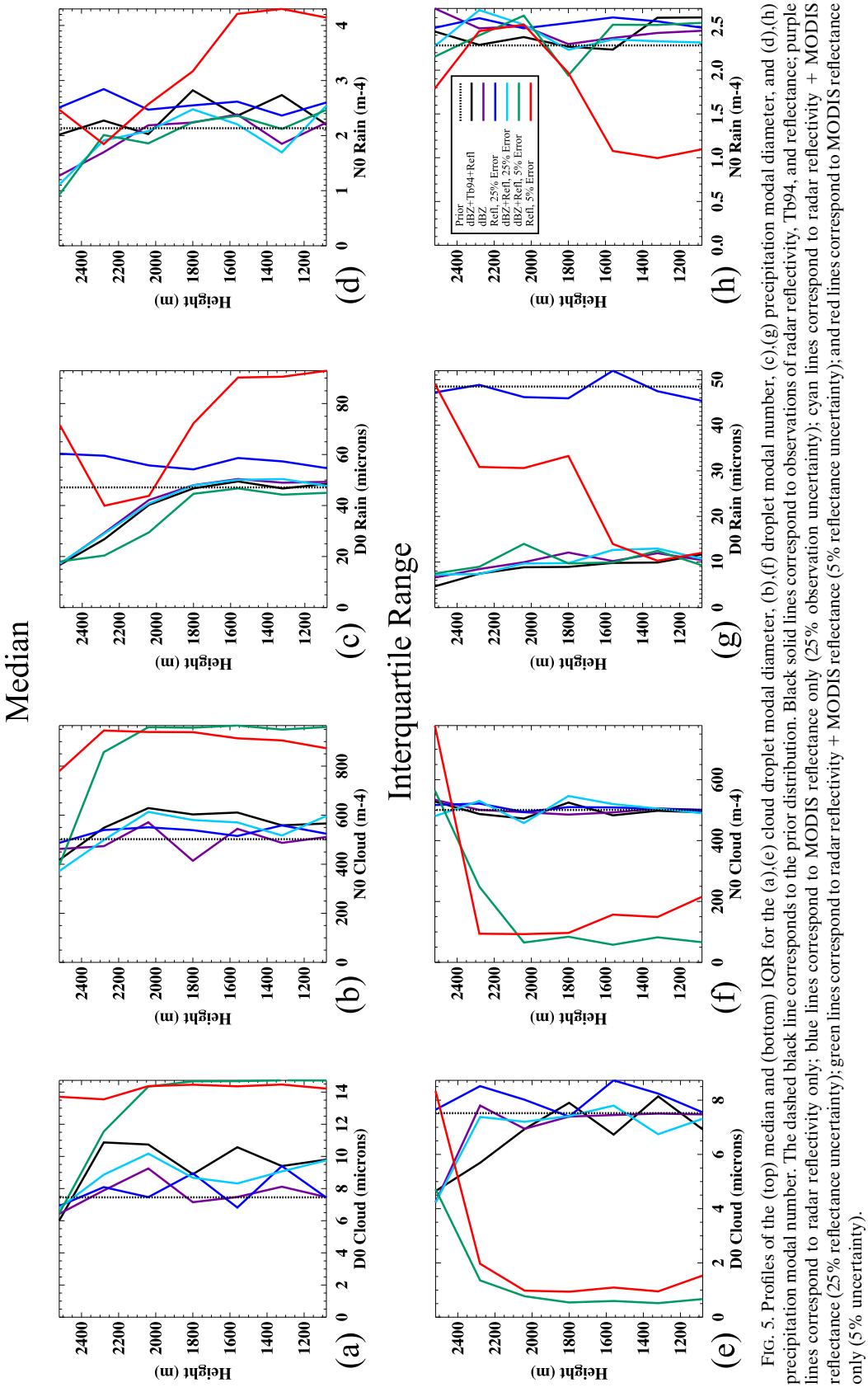

FIG. 5. Profiles of the (top) median and (bottom) IQR for the (a),(e) cloud droplet modal diameter, (b),(f) droplet modal number, (c),(g) precipitation modal diameter, and (d),(h) precipitation modal number. The dashed black line corresponds to the prior distribution. Black solid lines correspond to observations of radar reflectivity, Tb94, and reflectance; purple lines correspond to radar reflectivity only; blue lines correspond to MODIS reflectance only (25% observation uncertainty); cyan lines correspond to radar reflectivity + MODIS reflectance (25% reflectance uncertainty); green lines correspond to radar reflectivity + MODIS reflectance (5% reflectance uncertainty); and red lines correspond to MODIS reflectance only (5% uncertainty).

15 μm and 1000 cm^{-4}, respectively, while the interquartile range narrows considerably. When only the reflectance is used as an observation, and error is assumed to be 5%, the cloud droplet mode is nearly as well constrained (except near the cloud top). The drizzle-mode diameter is well constrained low in the cloud but not well constrained near the cloud top. The rain number near the cloud base is forced to increase when only the reflectance is used because the reflectance places an integral constraint on the retrieval in the absence of radar information. A number of questions naturally arise:

1) What is the appropriate uncertainty for MODIS reflectance in this case? As we shall see, specification of smaller uncertainty causes difficulty for the retrieval when we relax the prior constraints, possibly because of 3D radiative effects in the shallow cumulus clouds contained in this scene.
2) Would the 94-GHz brightness temperature have placed a larger integral constraint on both cloud and drizzle modes if it were more accurately observed? We shall see that it does indeed place a stronger constraint on the retrieval (in terms of reduction in posterior dispersion) when a less restrictive prior is used.

c. *Retrieval of vertically resolved cloud PSD properties using a climatological prior*

In all of the results reported above, it was assumed that the retrieval was restricted (via prior knowledge) to cloud properties consistent with the scene of interest. That is, we assumed that it was possible to specify in advance the realistic range of values for the modal diameter and number of both cloud and precipitation. We now explore the effect of relaxing these constraints, and instead consider the opposite extreme. In this case, we assume we only know the broad global and climatological distribution of cloud and drizzle and do not place any scene-consistent constraint on the retrieval. In practice, this amounts to allowing the modal cloud diameter to be as large as 30 μm and the modal number to be as large as 50 000 cm^{-4}. These numbers are intentionally extreme, and are designed to allow the retrieval to explore a very wide range of possible values of droplet and precipitation content and number. The maximum modal droplet diameter and number equate to a cloud liquid water content of 30.0 g m^{-3} and droplet number concentration of 18 000 cm^{-3}. We allow the modal precipitation diameter to be as large as 1000 μm and the modal rain number to be as large as 500 cm^{-4}, which correspond to rainwater content of 190 g m^{-3} and number concentration of 300 cm^{-3}. It is important to remember that these numbers are not expected to be in any sense typical but instead serve as the far upper bound on a large range of cloud water contents and number concentration values. As in the experiments with different observations, we run experiments for cloud 2, the profile that contains both cloud droplet and precipitation modes.

