



## Exploring the error characteristics of thin ice cloud property retrievals using a Markov chain Monte Carlo algorithm

D. J. Posselt,<sup>1</sup> T. S. L'Ecuyer,<sup>2</sup> and G. L. Stephens<sup>2</sup>

Received 22 July 2008; revised 15 September 2008; accepted 16 October 2008; published 24 December 2008.

[1] It is commonly assumed that ice cloud property retrievals are well-constrained so that a look-up table or an optimal estimation approach can be used to provide a unique solution. However, because of nonlinearities in radiative transfer models, error characteristics of retrieved ice cloud properties are generally not well-known. The Markov chain Monte Carlo (MCMC) approach, which represents information from prior knowledge, observations, and the forward model probabilistically, allows for an accurate assessment of the solution space and the nature of uncertainties in the retrieval. In this paper, an MCMC algorithm is used to examine the sensitivity of infrared split window retrieved ice water path and ice particle effective radius to changes in cloud top height, cloud geometric thickness, and the assumed ice crystal shape. In addition, we assess the effect of particle settling, changes to observation error magnitude, and the implementation of a log-normal error distribution on the retrieval results. It is found that, though the effects of uncertainty in cloud top height are not insignificant, uncertainty in the ice crystal shape contributes most to the uncertainty in the retrieval and gives rise to the potential for multiple solutions. Reduction of observation error and the assumption of particle settling serve to change the preferred combination of crystal shapes in the volume but do not eliminate the potential for a multimodal result. Application of the MCMC algorithm to a scene reveals that both the nature and magnitude of retrieval errors exhibit a strong dependence on cloud optical depth.

**Citation:** Posselt, D., T. S. L'Ecuyer, and G. L. Stephens (2008), Exploring the error characteristics of thin ice cloud property retrievals using a Markov chain Monte Carlo algorithm, *J. Geophys. Res.*, 113, D24206, doi:10.1029/2008JD010832.

### 1. Introduction

[2] In their most general sense, all retrievals exploit the fact that a set of observations contains information about a geophysical parameter of interest, and all employ a functional relationship between observations and retrieved parameters (e.g., a forward radiative transfer model, regression, etc. [Rodgers, 2000; Miller *et al.*, 2000, and references therein]). Though most retrieval techniques produce a single answer, all retrieval solutions are characterized by some uncertainty. The magnitude and characteristics of the uncertainty in the retrieval result depend on the characteristics of uncertainty in the observations, their sensitivity to the parameters of interest, forward radiative transfer model accuracy, and the quality of available prior information about the retrieved state [Rodgers, 2000; Cooper *et al.*, 2003; L'Ecuyer *et al.*, 2006; Cooper *et al.*, 2007]. Though the requirements of instrument design stipulate that uncer-

tainty in the measurements be well-understood, components of retrieval error owing to algorithm assumptions, whether explicit or implicit, and uncertainty in the prior knowledge are much more difficult to quantify. It has been shown that information regarding uncertainty can be conveniently stated in the form of a probability distribution so that the uncertainty in the solution is represented as the probability distribution that arises from the conjunction of probability distributions of observations, model uncertainty, and prior information on the parameters [Tamminen, 2004; Tarantola, 2005; Vukicevic and Posselt, 2008]. Such "probabilistic" approaches not only provide a solution to the retrieval problem but also a rigorous error analysis, however, the reliability of the resulting uncertainty estimates depends critically on an accurate specification of measurement, model, and a priori error distributions. Gaussian statistics are commonly used to assign uncertainty to a retrieval, but the problem is complicated by the fact that the probability distribution of the solution will not be Gaussian if the relationship between observations and retrieved parameters is nonlinear, as is the case with many radiative transfer problems [Jazwinski, 1970; Tarantola, 2005; Vukicevic and Posselt, 2008].

[3] In contrast to methods that rely on Gaussian statistics, the Markov chain Monte Carlo technique has been shown to robustly return the characteristics of the solution space and

<sup>1</sup>Department of Atmospheric, Oceanic, and Space Sciences, University of Michigan, Ann Arbor, Michigan, USA.

<sup>2</sup>Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado, USA.

to flexibly incorporate adjustments to assumptions on errors in observations and forward model. In previous examples, the MCMC approach has been used to retrieve atmospheric composition from the Global Ozone Monitoring by Occultation of Stars (GOMOS) instrument, and subsequently to assess errors in the operational retrievals [Tamminen and Kyrola, 2001; Tamminen, 2004]. In this study, the MCMC algorithm is used to diagnose the characteristics of ice cloud property retrievals from the well-known split window technique that relies on relative differences in absorption of infrared radiation by clouds at two channels in the infrared window (wavelengths between 8 and 14 microns [Inoue, 1985; Prabhakara et al., 1988]). Split-window retrievals provide global estimates of the optical properties of widespread thin cirrus clouds under both daytime and nighttime conditions through channels available on a number of both geostationary (GOES sounder) and polar orbiting (MODIS, VIRS on TRMM, AIRS, METEOSAT) satellites. Although the split-window technique has been extensively used [e.g., Inoue, 1985; Prabhakara et al., 1988; Ackerman et al., 1990] it is known to have large uncertainties due to errors in cloud top height (CTH) position [Miller et al., 2000; Cooper et al., 2003], cloud geometric thickness [Hong et al., 2007], and ice crystal shape [Cooper et al., 2003; Baum et al., 2005]. Though these sources of uncertainty have been identified, questions remain as to the relative magnitude of errors in cloud boundary information versus ice crystal shape. In addition, it is unknown whether an increase in accuracy of observations or atmospheric profiles can reduce these errors, or to what extent uncertainty is scene-dependent. A corollary to this question is whether errors are random (as is typically assumed) or systematic, and to what extent retrieval uncertainty conforms to a Gaussian distribution.

[4] We wish to emphasize at the outset that this paper does not attempt to outline a new operational retrieval technique, nor does it advocate the use of one type of algorithm over another. Our primary goals are first to demonstrate the robustness of the Markov chain Monte Carlo technique in the context of cloud property retrievals, and second to explore the characteristics of the infrared split window technique, which returns ice cloud property information relevant to the study of the role of cirrus clouds in climate. The importance of uncertainties in CTH, cloud geometric thickness, and ice crystal shape will be assessed in the context of their relative impacts on retrieved ice water path (IWP) and effective radius ( $R_e$ ). We then assess the effects of reducing observation errors, changes to the form of the CTH error distribution, and adding an explicit model of particle settling in the cloud on the character of the solution. Finally, we apply the MCMC algorithm to a cloud scene observed by the Moderate Resolution Imaging Spectroradiometer (MODIS) to demonstrate the scene-dependent characteristics of the solution space. We focus on the information content of the measurements, as well as the potential for non-Gaussianity in the solution probability density function (PDF), including the presence of skewness and multiple modes. Particular emphasis will be given to a demonstration of the utility of the MCMC algorithm for not only determining the effect of uncertainties on a retrieved estimate, but also on the relationship between variables.

[5] The remainder of this paper is organized as follows: the observations, forward model, and Markov chain Monte Carlo technique are briefly described in section 2. Section 3 details results of application of the MCMC algorithm to an ice cloud property retrieval for a single pixel that contains a cirrus cloud associated with a midlatitude warm front. Uncertainty associated with specification of cloud boundaries and ice crystal shape, and the effect of particle settling and changes to error assumptions on the retrieved PDFs of ice water path and effective radius are addressed. Variation in retrieval characteristics across a scene, as well as an examination of information content in the retrieval, are detailed in section 4, while a summary and conclusions are offered in section 5.

## 2. Data and Methods

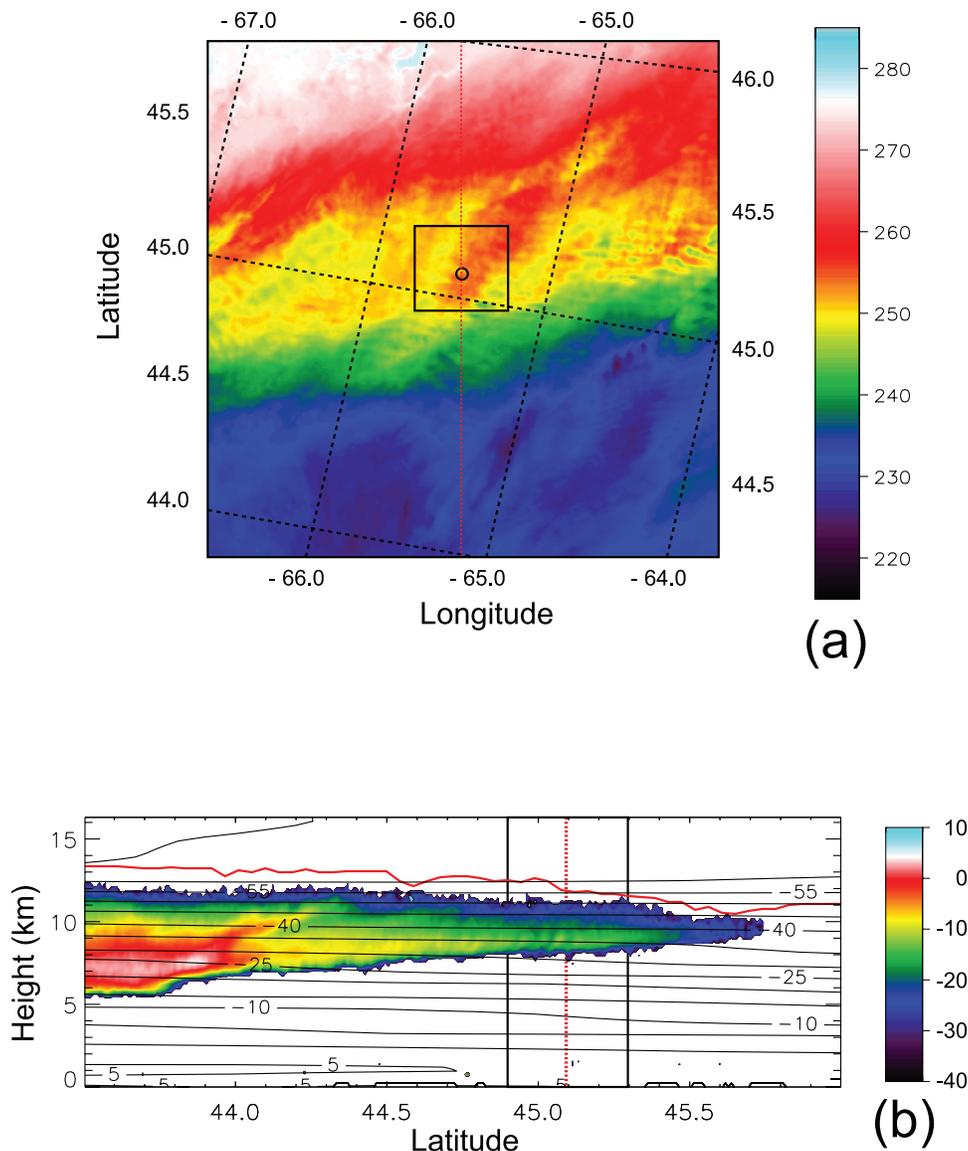
### 2.1. Cloud Observations

[6] The scene considered in this paper was observed by Aqua MODIS at approximately 1730 UTC 22 November 2006, and consists of a cirrus cloud associated with a warm front located off of the east coast of North America [Posselet et al., 2008]. Consistent with the warm-front archetype, the cloud decreases in thickness with increasing distance from the front's leading edge (increasing north latitude, in this case, Figure 1b), and exhibits evidence of size sorting, with higher reflectivities located near the bottom of the cloud and decreasing toward cloud top. Owing to its higher sensitivity to diffuse cloud [Weisz et al., 2007], the CALIPSO observed cloud top height is consistently 1–2 km higher than the CloudSat cloud top height over the entire scene. The collocated European Centre for Medium-range Weather Forecasts (ECMWF) analysis indicates that temperatures in the cloud were uniformly below  $-25.0$  degrees Celsius north of  $44.5$  degrees latitude, and we assume that the cloud in this scene is composed entirely of ice. Although the presence of supercooled water cannot be entirely ruled out, we assume that any liquid water is present in such low quantities as to cause it to be a negligible source of error. Because the scene is over ocean with an above-freezing skin temperature, we specify constant surface emissivities of 0.99 and 0.97 for the 11-micron and 13.3-micron channels, respectively.

[7] Eleven-micron brightness temperature for a portion of the swath around the CloudSat track is depicted in Figure 1a, and it can be seen that the cloud exhibits greater variability along the CloudSat track than in the cross-track direction. A subset of the full scene is examined in detail in section 4; this subset is indicated in the box in Figures 1a and 1b. Optical depth from the operational MODIS retrieval ranges from approximately 1–3 across this region, and the 11- to 12-micron brightness temperature difference ranges from 0.5 to 2.0. As such, the cloud geometric and optical thickness is in the range in which the split window technique would be expected to perform well. Note that we use the 13.3-micron channel in the MCMC-based retrieval results, as this channel has been shown to have slightly higher information content for retrievals of ice cloud effective radius [Cooper et al., 2007].

