https://doi.org/10.5194/egusphere-2023-2606
Preprint. Discussion started: 21 December 2023

Deriving cloud droplet number concentration from surface based remote sensors with an emphasis on lidar measurements.

Gerald G. Mace[1]

[1.] Department of Atmospheric Sciences, University of Utah, Salt Lake City Utah, USA

Correspondence to: Gerald G. Mace (jay.mace@utah.edu)

Abstract. Given the importance of constraining cloud droplet number concentrations (N_d) in low-level clouds, we explore two methods for retrieving N_d from surface-based remote sensing that emphasize the information content in lidar measurements. Because N_d is the zeroth moment of the droplet size distribution (DSD), and all remote sensing approaches respond to DSD moments are at least two orders greater than the zeroth moment, deriving N_d from remote sensing measurements has significant uncertainty. At minimum, such algorithms require extrapolation of information from two other measurements that respond to different moments of the DSD. Lidar, for instance, is sensitive to the second moment (cross-sectional area) of the DSD, while other measures from microwave sensors respond to higher-order moments. We develop methods using a simple lidar forward model that demonstrates that the depth to the maximum in lidar attenuated backscatter (r_{max}) is strongly sensitive to N_d when some measure of the liquid water content vertical profile is given or assumed. Knowledge of r_{max} to within 5 m can constrain N_d to within several 10's of percent. However, operational lidar networks provide vertical resolutions or >15 m, making a direct calculation of N_d from r_{max} prohibitively uncertain. Therefore, we develop a Bayesian optimal estimation algorithm that brings additional information to the inversion, such as lidar-derived extinction and radar reflectivity near cloud top. This statistical approach provides reasonable characterizations of N_d and effective radius (r_e) to within approximately a factor of 2 and 30%, respectively. By comparing surface-derived cloud properties with MODIS satellite and aircraft data collected during the Marcus and Capricorn 2 campaigns, we demonstrate the utility of the methodology.

Short Summary: The number of cloud droplets, N_d, in a cloud is important for understanding aerosol-cloud interaction. In this study we develop techniques to derive cloud droplet number concentration from lidar measurements combined with other remote sensing measurements such as cloud radar and microwave radiometer. We show that the deriving N_d is very uncertain although a synergistic algorithm seems to produce useful characterizations of N_d and effective particle size.

1 Introduction

The number of cloud droplets per unit volume (N_d) is essential for characterizing cloud properties. Particularly for lower tropospheric liquid-phase clouds, N_d forms a bridge between atmospheric aerosol and the earth's albedo by determining how condensed water is partitioned into droplet surface area. Higher droplet concentrations for a given condensed mass result in more surface area and more reflective clouds (Twomey, 1974). Thus, many cloud parameterizations used in models include N_d as one of the moments in multi-moment cloud schemes where the other moment is typically related to the mass mixing ratio (Gettelman and Morrison, 2015; Thompson and Eidhammer, 2014; Seifert and Beheng, 2005). Conceptually, using N_d as a baseline parameter makes sense since droplets typically condense on hygroscopic aerosol particles (hereafter cloud condensation nuclei or CCN), thereby fixing N_d as the water droplets grow in an updraft. The initial N_d at the cloud base would be an upper limit on N_d in the ascending updraft because coalescence processes would reduce N_d, and precipitation would further scavenge cloud droplets.

https://doi.org/10.5194/egusphere-2023-2606
Preprint. Discussion started: 21 December 2023

However, aircraft observations often show that for shallow clouds of less than 1 km in depth with minimal precipitation, N_d is reasonably constant with height (Miles et al., 2000).

In this paper, we revisit the methodology used in Mace et al., (2021; Hereafter M21) and attempt to extend that methodology with a focus on lidar measurements from below cloud. In M21, the method derived in M21 was applied to non-precipitating clouds since the layer-averaged radar reflectivity provides a primary source of information. Furthermore, while M21 used the lidar measurements at the cloud base to contribute to the first guess, M21 did not fully exploit the information content available in the lidar measurement near the cloud base. Here, we more thoroughly examine what the lidar can tell us about cloud properties near the cloud base in optically thick boundary layer clouds. Because the lidar backscatter is much larger at the cloud base than in sub-cloud drizzle, we apply the methodology to lightly precipitating and non-precipitating clouds.

We focus on data collected during the summer of 2018 from two ship-based campaigns on the Australian Research Vessel Investigator and the Australian Ice Breaker Aurora Australis during voyages between Hobart, Australia, and East Antarctica. These campaigns are known respectively as the second Clouds Aerosols Precipitation Radiation and Atmospheric Composition Over the Southern Ocean (Capricorn 2) and Measurements of Aerosols, Radiation, and Clouds over the Southern Ocean (Marcus) (McFarquhar et al. 2021). The key observations we include are vertically pointing depolarization lidars, W-band radars, microwave radiometers, and ancillary measurements provided by radiosondes and surface meteorological instruments.

2 Methods

2.1 Theory and Assumptions

The observed lidar attenuated backscatter β_{obs} can be combined with other measurements to derive N_d in fully attenuating liquid phase clouds when measured from the surface. Even though light precipitation may be present, we assume that β_{obs} is dominated by a droplet distribution ($N(D)$) describable by a modified gamma function. Following Appendix B in Posselt and Mace (2014):

$$\frac{dN(D)}{dD} = N_0 \left(\frac{D}{D_0}\right)^{\alpha} exp\left(-\frac{D}{D_0}\right) \qquad (1)$$

Where $\frac{dN(D)}{dD}$ is the droplet number concentration per unit size D with units of cm^{-4} in the cgs unit system. N_0 with units of cm^{-4}, D_0 with units of cm, and α (unitless) are respectively the characteristic number, diameter and the shape parameter of the $N(D)$ distribution function. This simple integrable function allows us to express the microphysical quantities, N_d, q (liquid water content), r_e (effective radius), σ (extinction), and Z (radar reflectivity in the Rayleigh limit), with the following expressions by integrating over all D,

$$N_d = N_0 D_0 \Gamma(\alpha + 1)$$

$$q = \rho \frac{\pi}{6} N_0 D_0^4\, \Gamma(\alpha + 4) \qquad (2)$$

$$r_e = \frac{D_0}{2}(\alpha + 3)$$

$$\sigma = \frac{\pi}{4} N_0 D_0^3\, \Gamma(\alpha + 3)$$

https://doi.org/10.5194/egusphere-2023-2606
Preprint. Discussion started: 21 December 2023