Before we compare the climatological prior with the scene-consistent results, we return to the question of the effect of fixed versus variable gamma width parameter. We show results in Fig. 6 for the same three levels as were shown in Fig. 3. The left column corresponds to an experiment that utilizes variable gamma width, while in the right-hand column the gamma width is fixed to 1.0 and 1.25 for the cloud droplet and drizzle modes, respectively. It is clear that a fixed alpha leads to a more well-constrained retrieval of the modal precipitation diameter. Values are concentrated around 20 μm at cloud top and around 30 and 40 μm lower in the cloud. However, close inspection of the 2D PDFs indicates that in fact there is larger uncertainty in the drizzle number when alpha is fixed. The range of values inside of the 99.7% probability contour extends to 300 cm^{-4} for a fixed alpha while it is less than 100, 250, and 80 cm^{-4} in layers at cloud top, top $-$ 480 m, and top $-$ 960 m, respectively. The cloud number exhibits the same range of variability in both experiments, but the droplet diameter is restricted to a smaller range of values, leading to a greater concentration of probability mass. While it may, in principle, be desirable to apply an a priori constraint on the gamma width parameter (resulting in an overall decrease in the number of degrees of freedom), our results indicate that variability in alpha allowed the MCMC algorithm to find values that better fit the observations than when alpha is artificially fixed.

Comparison between the posterior statistics for the scene-consistent and climatological priors for the physical space retrieval is shown in Fig. 7. These plots are similar to those presented in Fig. 4, but for simplicity we omit the 5% and 95% quantiles. Before we examine the results in detail, note that the uncertainty in the solutions (as quantified by the IQR) differs between the scene-consistent and climatological results by at least an order of magnitude. This in itself reflects the influence of the prior on the solution; *given a larger range of in the prior estimate, the observations alone are not able to constrain the solution*. The exception is the drizzle water content, which remains relatively well constrained even with a large prior range. If the interquartile range is taken as a measure of the uncertainty in the retrieval, the error in the retrieval increases by approximately 10 times and 100–1000 times for cloud water content and number, and by 40–100 times for drizzle number. In contrast, the errors for the drizzle water content increase only 2–5 times.

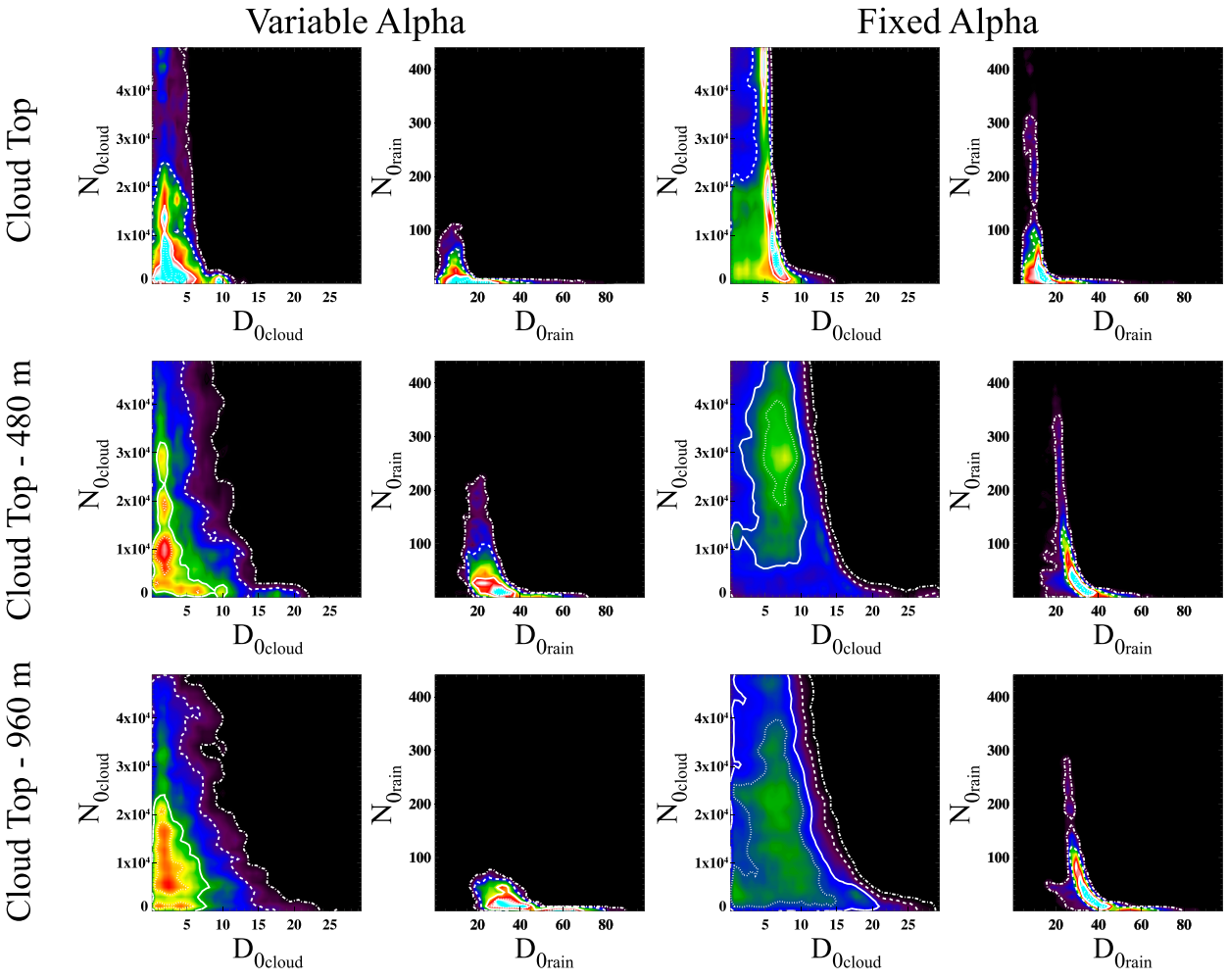

FIG. 6. As in Fig. 3, but for experiments in which the prior was allowed a climatological range. The two left columns correspond to an experiment in which the width of the gamma PSD was allowed to vary in the retrieval; alpha was set constant in the experiment represented in the two right columns.

