Tropical Composition, Cloud and Climate Coupling Experiment Validation for Cirrus Cloud Profiling Retrieval Using Cloudsat Radar and CALIPSO lidar

Min Deng  
*University of Wyoming, mdeng2@uwyo.edu*

G.G. Mace  
*University of Utah*

Z. Wang  
*University of Wyoming, zwang@uwyo.edu*

H. Okamoto  
*Kyushu University, Fukuoka, Japan*

Follow this and additional works at: [http://repository.uwyo.edu/atmospheric_facpub](http://repository.uwyo.edu/atmospheric_facpub)

**Publication Information**

Tropical Composition, Cloud and Climate Coupling Experiment validation for cirrus cloud profiling retrieval using CloudSat radar and CALIPSO lidar

Min Deng,1 Gerald G. Mace,2 Zhien Wang,1 and Hajime Okamoto3

Received 31 August 2009; revised 9 March 2010; accepted 2 April 2010; published 9 September 2010.

A profiling retrieval algorithm for ice cloud properties, such as effective radius \( r_e \), ice water content (IWC), and an extinction coefficient, has been developed to use combined CloudSat radar reflectivity factor \( Z_e \) and CALIPSO attenuated backscattering coefficient measurements based on an optimal estimation framework. Developed as an operational standard data product for the CloudSat project, the algorithm can treat a wide range of ice cloud situations from optically tenuous cirrus in the upper troposphere to geometrically and optically thick anvil clouds. It is designed to consider the attenuation of thick clouds in the radar and lidar forward model equations and multiple scattering in the lidar data. An optimal estimation approach allows for inversion of the forward model equations so that the uncertainty due to the assumptions can be evaluated. A sensitivity study shows that lidar multiple scattering has to be accounted for carefully. As for all ice cloud retrieval algorithms, assumptions regarding particle habits and size distribution shapes are critical to the accuracy of the results. The deviation in simulated \( Z_e \) among different size distribution assumptions is smaller than among different habit assumptions, which indicates that the uncertainty due to particle habits is larger than the size distribution assumption. Those uncertainties are included in the forward model error covariance matrix to analyze the retrieval error. The algorithm is applied to CloudSat-CALIPSO data as well as lidar and radar data collected by the ER-2 during the Tropical Composition, Cloud and Climate Coupling Experiment mission on 22 July 2007. The retrieved \( r_e \), IWC, and extinction are shown to compare favorably with coincident in situ measurements collected by instruments on the NASA DC-8. This algorithm is expected to be complementary to the set of standard data products that is already being produced by the CloudSat project.


1. Introduction

The NASA Tropical Composition, Cloud and Climate Coupling Experiment (TC4) field campaign took place during summer 2007 in the tropical Eastern Pacific [Toon et al., 2010]. One of the objectives of TC4 was to validate A-Train satellite observations and retrieval products. To accomplish this goal, three research aircraft were used. The NASA ER-2 carried primarily a remote sensing payload [McGill et al., 2004; Li et al., 2004] that is similar in many ways to the NASA A-Train satellite constellation [Stephens et al., 2008]. The NASA DC-8 and WB57 were used primarily as in situ platforms. The payloads of the research aircraft are described in detail by Toon et al. [2010].

Among the A-Train satellites, CALIPSO and CloudSat are designed to characterize the vertical profiles of aerosol and cloud properties. The lidar on CALIPSO (CALIOP [Winker et al., 2009]) provides backscatter profiles at 532 nm and 1064 nm through aerosol layers, subvisible cirrus and thin cirrus as well as a detailed mapping of the tops of optically thick clouds. The 94 GHz cloud profiling radar (CPR) on CloudSat [Stephens et al., 2008] measures reflectivity in cloud and provides the vertical structure of optically thick clouds. When taken together, the measurements provide an unprecedented set of coincident synergistic observations [Wang et al., 2008; Sassen et al., 2009; Mace et al., 2009] that allows for implementation of a new generation of global algorithms for retrieving cloud and
aerosol properties globally [i.e., Austin et al., 2009; Josset et al., 2008].

[4] Our goal is to develop and implement a new operational algorithm that exploits the synergy of the CPR and CALIOP to describe the microphysical properties of ice clouds. Several radar-lidar algorithms have been developed for ground-based data. A common way to invert the single wavelength elastic lidar equation is to assume a relationship between extinction and backscattering coefficient (i.e., lidar ratio) [Comstock and Sassen, 2001; Fernald, 1984; Klett, 1981]. Donovan et al. [2001] did not fix a value for the lidar ratio before the algorithm was run; rather their algorithm retrieved the value of this ratio such that it led to the combination of radar reflectivity and lidar extinction predicting the most plausible profile of particle size at the far end of the cloud. Wang and Sassen [2002] developed an ice cloud retrieval algorithm using Raman lidar retrieved extinction and radar reflectivity. Okamoto et al. [2003] (hereafter referred to as O03) developed a set of forward model equations that were used to retrieve effective size ($r_e$) and ice water content (IWC) using collocated radar reflectivity and lidar backscattering at 532 nm. Since it uses look-up tables to represent extinction and backscattering properties of ice clouds at the radar and lidar wavelengths based on detailed calculations, it is not necessary to introduce an assumed lidar ratio. The other unique feature of the O03 algorithm is the attenuation correction in both lidar and radar signals. This feature is also used by Tinel et al. [2005]. A common characteristic of these algorithms, however, is that they are designed for the region in the vertical column where concurrent backscatter from both the radar and lidar are received. Owing to the sensitivity of millimeter radars and the likelihood of attenuation with optical lidars, this overlap region typically does not extend through the depth of ice cloud layer.

[5] Currently, operational products from CloudSat include ice cloud microphysics that are derived using data from the CPR and temperature only [Austin et al., 2009], known as the radar-only, or RO, product). The RO algorithm depends heavily on the accuracy of IWC-radar reflectivity (Zr) relations from the work of Liu and Illingworth [2000]. Because at least two independent parameters are needed to describe a cloud particle size distribution (PSD), algorithms using one measured parameter are ill conditioned and have substantial uncertainty [Sassen et al., 2002].

[6] Exploiting the synergy of the A-Train is proving to be a powerful approach yet difficult to implement. An early innovation along this line is described by Delanoë and Hogan [2008], who use the combination of ground-based or spaceborne radar, lidar and infrared radiometer to retrieve profiles of visible extinction coefficient, IWC and $r_e$ in ice clouds. The retrieval is possible in regions of cloud detected by both radar-lidar. But for regions detected just by the radar, the presence of an a priori constraint on number concentration means that the values of IWC retrieved by the algorithm will tend seamlessly toward values similar to those that would be obtained from an empirical Zr-IWC relationship, but without such an empirical relationship being explicitly coded.