### 2.2. Radiative Transfer Models

[8] Simulated brightness temperatures at infrared wavelengths depend not only on cloud thickness and properties,



**Figure 1.** (a) MODIS 11-micron brightness temperature observations of the cirrus cloud of interest. The small circle indicates the pixel examined in section 3, while the box outlines the scene described in detail in section 4. The dotted red line indicates the CloudSat ground track. (b) Cross section of CloudSat radar reflectivities (color shaded) and CALIPSO retrieved cloud top height (red line). The vertical dotted line depicts the location of the single pixel examined in section 3, while, as in Figure 1a, the box delineates the subsce described in section 4.

but also on absorption/emission by  $\text{CO}_2$ , water vapor, and ozone. For the results presented in this paper, gaseous optical depths were computed using the OPTRAN forward radiative transfer model [Kleespies *et al.*, 2004], with temperature, water vapor, and ozone profiles obtained from the ECMWF operational analysis and AIRS response functions centered on the selected MODIS wavelengths. Because the 11- and 13.3-micron wavelengths lie in a region with little variation in absorption by  $\text{CO}_2$ ,  $\text{H}_2\text{O}$ , and  $\text{O}_3$ , the choice of response function should have a negligible effect on the characteristics of the retrieved PDFs.

[9] The effect of clouds on infrared brightness temperatures is treated using the Successive Order of Interaction (SOI) radiative transfer model, which has been shown to

perform well for a range of cloud thicknesses and scattering properties [Heidinger *et al.*, 2006; O'Dell *et al.*, 2006]. Ice crystal scattering properties for four different habits known to exist in warm front cirrus clouds [Heymsfield and Iaquinta, 2000]: solid columns, droxtals, bullet rosettes, and aggregates were computed using a combination of the finite-difference time domain method [Yang and Liou, 1996a] and an improved geometric optics method [Yang and Liou, 1996b]. The details of these data sets are reported by Yang *et al.* [2005].

[10] Our implementation of the infrared split window technique follows Cooper *et al.* [2003, 2006], and uses as observations the brightness temperature at the 11-micron wavelength, as well as the 11- to 13.3-micron brightness

temperature difference. Ice water path (IWP) and effective radius are the retrieved quantities, where the ice crystal effective radius is defined as

$$r_e = \frac{3}{4} \frac{(\text{Particle Volume})}{(\text{Projected Particle Area})}. \quad (1)$$

The SOI model requires inputs of optical depth, single scatter albedo, and asymmetry parameter. Optical depth is computed from IWP using the formulation of Heymsfield et al. [2003], while retrieved effective radius is used to assign values of the single scatter albedo and asymmetry parameter based on the precomputed single scattering properties of ice crystals mentioned above. With the exception of the particle settling experiments described below, clouds are assumed to be vertically homogeneous.

### 2.3. Markov Chain Monte Carlo Algorithm

[11] The most general statement of an inverse problem treats the solution as the conjunction of the set of information spaces that define our state of knowledge of the system [Rodgers, 2000; Tarantola, 2005; Vukicevic and Posselt, 2008]. The character and magnitude of the uncertainties in observations and forward model translate directly to the extent and shape of their corresponding information spaces, and the uncertainty in the resulting estimate is similarly characterized by the extent and shape of the solution information space. If uncertainties in measurements, model, and prior estimate are represented by probability distributions, then the solution is defined as the joint posterior distribution of the retrieved parameters. The conjunction of the probabilities of parameters  $\mathbf{x}$  and observations  $\mathbf{y}$  is typically expressed via Bayes' Theorem as a relationship between conditional probabilities

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x})p(\mathbf{x})}{p(\mathbf{y})}. \quad (2)$$

Here,  $p(\mathbf{x}|\mathbf{y})$  is the posterior probability density to be solved for,  $p(\mathbf{x})$  is the prior probability density of the set of parameters  $\mathbf{x}$ ,  $p(\mathbf{y}|\mathbf{x})$  is the probability density that describes the likelihood that the parameters are the true parameters (given the uncertainty in the forward model and observations), and  $p(\mathbf{y})$  is the prior probability density of the observations. The optimal parameter estimate is, by definition, the maximum likelihood point in the conditional distribution  $p(\mathbf{x}|\mathbf{y})$ , and the error characteristics for each retrieved parameter can be examined by considering each parameter's marginal probability distribution. The marginal probability distribution is computed by integrating the joint PDF over all parameters except the parameter of interest.

[12] Variational (least squares) retrieval methods are based in this framework [Cooper et al., 2006; Rodgers, 2000], and assume that the probability distributions for the prior estimate and observations are Gaussian so that an estimate of the maximum likelihood point can be obtained by minimizing a quadratic cost function

$$\Phi(\mathbf{x}) = [F(\mathbf{x}) - \mathbf{y}]^T \mathbf{S}_y^{-1} [F(\mathbf{x}) - \mathbf{y}] + [\mathbf{x} - \mathbf{x}_a]^T \mathbf{S}_a^{-1} [\mathbf{x} - \mathbf{x}_a], \quad (3)$$

where  $F(\mathbf{x})$  is the set of observations simulated using the parameters  $\mathbf{x}$ ,  $\mathbf{x}_a$  is the a priori parameter estimate, and  $\mathbf{S}_y$  and  $\mathbf{S}_a$  are the covariance matrices of the Gaussian observation and prior PDFs, respectively. While variational methods are capable of returning a robust estimate of the maximum likelihood point, they cannot return information about the presence of multiple solutions, and the retrieval uncertainty estimates are generally restricted to the values that lie immediately around the solution. A more complete description of a retrieval solution, including the potential for obtaining multiple solutions, can be obtained through examination of the full joint posterior PDF, which can be thought of as a map of the solution space.

[13] If the number of desired parameters is small, the simplest way to obtain a complete map of the solution space is to compute the (nonnormalized) posterior PDF directly. This can be done by first defining probability distributions for  $p(\mathbf{x})$  and  $p(\mathbf{y}|\mathbf{x})$ , then by running the forward model repeatedly for small increments in each variable parameter and computing the probability  $p(\mathbf{x}|\mathbf{y})$  for each set of simulated and observed measurements. However, the computational expense of doing so increases as  $N^M$ , where  $N$  is the desired resolution of the parameter space and  $M$  is the number of parameters allowed to vary. For joint probability distributions that contain a moderate amount of mass in each dimension, the theory of empty spaces [Tarantola, 2005] dictates that the amount of empty space in the solution space (e.g., the percentage of the space for which values of the parameters are associated with very low probability) increases exponentially with each additional dimension. It follows that an effective method of mapping the probability density function is one that avoids empty regions of the space in favor of regions with relatively high probability.

[14] Markov chain Monte Carlo methods address the problem of empty parameter spaces by sampling the probability distribution in such a way as to revisit those regions with high probability while avoiding regions with low probability [Tamminen and Kyrola, 2001; Tarantola, 2005]. Specifically, the MCMC algorithm is made up of a random walk that consists of multiple successive iterations. In each iteration, test values ( $\tilde{\mathbf{x}}$ ) of all parameters are randomly drawn from a "proposal" distribution ( $\tilde{\mathbf{x}} \sim q(\tilde{\mathbf{x}}; \mathbf{x}^{(i)})$ ) that depends on the current estimate ( $\mathbf{x}^{(i)}$ ). Proposed parameter values are then used in the forward model, and the resulting simulated measurements ( $\tilde{\mathbf{y}} = F(\tilde{\mathbf{x}})$ ) and prior ( $p(\tilde{\mathbf{x}})$ ) are compared with observations via a cost function derived from the assumed form of  $p(\mathbf{y}|\mathbf{x})$ . The proposed parameter values are accepted as a sample in the posterior distribution with probability

$$Q(\tilde{\mathbf{x}}; \mathbf{x}^{(i)}) = \min(\rho, 1), \quad (4)$$

where the acceptance ratio  $\rho$  is defined as

$$\rho = \frac{p(\tilde{\mathbf{x}})p(\mathbf{y}|\tilde{\mathbf{x}})q(\mathbf{x}^{(i)}; \tilde{\mathbf{x}})}{p(\mathbf{x}^{(i)})p(\mathbf{y}|\mathbf{x}^{(i)})q(\mathbf{x}^{(i)}; \tilde{\mathbf{x}})}. \quad (5)$$

If the prior parameter probability distribution is constant (e.g., Uniform), and the proposal distribution is symmetric

**Table 1.** Minimum and Maximum Values of the Retrieval Parameters Used in the Markov Chain Monte Carlo Algorithm

	Ice Water Path (g/m <sup>2</sup> )	Effective Radius (microns)	Cloud Top Height (m)	Cloud Base and Height (m)
Minimum	0.11	1.0	2000.0	1000.0
Maximum	800.0	100.0	20,000.0	20,000.0

(e.g.,  $q(\tilde{\mathbf{x}}; \mathbf{x}^{(i)}) = q(\mathbf{x}^{(i)}; \tilde{\mathbf{x}})$ ), then the acceptance ratio reduces to

$$\rho = \frac{p(\mathbf{y}|\tilde{\mathbf{x}})}{p(\mathbf{y}|\mathbf{x}^{(i)})}, \quad (6)$$

which easily be computed from the specified form of the observation uncertainty PDF applied to the current and proposed parameter values. Note that this also eliminates the need for an a priori term in the cost function.

[15] The value of the acceptance ratio determines whether the proposed set of parameters is saved as a sample of the posterior distribution. If the new parameters generate a better fit to observations ( $\rho > 1$ ), then the proposed set of parameters is saved as the next sample in the distribution ( $\mathbf{x}^{(i+1)} = \tilde{\mathbf{x}}$ ). If not ( $\rho < 1$ ), then a test value is drawn from a Uniform (0,1) distribution. If this value is less than  $\rho$ , then the proposed parameter set is also accepted as a sample in the distribution. Otherwise, the proposed set of parameters is rejected, and new proposed parameter values are drawn on the basis of the original set. The accept/reject procedure ensures that parameter sets that provide a better fit to the observations are immediately accepted, those that provide a similar fit are considered, and those that lead to simulated observations that are very different from the measurements are rejected. In the process, regions of the parameter space with relatively high probability are preferentially sampled, while regions with low probability are avoided, and a sample of the posterior distribution is produced using far fewer computations than brute force PDF mapping. A more complete description of the theoretical underpinnings of the MCMC algorithm is given by *Mosegaard and Tarantola* [1995] and *Tarantola* [2005].

[16] The sample of the PDF returned by the MCMC algorithm fully characterizes the retrieved state, and allows the sources of error that dominate the uncertainty in the result to be identified. Testing different error assumptions, including specification of not only different error magnitudes, but also different error probability distributions, is straightforward, as is introduction of new observations and evaluation of new physical models. Our implementation of the MCMC algorithm for the split window ice cloud property retrieval is based on a simplified version of the algorithm described in full by *Tamminen and Kyrola* [2001]. In our approach, we use a Gaussian proposal distribution centered on the current estimate and with variance that is adaptively changed during a burn in period to give the desired sampling rate. For a discussion on optimal sampling rates, the reader is referred to *Gelman et al.* [2004] and *Haario et al.* [1999]. It should be noted that samples of the posterior PDF obtained during burn-in are not used in the final analysis.

[17] Variable parameters consist of IWP, effective radius, cloud top height, cloud base height, and ice crystal shape,

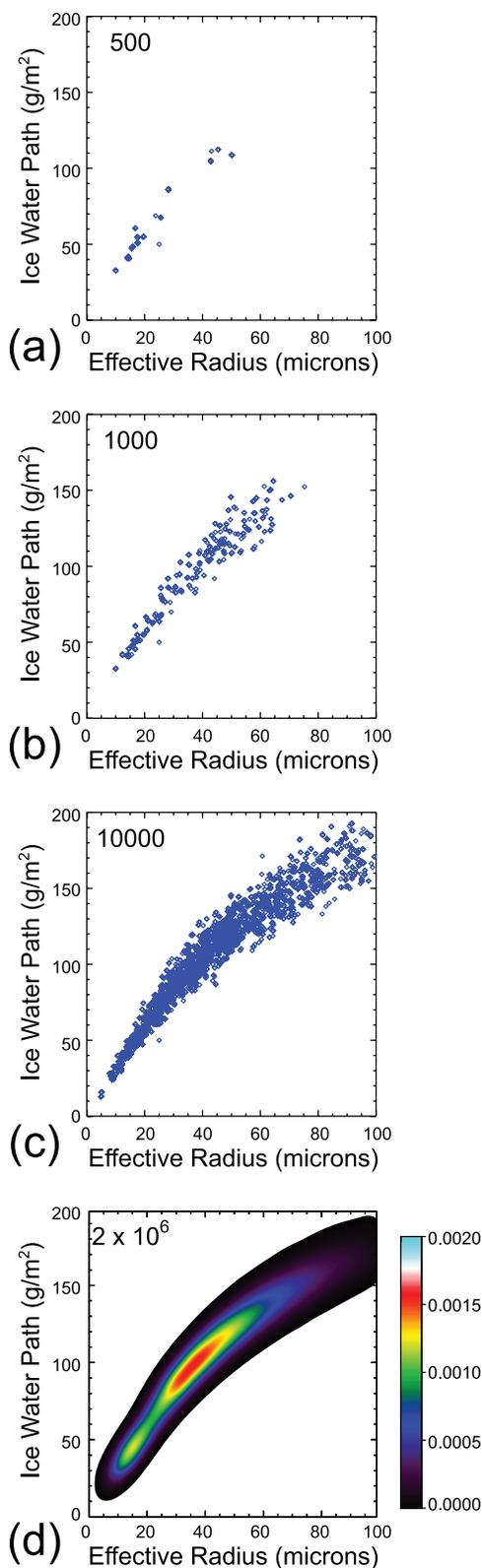
where the shape is allowed to be solid columns, droxtals, bullet rosettes, or aggregates. As such, the space to be sampled is five-dimensional. The prior distribution for each variable parameter is assumed to be Uniform, bounded by physically realistic values (Table 1). Note that we set the maximum IWP to 800 g/m<sup>2</sup> for convenience; tests in which the maximum ice water path was doubled to 1600 g/m<sup>2</sup> yielded identical results. Note also that the range of effective radius values was limited by the available lookup tables, in which effective radius varied between 1 and 100 microns. Results from in situ observations of midlatitude cirrus clouds [*Heymsfield and Iaquinta*, 2000] indicate that this should adequately span the appropriate range of particle sizes.