$$Z = N_0 D_0^7\, \Gamma(\alpha + 7)$$

Where ρ is the density of liquid water and Γ is the gamma function. r_e is derived as the ratio of the 3rd moment of $N(D)$ to the 2nd moment of $N(D)$ followed by application of the recursion relationship of the gamma function. For σ, we assume that the extinction efficiency can be approximated as 2 for integrations over typical water droplet distributions. The radar reflectivity Z is written as the sixth moment of the DSD consistent with the Rayleigh approximation which is valid for cloud droplets and radar wavelengths up to W-Band (~94 GHz or ~3mm wavelength). Conversion from conventional units of mm^6 m^{-3} to cgs requires multiplication of Z by 10^{-12}. Using Eqn. 2, we develop relationships among the variables:

$$N_d = \frac{3}{4}\frac{1}{k\pi\rho}\frac{q}{r_e^3}$$
$$Z = q r_e^3 C$$
$$\sigma = \frac{3}{2\rho}\frac{q}{r_e} \tag{3}$$

Where $k = \frac{(\alpha+2)(\alpha+1)}{(\alpha+3)^2}$, and $C = \frac{48\Gamma(\alpha+7)}{\pi\Gamma(\alpha+4)(\alpha+3)^3}$. The last expression in Eqn. 3 was first derived by Stephens (1978) and illustrates a pathway to deriving N_d from multi spectral satellite reflectance measurements. For instance, the bi spectral method applied to MODIS (Nakajima and King, 1990; Platnick et al. 2003) returns measurements of optical depth (τ) and r_e. Since τ is the vertical integral of σ, Eqn. 3 can be adapted for use with satellite retrievals. A full derivation and error analysis of deriving N_d and other quantities from bi spectral satellite retrievals is presented in Grosvenor et al. (2018; Hereafter G18).

Following Platt (1977) and extending through the work of Hu et al., (2007) and Li et al. (2011) among others, we express the observed lidar attenuated backscatter as

$$\beta_{obs}(z) = \beta(z) e^{-2\int \eta\sigma dz} \quad . \tag{4}$$

β_{obs} is the result of 2-way attenuation through the cloud to a point z in the layer and σ is the extinction coefficient with units of inverse length where σ is expressed in terms of the lidar ratio, $S = \frac{\sigma}{\beta}$. A factor η hereafter referred to as the multiple scattering factor accounts for the addition of photons to the observed signal due to multiple scattering in optically dense clouds. Defining the layer-integrated total attenuated backscatter as $\gamma = \int \beta_{\parallel+\perp}$ and the layer integrated depolarization ratio as $\delta = \frac{\int \beta_{\perp}}{\int \beta_{\parallel+\perp}}$ we express $\eta = \left(\frac{1-\delta}{1+\delta}\right)^2$ (Hu et al. 2009). Platt et al. (1999) relates S with η according to $S\eta = \frac{1-T^2}{2\gamma}$ and where T is the layer transmittance. When the layer is fully attenuating (T=0) and $S = \frac{1}{2\eta\gamma}$.

Figure 1 illustrates two examples of β_{obs} profiles measured by the micropulse lidar on board the Aurora Australis during MARCUS. We see the typically small β_{obs} below the cloud that is due to aerosol and molecular scattering in Fig. 1a, while in Figure 1b, there is a contribution from drizzle (observed by a collocated w-band radar, not shown). There is an immediate increase in β_{obs} at a height where condensed liquid water droplets near the cloud base activate, grow rapidly with height, and begin to dominate the lidar signal scattering. β_{obs} then increases exponentially according to Eqn. 4 until the two-way attenuation causes β_{obs} to reach a maximum value, which decays exponentially. We define the range from cloud base to the maximum in β_{obs} as r_{max}. Beyond r_{max}, β_{obs} gains more contribution by multiple-scattered light depending on the lidar field of view and, in liquid clouds, the signal becomes increasingly depolarized relative to the transmitted signal because the orientation of the electric field vector is modified

https://doi.org/10.5194/egusphere-2023-2606
Preprint. Discussion started: 21 December 2023

by the directionality of the scattering event even though each scattering event retains the polarization of the incident field. This effect is a function of the directionality of the scattering that is, in turn, a function of droplet size (Hu et al., 2009). The overall result is quantified by η. The logarithmic decay of β_{obs} was shown by Li et al. (2011) to be related to σ:

$$\eta\sigma = -\frac{\ln\beta(r_2)_{obs} - \ln\beta(r_1)_{obs}}{2(r_2 - r_1)} \tag{5}$$

Where $(r_2 - r_1)$ is the range over which the change in β_{obs} is calculated. Because we have estimated η from measurements, we can estimate σ in the optically thick part of the layer beyond the peak in β_{obs}using linear regression. Li et al. (2011) compare σ derived from this method to estimates of σ derived from passive reflectances and find an uncertainty of ~13% although we assume it to be higher (20%) below. This method's accuracy depends on calculating the rate at which the signal decays with depth in the layer. In practice, we fit a regression line to β_{obs} at ranges beyond r_{max} until the signal is a factor of 2 above the lidar noise floor. We determine the lidar noise level from the mean β_{obs}well beyond the point of full attenuation in the cloud layer. The goodness of the linear regression fit depends on the number of measurements in this range where the signal is decaying. The accuracy depends on the vertical resolution of the lidar measurements for a given σ. The accuracy of the fit is tracked and used to estimate uncertainty.

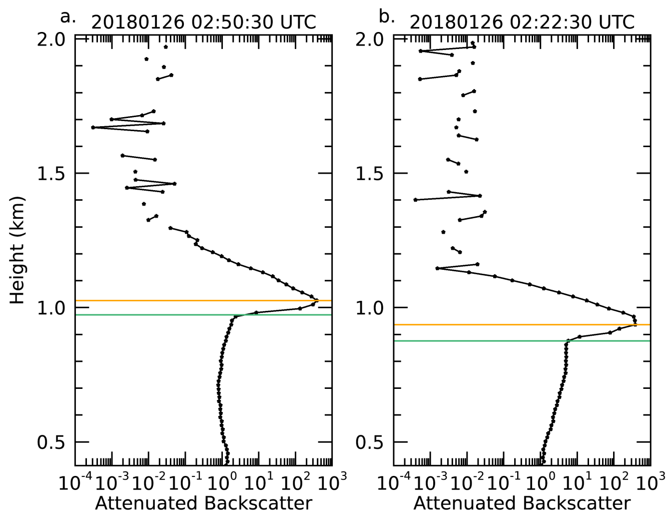

Figure 1. Two Examples of β_{obs} from Lidar data collected during Marcus collected on January 26, 2018 . a) shows a profile with in a non drizzling cloud. b) shows a profile that had sub cloud drizzle as indicated by the cloud radar. The green line indicates the height determined to be cloud base while the red line indicate the maximum in β_{obs}. The distance between the green and red lines is defined as r_{max}