[7] Our goal is to build incrementally on the work of Delanoë and Hogan [2008] to develop a fast forward model inversion of CloudSat radar reflectivity and CALIPSO lidar attenuated backscatter data. Use of radiance data is deferred in this initial implementation. Here we describe the theoretical basis of the profiling algorithm to retrieve both thick and thin ice cloud $r_e$, IWC, and extinction using CloudSat radar and CALIPSO lidar, which will be implemented as a CloudSat operational standard data product (known as 2C-ICE) that will be publicly available. In section 2, we first introduce the algorithm input data set and data preprocessing. Then we describe the theoretical basis for the algorithm and our approach to inverting a set of simple forward model equations. The algorithm error analysis follows. In section 3, algorithm tests are also done with simulated data using in situ measured PSD. In section 4, the algorithm is first applied to CloudSat-CALIPSO data and ER-2 lidar-radar measurements during the TC4 mission. The sensitivity tests of the solutions for CloudSat-CALIPSO case to the multiple scattering treatment and cloud microphysical parameter assumptions follow. Finally the retrieval results for the ER-2 case and the CloudSat-CALIPSO case are validated using DC-8 in situ measurements. A detailed discussion and summary is presented in section 5.

2. Algorithm Description

2.1. CloudSat CPR and CALIPSO Data Set and Preprocessing

[8] The CPR measures the backscattered power as a function of range every 240 m vertically with a 1.4 km cross track $\times$ 1.8 km along track footprint. It has a sensitivity of approximately $-30$ dBZ$_e$. The CloudSat algorithm team has developed Science algorithms and software for the CloudSat data. There are two types of CloudSat data products, standard data products and auxiliary data products. Because it provides the atmospheric state variables such as temperature, pressure along the CPR track, the ECMWF-AUX (European Centre for Medium Range Weather Forecasts forecast fields mapped to CloudSat profiles) data are used for the 2C-ICE algorithm. There are two standard data products required for ice cloud identification in our algorithm. The 2B-GEOPROF lidar (geographical profile from combined CPR and CALIPSO lidar [Mace et al., 2009]) data is produced by combining the 2B-GEOPROF data product [Mace et al., 2007] and collocated CALIPSO lidar data to determine the vertical locations of cloud layers. The 2B-CLDCLASS lidar (cloud classification from combined CPR and CALIPSO lidar measurements; refer to CloudSat standard data products) is under development to identify cloud phases and eight basic types of clouds, so that downstream retrieval algorithms or assumptions can be applied to the conditions for which they are considered valid. The 2B-GEOPROF is also required to provide the CPR cloud mask. The data input for 2C-ICE algorithm is shown in Figure 1.

[9] The CALIPSO lidar measures parallel and perpendicular backscattered laser energy at 532 nm and total backscattering at 1064 nm at altitude-dependent vertical resolutions and footprint (30 m vertically with about 0.075 km cross track $\times$ 1.0 km along track footprint above 8.2 km and 75 m vertically with about 0.075 km cross track $\times$ 0.3 km along track footprint below 8.2 km). As recorded in the level 1B files, the lidar data are calibrated against molecular backscatter profiles [Powell et al., 2009].
Figure 1. Schematic table of data input for standard 2C-ICE algorithm. The 1B–cloud profiling radar (CPR) and 1B-CALIPSO provides the radar reflectivity and attenuated backscattering at 532 nm, respectively. The 2B-GEOPROF lidar and 2B-CLDCLASS lidar provides the location of cloud layers and phase so that the ice cloud is identified, respectively. The 2B-GEOPROF lidar provides the CPR cloud mask. ECMWF-AUX provides the atmospheric state. The 1B-CALIPSO will be collocated and included in the new version of the 2B-GEOPROF lidar, and by then, we will no longer need the 1B-CALIPSO data.

[10] As the CloudSat and CALIPSO lidar products are on different vertical and horizontal resolutions, the first pre-processing step is to collocate CALIPSO 532 nm total backscattering profiles within the CloudSat CPR footprint and average the data horizontally and vertically to combine with the $Ze$ profile. According to the 2B-GEOPROF lidar, the lidar observations that could potentially overlap the spatial region enclosing the radar observational domain out to the CPR 2σ boundary will be considered as potentially contributing to the spatial description of the overlap region. Below 8.2 km, as many as 9–10 separate lidar profiles will be included while above 8.2 km, three to four profiles will potentially contribute to the hydrometeor description. At a given level above which the lidar beam becomes fully attenuated, a lidar cloud fraction within a radar footprint quantifies the partial filling of the radar volume by hydrometeors and is one of the output quantities of the 2B-GEOPROF lidar. Cirrus layers are then identified by cloud top and base temperatures test ($T_{\text{top}} < -40^\circ\text{C}, T_{\text{base}} < -10^\circ\text{C}$) based on the results of earlier studies using long-term ground-based data [Sassen and Campbell, 2001; Mace et al., 2007].

[11] Because the CPR and CALIOP lidar probe the atmosphere using such vastly separated frequencies, measurements provided by the two instruments contain information about different parts of the cloud particle size distribution in any given volume. The $Ze$ is more sensitive to the largest particles in the PSD, while the backscattering of the laser energy from the lidar is more dependent on the concentration of smaller particles which dominate the physical cross-sectional area within the volume. The characteristics of the instruments convolved on the physical properties of clouds in the upper troposphere require us to consider that three distinct lidar-radar regions could exist in any ice cloud layer. Cirrus clouds typically form from the top down where nucleation of small ice crystals often occurs in saturated updrafts near cloud top. Particle growth through vapor deposition and aggregation occurs in the middle regions of the ice layer as the particle population sediment through ice saturated layers while sublimation of crystal populations is found near the layer bottom [Heymsfield, 1975]. This schematic model is often seen in CloudSat and CALIPSO measurements where the lidar responds to the tenuous regions of small particles near cloud top that are below the detection threshold of the CPR (this portion of the layer is referred to as the lidar-only region). Often a middle region in the vertical profiles is found where backscatter from both lidar and radar are recorded (hereafter referred to as the lidar-radar overlap region). There is also often a region below an accumulated visible optical depth of about 3 where the lidar signal has been fully attenuated but the radar still provides reliable data (this region will be referred to as radar only). The radar signal generally is not fully attenuated in optically thick ice clouds; however, some attenuation correction for CPR measurements is needed in very optically dense anvils. A global operational algorithm to treat ice clouds must allow for the occurrence of any or all of these distinct regions in its formulation, recognize their presence in a given vertical profile, and provide specific information about the region. For instance, we do not know measured $Ze$ in the lidar-only region, but we do know the upper limit of $Ze$ in the lidar-only region (less than the minimum detectable signal of the CPR). For the radar-only region, although we do not measure attenuated backscatter, we do know that optical path above the region and the vertical structure of $Ze$. Combining this information offers a way to estimate extinction in the radar-only region [Matrosov et al., 2003]. The identification of three lidar-radar zones allows us to do different subsequent data handling in the different zones such as measurement error assignment.