[18] Because the ice crystal shape is allowed to assume one of four discrete forms, some discussion of how the MCMC algorithm chooses between discrete (integer) values is warranted here. Each allowed ice crystal shape is assigned a different integer index between 1 and 4. The MCMC algorithm is allowed to sample a range of real values between 0.5 and 4.5, and the proposed ice crystal shape is determined by rounding the result to the nearest integer value. When a proposed crystal shape is selected as a sample, the proposed (real) value of the parameter index is retained for use in the next proposal. The robustness of this technique was tested by assigning different integer values to each of the ice crystal shapes and rerunning the algorithm; the results were found to be completely insensitive to the integer value of the index assigned to each crystal shape.

[19] Except where noted below, we assume a Gaussian probability distribution function for the observation errors so that the misfit (cost) function consists of the sum of the squared differences between simulated and observed measurements, weighted by a specified error variance. For a given observation  $y_i$ , this can be written

$$\Phi(\mathbf{x}) = -\frac{1}{2} \left( \frac{F(\mathbf{x}) - y_i}{\sigma_{y_i}} \right)^2. \quad (7)$$

Error standard deviations ( $\sigma_{y_i}$ ) are based on the uncertainty analysis described in *Cooper et al.* [2006] that included both instrument noise and uncertainty from the specification of ECMWF temperature profiles, and are set equal to 1.5 K and 1.0 K for the 11.0-micron brightness temperature and 11.0– to 13.3-micron brightness temperature difference, respectively. The errors in observations of brightness temperature in 11– and 13.3-micron wavelengths are assumed to be uncorrelated, and all other sources of uncertainty (such as variability in ocean surface emissivity and the effects of cloud inhomogeneity) are assumed to be small by comparison.



**Figure 2.** Plots of successive numbers of samples of the joint PDF of IWP and effective radius for the single pixel results described in section 3.2. Number of samples depicted in each plot are (a) 500, (b) 1000, (c) 10,000, and (d) 2,000,000.

[20] According to the results of *Haario et al.* [1999], approximately 20,000 iterations is sufficient to sample a five-dimensional parameter space if the underlying target (posterior) distribution is multivariate Gaussian. In practice, PDFs of IWP and effective radius were found to be non-Gaussian, and approximately 100,000 iterations were necessary to sample the posterior PDF. Convergence of the algorithm was assessed by comparing distributions generated by drawing random samples of 100,000 from a much larger sample. For each single pixel experiment, we ran  $2 \times 10^6$  iterations to ensure a robust and complete sample. For the scene, this number was reduced to 500,000 to speed computation. In Figure 2, the convergent properties of the MCMC algorithm are demonstrated for the single MODIS pixel examined in section 3 (see also Figure 6 in section 3.2). This plot shows successively greater numbers of samples of the joint PDF of IWP and effective radius, demonstrating the manner in which the solution PDF is constructed.

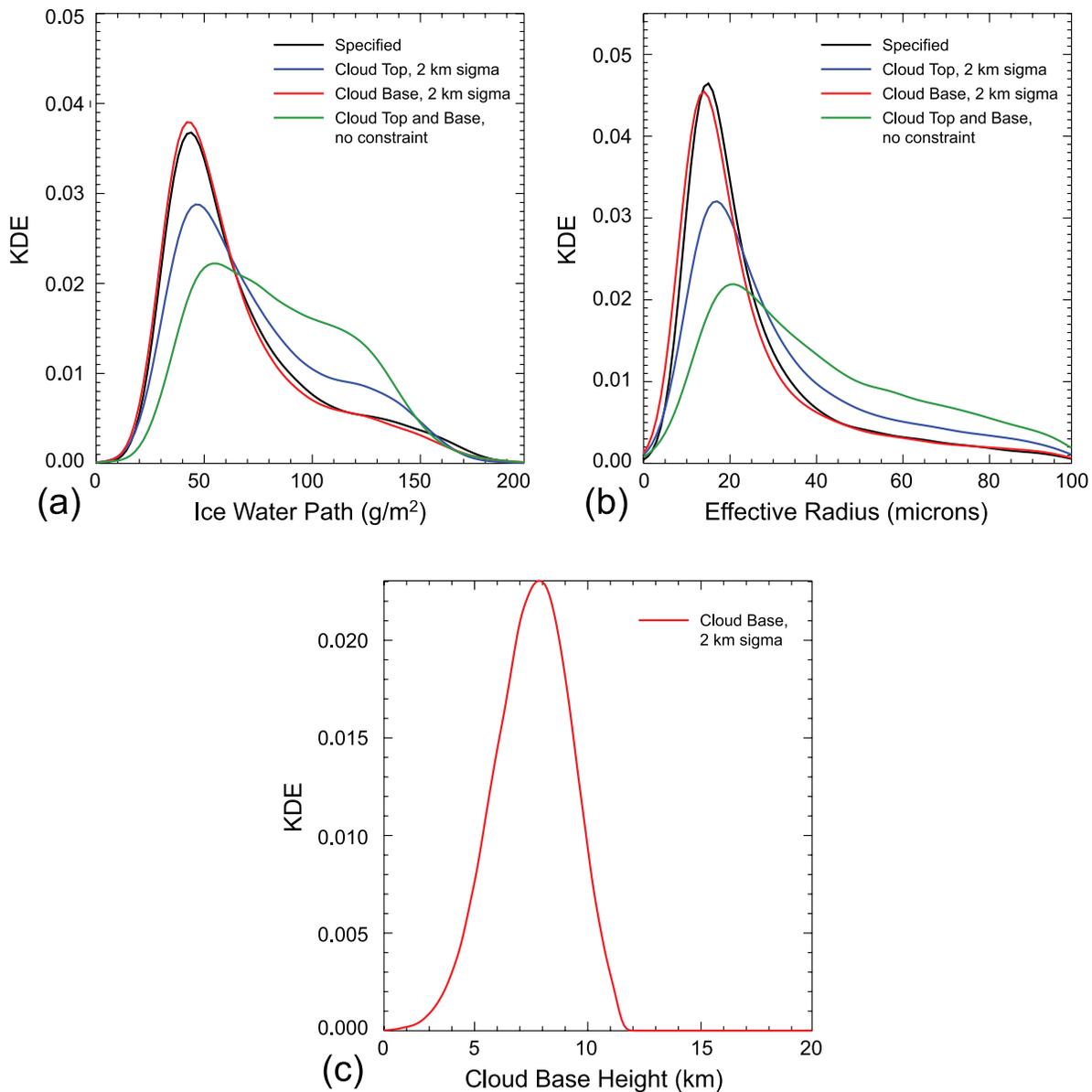
[21] Retrieved IWP and effective radius were estimated using the mode of the joint probability distribution returned by the MCMC algorithm. As such, the retrieved IWP and effective radius values correspond to the Maximum A Posteriori (MAP) estimate. While it is also possible to produce an estimate using marginal expectations, the MAP estimate is more robust for cases in which the posterior PDF exhibits skewness and/or multiple modes. For plotting and identification of modes in the PDF, we employ the kernel density estimate (KDE) with Gaussian kernel and variance (bandwidth) chosen via use of a data-driven bandwidth selector [*Wand and Jones, 1995*]. Because the sample of the PDF is relatively large, the characteristics of the PDF, including the location of the mode, are insensitive to the choice of bandwidth. The KDE has been shown to accurately characterize the properties of the underlying continuous PDF for discrete samples, and avoids problems associated with choice of histogram bin widths and start and end points.

### 3. Characteristics of Single Pixel Retrievals

[22] We first examine the characteristics of a single-pixel split-window retrieval PDF under different error assumptions. The pixel of interest is located at 45.1 degrees north latitude and  $-65.6$  degrees west longitude, and contains cirrus cloud 3–4 km thick, with cloud top located between 11 and 12 kilometers according to CloudSat and CALIPSO observations, respectively (Figure 1). The 11- and 13.3-micron brightness temperatures for this pixel are 251.1 K and 240.0 K, respectively, and results of the MCMC-based inversion documented below indicate an 11-micron optical depth of 1.77. Note that the 11.0- to 12.0-micron brightness temperature difference for this pixel is 1.52 K, so the observations are expected to contain sufficient information to constrain the split window retrieval. For experiments in which cloud top and base heights are fixed, cloud boundaries are specified to be the levels at which the CloudSat reflectivity drops below  $-30$  dBZ.

#### 3.1. Fixed Crystal Shape

[23] In the first set of experiments, we assume that all particles are solid columns. We first examine the general characteristics of the solution space for cloud top and base



**Figure 3.** Plots of the marginal PDFs of (a) IWP and (b) effective radius for cases in which cloud base and top are specified (black line), cloud top is allowed to vary with standard deviation of 2 km (blue line), cloud base is allowed to vary by 2 km (red line), and cloud top and base are allowed to vary without constraint (green line). (c) Marginal PDF of the cloud base height for the case in which cloud base is allowed to vary with standard deviation of 2 km.

height that are set to the CloudSat observed base and top height, then allow the cloud top and base heights to vary. Marginal PDFs of IWP and effective radius returned from the MCMC algorithm (Figures 3a and 3b, respectively) represent the probability that brightness temperatures in the single pixel being considered correspond to particular values of IWP and effective radius, and reveal a well-defined mode with skewness toward larger IWP/effective radius. Since the prior PDF was Uniform, the fact that the posterior PDF of both IWP and effective radius has a well-defined mode reflects the influence of information in the observations, however, the width of the resulting PDF demonstrates that a range of solutions are possible. If the

retrieval were perfectly constrained, the posterior PDF would take the shape of a delta function. Though skewness in both PDFs reflects nonlinearity in the model, the results suggest that, provided the crystal shape can be definitively identified as solid columns and cloud top and base are well-constrained, the optimal solution is also well-defined.

[24] In reality, however, a comparison of CloudSat and CALIPSO cloud top heights across the scene of interest (Figure 1b) indicates that CloudSat can be expected to constrain CTH to within  $\pm 1$  km. To test the effect of variations in cloud top and base height, we include CloudSat-observed top and base height as observations in the MCMC algorithm, with uncertainty modeled as a Gaussian distribution with

standard deviation of 2 km. The resulting marginal PDFs of IWP and effective radius (Figures 3a and 3b, respectively) reveal that uncertainty in cloud top height has a much greater effect on the distribution than variability in cloud base height, and has the effect of broadening the distribution of both IWP and effective radius toward larger values of IWP and effective radius.

[25] It is interesting to note that, when cloud top height is fixed and cloud base height is allowed to vary, the shape of the PDF of both IWP and effective radius is nearly identical to the PDF that results from fixed cloud top and base, but the mode shifts toward slightly lower values. Examination of the posterior marginal PDF of cloud base height (Figure 3c) explains why this happens. The prior PDF of cloud base height is a Uniform PDF, but the cloud base is constrained to lie below the fixed CTH. The PDF of the observed cloud base height is assumed to be Gaussian. The conjunction of these two PDFs produces the posterior PDF of cloud base height, which is Gaussian, truncated near the height of the cloud top. This effectively means that, while the mode of the posterior PDF of cloud base height lies at the observation, more of the mass in the PDF extends to lower heights (geometrically thicker cloud) because of the constraint of cloud base lower than cloud top. For a geometrically thicker cloud to produce the same brightness temperature, the model requires lower IWP and effective radius.

[26] We next allow the cloud top and base height to vary without bounds, as may be the case if no ancillary CloudSat or CALIPSO observations are available. We only impose the constraint that cloud top height be greater than cloud base height and cloud depth be greater than or equal to the width of a single CloudSat range bin. In comparison to the results obtained when cloud top and base height observations were used, both the variance and skewness of the posterior PDF continue to increase. This indicates that, with increasing uncertainty in cloud top height, larger values of IWP and effective radius become more likely (Figure 3). It should be pointed out that, with removal of the constraint on cloud top and base height, we have added another free parameter to the retrieval with no observational information to constrain it. The fact that the PDFs still exhibit a distinct mode in the vicinity of the solution obtained from fixed cloud top and base reflects the fact that information about cloud top height is present in the brightness temperature observations. However, the solution is now biased toward slightly larger values of IWP and effective radius. In addition, IWP appears to suffer more from the removal of constraint on cloud top and base height than does the effective radius. This is due to the fact that the IWP information is derived primarily from the magnitude of the cloud top brightness temperature, while effective radius is sensitive to the brightness temperature difference and hence more of the information in the observations is preserved under a change in cloud top height.