As in the earlier results, we compare the influence of different observations, but in this case we only show the effect of adding observations of 94-GHz brightness temperature difference. In contrast to the scene-consistent prior (Fig. 5), dTb94 observations lead to a significant reduction in the uncertainty for all variables considered. The IQR for the cloud water content is reduced by a factor of 2–10, bringing it to nearly the value of the scene-consistent prior in the lower portions of the cloud. Cloud number concentration uncertainty is also reduced by approximately a factor of 10; in this case reducing the error to approximately 100 times the scene-consistent value. Drizzle content is relatively well constrained by the radar, so is less affected by dTb94 observations. However, there is a reduction in uncertainty in the middle and lower portions of the cloud from an IQR of approximately 0.06 to 0.04, reducing the error to 4 times the scene-consistent value. Drizzle number IQR is unaffected by inclusion of Tb94 observations, and in both cases is approximately 30–100 times the error of the scene-consistent prior.

There are two key conclusions that may be drawn from these results:

1) There is not enough information in the observations alone to constrain any of the variables except the drizzle water content. For all others, a scene-consistent prior will be required.
2) In the case of cloud water content, a large portion of uncertainty introduced via unknown prior range can be mitigated if the 94-GHz brightness temperatures are used as an additional observational constraint.

4. Discussion

As we noted at the outset, radar observations alone cannot be expected to provide unique constraint on a

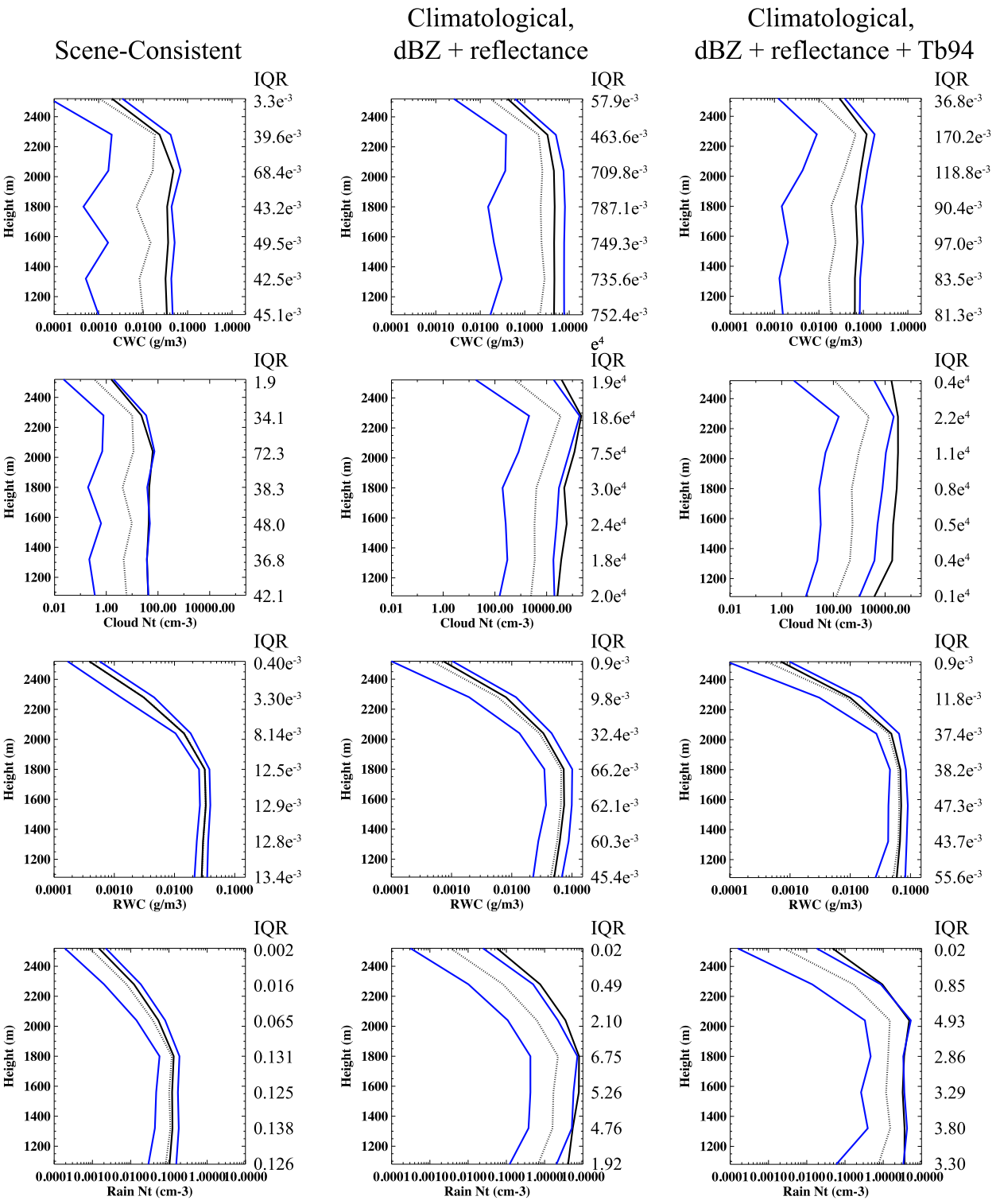

FIG. 7. As in Fig. 4, but in this case all three profiles correspond to the cloud that contains both droplets and precipitation and we have plotted only the 25th and 75th percentiles, mean, and median of the posterior distribution. The left column is a reproduction of the center column in Fig. 4. The center and right columns correspond to experiments in which a climatological prior was used. In the center column, the experiment used observations of MODIS reflectance and radar reflectivity, whereas the right column adds observations of Tb94.

bimodal particle size distribution, even in shallow clouds composed only of liquid. As has been shown by multiple recent studies (Lebsock et al. 2011; Lebsock and Suzuki 2016; Leinonen et al. 2016; Mace et al. 2016), at minimum a combination of active and passive measurements is required. Passive microwave observations place a constraint on total column water mass, visible optical depth constrains the total cloud-mode water path, and near-infrared observations contain information on particle size at cloud top. Retrievals that iterate to a solution probabilistically (e.g., Mace et al. 2016) utilize the passive observations to find a realistic first guess, then employ radar to retrieve vertical distributions of cloud properties. Our study has shown in detail the character of the retrieval space. It is clear that not only is a scene-dependent prior required, but additional observations will be necessary if multiple PSD modes are to be disentangled. Specifically, we hypothesize the following:

1) Knowledge of cloud dynamics, in the form of vertical velocities, could place a key constraint on the vertical distribution of cloud mass (cf. PM14). In addition, connecting passive forward models to a dynamic model that includes thermodynamics and updrafts may serve to produce an internally consistent cloud profile. Such experiments have already begun to be conducted in studies of cloud parameterization uncertainty (Posselt and Vukicevic 2010; Posselt 2016). Inclusion of cloud dynamics and environmental thermodynamic context moves the retrieval in the direction of dynamic data assimilation, which may be the most general way to retrieve cloud extensive and intensive properties.
2) In addition to information on cloud vertical velocity, additional radar frequency observations would almost certainly reduce the mass–number ambiguity in the particle size distribution in situations where rain is heavy enough to cause differential attenuation and different scattering signatures. Additional passive microwave frequencies (e.g., 31 GHz for liquid, higher frequency for ice) have been shown to place valuable integral constraints on the total solution. In addition to observations of 94-GHz brightness temperature, recent work has shown that the radar path-integrated attenuation may place a strong integral constraint on condensate profiles (Lebsock and Suzuki 2016).