2.2 Direct Calculation of N_d and r_e

The growth of the lidar signal from cloud base to r_{max} can be used to extract information about the cloud layer. Taking the natural logarithm of both sides of Eqn. 4, recognizing that $\beta S_c = \sigma$, and then differentiating with range r in the cloud layer, we can write,

$$\frac{\partial \ln\beta_{obs}}{\partial r} = \frac{\partial \ln\sigma}{\partial r} - 2\eta r. \tag{6}$$

Using Eqn. 3,

https://doi.org/10.5194/egusphere-2023-2606
Preprint. Discussion started: 21 December 2023

$$N_d = \left(\frac{\frac{2}{3}\frac{d\ln q}{dr} - \frac{d\ln \beta_{obs}}{dr}}{2\eta q^{\frac{2}{3}} B^3}\right)^{\frac{1}{3}} \tag{7}$$

We assume that $q = \Gamma_l f_{ad} r$ where f_{ad} is the adiabaticity of the layer (Albrecht et al., 1990) and Γ_l is the adiabatic liquid water lapse rate that is a function of temperature and pressure both of which are assumed as the mean over the cloud layer (G18). Substituting into Eqn. 7,

$$N_d = \left(\frac{\frac{2}{3r} - \frac{d\ln \beta_{obs}}{dr}}{2\eta (\Gamma_l r f_{ad})^{\frac{2}{3}} B^3}\right)^{\frac{1}{3}} \tag{8}$$

Recognizing that at r_{max}, $\frac{d\ln \beta_{obs}}{dr} = 0$, Eqn. 8 can be simplified:

$$N_d = \frac{1}{27B\eta^3\Gamma_l^2 r_{max}^5 f_{ad}^2} \tag{9}$$

N_d is a function of observable quantities with an assumption that the liquid water profile has an adiabatic shape. The DSD shape parameter α is also assumed and typically given a value that conforms to in situ data (see below). f_{ad}, which scales the adiabatic liquid water content, can be calculated as the ratio of the vertically integrated liquid water mass or LWP that is readily retrieved from measurements collected by a microwave radiometer (Turner et al., 2016) to the adiabatic LWP that can be derived by integrating Γ_l over the depth of the layer (G18). The depth of the layer must be determined from some means such as a vertically pointing cloud radar or perhaps from recent radiosonde soundings. Thus, N_d can be derived with a combination of a depolarization lidar, some means of determining cloud top, and a microwave radiometer. Neither the lidar nor the radar, if present, must be calibrated to derive N_d with Eqn. 9. With LWP and N_d, and a measure of layer depth, it is straightforward to estimate a characteristic cloud droplet size. Typically, the cloud top r_e is most representative of the layer reflectance and is derived from bispectral measurements such as MODIS to which we will compare later. Following G18,

$$r_e = \left(\frac{\frac{3h}{4\pi\rho_l}\Gamma_l f_{ad}}{kN_d}\right)^{1/3} \tag{10}$$

where h is the layer thickness and k is the cubed ratio of a volume weighted characteristic droplet size to the effective droplet size assumed constant at 0.8 following G18.

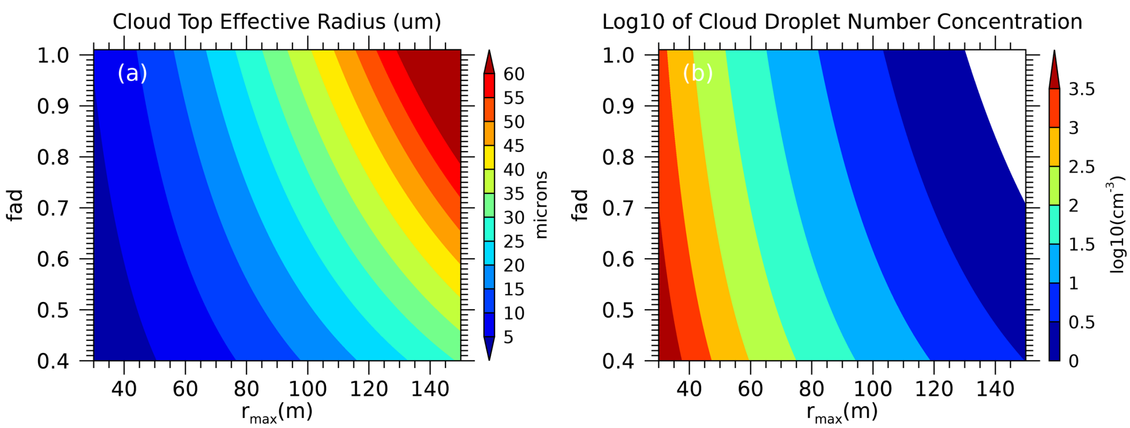

Figure 2. Response of Equation 10 (a) and Equation 9 (b) to typical values of zmax and fad.

https://doi.org/10.5194/egusphere-2023-2606
Preprint. Discussion started: 21 December 2023

Figure 2 shows the response of Equations 9 and 10 to typical ranges of r_{max} and f_{ad}. In these calculations, we fix η at 0.4 (a typical value for the lidar on CAPRICORN 2) and the cloud layer thickness at 500 m. We find that r_{max} contributes most significantly to the N_d calculation, given the fifth power exponent in the denominator of Eqn. 9. We find that N_d ranges from near 1000 cm^{-3} for low r_{max} values that would correspond to very opaque layers to values less than 10 cm^{-3} for layers with r_{max} exceeding 100 m. These correspond to the approximate typical extremes for r_{max} found in measurements. r_e ranges from 5 μm for small r_{max} to more than 50 μm for very large r_{max} corresponding to the change in N_d from high to low, respectively. For a given r_{max}, an increasingly adiabatic cloud layer causes N_d to decrease and r_e to increase. This tendency makes physical sense since for our simple conceptual model of an adiabatically increasing q profile, increasing f_{ad} for a given LWP and layer thickness (h) causes more liquid water in the profile. Therefore, for a given r_{max}, fewer and but larger droplets are required to achieve a given extinction profile that allows the lidar beam to penetrate the layer.