[12] In order to build up a uniform framework to cover the three regions, we consider the radar and lidar signal profiles for the full cloud layer. For the lidar-only zone typically at the tops of cirrus layers, the radar signal is below the minimum detectable signal. We parameterize the $Ze$ profile in this region using the statistics of Atmospheric Radiation Measurement (ARM) millimeter cloud radar (MMCR) measurements [Ackerman and Stokes, 2003]. The ARM MMCRs operate at a frequency of 34.86 GHz (Ka band) at the southern Great Plains (SGP) and tropical western Pacific (TWP) sites [Clothiaux et al., 1995] with a stated calibration uncertainty of 1 dB. The MMCRs are about 1 order of magnitude more sensitive than the CloudSat CPR. For thin cirrus clouds we expect that Rayleigh scattering is obeyed for both the 35 GHz frequency of the MMCR and the 94 GHz frequency of the CPR. We assume that if the CPR had the same sensitivity as the MMCR then the $Ze$ vertical distribution would be similar. Shown in Figure 2 is the MMCR radar $Ze$ distribution normalized at each height level from data collected at the TWP and SGP sites during 1999–2007. In general, the mean $Ze$ and the $Ze$ of the first percentile decrease with height until about 14.5 km, corresponding to about $-40 \text{dBZ}_{Ze}$. Above that the mean $Ze$ and minimum $Ze$ increase is probably due to missed thin cirrus clouds by the MMCR consistent with the finding of Comstock et al. [2002]. It seems reasonable to assume that with enough sensitivity, the minimum $Ze$ observed by the MMCR will monotonically decrease with height above 14.5 km (as extrapolated with the dashed line in Figure 2). For our purpose, we assume that the distribution of the $Ze$
below the minimum detectable signal would be similar to what is measured by the MMCR. Therefore, we propose a simple parameterization of the CPR Ze for the lidar-only region as

$$Ze(R) = Ze_{\text{min}}(R) + \frac{Ze_{\text{max}}(R) - Ze_{\text{min}}(R)}{\beta'_{\text{max}}(R) - \beta'_{\text{min}}(R)} \left( \beta'(R) - \beta'_{\text{min}}(R) \right),$$

where $Ze_{\text{max}}$ is $-31 \text{ dBZ}_e$ for CloudSat CPR and $Ze_{\text{min}}$ is the fitted lower $Ze$ boundary above which 99% data exist at a certain height ($R$) found from the ARM radar data. Variables $\beta'_{\text{max}}$ and $\beta'_{\text{min}}$ are the maximum and minimum of the lidar attenuated backscattering found in the lidar-only region at certain height ($R$) over a granule. Variable $\beta'(R)$ is the observed lidar attenuated backscatter. In practice, while this parameterization provides some information, the uncertainty is very large. This uncertainty is accounted for in the inversion framework.

[13] On the other hand, the lidar signal in the radar-only region at the bottom of thick cloud layers can be attenuated to below the lidar noise level. On the basis of the Ze profiles, we can estimate the lidar extinction [Matrosov et al., 2003] and hence the attenuated backscattering (ABS) assuming a lidar ratio. However, the estimated ABS might be easily over or under estimated given the uncertainty in the relationship between Ze and extinction. In the current practice, the estimated ABS is calibrated to the measured ABS above the radar-only region, because from lidar-radar overlapped region to radar-only region, the tendency of ABS profiles should be smooth. Of course the uncertainty in the parameterized data is much larger than where actual data exists resulting in about at least a factor 2 larger uncertainty in the solutions. The error variance and performance of the estimated ABS are discussed again in section 2.7.

2.2. Formulation of Variational Method

[14] The algorithm is built upon an optimal estimation framework [Rodgers, 2000; Livesey et al., 2003; Zhang and Mace, 2006; Deng and Mace, 2006; Austin et al., 2009; Delanoë and Hogan, 2008]. In this framework we consider a state vector, $x$ (i.e., $r_e$ and IWC), that describes the properties of the vertical profile of ice cloud microphysics that result in a set of $\beta$ and Ze measurements expressed as vector $y$:

$$x = \begin{bmatrix} r_{e1} \\ r_{e2} \\ r_{e3} \\ \vdots \\ r_{en} \\ \text{IWC}_1 \\ \text{IWC}_2 \\ \text{IWC}_3 \\ \vdots \\ \text{IWC}_n \end{bmatrix},$$

$$y = \begin{bmatrix} \ln \beta_1 \\ \ln \beta_2 \\ \ln \beta_3 \\ \vdots \\ \ln \beta_n \\ \ln Ze_1 \\ \ln Ze_2 \\ \ln Ze_3 \\ \vdots \\ \ln Ze_n \end{bmatrix},$$

where $n$ is the CloudSat range bins of a cloudy layer. The relationships linking the atmospheric state and the measurements are the forward model, i.e., $y = F(x, b) + e$, where $b$ are parameters used in the forward model. The inverse problem can be approximated using Bayes theorem where we maximize the a posteriori likelihood of $x$ given $y$. The solution is typically found by iteration by initializing a state vector $x$ with an a priori ($x_0$) estimation from extensive in situ measurements or empirical relations or algorithms in the literature. Using Gauss-Newton iteration, an expression for the state vector can be expressed as

$$x_{i+1} = x_i + \left( S_a^{-1} + K_i^T S_e^{-1} K_i \right)^{-1} K_i^T S_e^{-1}(y - F(x_i)) - S_e^{-1}(x_i - x_0),$$

where $K$ is the Jacobian matrix containing the derivatives of each observation with respect to each state vector. $S_a$ and $S_e$ are the error covariance matrices chosen to limit the amount of bias of a priori and measurement from the true state vector and ideal measurements, respectively. The error covariance matrices are assumed diagonal in the present implementation.

[15] Given the state vector, a forward model can be applied to predict what measurements the instruments would observe. By comparing these predictions with the measurements actually observed, and by making use of additional information provided by the forward model calculation (namely, derivatives of measurement with respect to the state vector, i.e., the Jacobian matrix), the retrieval algorithm iteratively computes a better estimate of the state vector, i.e., one for which the predicted measurements will be closer to observed. The retrieval algorithm adjusts the state vector until appropriate convergence has been achieved. An algorithm flowchart is shown in Figure 3. For a detailed explanation of each term in equation (2), please refer to the work of Rodgers [2000]. In the following, we describe the measurements, state vector, forward model, a priori estimation, as well as the algorithm error analysis in more detail.
2.3. Lidar Forward Model

The lidar equation for attenuated backscattering coefficient, $b'$, can be expressed generally as

$$
\frac{1}{C_1} \frac{R_0}{C_1} \beta(R) = \frac{1}{C_1} \frac{C_2}{C_1} b_c(R) + \frac{1}{C_1} \frac{C_2}{C_1} b_m(R) \exp \left[ \frac{1}{C_1} \frac{C_2}{C_3} Z \right],
$$

where $R$ is the range and $b_c(R), b_m(R)$, and $\sigma_{sl}(R)$ are the particle and molecular lidar backscattering coefficients and total extinction coefficient at the lidar wavelength, respectively. Variable $\eta$ is a multiple scattering correction factor [Platt et al., 1998]. The molecular lidar backscattering coefficient can be calculated according to the thermodynamic state of the atmosphere provided by the ECMWF-AUX data set.

The CALIOP lidar has a large footprint compared with ground-based or airborne lidars, which allow a greater contribution of the multiple-scattered light to the total return signal. The multiple scattering effects can influence the apparent extinction or transmittance of the medium. Therefore, multiple scattering effects must be accounted for in the algorithm to retrieve accurate cloud extinction and optical depth as well as IWC and $r_e$.