Recent field campaigns [e.g., IPHEX and the Olympic Mountain Experiment (OLYMPEX)] produced high-quality multifrequency active and passive observation datasets. We will report on the results of observation information content experiments in a forthcoming manuscript. Our research aimed at combining dynamic and thermodynamic information with cloud observations will also be reported elsewhere.

5. Summary and conclusions

This paper provides an analysis of the solution space for a combined radar–reflectance retrieval of cloud properties for shallow cumulus clouds over the ocean. Study of liquid-only clouds allows us to avoid the complexity associated with poorly known ice crystal shape and density. Several previous studies have examined the information content of active and passive measurements of shallow cumulus clouds, but have typically made assumptions about the uniqueness of the retrieval solution and the constraint of active versus passive observations. In addition, while previous work has made it clear that a unique and well-posed solution requires a prior constraint, it has not been obvious how such constraints should operate and in which situations they are necessary.

Bayesian algorithms produce a sample of the joint probability density of retrieved cloud properties given the information in observations and the prior constraint. They are flexible, in that they allow implementation of various retrieval assumptions, and provide a quantitative estimate of the information contained in observations. They are also capable of incorporating physical constraints that are more difficult to implement in optimal estimation retrievals (e.g., raindrop diameter $>$ cloud droplet diameter). In this study, we have used a Markov chain Monte Carlo algorithm to sample the joint posterior PDF for a bimodal particle size distribution (cloud droplets and rain) in shallow clouds over the Southern Ocean. A combination of W-band radar and visible and near-infrared reflectance measurements from three different cloudy columns (containing primarily cloud droplets, primarily rain, and a mixture) are used as observations. Two different prior constraints are tested: one that is scene dependent and restricts the retrieval to a range of PSD parameters consistent with the observed scene, and one that more encompasses a broader global climatology of droplet and rain PSD properties. In addition to an exploration of the solution space for low liquid cloud retrievals, this study implements a new three-stage delayed-reject MCMC algorithm, and explores its application to vertical profile observations. This study represents the highest-dimensional application of MCMC to cloud property retrievals to date, and paves the way for more sophisticated future analysis and for systematic Bayesian retrievals of cloud properties from active and passive sensors.

The major conclusions of this study include the following:

1) Comparison among the posterior PDFs for profiles containing primarily cloud, primarily rain, and a mixture of the two confirms that A-Train observations are able to effectively distinguish between precipitating and nonprecipitating cloud columns, even in the absence of any additional prior information.
2) As has been shown in other studies, single-frequency radar alone does not contain enough information to produce a unique estimate of the PSD properties, even in liquid-only clouds over the ocean. However, there is enough information contained in the radar reflectivity to provide reliable estimates of rainwater content.
3) Addition of visible and near-infrared reflectance observations bounds the cloud droplet size but only in the uppermost 500 m of the cloud and only if observation uncertainties are assumed to be rather low (e.g., 5%). If observation errors are set to ~25% (as is reasonable in scenes with broken and isolated clouds), the constraint provided by the passive measurements is negligible.
4) It is clear that a scene-consistent prior estimate is required if the retrieval is to have a unique solution. This is certainly true for algorithms that utilize data similar to what is employed in this study, and likely true for other combined active–passive cloud property retrievals. The intermittent availability of robust prior estimates (e.g., from field campaigns) means that proper specification of a scene-consistent prior is a major challenge. Data from measurement suites (e.g., A-Train) are necessary, but not sufficient, and investment is needed in expanding the observational knowledge of in situ cloud and precipitation properties.
5) Access to a sample of the full joint posterior distribution of the retrieved cloud properties allows us to compute a range of statistics on the retrieval solution. Examination of the posterior distributions of the cloud and rainwater content and number indicate the posterior distribution of rain may be effectively described using a centered single-mode (e.g., Gaussian) distribution, while the cloud number and water content exhibits significant skewness. As such, the mean of the solution PDF is not a useful measure of the true water content or number, and the standard deviation does not provide a useful estimate of uncertainty. In this case, it appears that a log transform may be necessary if the goal is to render a cloud-mode solution that is approximately Gaussian in form.
6) Passive 94-GHz brightness temperature observations provide a significant integral constraint on the retrieval in cases with a broad prior range of possible cloud PSD parameters, leading to more accurate estimates of the cloud and rainwater content in particular.
7) It is typically assumed that a reduction in the degrees of freedom in a retrieval, via specification of one or more of the free parameters in the scene or forward model, will result in a reduction in retrieval uncertainty. We tested this assumption by running experiments in which we alternately fixed and varied the width of the particle size distribution for cloud droplets and rain. We found that variability in the PSD width (effectively an increase in the number of degrees of freedom) in fact led to improved constraint (reduced uncertainty) on the raindrop number concentration. In this case, it appears that allowing a flexible distribution width allowed the retrieval to find a solution that was a closer match to the observations. This illustrates the fact that a smaller number of degrees of freedom in a retrieval is not necessarily desirable; while the retrieval may converge, the solution may not properly represent the true values of the free parameters when fixed parameters are incorrectly specified.
8) The three-stage delayed-reject MCMC algorithm is able to return a robust estimate of the retrieval space probability distribution in a relatively high-dimensional problem. The retrieval solution reflects known properties of oceanic cumulus clouds, while allowing a detailed exploration of the information contained in the measurements and the vertical structure in the clouds themselves.