While Eqns. 9 and 10 produce physically plausible results, the sensitivity of N_d to uncertainty in r_{max} is substantial. The resulting uncertainty in N_d then translates into uncertainty in r_e. Clearly, with the typical range in r_{max} between a few 10's of meters to values not much greater than about 100 m, the vertical resolution of the lidar has a significant bearing on how well we can know r_{max}. Lidars in operational networks typically operate with range bin spacing of between 10 and 15 m. The micropulse lidars operated by the DOE Atmospheric Radiation Measurement (ARM) program (Mather, 2021) use 15 m spacing while Vaisala laser ceilometers use a range bin spacing of 10 m. We use a bootstrap approach to evaluate the effect of this uncertainty in r_{max}. Fixing the uncertainty in f_{ad} and η at 20% and allowing a variable r_{max} uncertainty of 1m, 5m, 10m, and 15m, we use a normally distributed set of random numbers to perturb the r_{max}, f_{ad}, and η about their assumed values prior to implementation of Eqns. 9 and 10. 25000 iterations are used to compute the frequency distribution of the resulting N_d and r_e (Fig. 3) for each r_{max} uncertainty. We find that range bin spacing that is typical of operational lidars and ceilometers is inadequate for calculating N_d. A 15 m range bin spacing results in a normalized standard deviation in the N_d distribution for the example shown here of ~3 for a fixed value of 102 cm^{-3}. The r_e normalized standard deviation is approximately 29% in this case. The uncertainty in N_d and r_e decrease as the uncertainty in r_{max} is reduced from 15 m to 1 m. At 1 m and 5 m uncertainty in r_{max}, N_d (r_e) has uncertainties of 0.16(0.16) and 0.55 (0.18), respectively.

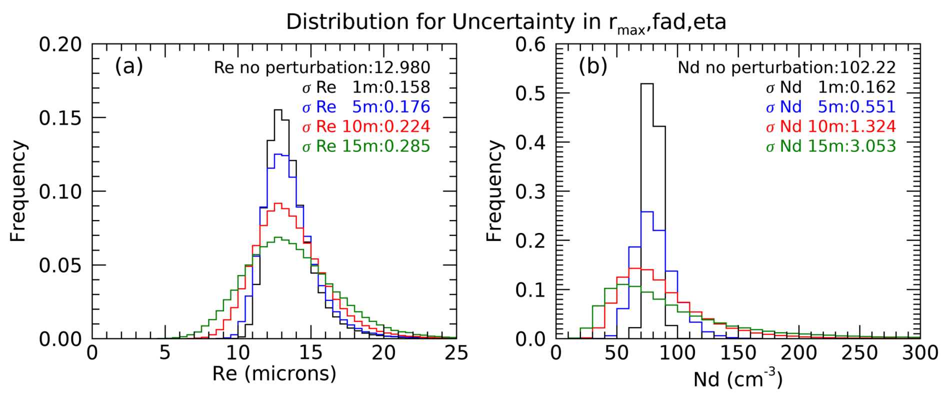

Figure 3. Sensitivity of Eqns. 9 and 10 to uncertainty in input parameters. Inset lists the resulting uncertainties corresponding to the color-coded frequency distributions. Insets list normalized standard deviations for an assumed uncertainties in r_{max} of 1, 5, 10, and 15 m.

https://doi.org/10.5194/egusphere-2023-2606
Preprint. Discussion started: 21 December 2023

These levels of uncertainty would convey useful information about a cloud layer, whereas the typical ranges of uncertainty that we encounter with operational lidars and ceilometers are only marginally to insignificantly informative.

2.3 An Optimal Estimation Algorithm

To lessen the effects of uncertainty in r_{max}, we attempt to bring additional information to bear by developing a Bayesian optimal estimation (OE) inversion algorithm (Maahn et al., 2019) to retrieve N_d and r_e. This methodology allows us to use additional data sources that contribute to our understanding of droplet N_d and r_e while balancing the observational and forward modeling errors that contribute to retrieval uncertainty. In addition to the independent variables in equations 9 and 10, we also use the layer σ derived from the lidar data (Eqn. 5) and the radar reflectivity near cloud top (Z_{top}) from a collocated millimeter radar. We choose to use the radar reflectivity near the cloud top to avoid, to the extent possible, multimodal droplet distributions that often occur as drizzle or snow sediments through a cloud layer. Near layer top, at least for reasonably shallow and not strongly convective clouds, we assume the precipitation droplet mode to be nascent and the cloud droplet distribution to be approximately unimodal. Inspection of aircraft in situ drop size distributions collected over multiple campaigns reasonably support this assumption (Lawson et al., 2017). Z_{top} provides a useful constraint on the liquid water profile's shape and conveys information on f_{ad} and r_e. We define an observational vector,

$$y = [r_{max} \quad \sigma \quad LWP \quad Z_{top}] \tag{11}$$

An observational error covariance matrix, Sy, is a 4x4 element matrix that records the uncertainty of the measurements in y due to random noise and uncertainties in forward modeling of that quantity along the diagonal. We allow for covariance among the observations as listed in Table 1. These correlations are derived from the Capricorn 2 and Marcus combined data set. We find significant correlations among the measurements in y. These correlations show that the measurements in y are not independent and are not, therefore, unique in terms of information. We address the information content below.

Table 1. Sources of uncertainty estimates (diagonal) and correlations (off diagonal) among measurements in y (Eqn. 11) used in the OE algorithm. Correlations are derived from the combined Marcus and Capricorn 2 data sets.

	r_{max}	σ	LWP	Z_{top}
r_{max}	Lidar Range Bin Space			
σ	-0.58	20% (Lin et al. 2011)		
LWP	+0.24	-0.22	20 g m-2 (LWP<100) 30% (LWP>100) (Turner et al., 2016)	
Z_{top}	+0.23	+0.48	+0.47	1 dB Capricorn, 4 dB Marcus (Kollias et al., 2019)

The quantities to be estimated and their covariance are denoted in the state vector x respectively:

$$x = [N_d \quad r_e] \tag{12}$$

https://doi.org/10.5194/egusphere-2023-2606
Preprint. Discussion started: 21 December 2023

And S_x is a 2x2 element matrix that records the uncertainties of x along the diagonal. r_e is assumed to be near the layer top as defined in Eqn. 10.