[27] As in the marginal PDFs, the joint PDF of IWP and effective radius for specified cloud top and base shown in Figure 4a exhibits a well-defined mode. In addition, the joint PDF also clearly shows a positive correlation between IWP and effective radius, with nonlinearity in the relationship indicated in the curve of the PDF. The joint skewness toward larger values of IWP and effective radius reflects the

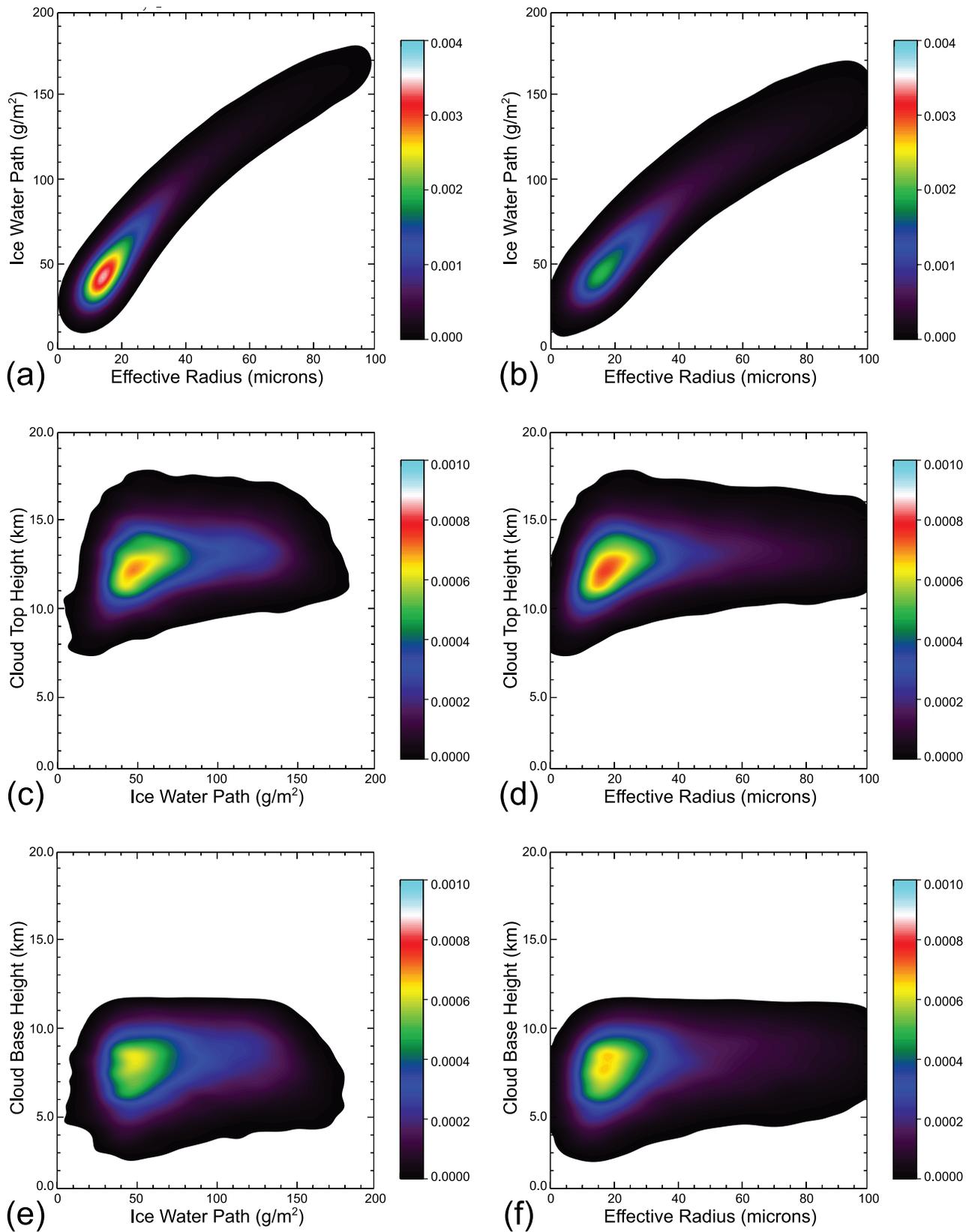
saturation of information at higher IWP noted by *Cooper et al.* [2007]. The relationship between an increase in the likelihood of higher cloud top and base heights and an increase in the probability of larger IWP/Re described above is evident in the remaining joint PDFs in Figure 4; as values of cloud top and base increase, the PDF spreads in the direction of greater IWP and effective radius. The form of the joint PDFs of IWP and effective radius with cloud top and base height reveals the reason for the selective shift toward larger IWP/Re: the conjunction of the skewed PDF of effective radius and IWP with the Gaussian PDF of the cloud top and base height leads to favoring a higher solution to the cloud top height at larger values of IWP and effective radius.

[28] The marginal and joint PDFs of the solution contain valuable information on the relationships between variables, nonlinearity in the forward model, and error characteristics of the solution. What is encouraging is that the solution space is well-behaved for the sources of error examined; there exists a single most likely solution for a given set of measurements. The results are not surprising, given the fact that OE and lookup table split window retrievals have been extensively used in the case of a fixed ice crystal shape and well-constrained cloud top and base. What the MCMC approach provides is the flexibility to examine the effect of assumptions on the solution. In the case of cloud top height uncertainty, the relationship between IWP and effective radius leads to a bias toward large IWP and effective radius. In addition, skewness in the PDF indicates that methods that assume Gaussian PDFs and return the conditional mean estimate would be biased, whereas the conditional mode estimate should work well. Examination of the region immediately around the modes of the marginal PDFs of IWP and effective radius (Figure 3) indicates a Gaussian PDF can be assumed locally, indicating that optimal estimation algorithms that return the conditional mode will produce a reasonable error estimate around the solution, provided the ice crystal shape is known.

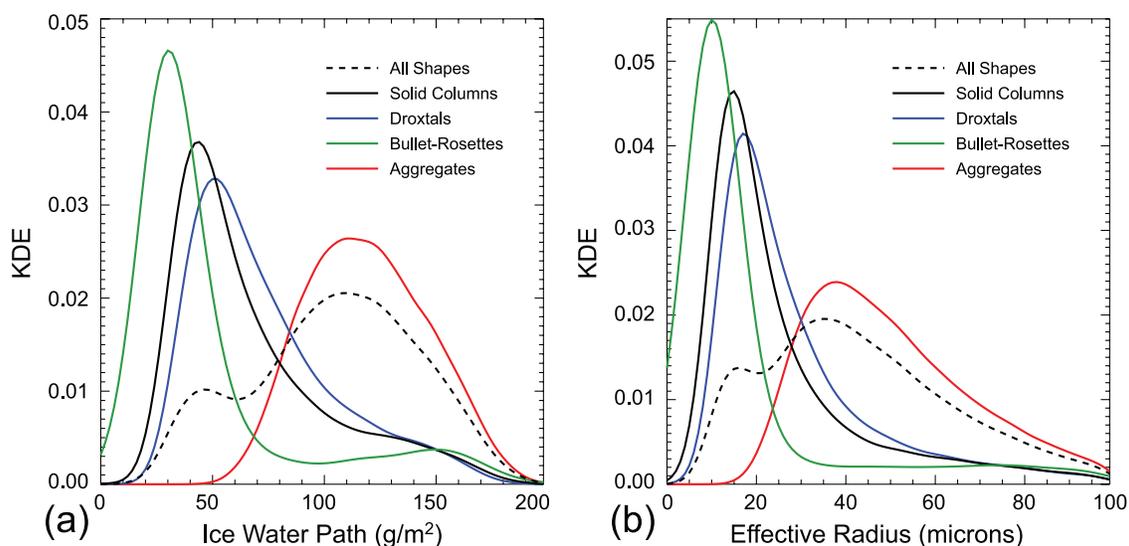
### 3.2. Variation in Crystal Shape

[29] The results presented in section 3.1 assumed a fixed ice crystal shape (solid columns). Observational studies [*Heymsfield and Iaquina, 2000*] have shown that a mix of crystal shapes are typically found in midlatitude cirrus clouds, and that retrieved IWP and effective radius are sensitive to the assumed crystal shape [*Baum et al., 2005*]. While it is possible to use temperature and humidity to crudely predict the most probable ice crystal habit or mixture of habits in a statistical sense, it is not generally possible to prescribe the precise shape of ice crystals that comprise any given cloud from infrared radiance measurements alone. To simulate the impact of imprecise prior knowledge of crystal habit, the MCMC algorithm is allowed to adaptively choose between four commonly observed forms: droxtals, solid columns, bullet rosettes and aggregates. The results indicate the relative magnitudes of uncertainty due to changes in crystal shape versus changes in cloud top and base height as well as identifying the crystal shape that provides the best fit for this pixel.

[30] Marginal PDFs of IWP and effective radius shown in Figure 5 reveal that different crystal shapes produce posterior PDFs with very different modes and shapes. Bullet rosettes exhibit the smallest values of IWP and effective



**Figure 4.** Plots of the joint PDFs of (a and b) IWP and effective radius, (c) IWP and cloud top height, (d) effective radius and cloud top height, (e) IWP and cloud base height, and (f) effective radius and cloud base height. In Figure 4a, cloud top and base height are fixed, while cloud top and base height are allowed to vary by  $\pm 1$  km in Figures 4b–4f.



**Figure 5.** Marginal PDFs of (a) IWP and (b) effective radius in a single pixel for the case in which crystal shape is allowed to vary. In each plot, the lines correspond to the sum of all crystal shapes (black dashed line), solid columns (solid black line), droxtals (blue line), bullet rosettes (green line), and aggregates (red line).

radius, with size and ice water content increasing as the crystal shape is changed to solid columns, droxtals, and aggregates. Droxtals exhibit a PDF with very similar shape to solid columns, with a single well-defined mode and positive skewness. In contrast, the PDF of ice water path for aggregates has little skewness, but much larger variance, while the PDF of effective radius for aggregates is skewed toward large effective radii. The PDF for bullet rosettes is bimodal in ice water path and has a very long tail in effective radius. Note that, for this pixel, bullet rosettes are not a favored crystal shape; the total PDF exhibits little mass at the IWP and effective radius mode relative to the other crystal types. Solid columns and droxtals are more likely, but aggregates provide the best fit to the observed brightness temperatures. These differences reflect the differences in radiative properties of the various crystal shapes, and it is not unexpected that the most likely values of ice water path and effective radius will differ between them.

[31] When all shapes are considered in a single retrieval (Figure 6) the joint PDF is bimodal, with a mode at smaller Re/IWP that corresponds to the combination of solid columns and droxtals, and a mode at higher values of Re/IWP that corresponds to aggregates. Examination of the joint PDF for each individual habit reveals some similarities: there is correlation between IWP and effective radius for each crystal shape, but the degree of correlation varies significantly between the different habits. As was noted above, the joint PDF for solid columns and droxtals demonstrates a degree of correlation between IWP and effective radius, but by contrast, the correlation between

IWP and effective radius is very weak for bullet rosettes and quite strong for aggregates.

### 3.3. Effect of Changes to Error Assumptions and Retrieval Characteristics

[32] We now look at the effect of changes to the retrieval assumptions on the characteristics of the solution space. This is done in three different ways. First, we examine the effect of changes to the uncertainty in the observations themselves. This is designed to simulate an improvement in either instrument quality, analysis temperature profile, or both. To do this, we reduce the errors in the observations to 1/2 of their original value, specifying an error of 0.75 K for the 11-micron brightness temperature, and an error of 0.5 K for the 11.0– to 13.3-micron brightness temperature difference. The second change is to the errors on cloud top height. In the above discussion of the MODIS scene, we noted that the cloud top height observed by CALIPSO was consistently higher than that observed by CloudSat. Because it is unlikely that the true (radiative) cloud top height will be lower than that observed by CloudSat [Weisz *et al.*, 2007], it is reasonable to model the cloud top height uncertainty using a distribution that allows for greater likelihood of a higher cloud top height and relatively smaller likelihood of a lower cloud top height. To test the effect of such a change, we include observations of cloud top height from CloudSat in the retrieval, and model the errors using a lognormal probability distribution. This introduces a modification to the objective function used in the MCMC algorithm, and we shall illustrate this briefly here. For a more detailed discussion of the properties of the lognormal PDF when used

**Figure 6.** Joint PDFs of IWP and effective radius in a single pixel for the case in which cloud top and base height are fixed, and crystal shape is allowed to vary. Plots correspond to cases in which (a) all crystal shapes are considered, (b) only solid columns are allowed, (c) only droxtals are allowed, (d) only bullet rosettes are allowed, and (e) only aggregates are allowed. Note that a white line has been added to Figure 6d to highlight the secondary mode in the joint PDF of bullet rosettes.