The experiments described in this manuscript set the stage for a more detailed trade study of which measurements are necessary to be able to uniquely retrieve cloud PSD properties in liquid, mixed-phase, and ice clouds. Such studies are critical for planning the next generation cloud observing system. We are using recent observations from the IPHEX and OLYMPEX field campaigns and are specifically exploring the relative influence of multiple-frequency radar observations of both reflectivity and the Doppler spectrum. We are quantifying the information provided by observations in multiple different passive microwave, visible, and infrared channels, and will report on the results of this study in a forthcoming manuscript. In addition, we are actively working to utilize the MCMC algorithm to produce long-term estimates of liquid cloud properties over the U.S. Department of Energy Atmospheric Radiation Measurement (DOE ARM) sites. Results of each of these studies will be reported in forthcoming manuscripts.

Acknowledgments. This work was supported by the NASA Aerosol, Clouds, and Ecosystems Mission under

Grant NNX15AK17G. Matt Lebsock provided valuable feedback that aided in interpretation of the results. The comments of two anonymous reviewers helped to focus and improve the narrative.

APPENDIX A

Vertical Correlation Function for Cloud Property Retrievals

In previous work, our MCMC algorithm utilized a proposal distribution that treated each control variable as independent of the others. While this is less efficient for posterior distributions that exhibit significant covariability between two or more variables (Tamminen 2004), it allows greater flexibility when sampling distributions that have a nonstatic covariance structure (e.g., covariance between two variables changes sign). Clouds are known to strongly covary in both the vertical and horizontal directions. If the vertical structure of clouds is not taken into account (e.g., PSD properties for each level in the cloud are perturbed independently of the other levels), this will cause the algorithm to converge more slowly, and will introduce noise into the retrieval.

The algorithm used in this paper applies a Gaussian vertical structure function to the proposal distribution, with covariance set according to the observed vertical variability of radar reflectivity in the cloud. In practice, in each MCMC iteration, a cloud layer is selected at random, and all variables in this layer are randomly perturbed according to the proposal variance tuned for the layer (set during burn-in). After the resulting perturbations are inspected to ensure they are within the range of values allowed in the prior, PSD properties in adjacent layers are perturbed in the same direction (with the same sign), and with magnitude that is weighted by the distance from the layer perturbed. Specifically, consider a perturbation dp applied at level z_0. The perturbation applied at another level z_i will be

$$dp_i = dp_0 \exp\left[-\frac{1}{2}\frac{(z_0 - z_i)^2}{\sigma_z^2}\right], \tag{A1}$$

where dp_i is the parameter perturbation at level i, dp_0 is the perturbation applied at level z_0, and σ_z is the vertical decorrelation length.

We have found the results to be insensitive to the precise specification of σ_z, as long as it is smaller than the depth of the cloud (so that perturbations are not applied uniformly over the cloud depth) and large enough so that more than one layer is modified at a time. Use of a static covariance length also allows us to retain the relatively simple Metropolis–Hastings accept–reject procedure. Allowing the covariance to vary with height, or applying a more sophisticated vertical structure function, is tantamount to including additional information in the prior. While this will almost certainly need to be done if this algorithm is used to generate routine retrieved products (see discussion above), specification of such a structure is beyond the scope of this paper.

APPENDIX B

Three-Stage Delayed-Reject MCMC Algorithm

MCMC algorithms are intended to produce a sample of the Bayesian posterior joint probability distribution given a prior, likelihood, and a model that relates them. The algorithm works by producing a Markov chain that consists of successive samples from the posterior distribution. Each new sample is generated by first computing a set of candidate (or test) values $\hat{x}$ of each element in the control vector $\mathbf{x}$, then using a Metropolis–Hastings (MH) accept–reject test to determine whether the candidate is indeed a sample from the posterior PDF. The efficiency with which the algorithm will sample the posterior distribution $P(\mathbf{x}|\mathbf{y})$ depends on how well the *proposal distribution* used to generate new candidate values approximates the *posterior distribution*. The ideal proposal distribution is identically the posterior distribution, but since the posterior distribution is generally not known in advance (else, why bother to sample it?), it is in practice not possible to construct an optimal proposal distribution. Instead, in most applications the proposal distribution is assumed to have a well-defined structure (e.g., multivariate Gaussian), and the dispersion (and higher moments, if applicable) are tuned in an iterative procedure so that a reasonable fraction of the proposals are accepted. If the target density is Gaussian, and every variable is independent and identically distributed (i.i.d.), the optimal accept rate is 0.234 (Gelman et al. 1997; Roberts and Rosenthal 2001). This fractional accept rate strikes a balance between sampling rapidly in a high-dimensional space, and resolving local features of the PDF. The proposal density for the multivariate Gaussian target distribution has covariance Σ_p:

$$\Sigma_p = \left[\frac{(2.38)^2}{d}\right]\Sigma, \tag{B1}$$

where Σ is the covariance of the target (posterior) distribution and d is the dimension of the state vector. If the

model is nonlinear, the posterior density will not be Gaussian, and the proposal will be a suboptimal approximation. The result will be less than perfect sampling efficiency. For example, let the true target density be unimodal and skewed, and consider a case in which the proposal variance is tuned to produce the Gaussian-optimal accept rate while the chain is located near the mode of the distribution. Proposals will be close to optimal when the chain is near the mode but will be increasingly suboptimal (undervariant) when the chain explores the tail of the posterior distribution. An even more difficult case arises when the posterior distribution has more than one mode, and the modes are distinct (separated by regions with very low probability). In this case, if the algorithm is tuned to optimally sample a single low-variance mode, there will be a low likelihood that it will be capable of spanning the gap between modes. One way to combat this problem is to build an adaptive algorithm that adjusts to the various regions of the space (e.g., Haario et al. 1999). However, it is difficult to construct such an algorithm so that it is properly ergodic. Another option is to *delay rejection* when the MH criterion is not met. Instead, in this case, a second (or third, fourth, etc.) proposal is conducted with modified (smaller) variance. As long as the secondary proposal is properly incorporated into the MH accept–reject test, the chain will be ergodic. Delayed rejection allows the proposal distribution to have a variance that is large enough to span the range of finite probability in the space but may be adapted to explore locally if there are well-defined modes.