We use x and additional observations and assumptions to derive a forward calculation of y or F(x) based on initial and incremental x guesses (see below) with a simple forward model. Our forward model begins with the observed thermodynamics, cloud base height, and layer thickness. With an observed or simulated LWP and a temperature-dependent Γ_l, we create a vertical profile of liquid water that varies with an adiabatic shape scaled by f_{ad}. Using an assumed shape parameter (α=2, justified below), we then calculate profiles of r_e and N_d allowing us to estimate the terms in y using the simple lidar equation (Eqn. 4) and the expressions for Z and σ in Eqn. 3.

To derive x from y using OE, we express the first order derivatives of y with respect to x in a Jacobian matrix, K_x, that has dimensions of the number of elements in y (4) by the number of elements in x (2):

$$K_x = \begin{matrix} \frac{\partial r_{max}}{\partial N_d} = -0.29 & \frac{\partial \sigma}{\partial N_d} = 0.24 & \frac{\partial LWP}{\partial N_d} = 0 & \frac{\partial Z_{top}}{\partial N_d} = 0.01 \\ \frac{\partial r_{max}}{\partial r_e} = 0.92 & \frac{\partial \sigma}{\partial r_e} = -2.9 & \frac{\partial LWP}{\partial r_e} = 0.44 & \frac{\partial Z_{top}}{\partial r_e} = 1.2 \end{matrix}$$

These terms are calculated analytically using the expressions in Eqns. 2, 3, 9, and 10. Also, we set $\frac{\partial LWP}{\partial N_d} = 0$ because we assume that the amount of water made available for condensation is the result of thermodynamics while how that water is distributed into droplets depends more on the CCN that is available for the water to condense onto. The quantities listed in the K_x matrix show typical values of the terms for Case 5 listed in Table 3 below in terms of $\frac{\partial \ln(y)}{\partial \ln(x)}$. We find that r_e influences σ, LWP, and Z_{top} in predictable ways. For instance, the derivative is negative in the r_e - σ relationship. The sensitivities of the observations in y are much more sensitive to r_e than to N_d illustrating the challenge of retrieving N_d with remote sensing observations as discussed earlier.

The OE formalism derives x by balancing the uncertainties and information in the measurements with what is known about the statistical properties of x given the atmospheric state. The information from prior knowledge is contained in an a priori vector of statistical estimates of the quantities in x (Eqn. 12) or x_a and their covariance, S_a. For the prior estimate of N_d, we reason that coincident cloud condensation nuclei (CCN) measurements provide an upper limit on the droplet number in each situation. These measurements were collected during Marcus and CAPRICORN 2 and are available hourly when the wind direction was favorable by not contaminating aerosol inlets with ship exhaust (Humphries et al. 2021). These hourly CCN measurements at 0.2% supersaturation are simply multiplied by 0.8 to account for coalescence processes and used in x_a. The hourly standard deviation of the CCN is then used along the diagonal of S_a. When CCN are not available, within the previous 6 hours, we use averages of the surface-based CCN measurements for the latitudinal bands from 40°S-50°S, 50°S-60°S, and >60°S (Humphries et al., 2023). For the prior value of r_e, we use the 0.8*CCN, the LWP, and layer thickness in Eqn. 10. For r_e when CCN data are not available, we use in situ aircraft data collected during the Southern Ocean Cloud Radiation and Aerosol Transport Experiment (SOCRATES; McFarquhuar et al., 2021) that was conducted in the Southern Ocean region south of Hobart Australia during the Austral Summer of 2018 by the NSF/NCAR HIAPER Gulfstream V (GV) aircraft. In this campaign, the GV completed 15 research flights. We combine the CDP and 2DS measurements into a single droplet size distribution (DSD) and use a moments minimization method (Zhao et al., 2011) to estimate of Eqn. 1 for each low-level cloud 1-second DSD. W-Band radar reflectivity is then calculated using Eqn. 3. For a particular retrieval where we have a measured Z_{top}, the Socrates data set is searched for all instances where Z is within 1.5 db of the measurement and the prior r_e is then estimated from the mean of the in-situ measurements. For the covariance among

https://doi.org/10.5194/egusphere-2023-2606
Preprint. Discussion started: 21 December 2023

the quantities in S_a, we know that r_e and N_d are strongly correlated (G18) so we use a correlation of 0.7 among those terms based on in situ data.

The OE formalism also allows us to quantify the added uncertainty in our forward model calculations due to model parameters and assumptions (Maahn et al. 2019; Austin and Stephens, 2001) which we take to include α (droplet distribution function shape parameter), f_{ad} (the adiabaticity of the column) and η. We find that a value of α=2 with a standard deviation of 1.5 reasonably characterizes the in-situ cloud collected during Socrates. f_{ad} is estimated by taking the LWP and cloud thickness observations collected over the Marcus and CAPRICORN 2 voyages and deriving a linear regression of f_{ad} in terms of LWP following Miller et al., (1998) to wit, $f_{ad} = 1.-(0.002 * LWP)$. With LWP in g m^{-2}, this equation returns f_{ad}=0.6 for LWP=200 gm^{-2} and 0.5 for LWP=250 g m^{-2}. The scatter in the LWP-f_{ad} observations suggest an uncertainty in this estimate of 0.15. η is derived from the depolarization lidar data following the method described in Hu et al. (2007). While the uncertainty of this quantity is difficult to assess, examining the consistency of the estimates over periods of persistent cloud cover we determined that an uncertainty of 30% is reasonable. A term of the form $K_b S_b K_b^T$ is added to the instrumental uncertainties where K_b is a Jacobian matrix that contains the first derivatives of the measurements in y with respect to α, f_{ad}, and η determined through finite differences in the forward model:

$$K_b = \begin{matrix} \frac{\partial r_{max}}{\partial \alpha} = -0.08 & \frac{\partial r_{max}}{\partial f_{ad}} = -0.60 & \frac{\partial r_{max}}{\partial \eta} = -0.63 \\ \frac{\partial \sigma}{\partial \alpha} = 0.11 & \frac{\partial \sigma}{\partial f_{ad}} = 0.55 & \frac{\partial z_{max}}{\partial \eta} = 0.03 \\ \frac{\partial LWP}{\partial \alpha} = 0.20 & \frac{\partial LWP}{\partial f_{ad}} = 1.0 & \frac{\partial LWP}{\partial \eta} = 0.0 \\ \frac{\partial Z_{top}}{\partial \alpha} = -0.35 & \frac{\partial Z_{top}}{\partial f_{ad}} = 2.0 & \frac{\partial Z_{top}}{\partial \eta} = 0.0 \end{matrix}$$

The numbers in the K_b expression are in terms of $\frac{\partial \ln(y)}{\partial \ln(x)}$ and are derived from the forward model over the physically reasonable ranges of the parameters. We find that these numbers vary by less than 20% in the Capricorn and Marcus data sets and are used as written in the inversion algorithm. S_b contains the variance of α, fad, and η and we assume that the covariance among these quantities can be neglected.