In principle, the multiple scattering contributions to the lidar signal [Battaglia et al., 2007; Hogan, 2006; Eloranta, 1998] usually is derived from either a Monte Carlo method or is computed from analytical approximation. However, these approaches are computationally intensive and the results vary among different approaches depending on the truncation error and lidar resolutions [Hogan, 2008]. On the other hand, statistical techniques can be used to increase computational efficiency but involve approximations. For example, the multiple scattering effects can be parameterized using a range-dependent multiple scattering function [Platt et al., 1998; Comstock and Sassen, 2001; Okamoto et al., 2003; Winker, 2003]. For algorithms developed for CALIPO, Winker [2003] did an extensive analysis of the multiple scattering impact on cirrus and aerosol extinction retrievals using a Monte Carlo method. For cirrus, he found that a single $\eta$ can be used, which ranges from 0.6 to 0.8 depending on the penetration level of clouds. In the current algorithm, we adopt a single $\eta$ method. Although there is error associated with this simple approach, it makes the lidar forward model and hence the entire algorithm faster, which is necessary for an operational algorithm that is to be applied to a large amount of global data. In the next version of the 2C-ICE algorithm we plan to incorporate more sophisticated treatments of multiple scattering [Hogan, 2008]. The sensitivity of the algorithm solution to this parameter is investigated in section 3.

For a layer with finite thickness, the attenuation term is modified, following O03, as

$$
\beta(R_i) = [\beta_c(R_i) + \beta_m(R_i)] \exp \left\{ -2 \int_0^{r_i} \eta(r) \sigma_{sl}(r) \, dr \right\} \frac{\exp \left\{ -2 \eta(R) \Delta R \right\} - 1}{-2 \eta(R) \Delta R},
$$

where $R_i$ is the range of the layer, $\eta$ is the multiple scattering correction factor, $\sigma_{sl}(R)$ is the single scattering extinction coefficient, and $\Delta R$ is the thickness of the layer.

Figure 3. Algorithm flowchart to retrieve ice cloud $r_e$ and ice water content (IWC) from CloudSat CPR and CALIPSO lidar.
where \( R_i \) denotes the altitude of the center of the layer \( i \). \( R_{i-1/2} \) is the top boundary of the cloud layer \( i \). Variable \( \Delta R \) is the vertical resolution. The first term on the right is the particle and molecular backscatter before attenuation. The second term accounts for two-way attenuation due to gasses and particles and accounted for multiple scattering effect with a \( \eta \). The third term accounts for attenuation and multiple scattering within the range resolution volume, a potential issue for the coarse vertical resolutions of CloudSat and CALIPSO. The particle backscatter and extinction are calculated for certain PSD and certain particle habits as a part of the look-up table discussed in section 2.5.

2.4. Radar Forward Model

[20] For the CloudSat CPR signal, we apply a similar analysis as for the lidar following O03:

\[
Z_e'(R_i) = Z_{true}(R_i) \exp \left\{ -2 \int_{R_{i-1/2}}^{R_i} \sigma_{ra}(r) \, dr \right\} \left\{ \exp \left[ -2 \sigma_{ra}(R_i) \Delta R - 1 \right] / -2 \sigma_{ra}(R_i) \Delta R \right\}.
\]

(5)

\( Z_e' \) and \( Z_{true} \) are the observed \( Z_e \) and true \( Z_e \) (the radar reflectivity that would be observed without particle attenuation), respectively. Variable \( \sigma_{ra} \) is the extinction coefficient at radar wavelength. \( Z_{true} \) and \( \sigma_{ra} \) are calculated for certain PSD and certain particle habits as a part of the look-up table discussed in section 2.5.\[21\] The last part of the forward model includes the calculation of the Jacobian that contains the derivative of \( X_r \) related to the particle size distribution function as the sectional area and particle size. The backscattering and extinction properties of nonspherical ice crystals at 532 nm are calculated for certain PSD and certain particle habits as a part of the look-up table discussed in section 2.5.\[22\] Equations (4) and (5) represent the forward model \( F(X) \). Because the solution algorithm assumes a nonlinear relationship between \( Y \) and \( X \), we express the forward model equations (4) and (5) in terms of the natural logarithm of \( Z_e \) and \( \beta_e \). The use of logarithms in \( Y \) also results in faster convergence to the correct solution as indicated by Delanoë and Hogan [2008].

2.5. Look-Up Table and State Vector

[23] The look-up table contains the bulk microphysical relations between IWC, \( r_e \), \( \beta_e \), \( \alpha_{li} \), \( \sigma_{li} \), and \( Z_{true} \) of ice particles in the CPR sample volume. It depends on a microphysical model, which describes the functional shape of the PSD, and relationships between particle mass, cross-sectional area and particle size. The backscattering and extinction coefficients at lidar and radar wavelengths are related to the particle size distribution function as the following:

\[
\beta_e(R) = \frac{D_{max}}{4\pi} \int_{D_{min}}^{D_{max}} N(D) C_{ext,i}(D, R) \, dD,
\]

(6)

\[
\sigma_{ra}(R) = \int_{D_{min}}^{D_{max}} N(D) C_{ext,ra}(D, R) \, dD,
\]

(7)

\[
\sigma_{ra}(R) = \int_{D_{min}}^{D_{max}} N(D) C_{ext,ra}(D, R) \, dD,
\]

(8)

\[
Z_{true}(R) = \frac{A^4}{\pi^2 |K|^2} \int_{D_{min}}^{D_{max}} N(D) C_{li}(D, R) \, dD,
\]

(9)

where \( C_{li}, C_{ra}, C_{ext,li} \) and \( C_{ext,ra} \) are the backscattering and extinction cross sections at the lidar and radar wavelengths, respectively. \( K \) is estimated from the complex refractive index of water. \( N(D) \) is the PSD, which is commonly represented by either a modified gamma, lognormal, or exponential distributions with two parameters [McFarquhar and Heymsfield, 1997; Mace et al., 2002; Okamoto et al., 2003; Deng and Mace, 2006]. For the 2C-ICE algorithm, a modified gamma PSD is assumed:

\[
N(D) = N_e D^{\alpha} \left( \frac{D}{D_g} \right)^\alpha \exp \left[ -\frac{D}{D_g} \right],
\]

(10)

where \( D \) is the particle maximum length, \( N_e \) is a proportionality constant, \( D_g \) is the size where the function \( N(D) \) maximizes, and \( \alpha \) indicates the breadth of the spectrum. Instead of retrieving particle size distribution parameters, we retrieve \( r_e \) and IWC. Variable \( r_e \) is defined as

\[
r_e = \frac{3}{4} \frac{\text{Volume}}{\text{Area}} = \frac{3}{4} \frac{D_{max}}{D_{min}} \int_{D_{min}}^{D_{max}} \alpha_{li} D^{h_i} N(D) \, dD.
\]

(11)

where \( \alpha_{li} \) and \( h_i \) are the parameters in a mass and size power law relation, \( \alpha_{ra} \) and \( h_r \) are the parameters in an area and size power law relation [Brown and Francis, 1995; McFarquhar et al., 1999; Baum et al., 2005; Deng and Mace, 2006], and \( \rho_{ice} \) is the density of solid ice. The dimensional power law relation parameters are different for different crystal habits and shapes. Those can be fitted from in situ measurements of mass, extinction, and maximum length [Brown and Francis, 1995; McFarquhar et al., 1999] or from directed calculation assuming certain particle habits or shape [Yang et al., 2000]. In our algorithm we use the later. To test the effects of particle habit assumptions on retrieval accuracy, we build up several look-up tables with different particle habits (sphere, hexagonal plate, column, bullet rosette or aggregate), which are assumed to be oriented randomly. A look-up table of backscattering and extinction at 532 nm and 94 GHz for spherical particles is adapted from O03 using Mie theory, so that the lidar ratio is implicitly specified, which varies between 15 and 30 depending on the particle size [Chen et al., 2002; Eloranta et al., 2000; Immler et al., 2007; Vaughan et al., 2008; Whitman et al., 2004]. The extinction properties of nonspherical ice crystals at 532 nm
are computed by accurate light scattering calculations and parameterized by Yang et al. [2000]. The backscattering properties are calculated from extinction assuming a lidar ratio ranging from 15 to 30.