Here, we document a three-stage delayed-reject algorithm that has proven to work well for the cloud retrievals considered here. Our derivation follows the work of Green and Mira (2001), adapted for our specific application. First, let the numerator in Bayes's relationship be denoted $\pi(x)$:

$$\pi(x) = P(x)P(y\,|\,x). \tag{B2}$$

Note that this is simply the nonnormalized right-hand side of (2). Given a set of proposed (candidate) control values $\hat{x}_1$, the MH accept ratio (Tamminen and Kyrölä 2001; Posselt 2013) is computed as

$$\rho(x,\hat{x}_1) = \frac{P(y\,|\,\hat{x}_1)P(\hat{x}_1)q(\hat{x}_1,x)}{P(y\,|\,x)P(x)q(x,\hat{x}_1)}. \tag{B3}$$

Using the notation of Green and Mira, we can write (B3) as

$$\rho(x,\hat{x}_1) = \frac{\pi(\hat{x}_1)q(\hat{x}_1,x)}{\pi(x)q(x,\hat{x}_1)}, \tag{B4}$$

where $q(x, \hat{x}_1)$ is the proposal distribution used to generate new candidate values $\hat{x}_1$, and $q(\hat{x}_1, x)$ represents the probability of randomly drawing x from $\hat{x}_1$. The (first stage) acceptance probability is

$$\alpha_1(x,\hat{x}_1) = \min\left\{1, \frac{\pi(\hat{x}_1)q(\hat{x}_1,x)}{\pi(x)q(x,\hat{x}_1)}\right\}. \tag{B5}$$

What this means, in practice, is that if the candidate point $\hat{x}_1$ has a higher density than the previous point x, it is automatically accepted as a sample. If not $[\alpha_1(x, \hat{x}_1) < 1]$, then α_1 is compared with a uniform random number on [0, 1]. If α_1 is larger than the randomly drawn value, $\hat{x}_1$ is accepted as a sample of the posterior distribution. Now, if (B5) is evaluated and the candidate point $\hat{x}_1$ is rejected (α_1 is less than U[0, 1]), *delay rejection* and instead draw another candidate point $\hat{x}_2$, centered on x and using a proposal distribution with smaller variance (e.g., explore the state space in a region closer to the current point). We have chosen to aggressively decrease the proposal variance in each delay iteration so that the proposal distribution in delay step n is

$$q(x,\hat{x}_n) \sim N\left[x, \frac{\sigma_x}{2^{(n-1)}}\right], \tag{B6}$$

where $n = 2$ in the first delay step, and σ_x is the variance of the proposal distribution. Clearly, the proposal standard deviation is halved in each subsequent delay step. The new (stage 2) acceptance probability must now be evaluated taking into account both proposal distributions and the original (stage 1) acceptance probability:

$$\alpha_2(x,\hat{x}_1,\hat{x}_2) = \min\left\{1, \frac{\pi(\hat{x}_2)}{\pi(x)}\frac{q_1(\hat{x}_2,\hat{x}_1)q_2(\hat{x}_2,\hat{x}_1,x)[1-\alpha_1(\hat{x}_2,\hat{x}_1)]}{q_1(x,\hat{x}_1)q_2(x,\hat{x}_1,\hat{x}_2)[1-\alpha_1(x,\hat{x}_1)]}\right\}. \tag{B7}$$

Here, $\alpha_1(\hat{x}_2, \hat{x}_1)$ represents the probability of accepting a transition from $\hat{x}_2$ back to $\hat{x}_1$, and may be computed from (B5); $q_1(\hat{x}_2, \hat{x}_1)$ is the (stage 1; $n = 1$) proposal distribution that generates $\hat{x}_1$ from $\hat{x}_2$ using the proposal in (B6) with $n = 1$, while q_2 is the (stage 2; $n = 2$) proposal distribution that generates $\hat{x}_2$ from x (through $\hat{x}_1$; denominator) or x from $\hat{x}_2$ (numerator). If the stage 2 proposal is also rejected, a third-stage proposal to $\hat{x}_3$ may be attempted using a yet smaller proposal variance [$n = 3$ in (B6)]. This third set of proposed parameter values may be accepted with probability

$$\alpha_3(x,\hat{x}_1,\hat{x}_2,\hat{x}_3)=\min\left\{1,\frac{\pi(\hat{x}_3)}{\pi(x)}\left[\frac{q_1(\hat{x}_3,\hat{x}_2)q_2(\hat{x}_3,\hat{x}_2,\hat{x}_1)q_3(\hat{x}_3,\hat{x}_2,\hat{x}_1,x)}{q_1(x,\hat{x}_1)q_2(x,\hat{x}_1,\hat{x}_2)q_3(x,\hat{x}_1,\hat{x}_2,\hat{x}_3)}\right]\cdots\left[\frac{[1-\alpha_1(\hat{x}_3,\hat{x}_2)][1-\alpha_2(\hat{x}_3,\hat{x}_2,\hat{x}_1)]}{[1-\alpha_1(x,\hat{x}_1)][1-\alpha_2(x,\hat{x}_1,\hat{x}_2)]}\right]\right\}. \tag{B8}$$

As above, the third-stage proposal density q_3 generates $\hat{x}_3$ from x through intermediate stages $\hat{x}_1$ and $\hat{x}_2$ (and the reverse). The higher-order proposal densities [e.g., $q_3(x, \hat{x}_1, \hat{x}_2, \hat{x}_3)$] are simply a combination of the proposals required to transition from the initial state to proposed states $\hat{x}_1, \hat{x}_2, \hat{x}_3$, and so on. In principle, an n-stage delayed-reject algorithm may be constructed recursively so that the probability of acceptance of proposal i is

$$\alpha_i(x,\hat{x}_1,\ldots,\hat{x}_i)=\min\left\{1,\frac{\pi(\hat{x}_i)}{\pi(x)}\left[\frac{q_1(\hat{x}_i,\hat{x}_{i-1})q_2(\hat{x}_i,\hat{x}_{i-1},\hat{x}_{i-2})\cdots q_i(\hat{x}_i,\ldots,\hat{x}_1,x)}{q_1(x,\hat{x}_1)q_2(x,\hat{x}_1,\hat{x}_2)\cdots q_i(x,\hat{x}_i,\ldots,\hat{x}_{i-1})}\right]\times\left[\frac{[1-\alpha_1(\hat{x}_i,\hat{x}_{i-1})][1-\alpha_2(\hat{x}_i,\hat{x}_{i-1},\hat{x}_{i-2})]\cdots[1-\alpha_{i-1}(\hat{x}_i,\hat{x}_{i-1},\ldots,\hat{x}_1)]}{[1-\alpha_1(x,\hat{x}_1)][1-\alpha_2(x,\hat{x}_1,\hat{x}_2)]\cdots[1-\alpha_{i-1}(x,\hat{x}_1,\ldots,\hat{x}_{i-1})]}\right]\right\}. \tag{B9}$$