2.4 Evaluation

Inversion of y for x then follows a standard iterative approach by applying a Gauss-Newton minimization technique derived in Rodgers (2000). See also Maahn et al., (2019). In this approach, successive guesses of x are derived using the well-known expression,

$$\delta x = \left(S_a + K_x S_y K_x^T\right)^{-1}\left[S_a^{-1}(\hat{x} - x_a) + K_x^T S_y^{-1}\left(y - F(\hat{x})\right)\right] \quad 13$$

Where $\hat{x}$ is a present guess, $F(\hat{x})$ is the forward estimate of the measurements in y using the present guess. δx then becomes the next increment on $\hat{x}$. Eqn. 13 is iterated until either a convergence criterium is met or divergence of the result occurs. Typically, less than 10 iterations are necessary if the algorithm converges which it does > 90% of the time in non-precipitating conditions while convergence occurs less frequently as drizzle and light snow increase due to the inability to accurately estimate r_{max}.

The response of the OE algorithm is equivalent to the results presented in Fig. 3, except that additional information is used to lessen the effects of uncertainty in r_{max}. In Table 2, we list 6 cases that we use to examine the response of the OE algorithm in terms of

https://doi.org/10.5194/egusphere-2023-2606
Preprint. Discussion started: 21 December 2023

the retrieved quantities and their uncertainties. The cases 1 and 2 are designed to illustrate a situation that might be found in a heavy aerosol environment with a low r_{max}, high σ, and low Z_{top} that produces high N_d, small cloud drops and moderately high LWP. Cases 3 and 4 show the opposite with a rather large r_{max} and lower σ. Z_{top} is set higher with a larger LWP. The algorithm returns a low cloud N_d and large r_e in cases 3 and 4. Cases 5 and 6 are in between the two extremes. F_{ad} in these cases range from 0.8 to 0.9, and this is by design as the cloud depth is specified. The uncertainties listed in Table 3 are used in Cases 1, 3, and 5; except for Z_{top} which is listed in dB, the uncertainties are a fraction of the measurement. Cases 2, 4, and 6 use twice the listed uncertainties in Cases 1, 3, and 5. As a fraction of the returned values, the 1 standard deviation uncertainties do not change significantly from case to case, and they respond predictably to a doubling of the observational errors increasing approximately by a factor of 2. We also test the OE uncertainty by randomly perturbing the observations about their stated uncertainties until the error statistics converge. These are reported in Table 3 in the "Bootstrapping" column. The bootstrap experiment generally returns uncertainty in r_e that is equivalent to or slightly smaller than the OE results. For N_d, the bootstrap experiment returns marginally larger uncertainties than the OE results.

The Shannon information content measures the extent to which the observations reduce the uncertainty in the prior. The studies of L'Ecuyer et al. (2006) and Cooper et al. (2006) provide detailed discussions of this concept. Doubling the observational uncertainty reduces the information content by approximately 1/3. The number of independent parameters is less than the number of elements in y (the observations) because the observations are correlated. For instance, as shown in Table 2, r_{max} and σ both constrain N_d while LWP and Z_{top} constrain r_e. Even in the lower error cases, the observations do not provide sufficient information to retrieve three independent quantities, suggesting that the results are correlated and not independent.

The uncertainty in r_e remains roughly equivalent to the results shown in Fig. 3, although we consider the results of the OE to be more accurate because a better accounting of information is used. Notable is the magnitude of the uncertainties for the retrieved N_d. We find that it remains large, although the additional information provided by the other observations reduces the uncertainty compared to the results in Fig. 3. We also tested how well the OE algorithm without r_{max} would do where just extinction is the primary constraint on N_d. This was accomplished by setting the Kx term $\frac{\partial r_{max}}{\partial N_d} = 0$. We found that for the uncertainties in the other quantities listed in Table 3, the uncertainty in N_d was approximately 150%, showing that r_{max} is a useful quantity in this regard. However, retrieval of N_d remains highly uncertain when lidar range bin spacing exceeds 5 m.

Table 2: Cases used to demonstrate the response of the OE algorithm. 1 standard deviation uncertainties are listed in parentheses. Cases 2, 4, and 6 use uncertainties a factor of 2 larger than those listed for cases 1, 2, and 3.

	r_{max}(m)	σ (km-1)	Z_{top} (dBZ)	LWP (g m-2)
Case 1	38 (4)	28 (4.5)	-19 (2)	126 (30)
Case 3	62 (6)	16 (2.5)	-12 (2)	101 (25)
Case 5	56 (5.5)	23 (3.5)	-15 (2)	150 (37)

Table 3: The retrieved parameters and their 1 standard deviation uncertainties as fractional values in parentheses for cases 1-6 listed in Table 3. We also list the Shannon information content in bits, and the number of independent observations in

https://doi.org/10.5194/egusphere-2023-2606
Preprint. Discussion started: 21 December 2023

the retrieval as derived from the OE formalism – see Rodgers, (2000). Cases 2, 4, and 6 have observational uncertainties a factor of 3 greater than listed in Table 3.