[24] For 94 GHz, the backscattering properties were reported by Aydin and Walsh [1999] and Hong [2007]. In our calculation we refer to Hong [2007], which just uses the consistent particle habits and maximum length with Yang et al. [2000]. Assuming a modified gamma PSD and using equations (6)–(11), we can calculate \( r_e \), \( \beta_c \), \( \sigma_{bb} \), \( \sigma_{rr} \), and \( Z_{true} \). According to equation (11), \( r_e \) is a function of \( D_g \). For one \( r_e \) or \( D_g \), IWC, \( \beta_c \), and \( Z_{true} \) change linearly with \( N_g \). Thus, we can use one dimensional look-up tables normalized by IWC to relate \( r_e \) to \( \beta_c \), \( \sigma_{bb} \), \( \sigma_{rr} \), and \( Z_{true} \). These look-up tables for different habits are plotted in Figures 4a–4c. First we can see that the extinction coefficients are very close for different habits for the same \( r_e \). Second, the radar extinction is more than 1 order smaller than the lidar extinction; therefore the radar attenuation term is normally neglected in thin ice cloud retrievals but is not necessarily negligible in thick cirrus anvils. Third, the radar reflectivity deviates by 1–5 dB for different habits with the same \( r_e \). Therefore a habit mixture is used as in Baum 2005 to improve the comparison of derived with modeled microphysical properties. The sensitivity of the retrieval accuracy to this assumption is examined section 4.

2.6. A Priori Estimation and Variance Error

[25] The a priori estimation of \( r_e \) and IWC and the associated error covariance matrix represent prior knowledge that can be incorporated to improve our estimation and the deviation of the element of the true state from the a priori. The prior knowledge can come from empirical relations derived from extensive in situ measurements, such as IWC and \( Z_e \) relations [Liu and Illingworth, 2000] or normalized number concentration with temperature [Delanoë and Hogan, 2008], or from other validated algorithm estimations in the literature. In the current algorithm, we use empirical relations and retrievals from published algorithms for prior knowledge. Because CloudSat and CALIPSO measurements contain enough information to constrain \( r_e \) and IWC retrievals, the retrieval results are less dependent on the a priori. Thus, the a priori covariance matrix \( S_p \), the diagonal values are simply set to as twice as the a priori estimation, which make the a priori information weighs less than the measurements. Our goal to set up a priori to speed up the retrieval (converge in less iteration).

Figure 4. (a–c) Look-up table comparison assuming gamma particle size distribution (PSD) for different habits. (d–f) Simulated IWC, IWC normalized extinction, and radar reflectivity using in situ PSD are shown assuming four nonspherical particle habits, as in Figure 4a–4c. The symbols denote quantities calculated from particle spectra measured in situ, assuming different particle habits. See the text for more details.
The extinction retrieval using the lidar return below the layer base. This penetrate the entire layer and the layer transmittance can be estimated from the covariance matrices of the measurements and the a priori combined with the Jacobian following Austin et al. [2006] from the noise scale factor, the lidar error is parameterized using equation (1), as illustrated in Figure 5. As it shows, the measurement profiles of CPR and CALIPSO lidar shows there is a lidar-only region (where the Ze is less than \(-31 \text{ dBZ}_e\)) above the lidar-radar overlapped region. Below the overlap portion of the profile is a radar-only region where the lidar signal is attenuated below the noise level. With equation (1), Ze is parameterized and plotted in Figure 5 as the dotted line. Given the geographical variation of cirrus clouds and uncertainty in equation (1), the corresponding Ze error in the lidar-only region is assigned to be \(5 \text{ dB}\). This large error assignment ensures that the parameterized radar reflectivity value has minimal contribution on the actual retrieved ice cloud properties in this portion of the layer although the actual information that the radar reflectivity is lower than the minimum detectable by the CPR is retained.

For the lidar-only region, Ze is parameterized using equation (1), as illustrated in Figure 5. It shows the measurement profiles of CPR and CALIPSO lidar shows there is a lidar-only region (where the Ze is less than \(-31 \text{ dBZ}_e\)) above the lidar-radar overlapped region. Below the overlap portion of the profile is a radar-only region where the lidar signal is attenuated below the noise level. With equation (1), Ze is parameterized and plotted in Figure 5 as the dotted line. Given the geographical variation of cirrus clouds and uncertainty in equation (1), the corresponding Ze error in the lidar-only region is assigned to be \(5 \text{ dB}\). This large error assignment ensures that the parameterized radar reflectivity value has minimal contribution on the actual retrieved ice cloud properties in this portion of the layer although the actual information that the radar reflectivity is lower than the minimum detectable by the CPR is retained.

2.7. Retrieval Error Estimation

The covariance of the state vector, \(x\), can be calculated from the covariance matrices of the measurements and the a priori combined with the Jacobian following Rogers [2000]:

\[
S = (K^T S_x^{-1} K + S_{\eta}^{-1})^{-1}.
\]

Here \(S\) represents the total measurement error which includes forward model error, random measurement error, and systematic error. If parameters used in the forward model have large uncertainties, then extra sources of error associated with these parameters must be considered. In the following, we discuss each term for CPR and CALIPSO lidar measurements, respectively.

[28] The radar forward model as described by equations (5), (8), and (9) is sensitive to assumptions regarding the PSD functional form and the particle habit. According to look-up table calculations discussed in section 2.5, for a set of IWC and \(r_c\), the radar forward model can produce different Ze (varies by 1–5 dB) for different habits. As we choose a habit mixture, an error of about \(2.5 \text{ dB}\) in Ze due to the habit assumption is assigned. The CloudSat CPR is well calibrated and validated with CALIPSO-CloudSat Validation Experiments [Tanelli et al., 2008], therefore, the systematic error is less than \(1 \text{ dB}\). The random measurement error is due to the limited sample and background noise can be computed from the number of pulses averaged and the linear signal-to-noise ratio (SNR) of CPR. This term is small compared to the forward model error.

[29] For the lidar-only region, Ze is parameterized using equation (1), as illustrated in Figure 5. As it shows, the measurement profiles of CPR and CALIPSO lidar shows there is a lidar-only region (where the Ze is less than \(-31 \text{ dBZ}_e\)) above the lidar-radar overlapped region. Below the overlap portion of the profile is a radar-only region where the lidar signal is attenuated below the noise level. With equation (1), Ze is parameterized and plotted in Figure 5 as the dotted line. Given the geographical variation of cirrus clouds and uncertainty in equation (1), the corresponding Ze error in the lidar-only region is assigned to be \(5 \text{ dB}\). This large error assignment ensures that the parameterized radar reflectivity value has minimal contribution on the actual retrieved ice cloud properties in this portion of the layer although the actual information that the radar reflectivity is lower than the minimum detectable by the CPR is retained.