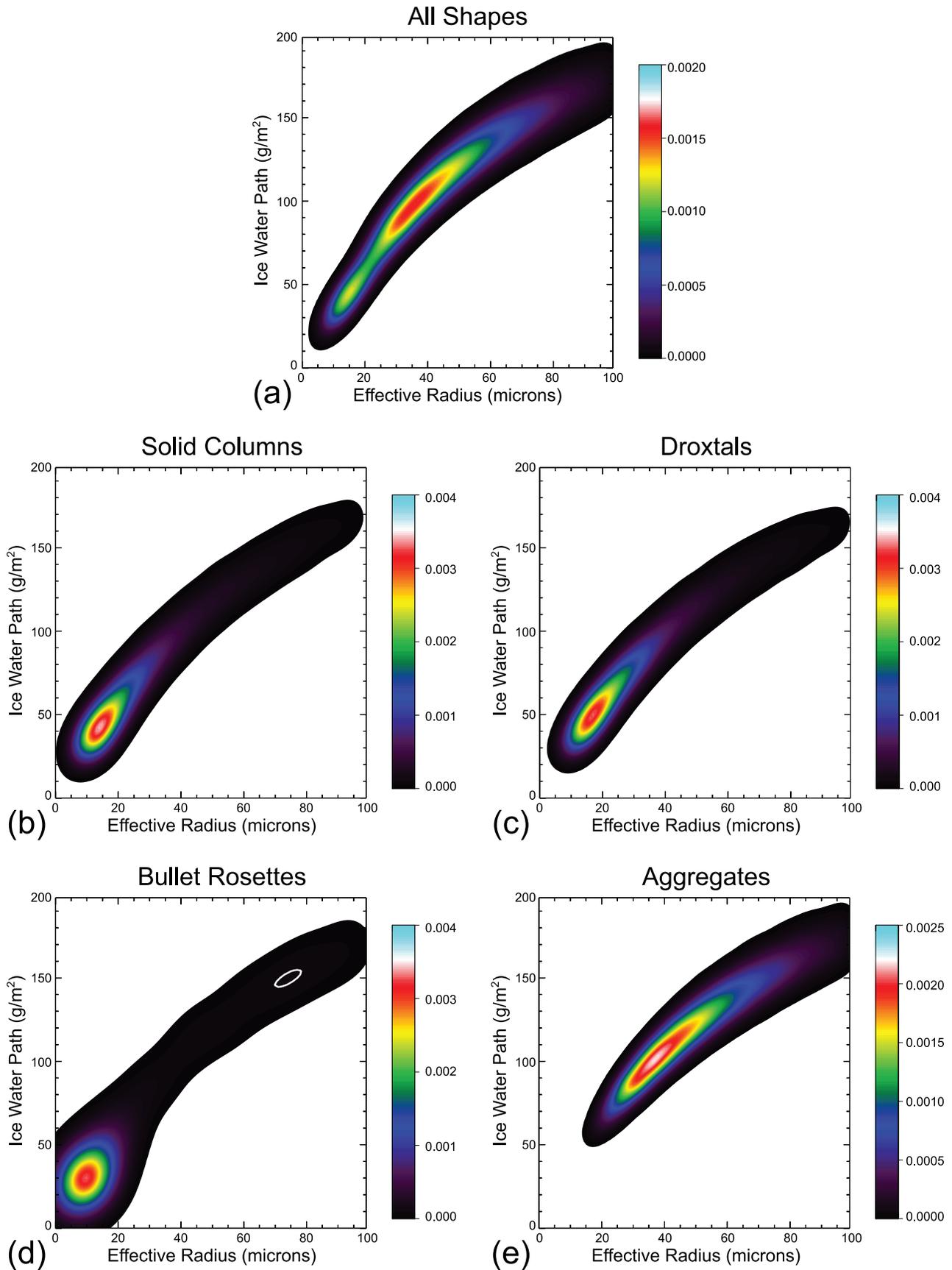


Figure 6

in inverse problems, the reader is referred to *Cohn* [1997] and *Fletcher and Zupanski* [2006].

[33] The implementation of a lognormal cost function in the MCMC algorithm starts from the definition of the (univariate) lognormal PDF

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma_x}\right)^2\right\}. \quad (8)$$

In this case, the error standard deviation  $\sigma_x$  is defined as the standard deviation of the log of the normally distributed error, which can be written

$$\sigma_x = \left[ \ln\left(1 + \left(\frac{\sigma_{Gx}}{\mu_x}\right)^2\right) \right]^{\frac{1}{2}}, \quad (9)$$

where  $\sigma_{Gx}$  is the error standard deviation for the corresponding Gaussian uncertainty, and  $\mu_x$  is the mean of the sample.

[34] Since the lognormal is a geometric distribution, errors are most properly expressed as a ratio [*Cohn*, 1997; *Fletcher and Zupanski*, 2006]. Hence, for a given observation  $y$  and forward observation  $F(\mathbf{x})$ , the error distribution takes the form

$$f\left(\frac{y}{F(\mathbf{x})}\right) = \frac{F(\mathbf{x})}{y\sigma_y\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\ln y - \ln F(\mathbf{x}) - \mu}{\sigma_y}\right)^2\right\}. \quad (10)$$

Now, if we assume unbiased errors so that  $\mu = 0$ , we obtain the cost function as the negative log of the error distribution

$$\Phi(x) = \ln y - \ln F(\mathbf{x}) + \frac{1}{2} \left( \frac{\ln y - \ln F(\mathbf{x})}{\sigma_y} \right)^2, \quad (11)$$

where we have neglected the constant scale factor. The MCMC algorithm then includes equation (11) as a term in the total misfit function (assuming observation errors are uncorrelated).

[35] Finally, there is evidence from in situ observations [*Heymsfield and Iaquinta*, 2000], and an examination of the CloudSat reflectivity depicted in Figure 1a bears this out, that larger particles tend to settle to lower levels within a midlatitude cirrus cloud. We model this size sorting effect in the retrieval by allowing layer ice water contents and effective radii to vary by  $\pm 25\%$  between cloud top and base while holding the IWP of the entire cloud fixed. The retrieved effective radius is assumed to be characteristic of the middle of the cloud and to increase toward the cloud base and decrease toward the cloud top.

[36] Marginal PDFs under these three alternate sets of error assumptions are collectively presented in Figure 7. The use of a lognormal PDF for the cloud top height (Figures 7a and 7b) both shifts the mode of the resulting IWP and effective radius distributions toward slightly higher values, and adds mass in the tail of the PDFs, indicating higher likelihood of larger IWP and effective radius for the lognormal cloud top versus specified cloud top or Gaussian error. It is interesting to note that, even with the assumption of

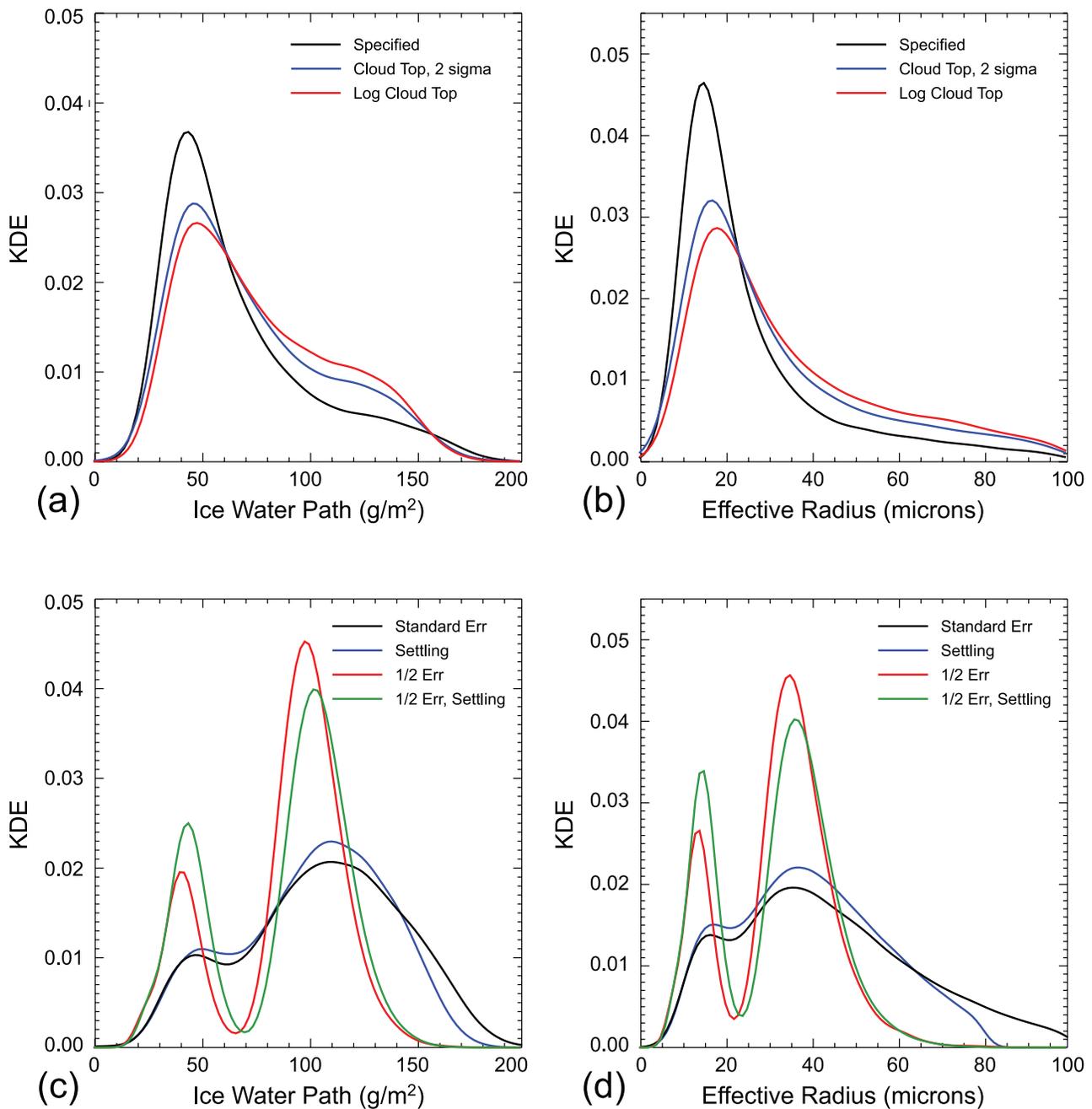
lognormal cloud top height error, the likelihood of obtaining an ice water path greater than  $150 \text{ g/m}^2$  is near zero for this profile, reflecting the constraint of the brightness temperature observations on the cloud optical thickness. This stands in contrast to the effective radius, where increased variability in the cloud top height increases the likelihood of larger effective radius across the entire range of allowed values.

[37] Effects of decreased observation error and particle settling are shown in Figures 7c and 7d for IWP and effective radius, respectively. The full PDF for all shapes is shown here to illustrate the effects of changes to errors on the PDFs of IWP and effective radius for different shapes. The primary effect of increased accuracy in the observations is a decrease in variance for the solution and a strong sharpening of the modes associated with solid columns/droxtals and aggregates. Increased separation in the solution modes means that it will be easier for iterative solution methods (e.g., optimal estimation) to get “stuck” in the secondary mode, and not find an estimate of the true maximum likelihood solution. In the case of a retrieval that is based on simple conditional mean estimation, the retrieved values would lie between the two solution modes in the true underlying PDF. In addition, error characteristics that are based on Gaussian statistics will represent only the sensitivity around the local mode, and hence will underestimate the true solution variance.