Note, however, that 1) this recursive algorithm requires evaluation of successively greater numbers of α_i, and 2) the forward model(s) that map from $\mathbf{x}$ to $\mathbf{y}$ must be run each time a new $\hat{x}_i$ is tested, so there are rapidly diminishing returns. Our algorithm is a three-stage delayed-reject in which q_1, q_2, and q_3 are all symmetric, with proposal variance that varies only according to stage. When all proposal densities are symmetric their ratios cancel. In this case, the accept ratio in the first stage of the three-stage delayed-reject algorithm reduces to

$$\alpha_1(x,\hat{x}_1)=\min\left\{1,\frac{\pi(\hat{x}_1)}{\pi(x)}\right\}, \tag{B10}$$

and the stage 2 accept ratio is

$$\alpha_2(x,\hat{x}_1,\hat{x}_2)=\min\left\{1,\frac{\pi(\hat{x}_2)}{\pi(x)}\frac{[1-\alpha_1(\hat{x}_2,\hat{x}_1)]}{[1-\alpha_1(x,\hat{x}_1)]}\right\}, \tag{B11}$$

where $\alpha_1(\hat{x}_2, \hat{x}_1)$ is simply (B10) evaluated with $\pi(\hat{x}_1)$ in the numerator and $\pi(\hat{x}_2)$ in the denominator. The stage 3 accept ratio is

$$\alpha_3(x,\hat{x}_1,\hat{x}_2,\hat{x}_3)=\min\left\{1,\frac{\pi(\hat{x}_3)}{\pi(x)}\frac{[1-\alpha_1(\hat{x}_3,\hat{x}_2)][1-\alpha_2(\hat{x}_3,\hat{x}_2,\hat{x}_1)]}{[1-\alpha_1(x,\hat{x}_1)][1-\alpha_2(x,\hat{x}_1,\hat{x}_2)]}\right\}, \tag{B12}$$

where, as above, $\alpha_1(\hat{x}_3, \hat{x}_2)$ is (1) evaluated with $\pi(\hat{x}_2)$ in the numerator and $\pi(\hat{x}_3)$ in the denominator. $\alpha_2(x, \hat{x}_1, \hat{x}_2)$ has already been evaluated [(B11)], and $\alpha_2(\hat{x}_3, \hat{x}_2, \hat{x}_1)$ can be obtained using the appropriate inputs to (B11). In practice, log probabilities (prior and likelihoods) are computed in the MCMC algorithm, and in the application here the prior is uniform so that $\pi(\cdot)$ in each case reduces to the likelihood. In closing, we note that there is no established theory that dictates the ideal accept rate for the delayed-reject algorithm. In the limit of an infinite number of delay steps, the accept rate should asymptotically approach 1.0, and between the single stage and the infinite recursive limit the accept rate should lie somewhere between 0.23 and 1.0. In practice, we aim for an accept rate between 0.4 and 0.7; higher than the single-stage target, but lower than 1.0.

APPENDIX C

Assessing MCMC Convergence Using the $\hat{R}$ Statistic

A key aspect of any experiment that involves sampling using an MCMC algorithm is the diagnosis of convergence of the Markov chain to sampling the true posterior PDF. In cases for which the shape of the true posterior PDF is unknown, it is not possible to know with absolute certainty that the Markov chain has converged to sampling the true target distribution. However, there are several diagnostic tools that can be brought to bear in assessing whether the chain has converged to sampling a stationary distribution. In the case of the R statistic (Gelman et al. 2004), the diagnosis of convergence is paired with an assessment of whether multiple Markov chains are all sampling a similar distribution. This method leverages the information contained in the differences between chains of a multichain MCMC simulation. It is based on a comparison between the variance (or other moments) *within* each chain to the variance *between* chains for each estimated parameter, and is done in the following manner.

Consider m chains, each of length n samples. First, the within-chain variance is computed for each parameter x as

$$W = \frac{1}{m}\sum_{j=1}^{m}\left[\frac{1}{n}\sum_{i=1}^{n}(x_{ij}-\bar{x}_j)^2\right], \tag{C1}$$

where $\bar{x}_j$ is the mean of each parameter x within each chain, defined as

$$\bar{x}_j = \frac{1}{n}\sum_{i=1}^{n}x_{ij}. \tag{C2}$$

The between-chain variances are computed as

$$B = \frac{n}{m-1}\sum_{j=1}^{m}(\bar{x}_j-\bar{x})^2, \tag{C3}$$

where $\bar{x}$ is the mean of the given parameter across all chains, defined as

$$\bar{x} = \frac{1}{m}\sum_{j=1}^{m}\bar{x}_j. \tag{C4}$$

An unbiased estimate of the marginal posterior variance of each variable x conditioned on the set of observations $\mathbf{y}$ can be obtained from a weighted combination of B and W as

$$\widehat{\mathrm{var}}^{+}(x\,|\,\mathbf{y}) = \frac{n-1}{n}W + \frac{1}{n}B. \tag{C5}$$

This quantity tends to overestimate the true marginal posterior variance, but converges to the true variance as $n \to \infty$. Proper chain mixing is assessed by comparing the variance estimate with the within-chain variance, and computing the R statistic $\hat{R}$, an estimate of the factor by which the dispersion in the current sample would be reduced if each chain were allowed an infinite length:

$$\hat{R} = \sqrt{\frac{\widehat{\mathrm{var}}^{+}(x\,|\,\mathbf{y})}{W}}. \tag{C6}$$

It can be readily seen from (C6) that this estimate will converge to 1 in the limit as $n \to \infty$. According to Gelman et al. (2004), there is no specific value of the R statistic for which chains can be said to have sufficiently mixed, though a value of $\hat{R}$ less than 1.1 for each parameter is generally deemed acceptable.

REFERENCES

Bohren, C. F., and D. R. Huffman, 1983: *Absorption and Scattering of Light by Small Particles*. Wiley-Interscience, 530 pp.

Christi, M., and P. Gabriel, 2003: Radiant 2.0: A user's guide. Colorado State University, Fort Collins, CO, 39 pp.

Gelman, A., W. R. Gilks, and G. O. Roberts, 1997: Weak convergence and optimal scaling of random walk Metropolis algorithms. *Ann. Appl. Probab.*, **7**, 110–120, doi:10.1214/aoap/1034625254.

——, J. B. Carlin, H. S. Stern, and D. B. Rubin, 2004: *Bayesian Data Analysis*. 2nd ed. Chapman and Hall/CRC, 690 pp.

Green, P. J., and A. Mira, 2001: Delayed rejection in reversible jump Metropolis–Hastings. *Biometrika*, **88**, 1035–1053, doi:10.1093/biomet/88.4.1035.

Haario, H., E. Saksman, and J. Tamminen, 1999: Adaptive proposal distribution for random walk Metropolis algorithm. *Comput. Stat.*, **14**, 375–395, doi:10.1007/s001800050022.