[30] The CALIPSO lidar instrument error includes systematic error due to calibration uncertainties as well as random error. The lidar calibration uncertainty is reported to better than 5% using the molecular signal by the CALIPSO science team [Powell, 2009]. The random error is evaluated by Liu et al. [2006] from the noise scale factor, the lidar range distance, and mean and standard deviation of background signal power. Since we average the lidar signal to the CPR vertical resolution, the SNR is improved. For our calculations, the random error is evaluated from the standard deviation of backscattering and SNR in the averaged bins.

[31] For the lidar forward model, the principal sources of uncertainty include the assumptions regarding the microphysics (habit and PSD functional form) as well as uncertainty due to the multiple scattering and lidar ratio approximations. The lidar extinction coefficients vary little for different habits as a function of \(r_c\), since \(r_c\) is defined to account for shape sensitivity. So the microphysical assumption error is small. The perturbation caused by the \(\eta\) and lidar ratio is significant (which will be shown in section 3), then extra sources of error needed to be included in the measurement error covariance by \(S_x^f + K_b^T S_b K_b\), where \(S_b\) is the parameter error covariance matrix, and \(K_b\) is the sensitivity of the forward model to parameters.

[32] Special treatment of lidar error is needed in the radar-only region as shown in Figure 5. In that region, the measured lidar signal is not valid because it is below the noise level; therefore a large error in the natural logarithm of estimated attenuated backscattering derived from the Ze profile is applied as shown in Figure 5, so that the retrieval mainly relies on radar observations in this portion of the
layer. Like the parameterized Ze in the radar-only region, this approach allows us to maintain a consistent set of forward model equations throughout the layer regardless of the layer properties.

[33] One profile of retrieval results is illustrated in Figure 5. As expected, the top and bottom portions of the layer where only lidar and radar data are available, respectively, are much more uncertain compared to the portion of the layer where both measurements are available. The simulated radar reflectivity and attenuated backscattering are also plotted in Figure 5. These are the values that emerge from the retrieval algorithm that are quantitatively consistent with the retrieved IWC, \( r_e \), and assumed microphysics. They line up nicely with the measurements and are within a reasonable approximation of the uncertainties of the parameterized quantities within the respective regions of the cirrus layer.

3. Simulation Studies of the Algorithm Performance

[34] To test the algorithm performance we use the simulation method as used by Hogan et al. (2006), Donovan et al. (2001), and Tinel et al. (2005). First, we produce synthetic profiles of apparent lidar backscattering and radar reflectivity using the in situ measured particle size distribution and our forward model. Then we examine the retrievals from the algorithm and investigate the effect of a priori, lidar ratio, and parameterized lidar signal in the radar-only region.

[35] The algorithm dependence on the a priori is shown in Figure 6. When the a priori of \( r_e \) and IWC are far from the true solution and given large variance, the algorithm successfully find a solution close to the true solution through several iterations (the first guess is assumed to be same as the a priori). This simulation shows that the algorithm does not rely heavily on the a priori because lidar and radar measurements contain significant information to constrain IWC and \( r_e \). However, as discussed above, beginning the iteration close to the true solution speeds the convergence of the algorithm.

[36] In the algorithm, the lidar ratio is assumed to 20 for nonspherical particles. We first simulate the apparent lidar backscatter using a lidar ratio of 40, then apply the algorithm assuming a lidar ratio of 20. The results are shown in Figure 7. It shows that at the bottom of cloud layer, the effective size is a little underestimated and IWC and extinction are overestimated. However, in general, the algorithm reliably retrieves the cloud properties.

[37] There are several reasons to motive us use parameterized radar and lidar information in the lidar-radar-only region is useful. First, this approach allows the framework to be consistent throughout the profile so that we could have \( n \times 2 \) inputs and \( n \times 2 \) outputs. Second, there is radar information in the lidar-only region (Ze is less than the CloudSat detection threshold) and there is information about the optical depth of the overlying profile in the radar-only region. It is our goal to use this information without artificially forcing the solution. This is accomplished by allowing the covariance to be appropriately large in the regions where parameterized data are used. Figure 8 shows the use of the parameterized data in the radar-only region. Since lidar signal is attenuated to the background level at about 12 km, we parameterize \( \beta \) signal according to a normalized Ze-extinction relationship discussed in section 2.1. The retrieval with parameterized \( \beta \) is close to the control one (true solution). However, if we discard the parameterized data totally and rely on a Ze-IWC relationship only and converge to reflectivity only, then the resulting \( r_e \) is significantly underestimated and the IWC and extinction are significantly underestimated in the radar-only region. We would find similar sensitivities in the lidar-only region near cloud top.

If we parameterize the Ze profile from the ARM MMCR data, it provides constraints in particle size. For our current algorithm which uses radar and lidar but not radiance, the parameterized data is very useful and appropriately provides important information about the state of the profile.

4. Algorithm Application and Evaluation During the TC4

[38] In this section, we first apply the algorithm to lidar-radar observation on ER-2 and CloudSat-CALIPSO measurement during the TC4 field campaign. The sensitivity study based on a TC4 case is followed. Then in situ measurement of PSD is introduced. Finally the retrieved results for the ER-2 and CloudSat cases are validated with cloud properties measured or calculated from in situ measurement.
Data collected during TC4 provide two kinds of situations for validation of remote sensing algorithms. On several days, the DC-8 flew along the CloudSat-CALIPSO track during the overpass time allowing for direct comparison of cloud properties derived from CloudSat-CALIPSO data with in situ measurements. This allows us to evaluate the algorithm assumptions such as shape, mass area, and size relationships, Ze-IWC relations, etc. In addition to this direct comparison with the satellites, coordinated flight patterns between the remote sensing ER-2 and the in situ DC-8 and WB57 were conducted. The differences between the ER-2 measurements and A-Train measurements, such as vertical and horizontal resolution, and fields of view, can be exploited to test the treatment of these sources of uncertainty in the algorithms applied to the A-Train data. Data collected on 22 July 2007 happen to provide a segment of data collected during the A-Train overpass as well as a portion of the flight when the ER-2 and DC-8 were flying coordinated patterns.

4.1. Optically Thick Cloud Retrieval With ER-2 Radar-Lidar Measurements

Early in the flight on 22 July 2007, the DC-8 sampled an anvil outflow between about 7.6 and 11.6 km that were streaming southwestward from dissipating convective sources over the Eastern Pacific. Two legs were coordinated with the ER-2 that was flying coordinated racetracks from a flight level near 20 km. In situ observations on board the DC-8 at about 10 km show crystals as large as 2 mm and IWCs as high as 0.3 g/m³. The CPL backscattering at 532 nm and CRS reflectivity are plotted in Figure 9. The lidar signals show the anvil is optically thick enough to fully attenuate the lidar near the middle of the cloud layer, causing a large radar-only region in the lower portion of the layer. The CRS reflectivity indicates the cloud layer is about 4 km deep with pronounced vertical structure and horizontal variability probably due to the variable convective sources that were feeding the fresh anvil. Assuming \( \eta = 1.0 \) because of the narrow field of view and small footprint of CPL, the retrieved \( r_e \), IWC, extinction coefficient and optical depth from this algorithm are shown in Figure 10. The figures show that this is an optically and physically thick anvil cloud with large particle size and water content, hence large extinction coefficients. In Figure 9, the simulated lidar backscattering and reflectivity from the retrieved cloud properties in Figure 10 are also plotted. Comparing the panels in Figure 9, we can see that the observations are in reasonable agreement with the simulated observations from the forward model.