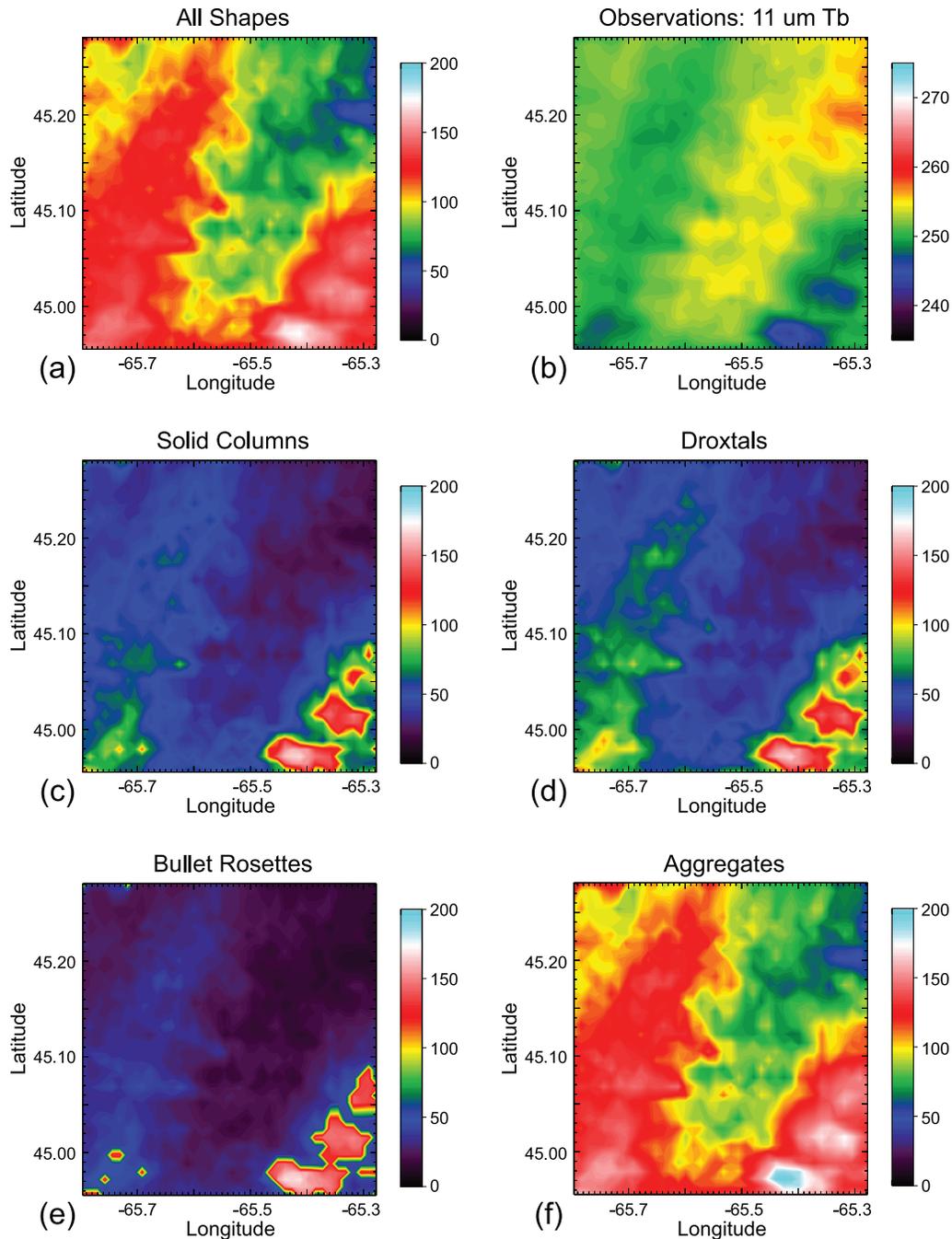
[38] The effects of particle settling are also shown in Figures 7c and 7d. Under standard error assumptions, the probability distributions of effective radius and IWP do not appreciably change, though there is a slight shift in the IWP PDF toward higher values, consistent with a decrease in IWP and optical depth at the top of the cloud: the total ice water path increases so that the brightness temperature remains the same. When errors are artificially reduced by 50%, settling again causes a slight shift toward higher IWP and effective radius but there is also notable redistribution of mass in the PDF. For the reduced error case, the likelihood of solid columns/droxtals increases relative to aggregates, and it appears that the increased likelihood of solid columns and droxtals mirrors the assumed decrease in the particle size at the top of the cloud.

[39] The most significant implication of the above experiments is that changes to error assumptions produce changes to the solution space that are not easily predictable, and that differ across different crystal shapes and in combination with changes to other error assumptions. The implementation of particle settling is a key example. It has relatively little effect on the marginal and joint PDFs when standard errors are assumed on the observed brightness temperature, however, when the errors are reduced, the assumption of settling causes the solution space to change in character so that solid columns and droxtals become a more preferred solution in comparison to the case without settling.

[40] The general picture that emerges from a thorough examination of the statistics of a single pixel is the large effect that assumptions about crystal shape, particle settling, and observation error magnitude have on retrievals of IWP and effective radius. These results imply that a very careful examination of the effect of new assumptions on the solution should be performed before any change is implemented so that the solution can be shown to be unique and



**Figure 7.** Plots of the marginal PDFs of (a and c) IWP and (b and d) effective radius for various retrieval uncertainty assumptions. In Figures 7a and 7b, the crystal shape is specified as solid columns. The solid black line corresponds to the case in which cloud top and base are specified. The blue line corresponds to the case in which cloud top is allowed to vary according to a Gaussian distribution with 2-km standard deviation. The red line corresponds to the case in which the cloud top is allowed to vary according to a log-normal distribution with 2-km standard deviation. In Figures 7c and 7d, the results for all crystal shapes are shown. Here the solid black line corresponds to standard error assumptions and no assumed particle settling. The blue line corresponds to standard error assumptions and assumed particle settling. The red line corresponds to an assumption of half of the retrieval error in both channels. The green line corresponds to an assumption of half of the retrieval error in both channels in combination with assumed particle settling.



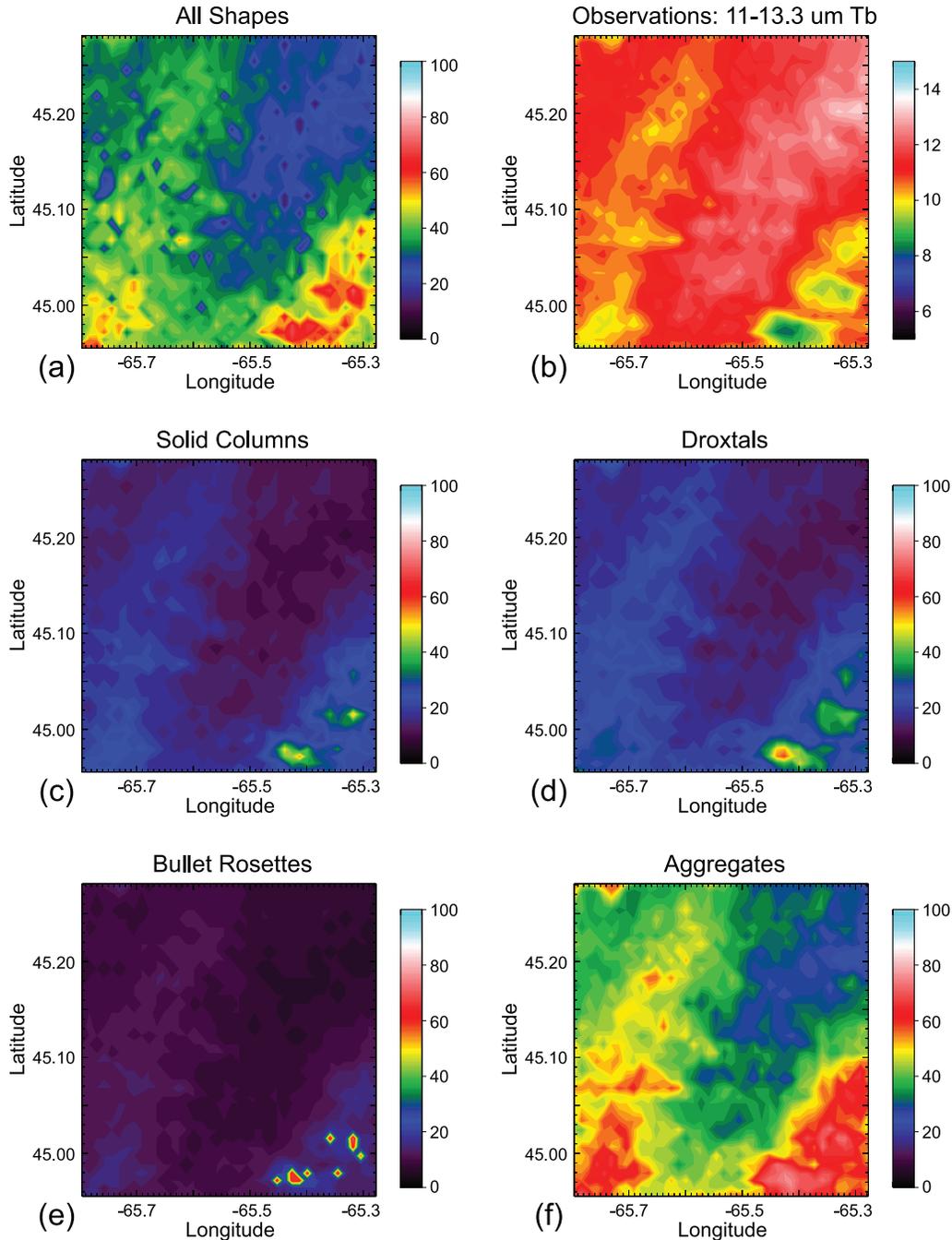
**Figure 8.** Plots of (a) 11-micron MODIS brightness temperatures, and (b–f) retrieved ice water path ( $\text{g/m}^2$ ) for the MODIS scene outlined in the box in Figures 1a and 1b. IWP corresponds to all crystal shapes (Figure 8b), solid columns (Figure 8c), droxtals (Figure 8d), bullet rosettes (Figure 8e), and aggregates (Figure 8f).

well-defined, and so that a correct specification of error statistics can be made.

#### 4. Scene Dependence of Retrieval Error Characteristics

[41] The MCMC algorithm is now applied to a scene observed by MODIS to explore variations in solution space as characteristics of the target clouds such as geometric and optical thickness change. The scene of interest is depicted in the box in Figure 1a, and contains cloud that transitions

from brightness temperature of 240 K to 265 K and from geometric thickness of 4 km to 2 km. Results consist of the full sample of the PDF for each pixel, from which we can infer the maximum a posteriori (MAP) estimate for each crystal shape, the variance of this estimate, the departure of the solution PDF from Gaussian, i.e., skewness and/or multiple modes, and the information content, which is inversely related to the variance in the posterior estimate. As in the single pixel experiment, IWP, effective radius, crystal shape, cloud top and base heights are all allowed to



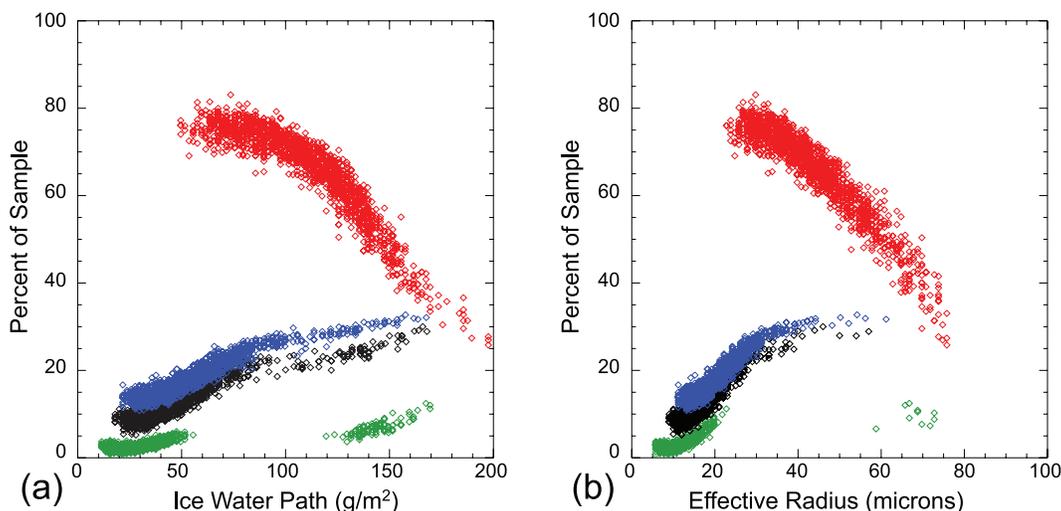
**Figure 9.** As in Figure 8, but for (a) 11- to 13.3-micron MODIS brightness temperatures, and (b–f) retrieved effective radius.

vary. Uncertainties of 1.5 K and 1.0 K are specified for 11.0-micron Tb and 11– to 13.3-micron Tb, respectively, and observations of cloud top and base height from Cloud-Sat are used to constrain the cloud boundaries. Log-normal probability distributions are used for cloud boundaries with  $\pm 1$  km errors assumed for CTH to accommodate the presence of diffuse cloud near the radiative cloud top. The uncertainty in cloud base is increased to  $\pm 2$  km to account for possible misinterpretation of radar returns in the pres-

ence of virga that may artificially lower the estimated cloud base height.

#### 4.1. Maximum a Posteriori Estimate

[42] The MAP estimates of IWP and effective radius for the scene, defined as the mode of the posterior PDF, are plotted in Figures 8 and 9, respectively. Retrievals employing all crystal shapes are compared to those that would be obtained if prior information was available to constrain crystals to one of the four individual shapes defined above. The differences in the mode of the PDF for each crystal



**Figure 10.** Plots of the percent of the MCMC-derived sample in each pixel that corresponds to solid columns (black), droxtals (blue), bullet rosettes (green), and aggregates (red) for (a) IWP and (b) effective radius.

shape are striking. With the exception of the region of largest IWP in the southeast corner of the domain, the assumption of bullet rosettes (Figure 8e) leads to IWP estimates that are a factor of three smaller than those obtained with aggregates (Figure 8f), indicative of the large uncertainties incurred owing to the specification of crystal shape in thin ice cloud retrievals. When the crystal shape assumption is relaxed and the algorithm allowed to choose the shape that provides the best fit to the observations (Figure 8a), aggregates appear to make up the bulk of the solution again, with the exception of the thickest clouds in the region.