	Nd (cm^{-3})	Nd OE Fractional Uncert. (Fraction)	Nd Bootstrap Uncert. (Fraction)	Re (um)	Re OE Uncert. (Fraction)	Re Bootstrap Uncert. (Fraction)	Info (bits)	# Ind Params
Case 1	229	0.69	0.77	9.8	0.24	0.23	3.1	1.7
Case 2	231	0.83	0.93	9.9	0.42	0.35	1.2	1.4
Case 3	36	0.70	0.88	16	0.19	0.28	3.6	1.7
Case 4	37	0.84	1.2	15	0.40	0.32	1.2	1.4
Case 5	95	0.70	0.95	13	0.18	0.27	3.5	1.7
Case 6	91	0.84	1.2	12	0.40	0.34	1.3	1.5

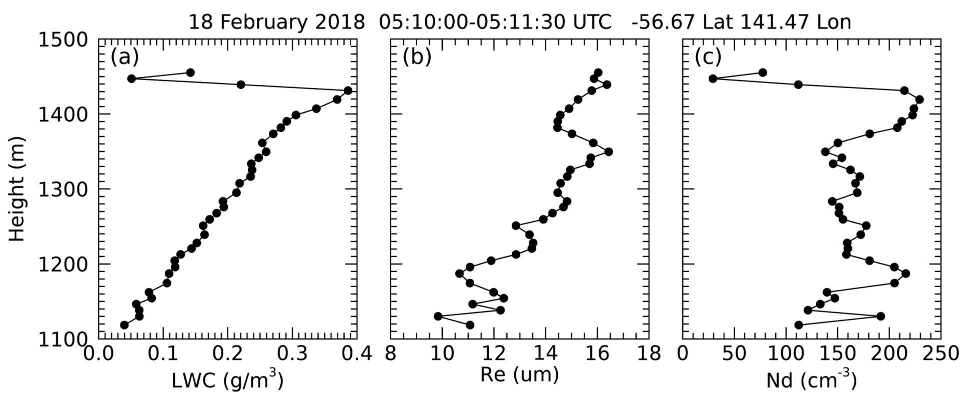

Figure 4. Ramp through an MBL cloud layer on 18 February 2018 collected by instruments on the NCAR Gulfstream V during Socrates. This ramp was conducted near the RV Investigator ship during Capricorn 2.

To provide a more realistic evaluation of the OE algorithm performance, we use data collected during the Socrates campaign, where ramps through low-level cloud layers were conducted. Such a ramp is depicted in Fig. 4 which was collected on February 18, 2018 (hereafter 2/18) at 0510 UTC when the GV was conducting a mission near the R/V Investigator at 57°S and 142°E. We will expand on the February 18 case study below. For this analysis, we focused on 1-second data collected by the Cloud Droplet Probe (CDP) that recorded droplet spectra in 2 μm size bins up to 50 μm. The aircraft entered the cloud layer with a temperature near -5°C at 1100 m. LWC and r_e steadily increased as the GV ascended and exited the cloud layer approximately 90 seconds later at an altitude of 1450 m where q reached a maximum of 0.4 g m^{-3} and the re near cloud top was ~15 microns. We note an interesting structure in the vertical r_e profile with a sudden decrease near 1375 m. During this ascent, N_d was quite variable but averaged 150 cm^{-3} through most of the ramp until 1375 m here there is an abrupt increase in N_d to ~225 cm-3 in conjunction with the decrease in r_e. Summing q vertically through the layer, the LWP was 65 g m^{-2} with an adiabatic LWP of 80 g m^{-2}, suggesting a sub-adiabatic layer with f_{ad} of 0.8. The radar reflectivity time series (discussed later) shows that drizzle was occurring sporadically during this

https://doi.org/10.5194/egusphere-2023-2606
Preprint. Discussion started: 21 December 2023

case. We used the cloud droplet concentrations collected during the ramp to get r_{max} (32 m), the expression for Z (Eqn. 3) to estimate Z_{top} (-15 dBZe), and the cross-sectional area of the droplet distribution to estimate σ (layer mean of 30 km^{-1} and layer optical depth (τ) of 14). These values were used to drive the lidar forward model. We implement the OE algorithm with f_{ad} and LWP to get a retrieved N_d of 165 cm^{-3} and r_e of 14 μm in reasonable agreement with the input data.

We repeated this exercise for other ramps collected during Socrates, excluding ramps that were super-adiabatic or had non-adiabatic structure in the vertical profile, reasoning that the finite distance over which the ramps occurred (~10-20 km) had the potential to sample cloud elements of varying properties. For instance, on 2/18 three additional ramps were not considered. The observational uncertainties used in the inversion are as discussed above for Cases 1, 3, and 5. Figure 5 shows the relationship between observed and retrieved N_d and r_e., showing that the OE algorithm can reasonably capture the characteristics of the cloud layers. While we would expect the algorithm to provide a reasonable comparison of the retrieved and observed N_d and r_e, we note that the OE uncertainty, for the most part, extends over the 1:1 line, suggesting that the characterization of uncertainty in the retrieved quantities is a reasonable estimation of the actual uncertainty of the algorithm.

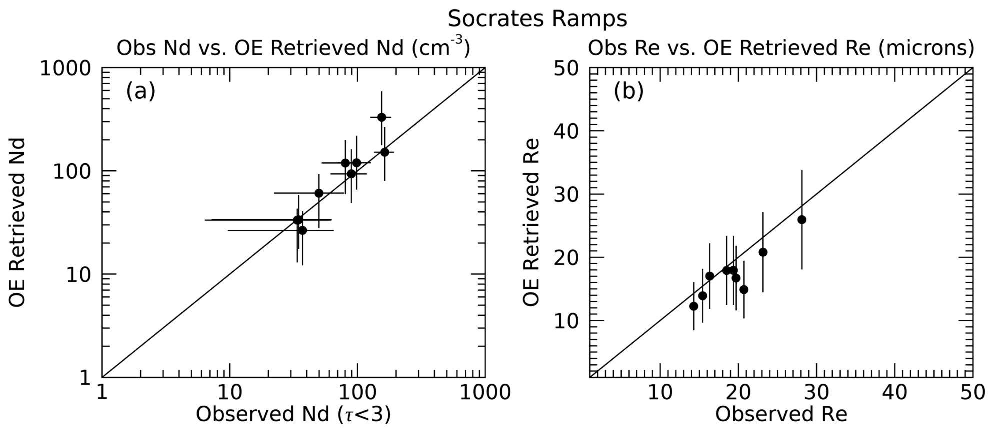

Figure 5. Comparison of Observed and Derived N_d and r_e from Socrates ramps. The error bars on the retrieved quantities are as derived from the optimal estimation.