4.2. Optically Thin Cirrus Retrieval With CloudSat-CALIPSO Measurements

After sampling the anvil clouds over the eastern Pacific, the ER-2 departed for base and the DC-8 headed
northbound along the CloudSat-CALIPSO track, flying through a persistent cirrus layer over the Southwestern Caribbean Sea shown in Figure 11. CloudSat-CALIPSO observed thicker cirrus at the beginning and end of this track with physical thicknesses up to 3 km and optical thicknesses to 4. This case also contained an optically thin cirrus layer above 12 km between 13°N and 14.5°N latitude, whose top (totally missed by CloudSat radar) is relatively uniform at \( \sim 15 \) km while its base is more variable (12–14 km), with physical thicknesses ranging from 1 to 3 km but optical thickness less than 1 (Figure 12). Satellite imagery suggests that this entire region of cirrus was approximately 24 hours removed from a deep convective event over Northern Columbia. In situ Particle Images show primarily pristine particle habits composed mostly of bullet rosettes. Using a multiple scattering factor equal to 0.7 and a gamma PSD of bullet rosette particles, the retrieval results are plotted in Figure 12. Optical depth from the lidar transmittance method when the lidar signal penetrates the cirrus layer is also plotted in Figure 12, which is very consistent with the optical depth.

**Figure 9.** Retrieval example of 22 July 2007 ER-2 case during the TC4 experiment. (a, b) Measured and simulated radar reflectivity, respectively. (c, d) Measured and simulated lidar backscattering, respectively.

**Figure 10.** Retrieved cloud properties for the cloud in Figure 9.
retrieved from the algorithm herein described. The simulated lidar backscattering and reflectivity from retrieved cloud properties are also plotted in Figure 11.

4.3. Algorithm Sensitivity Study

[42] The algorithm sensitivity to assumptions is investigated using a perturbation technique on the TC4 case study observed by CloudSat-CALIPSO on 22 July 2007. Taking the results described in section 4.2 as a control, we perform the retrievals sequentially changing the look-up table or $\eta$ or $Ze_{\text{min}}$ in equation (1) one at a time to test the algorithm sensitivity to the microphysical model assumptions, multiple scattering correction, and $Ze$ estimation in the lidar-only region by comparing with the control run. The statistic results of these sensitivity studies are shown in Table 1.

4.3.1. Multiple Scattering Correction

[43] The results from $h = 0.6$ and $h = 1.0$ are compared with the control run and shown in Figure 13, respectively. IWC and $re$ follow closely to the 1:1 line, but the extinction and optical depth have systematic biases. When $h = 1.0$, which means neglecting the multiple scattering effect, the extinction and optical depth becomes smaller by $\sim 35\%$.
because the multiple scattering allows more light scattered into the lidar return signal and underestimates the attenuation, i.e., optical depth. If $h = 0.6$ (i.e., multiple scattering is overestimated), then the optical depth is overestimated by $\sim 20\%$. Hogan et al. [2006] also found that if multiple scattering is neglected then the optical depth is underestimated by around 35%. According to Winker [2003], $h$ mainly ranges from 0.6 to 0.8. Hereafter we choose $h = 0.7$ with 20% variance in the current algorithm.

### 4.3.2. Microphysical Model Sensitivity

The look-up table depends on a microphysical model, which describes the shape of the particle size distribution, relationships between particle mass, cross-sectional area and size. To test the algorithm sensitivity to these assumptions, we substitute one look-up table with the lookup table assuming a modified gamma PSD of randomly oriented hexagonal column ice particles. The results using this look-up table are plotted against the control in Figure 14. For these two microphysical models, most data are scattered closely around the 1:1 line while some retrieved $r_e$, IWC and extinction become smaller when assuming hexagonal columns. The mean percentage difference is about $-14\%$ for $r_e$.

#### Table 1. Results From the Sensitivity Analysis for Effect Radius, Ice Water Content, Extinction, and Optical Depth*

<table>
<thead>
<tr>
<th>Microphysical Model</th>
<th>$\eta = 0.6$</th>
<th>$\eta = 1.0$</th>
<th>$\Delta Z_{\text{min}} = -6 \text{ dB}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_e$</td>
<td>$-14$</td>
<td>$+1$</td>
<td>$-4$ ($-8$)</td>
</tr>
<tr>
<td>Ice water content</td>
<td>$-21$</td>
<td>$+3$</td>
<td>$-10$ ($-2$)</td>
</tr>
<tr>
<td>Extinction</td>
<td>$-16$</td>
<td>$+20$</td>
<td>$-37$ ($-3$)</td>
</tr>
<tr>
<td>Optical depth</td>
<td>$-14$</td>
<td>$+20$</td>
<td>$-35$ ($-2%$)</td>
</tr>
</tbody>
</table>

*The relative changes are calculated from the whole 22 July case. For $Ze$ estimation sensitivity, relative changes for the lidar-only region are included in parentheses. Values are in percentage.

-21% for IWC, and $-15\%$ for extinction and optical depth. In referring to Figure 4, for the same backscattering and extinction coefficients and radar reflectivity, $r_e$ of hexagonal column particles is smaller (up to 20%–50%) than that of solid spherical particles.

Figure 13. The sensitivity test of retrieval extinction, $r_e$, IWC, and optical depth to multiple scattering factors. The $x$ axes are results assuming lognormal solid spherical particles and $\eta = 0.7$. The $y$ axes are results assuming either $\eta = 0.6$ (dots) or $\eta = 1.0$ (diamonds).

Figure 14. The sensitivity test of retrieval extinction, $r_e$, IWC, and optical depth to microphysical modes (i.e., look-up table). The $x$ axes are results assuming gamma PSD of mixed particles and $\eta = 0.7$. The $y$ axes are results assuming column particles in gamma PSD.

Figure 15. Particle size distribution measured by 2D-S on board the DC-8. (a) ER-2 case in Figures 9 and 10. (b) CloudSat-CALIPSO case in Figures 11 and 12.
The functional form of the PSD is assumed to be a gamma distribution. In situ PSD measurements during the TC4 are available from probes on the DC-8 and WB57. Figure 15 shows the PSD measured by 2D-S during two separate cases on 22 July. Figure 15a shows the PSD collected in the thick anvil cloud observed by lidar-radar on ER-2 shown in Figure 9, while Figure 15b shows measured PSDs while sampling a high cirrus layer observed by CloudSat-CALIPSO in Figure 11. For the thick anvil case, there are more large particles than the high-cirrus case, so that the PSD suggests some bimodality in the distribution. For the high-cirrus case, there is a higher concentration of small particles with little evidence of bimodal particle spectra.

Using the measured PSD and the parameterization of volume-equivalent and area-equivalent size, backscattering, and extinction for different habits from Yang [2000] and Hong [2007] we calculate IWC, Ze and lidar attenuated backscatter using equation (1), from which, Ze min is derived from the ARM MMCR observations. Even though the Ze error variance in this region is larger than other regions, Ze is still important in the retrieval. In essence, we are using the knowledge that Ze in the lidar-only region has an upper bound. So we must understand how the algorithm is sensitive to the Ze estimation. The results using Ze min equal to −39 dB is plotted against the control which assumes Ze min equal to −33 dB in Figure 16. As we can expect, only the data in the lidar-only region will change while data in the other regions should follow the 1:1 line. The extinction, re and IWC generally become smaller by about 13%, 8%, and 17%, respectively. Errors in Ze estimations could be random rather than systematically biased as we assigned here. This uncertainty is included in the measurement error covariance matrix as discussed in section 2.7.