[43] In contrast to IWP, in which the solution tends to be dominated by aggregates, retrieved effective radii (Figure 9) tend to exhibit contributions from all crystal shapes and lie between the larger values that correspond to aggregates (Figure 9f) and smaller values obtained when solid columns or droxtals are assumed (Figures 9c and 9d, respectively). This can be more clearly seen by plotting the fraction of the MCMC-derived sample for all pixels in the scene that corresponds to each crystal shape as a function of retrieved values of IWP and effective radius (Figures 10a and 10b, respectively). It is clear that the sample consists primarily of aggregates over most of the range of retrieved IWP and effective radius, with droxtals the next most common shape followed by solid columns then bullet rosettes. It is interesting that the fraction of the sample made up of aggregates decreases with increasing IWP and effective radius, while the fraction of other crystal shapes increases. With the exception of bullet rosettes, all crystal shapes are equally likely at relatively large values of IWP/effective radius. It is also interesting that, out of the four crystal shapes considered, bullet rosettes alone exhibit a discontinuity in the range of most likely values of IWP and effective radius, reflecting the existence of two distinct sets of solutions.

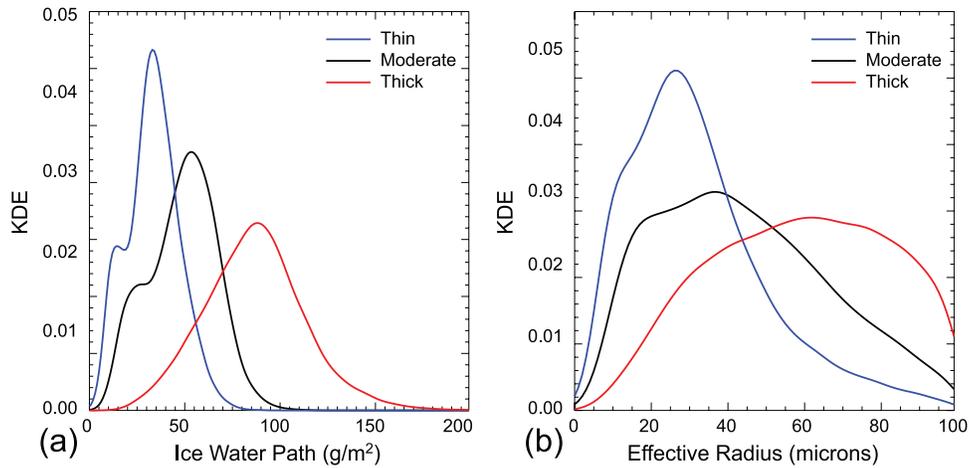
#### 4.2. Characteristics of the Solution Space

[44] Having thoroughly examined the characteristics of retrieved PDFs for a pixel containing moderately thick

cloud (section 3), we now examine how the characteristics of the solution space might differ with respect to changes in cloud physical and optical thickness. We first examine the results for three different pixels, which contain thin, moderate, and relatively thick cloud, from the perspective of split-window retrievals (Figure 11).

[45] Several features of the retrieved PDFs are worth noting. First, in addition to the shift in MAP estimate to higher values of IWP and effective radius with increasing cloud thickness, the distributions broaden. This is especially true for the effective radius. The fact that there is a relatively high probability of obtaining an effective radius anywhere in the range between 30 and 90 microns, combined with the lack of a well-defined mode indicates the relative lack of information in observations of effective radius in thicker clouds. Second, for IWP (Figure 11a), the presence of multiple modes in the solution is much more likely for thin and moderate cloud than for relatively thick. For effective radius (Figure 11b), we can see that none of the pixels has a multimode solution, though there is a hint of a secondary mode for the thin and moderate pixels. In the case of the solution that considers all habits, PDFs are multimodal for 50% of the pixels in the scene. The solution is uniformly unimodal for solid columns, droxtals, and aggregates, while bullet rosettes produce a multimodal solution over 57% of the scene. Further examination reveals that the all-shape solution is multimodal for thin cloud and unimodal for thick, while the reverse is true for bullet rosettes.

[46] The nature of the variability in the solution can be further explored by examining plots of standard deviation and skewness across the domain (Figure 12). Variance in the posterior PDF is relatively low for both IWP and effective radius in regions of relatively thin cloud, and PDFs of IWP exhibit a larger range of variance across the entire scene than PDFs of effective radius. It is interesting that the skewness is much larger for effective radius than for IWP over much of the domain. It is also interesting to note that the skewness is at a maximum in regions of thin cloud,

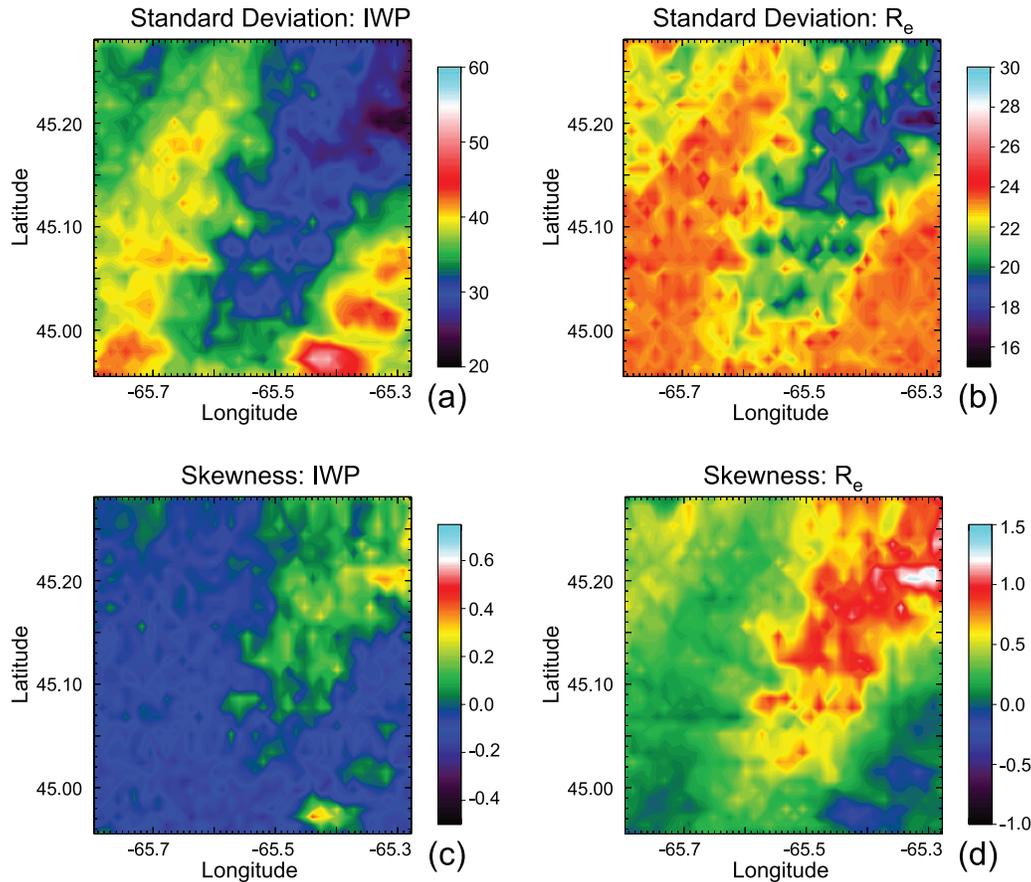


**Figure 11.** Plots of the marginal PDF of (a) IWP and (b) effective radius for three pixels containing relatively thin (blue line), moderate (black line), and thick (red line) cloud.

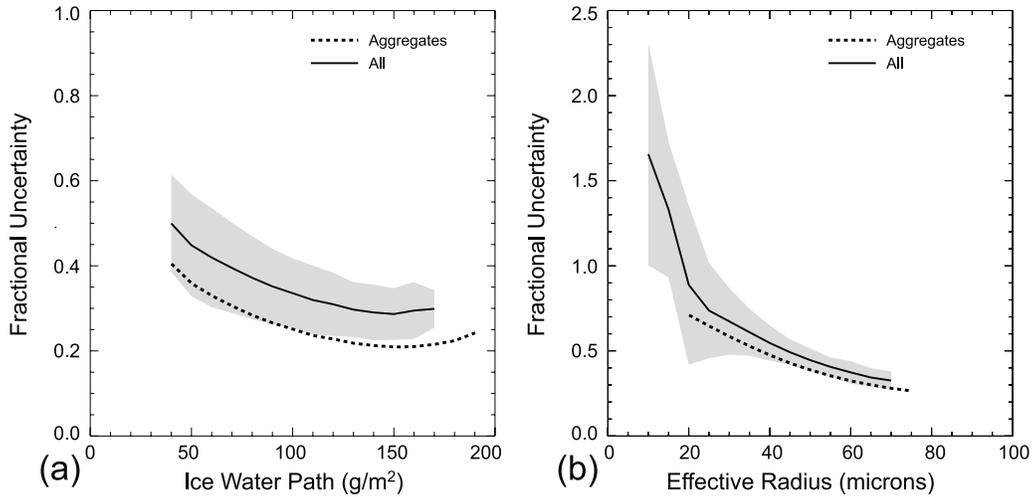
while distributions for both IWP and effective radius tend toward symmetry where the cloud is thicker.

[47] An estimate of the uncertainty in the retrieval relative to the retrieved value can be obtained by dividing the standard deviation of the retrieved PDF of IWP and effective radius by their MAP estimates. Plots of the resulting normalized retrieval error (Figure 13) reveal that the frac-

tional uncertainty decreases with an increase in IWP at low values of IWP, then increases again as IWP values increase beyond about  $150 \text{ g/m}^2$ . In contrast, the fractional retrieval error decreases monotonically over the range of effective radius values. For reference, the fractional uncertainty associated with the PDF derived from a single crystal shape, in this case, aggregates, is also plotted, and it can be seen



**Figure 12.** Plots of (a and b) standard deviation and (c and d) skewness of the retrieved PDF of IWP (Figures 12a and 12c) and effective radius (Figure 12b) and over the MODIS scene (Figure 12d).



**Figure 13.** Fractional uncertainty for retrieved (a) IWP and (b) effective radius as a function of the retrieved value. The solid black line depicts the result for all crystal shapes while the gray shading indicates  $\pm 1$  standard deviation around the fractional uncertainty. The dashed black line depicts the fractional uncertainty for the case in which the crystal shape is specified to be aggregates.

that specification of a discrete shape reduces the retrieval error across the range of values for both IWP and effective radius.

#### 4.3. Information Content

[48] Another useful diagnostic of retrieval error is the Shannon information content [Shannon and Weaver, 1949; Rodgers, 2000; Cooper et al., 2006], which provides a metric for determining the contribution of observations to a retrieval solution. The information content of an observation or set of observations is computed as the reduction in the entropy of the estimate due to the addition of information from observations. In discrete form, the entropy of state  $P$  is defined as

$$S(P) = \sum_{i=1}^N p_i \log_2(p_i), \quad (12)$$

where  $p_i$  is the discretized PDF and  $N$  is the number of discrete bins the PDF is divided into. Shannon information content is then defined as the difference between the entropy of the a priori state and the entropy of the retrieved state

$$\begin{aligned} H &= S(\mathbf{x}_a) - S(\hat{\mathbf{x}}) \\ &= \left[ -\sum_{i=1}^N p_i(\mathbf{x}_a) \log_2(p_i(\mathbf{x}_a)) \right] - \left[ -\sum_{i=1}^N p_i(\hat{\mathbf{x}}) \log_2(p_i(\hat{\mathbf{x}})) \right], \end{aligned} \quad (13)$$

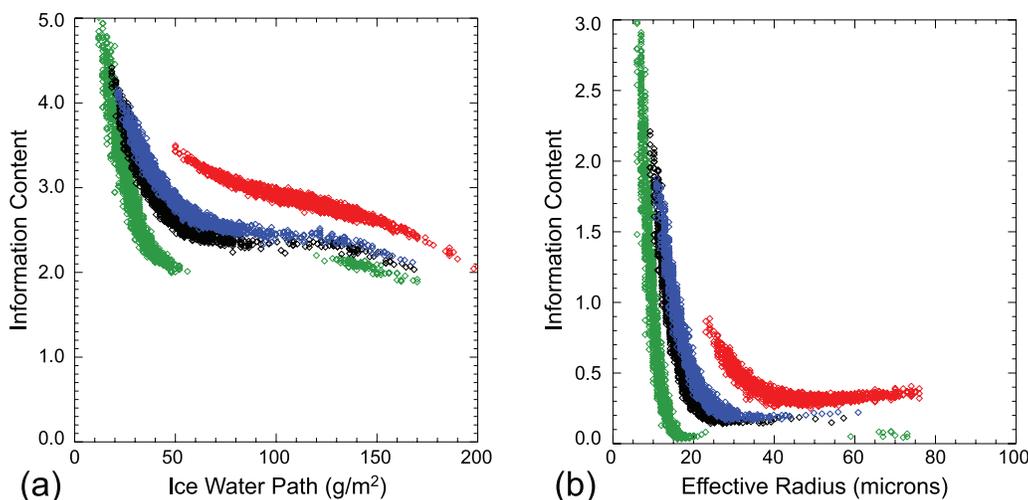
which can be interpreted as the extent to which the number of allowable states is reduced by the addition of information from the measurements. Because the MCMC algorithm returns a sample of the full PDF, we may compute  $H$  directly from the above relationship, and infer the types of clouds for which the retrieval is returning useful information on the physical scene.