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

Hegg, D. A., D. S. Covert, H. H. Jonsson, and R. K. Woods, 2012: A simple relationship between cloud drop number concentration and precursor aerosol concentration for the regions of Earth's large marine stratocumulus decks. *Atmos. Chem. Phys.*, **12**, 1229–1238, doi:10.5194/acp-12-1229-2012.

Hou, A. Y., and Coauthors, 2014: The Global Precipitation Measurement mission. *Bull. Amer. Meteor. Soc.*, **95**, 701–722, doi:10.1175/BAMS-D-13-00164.1.

Kummerow, C., 1993: On the accuracy of the Eddington approximation for radiative transfer in the microwave frequencies. *J. Geophys. Res.*, **98**, 2757–2765, doi:10.1029/92JD02472.

——, and Coauthors, 2000: The status of the Tropical Rainfall Measuring Mission (TRMM) after two years in orbit. *J. Appl. Meteor.*, **39**, 1965–1982, doi:10.1175/1520-0450(2001)040<1965:TSOTTR>2.0.CO;2.

Lebsock, M. D., and K. Suzuki, 2016: Uncertainty characteristics of total water path retrievals in shallow cumulus derived from a spaceborne radar/radiometer integral constraints. *J. Atmos. Oceanic Technol.*, **33**, 1597–1609, doi:10.1175/JTECH-D-16-0023.1.

——, T. S. L'Ecuyer, and G. L. Stephens, 2011: Detecting the ratio of rain and cloud water in low-altitude shallow marine clouds. *J. Appl. Meteor. Climatol.*, **50**, 419–432, doi:10.1175/2010JAMC2494.1.

Leinonen, J., M. D. Lebsock, G. L. Stephens, and K. Suzuki, 2016: Retrieval of cloud liquid water from *CloudSat* and MODIS. *J. Appl. Meteor. Climatol.*, **55**, 1831–1844, doi:10.1175/JAMC-D-16-0077.1.

Mace, G. G., and Q. Zhang, 2014: The CloudSat radar-lidar geometrical profile product (RL-GeoProf): Updates, improvements, and selected results. *J. Geophys. Res. Atmos.*, **119**, 9441–9462, doi:10.1002/2013JD021374.

——, S. Avey, S. Cooper, M. Lebsock, S. Tanelli, and G. Dobrowalski, 2016: Retrieving co-occurring cloud and precipitation properties of warm marine boundary layer clouds with A-Train data. *J. Geophys. Res. Atmos.*, **121**, 4008–4033, doi:10.1002/2015JD023681.

Marchand, R. T., G. G. Mace, T. P. Ackerman, and G. Stephens, 2008: Hydrometeor detection using *CloudSat*—An Earth-orbiting 94-GHz cloud radar. *J. Atmos. Oceanic Technol.*, **25**, 519–533, doi:10.1175/2007JTECHA1006.1.

Platnick, S., 2000: Vertical photon transport in cloud remote sensing problems. *J. Geophys. Res.*, **105**, 22 919–22 935, doi:10.1029/2000JD900333.

——, M. D. King, S. A. Ackerman, W. P. Menzel, B. A. Baum, J. C. Riedi, and R. A. Frey, 2003: The MODIS cloud products:

Algorithms and examples from *Terra*. *IEEE Trans. Geosci. Remote Sens.*, **41**, 459–473, doi:10.1109/TGRS.2002.808301.

Posselt, D. J., 2013: Markov chain Monte Carlo methods: Theory and applications. *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications*, 2nd ed. S. K. Park and L. Xu, Eds., Springer, 59–87.

——, 2016: A Bayesian examination of deep convective squall line sensitivity to changes in cloud microphysical parameters. *J. Atmos. Sci.*, **73**, 637–665, doi:10.1175/JAS-D-15-0159.1.

——, and T. Vukicevic, 2010: Robust characterization of model physics uncertainty for simulations of deep moist convection. *Mon. Wea. Rev.*, **138**, 1513–1535, doi:10.1175/2009MWR3094.1.

——, and G. G. Mace, 2014: MCMC-based assessment of the error characteristics of a surface-based combined radar–passive microwave cloud property retrieval. *J. Appl. Meteor. Climatol.*, **53**, 2034–2057, doi:10.1175/JAMC-D-13-0237.1.

——, X. Li, S. A. Tushaus, and J. R. Mecikalski, 2015: Assimilation of dual-polarization radar observations in mixed- and ice- phase regions of convective storms: Information content and forward model errors. *Mon. Wea. Rev.*, **143**, 2611–2636, doi:10.1175/MWR-D-14-00347.1.

Rauber, R. M., and Coauthors, 2007: Rain in shallow cumulus over the ocean. *Bull. Amer. Meteor. Soc.*, **88**, 1912–1928, doi:10.1175/BAMS-88-12-1912.

Roberts, G. O., and J. S. Rosenthal, 2001: Optimal scaling for various Metropolis-Hastings algorithms. *Stat. Sci.*, **16**, 351–367, doi:10.1214/ss/1015346320.

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

Stephens, G. L., 2005: Cloud feedbacks in the climate system: A critical review. *J. Climate*, **18**, 237–273, doi:10.1175/JCLI-3243.1.

——, and Coauthors, 2008: *CloudSat* mission: Performance and early science after the first year of operation. *J. Geophys. Res.*, **113**, D00A18, doi:10.1029/2008JD009982.

Tamminen, J., 2004: Validation of nonlinear inverse algorithms with Markov chain Monte Carlo method. *J. Geophys. Res.*, **109**, D19303, doi:10.1029/2004JD004927.

——, and E. Kyrölä, 2001: Bayesian solution for nonlinear and non-Gaussian inverse problems by Markov chain Monte Carlo method. *J. Geophys. Res.*, **106**, 14 377–14 390, doi:10.1029/2001JD900007.

Wilks, D. S., 2011: *Statistical Methods in the Atmospheric Sciences*. 3rd ed. Academic Press, 704 pp.

Wood, N. B., T. S. L'Ecuyer, A. J. Heymsfield, G. L. Stephens, D. R. Hudak, and P. Rodriguez, 2014: Estimating snow microphysical properties using collocated multisensor observations. *J. Geophys. Res. Atmos.*, **119**, 8941–8961, doi:10.1002/2013JD021303.

Wood, R., T. L. Kubar, and D. Hartmann, 2009: Understanding the importance of microphysics and macrophysics for warm rain in marine low clouds. Part II: Heuristic models of rain formation. *J. Atmos. Sci.*, **66**, 2973–2991, doi:10.1175/2009JAS3072.1.