**4.3.3. Ze Estimation Sensitivity in the Lidar-Only Region**

For the lidar–only region, Ze is scaled by the attenuated backscatter using equation (1), from which, Ze min is derived from the ARM MMCR observations. Even though the Ze error variance in this region is larger than other regions, Ze is still important in the retrieval. In essence, we are using the knowledge that Ze in the lidar-only region has an upper bound. So we must understand how the algorithm is sensitive to the Ze estimation. The results using Ze min equal to −39 dB is plotted against the control which assumes Ze min equal to −33 dB in Figure 16. As we can expect, only the data in the lidar-only region will change while data in the other regions should follow the 1:1 line. The extinction, re and IWC generally become smaller by about 13%, 8%, and 17%, respectively. Errors in Ze estimations could be random rather than systematically biased as we assigned here. This uncertainty is included in the measurement error covariance matrix as discussed in section 2.7.

**4.4. Validation of the Retrieval Results With In Situ Data**

The payload of the DC-8 included the 2D-S probe [Lawson et al., 2006] measuring particle size distributions,
from which extinction and ice water content are derived, and the NCAR CVI (counterflow virtual impactor [Twohy et al., 1997]) measuring total condensed water content. The 2D-S has two 128-photodiode linear arrays working independently at high-speed and high-resolution optical imaging probes, which improves the three-dimensional properties (such as IWC and extinction) estimation compared to 2D-C [Lawson et al., 2006].

4.4.1. Optically Thick Cloud Retrieval With ER-2 Radar-Lidar Measurements

The DC-8 flight leg flown in conjunction with the CloudSat-CALIPSO overpass on 22 July is an extremely valuable source of information, since a thin cirrus layer existed above a mostly cloud-free ocean surface, and the layer is well sampled by the CALIPSO lidar with optical depth less than or about 1 as shown in Figure 9. Since the DC-8 observations led the satellite observations, collocated CloudSat-CALIPSO retrieval results within 1.5 km (along the track) × 1 km (vertically) boxes are compared with in situ validations of retrieved particle size, IWC, and extinction with measurement from 2D-S and CVI IWC on board the DC-8, which coincides with CloudSat-CALIPSO at 15.0°. Vertical bars are the standard deviations of retrieval.

Figure 19. In situ validations of retrieved particle size, IWC, and extinction with measurement from 2D-S and CVI IWC on board the DC-8, which coincides with CloudSat-CALIPSO at 15.0°. Vertical bars are the standard deviations of retrieval.
Figure 20. In situ validations of retrieved particle size, IWC, and extinction of the CloudSat-CALIPSO case in Figure 12 with measurement from 2D-S on board the DC-8. (a) Ratio of retrieved IWC with 2D-S measured as a function of 2D-S measured IWC; (b) ratio of retrieved extinction with 2D-S measured as a function of 2D-S measured extinction; (c) ratio of retrieved \( r_e \) with 2D-S calculated \( r_e \) assuming a habit as bullet-rosette as a function of 2D-S. The vertical bars are the standard deviations of the ratios within given bins. The black asterisks represent data collected within 5 km and 30 min. The blue asterisks represent data collected within 3 km and 5 min.

Situ data and shown in a time series (Figure 19). The in situ measurements and retrievals show the cloud properties are quite variable and the retrievals capture that tendency very well. For the relatively thick cirrus at the beginning and the end, IWC and extinction agree well. For the thin cirrus cloud in the middle, the extinction and IWC seems to be biased smaller than in situ measurements. The vertical inhomogeneity of the thin cirrus or the time offset between measurements (causing horizontal mismatch) may also contribute to the biases. In situ \( r_e \) from 2D-S assuming hexagonal columns and bullet rosettes is plotted with retrieved \( r_e \). In Figure 20, we plotted the ratio of retrieved properties with 2D-S measured as a function of 2D-S measured for data collected within 5 km and 30 min (black asterisks) and data collected within 3 km and 5 min (blue asterisks). It is clear that the data collected within close temporal and spatial collocation have better agreement with the in situ measurement from 2D-S.

5. Summary

A profiling retrieval algorithm for ice cloud properties such as \( r_e \), IWC and extinction coefficients has been developed to combine CloudSat-CALIPSO measurements using an optimal estimation framework, which will be used as the operational algorithm to produce the CloudSat 2C-ICE standard data product. Taking advantage of the a priori estimation in the framework based on empirical relations and previous retrieval algorithms in the literature, the algorithm is developed to treat a wide range of ice cloud situations from optically tenuous cirrus in the upper troposphere to optically thick anvil clouds. The algorithm considers attenuation in both radar and lidar backscattering signals and multiple scattering in CALIPSO lidar measurements due to its large footprint. The inevitable uncertainty associated with particle habits and PSD shape is accounted for in the forward model error covariance to estimate the retrieved errors. The main algorithm uncertainty comes from the assumptions of the ice crystal microphysical model, the multiple scattering estimation, and the \( Z_e \) parameterization in the lidar-only region. The sensitivity study shows that the lidar multiple scattering has to be included. Over or under estimation of the multiple scattering factor can cause 10–15% uncertainties in optical depth. As is common to all retrieval algorithms applied to ice clouds, the retrieved cloud properties are very sensitive to assumptions regarding particle habit. These assumptions are the largest source of uncertainty.

Two case studies with CloudSat-CALIPSO data and ER-2 CRS-CPL data during the TC4 mission on 22 July 2007 illustrated the capability of the algorithm for retrieving both optically thin and thick ice cloud property profiles (CloudSat-CALIPSO case given in Figures 11 and 12 and ER-2 case given in Figures 9 and 10). The retrieved \( r_e \), IWC, and extinction are shown to be in reasonable agreement with measurements collected in situ by the DC-8. The current algorithm is expected to be complimentary to the set of standard data products that is already being produced by the CloudSat project.

Acknowledgments. This research is supported by NASA under contract from the Jet Propulsion Laboratory (CloudSat). The authors thank the members of the CloudSat and CALIPSO science teams for their hard efforts providing the corresponding data set. The ER-2 lidar is prepared by Matthew McGill. CRS radar data are provided by Gerry Heymsfield. The 2D-S is provided by Paul Lawson. CVI data are prepared by Cynthia Twothy. The data are downloaded from the TC4 project. The optical depth of lidar transmittance is from Dong Liu, University of Wyoming. The authors would also like to thank the anonymous reviewers for their important comments and suggestions. We are very grateful for R. J. Hogan’s constructive comments about this paper and for suggestions about the simulation studies of algorithm performance during the review procedure.

References

Ackerman, T. P., and G. Stokes (2003), The atmospheric radiation measurement program, Phys. Today, 56, 38–45.


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: Micro-


Twomey, C. H., A. J. Schanot, and W. A. Cooper (1997), Measurement of condensed water content in liquid and ice clouds using an airborne coun-