[49] We present plots of the relationship between information content and IWP and effective radius in Figures 14a and 14b, respectively. Since a uniform prior distribution is

assumed, the geographic pattern of information content across the scene is inversely proportional to that of the retrieval variance shown in Figures 12a and 12b; for example, where retrieval variance is large, information content is low, and vice versa. The distribution of information content from the MCMC-derived sample reflects the fact that the split window retrieval is generally more applicable in regions of relatively thin cloud. The decrease in information content with increasing IWP and effective radius is clearly evident, as is the increase in general information content from bullet rosettes to solid columns, droxtals, and aggregates. It is also clear that there is little information on the effective radius in observations of clouds with effective radius greater than about 20 microns. What is particularly interesting is the difference in variability of information content for each of the different crystal shapes. Bullet rosettes exhibit very high information content in thin cloud regions and much lower information content for thick clouds. The variability is somewhat lower for solid columns, lower still for droxtals, and lowest for aggregates. Overall examination of the SIC reinforces the result that when ice crystal habit is allowed to vary there is very little effective radius information in observations of thicker clouds. In contrast, the measurements continue to provide limited information about IWP up to ice water path values of approximately 200  $\text{g/m}^2$ , which corresponds to an optical depth of approximately 3.

#### 5. Conclusions

[50] In this paper, the MCMC algorithm is shown to effectively characterize the solution space for a retrieval of ice cloud properties using the split window technique. It is found that the shape of the posterior PDF differs for different assumed crystal shapes, but is generally skewed, with a tail toward higher values of IWP and effective radius and a positive correlation between IWP and effective radius. In typical split-window retrieval conditions where crystal



**Figure 14.** Plots of information content versus retrieved (a) IWP and (b) effective radius for solid columns (black), droxtals (blue), bullet rosettes (green), and aggregates (red).

shape is difficult to constrain in advance, PDFs of IWP and effective radius generally exhibit a multimodal structure with one mode corresponding to droxtals and solid columns and the other to aggregates. As previous studies have shown, uncertainty in cloud top and base height also has a strong effect on the retrieval, but the MCMC algorithm used here reveals that variations in crystal shape generally drive the accuracy of retrieved products. Changes in crystal shape led to changes not only in the relationship between ice water path and effective radius, but also in the relationship between IWP and effective radius and the cloud geometry.

[51] Modeling the effects of particle settling is found to result in lower ice content at the top of the cloud and a consequent shift in the radiative cloud top downward into the interior of the cloud. Reducing observation errors by 50% decreases posterior variances but results in more widely distributed modes. This may be an indication of why errors have had to be inflated for optimal estimation type retrievals to converge. When these two effects are combined, the modes of the IWP and effective radius remain similar to those under standard error conditions but with lower error values. Droxtals and solid columns become more likely relative to aggregates when settling is modeled, a result that would not be evident with many other retrieval methods.

[52] Application of the MCMC algorithm to a portion of a MODIS scene suggests that aggregates consistently provide the best fit to observed brightness temperatures for thin clouds while a combination of aggregates, solid columns, and droxtals made up a nearly equal percentage of the sample for thick clouds. Characteristics of the PDF changed with changing cloud ice content; thin clouds exhibit more highly skewed PDFs with lower variance, i.e., larger information content. PDFs in thicker cloud regions exhibited less skewness, but larger variance, i.e., lower information content. Normalized retrieval error is found to decrease with increasing IWP for retrieved IWP values less than about  $150 \text{ g/m}^2$ , then increase again with increasing IWP, while the fractional error in effective radius decreases monoton-

ically with increasing values of retrieved effective radius. Plots of information content versus retrieved IWP and effective radius indicate that, when ice crystal habit is allowed to vary, there is very little effective radius information in observations of thicker clouds, while the measurements continue to provide limited information about IWP up to optical depths of approximately 3.

[53] The results presented in this paper have implications for both ice cloud property retrievals and data assimilation in the presence of clouds. In either case, it would be convenient to assume a single form for the PDF so that uncertainty can be assigned to observations and to the results of the forward model calculation. Comparisons of pixels from relatively thin, moderate, and thick clouds indicate that the shape of the PDF changes with changes in cloud morphology, and it is evident that uncertainty in the crystal shape dominates retrieval errors for this case. Recall that the cloud of interest was selected because warm-frontal stratiform clouds are assumed to be relatively uniform and well-understood. For other types of ice clouds, in particular those associated with deep convection, ice crystal shapes are not as easily characterized. The problem is further complicated by the possibility of multiple cloud layers, for which the infrared brightness temperature may contain information from more than one cloud layer.

[54] Our results indicate that there is complexity in retrievals of ice cloud properties that needs further investigation, and questions that remain to be answered. First among these is the question of how to determine the appropriate ice crystal shape. Specifically, in the case of a cloud that is expected to contain multiple crystal shapes, which one is most appropriate? It would be reasonable to select the shape that is typically found near the cloud top, but this raises questions of the definition of the cloud top itself. In the case of thick clouds, the radiative cloud top is close to the physical top of the cloud, which, in and of itself, is difficult to define, however, the radiative cloud top is more difficult to define in the case of thin clouds or multiple cloud layers.

[55] **Acknowledgments.** Tom Greenwald provided lookup tables with ice crystal properties, and gave valuable assistance with the implementation of the SOI model. Steve Cooper computed the estimates of observation uncertainty used in this study, and provided feedback on the retrieval results. The comments of three anonymous reviewers were helpful in refining the content and flow of the manuscript. This work was funded under NASA NMP contracts NAS5-99237, NNG06GB416, and NNG06GC10G.

## References

- Ackerman, S. A., W. L. Smith, J. D. Spinhirne, and H. E. Revercomb (1990), The 27–28 October 1986 FIRE IFO cirrus case study: Spectral properties of cirrus clouds in the 8–12  $\mu\text{m}$  window, *Mon. Weather Rev.*, *118*, 2377–2388.
- Baum, B. A., A. J. Heymsfield, P. Yang, and S. T. Bedka (2005), Bulk scattering properties for the remote sensing of ice clouds: Part I: microphysical data and models, *J. Appl. Meteorol.*, *44*, 1885–1895.
- Cohn, S. E. (1997), An introduction to estimation theory, *J. Meteorol. Soc. Jpn.*, *75*, 257–288.
- Cooper, S. J., T. S. L'Ecuyer, and G. L. Stephens (2003), The impact of explicit cloud boundary information on ice cloud microphysical property retrievals from infrared radiances, *J. Geophys. Res.*, *108*(D3), 4107, doi:10.1029/2002JD002611.
- Cooper, S., T. L'Ecuyer, P. Gabriel, K. A. J. Baran, and G. Stephens (2006), Objective assessment of the information content of visible and infrared radiance measurements for cloud microphysical property retrievals over the global oceans. Part 2: Ice clouds, *J. Appl. Meteorol.*, *45*, 42–62.
- Cooper, S. J., T. S. L'Ecuyer, P. Gabriel, A. J. Baran, and G. L. Stephens (2007), Performance assessment of a five-channel estimation-based ice cloud retrieval scheme for use over the global oceans, *J. Geophys. Res.*, *112*, D04207, doi:10.1029/2006JD007122.
- Fletcher, S. J., and M. Zupanski (2006), A data assimilation method for log-normally distributed observational errors, *Q. J. R. Meteorol. Soc.*, *132*, 2505–2519.
- Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin (2004), *Bayesian Data Analysis*, 2nd ed., Chapman and Hall, New York.
- Haario, H., E. Saksman, and J. Tamminen (1999), Adaptive proposal distribution for random walk Metropolis algorithm, *Comput. Stat.*, *14*, 375–395.
- Heidinger, A., C. W. O'Dell, T. Greenwald, and R. Bennartz (2006), The successive-order-of-interaction radiative transfer model. Part I: Model development, *J. Appl. Meteorol. Climatol.*, *45*, 1388–1402.
- Heymsfield, A. J., and J. Jaquinta (2000), Cirrus crystal terminal velocities, *J. Atmos. Sci.*, *57*, 916–938.
- Hong, G., P. Yang, H.-L. Huang, B. A. Baum, Y. Hu, and S. Platnick (2007), The sensitivity of ice cloud optical and microphysical passive satellite retrievals to cloud geometrical thickness, *IEEE Trans. Geosci. Remote Sens.*, *45*, 1315–1323.
- Inoue, T. (1985), On the temperature and effective emissivity determination of semi-transparent cirrus clouds by bispectral measurements in the 10 mm window region, *J. Meteorol. Soc. Jpn.*, *63*, 88–99.
- Jazwinski, H. (1970), *Stochastic Processes and Filtering Theory*, *Math. Sci. Eng.*, vol. 64, 376 pp., Academic Press, San Diego, Calif.
- Kleespies, J., P. van Delst, L. M. McMillin, and J. Derber (2004), Atmospheric transmittance of an absorbing gas. 6. OPTRAN status report and introduction to the NESDIS/NCEP Community Radiative Transfer Model, *Appl. Opt.*, *43*, 3103–3109.
- L'Ecuyer, T. S., P. M. Gabriel, K. Leesman, S. J. Cooper, and G. L. Stephens (2006), Objective assessment of the information content of visible and infrared radiance measurements for cloud microphysical property retrievals over the global oceans. Part I: Liquid clouds, *J. Appl. Meteorol.*, *45*, 20–41.
- Miller, S. D., G. L. Stephens, C. K. Drummond, A. K. Heidinger, and P. T. Partain (2000), A multisensor diagnostic satellite cloud property retrieval scheme, *J. Geophys. Res.*, *105*, 19,955–19,971.
- Mosegaard, K., and A. Tarantola (1995), Monte Carlo sampling of solutions to inverse problems, *J. Geophys. Res.*, *100*, 12,431–12,447.
- O'Dell, C. W., A. K. Heidinger, T. Greenwald, P. Bauer, and R. Bennartz (2006), The successive-order-of-interaction radiative transfer model. Part II: Model performance and applications, *J. Appl. Meteor. Climatol.*, *45*, 1403–1413.
- Posselt, D. J., G. L. Stephens, and M. Miller (2008), CloudSat: Adding a new dimension to a classical view of extratropical cyclones, *Bull. Am. Meteorol. Soc.*, *89*, 599–609.
- Prabhakara, C., R. S. Fraser, G. Dalu, M. L. C. Wu, and R. J. Curran (1988), Thin cirrus clouds: Seasonal distribution over oceans deduced from Nimbus-4 IRIS, *J. Appl. Meteorol.*, *27*, 379–399.
- Rodgers, C. D. (2000), *Inverse Methods for Atmospheric Sounding, Theory and Practice*, World Sci., Singapore.
- Shannon, C. and W. Weaver (1949), *The Mathematical Theory of Communication*, 144 pp. Univ. of Ill. Press, Champaign.
- Tamminen, J. (2004), Validation of nonlinear inverse algorithms with Markov chain Monte Carlo method, *J. Geophys. Res.*, *109*, D19303, doi:10.1029/2004JD004927.
- Tamminen, J., and E. Kyrola (2001), Bayesian solution for nonlinear and non-Gaussian inverse problems by Markov chain Monte Carlo method, *J. Geophys. Res.*, *106*, 14,377–14,390.
- Tarantola, A. (2005), *Inverse Problem Theory and Methods for Model Parameter Estimation*, Soc. of Ind. and Appl. Mech., Philadelphia, Pa.
- Vukicevic, T., and D. J. Posselt (2008), Analysis of the impact of model nonlinearities in inverse problem solving, *J. Atmos. Sci.*, *65*, 2803–2823.
- Wand, M. P., and M. C. Jones (1995), *Kernel Smoothing*, Chapman and Hall, London.
- Weisz, E., J. Li, W. P. Menzel, A. K. Heidinger, B. H. Khan, and C.-Y. Liu (2007), Comparison of AIRS, MODIS, CloudSat, and CALIPSO cloud top height retrievals, *Geophys. Res. Lett.*, *34*, L17811, doi:10.1029/2007GL030676.
- Yang, P., and K. N. Liou (1996a), Finite-difference time domain method for light scattering by small ice crystals in three-dimensional space, *J. Opt. Soc. Am. A Opt. Image Sci.*, *13*, 2072–2085.
- Yang, P., and K. N. Liou (1996b), Geometric-optics-integral-equation method for light scattering by nonspherical ice crystals, *Appl. Opt.*, *35*, 6568–6584.
- Yang, P., H. Wei, H.-L. Huang, B. A. Baum, Y. X. Hu, G. W. Kattawar, M. I. Mishchenko, and Q. Fu (2005), Scattering and absorption property database for nonspherical ice particles in the near- through far-infrared spectral region, *Appl. Opt.*, *44*, 5512–5523.

T. S. L'Ecuyer and G. L. Stephens, Department of Atmospheric Science, Colorado State University, 1371 Campus Delivery, Fort Collins, CO 80523, USA.

D. J. Posselt, Department of Atmospheric, Oceanic, and Space Sciences, University of Michigan, Space Research Building 2517D, 2455 Hayward Street, Ann Arbor, MI 48109-2143, USA. (dposselt@umich.edu)