3. Independent Comparisons

The 2/18 case study provides a unique opportunity for independent comparisons of the algorithm with data collected while the GV aircraft operated in the vicinity of the R/V Investigator and with an overpass of the Terra satellite that provided independent retrievals of τ and r_e (Platnick et al., 2004) from which we can derive LWP and N_d (G18) using the MODIS τ and r_e. During this case study period, the ship remained stationary at 56.6°S and 141.5°E to facilitate coordination with the GV. Figure 6 illustrates the data collected from the surface-based onboard instruments. The lidar attenuated backscatter indicates a fully attenuating layer through the entire period. With a cloud base temperature near -5C, the lidar depolarization ratio data suggest that the cloud base phase and the sub cloud precipitation were liquid. The W-Band radar on the RV Investigator indicated episodic drizzle events of 10-20 minute duration roughly every hour, some of it rather heavy. Intervening periods without drizzle had radar reflectivity near the detection threshold of the radar (~-25 dBZe during Capricorn 2). The radar and sounding data collected at the ship showed

https://doi.org/10.5194/egusphere-2023-2606
Preprint. Discussion started: 21 December 2023

that the layer was topped by a strong marine inversion near 1.5 km in agreement with the GV ramp in Figure 4. The LWP was variable between 50-60 g m^{-2} during periods without drizzle to value near 250 g m^{-2} during periods of drizzle. The retrieved cloud properties varied depending on the proximity of a drizzle event. While the algorithm did not converge in regions of heavier drizzle,

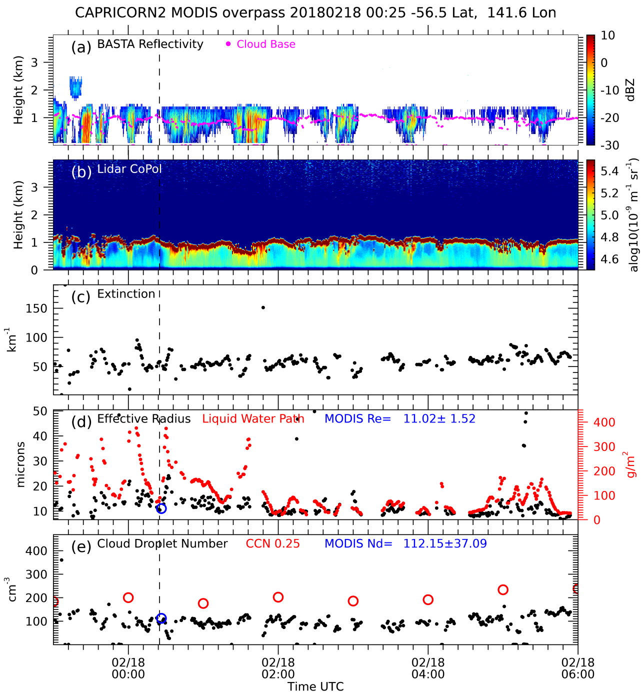

Figure 6. Surface-based measurements and derived properties from data collected on February 18, 2018 on the R/V Investigator near 55.6S and 141.5E. a) radar reflectivity with cloud base, b) lidar attenuated backscatter, c) extinction derived from the lidar attenuated backscatter, d) effective radius and liquid water path, e) cloud droplet number concentration. The blue circles and inset values are from an overpass at 0025 UTC of MODIS on Terra. CCN at 0.25% supersaturation is shown on e.

we find near the boundaries of several drizzle events that the N_d decreased to 20-30 cm^{-3} and r_e increased to be more than 20 μm. Otherwise, the algorithm tended to produce N_d in the range of 100 cm^{-3} and r_e in the 10 μm range.

A Terra MODIS overpass occurred at 0025 UTC. We collect the Level 2 retrieval of τ and r_e in a region of 50 km diameter centered on the ship and the ship data are collected between 23 UTC on 17 February and 0130 UTC on 2/18. The comparison results are shown in Fig. 7. A broad distribution of LWP is demonstrated during this period that has a similar character in both data sets. The ship has an LWP mode near 160 g m^{-2}, that is due to the drizzle event that is evident near 00 UTC in Fig. 6. The mean LWP of the ship is slightly larger than MODIS but the two are in broad agreement. The distributions of r_e in the two data sets overlap with the surface data skewed to larger values, likely because of the predominance of the drizzle event. The N_d retrievals also demonstrate broad agreement with quite wide distributions even though the ship N_d is skewed to smaller values. The ship τ distribution is

https://doi.org/10.5194/egusphere-2023-2606
Preprint. Discussion started: 21 December 2023

skewed to smaller values than MODIS, consistent with larger effective radii and smaller cloud droplet number. It is worth noting that the τ and r_e are the quantities that are most directly retrieved from the MODIS algorithm, whereas the LWP and especially Nd require additional assumptions in their derivation from τ and re.

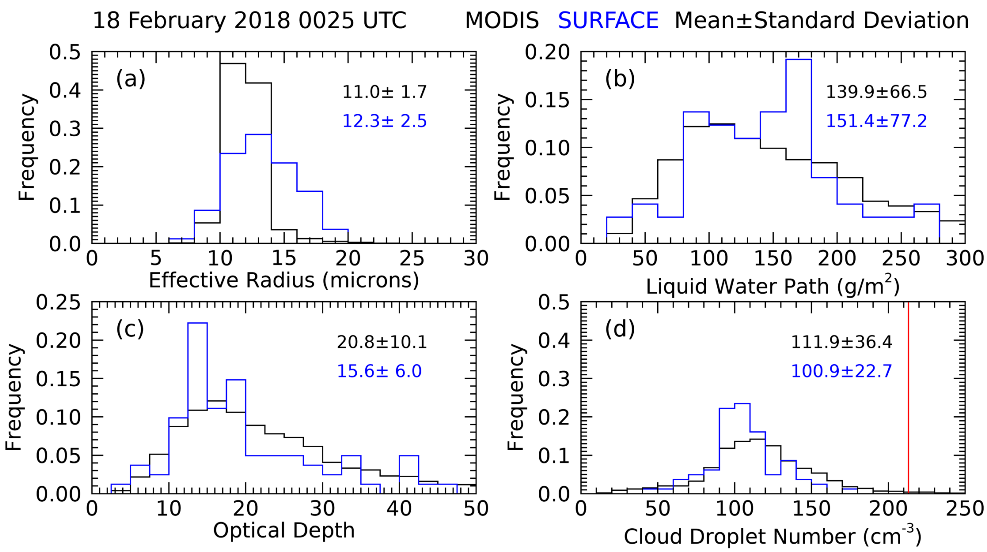

Figure 7. Comparison of properties observed and derived from data collected on the RV Investigator (blue) with cloud properties derived from a Terra MODIS overpass at 00:25 UTC on February 18, 2018. a) Effective Radius, b) LWP, c) Optical depth, d) cloud droplet number concentration.

On the other hand, the surface data LWP is independent of the radar, lidar, and other measurements and requires a minimum of assumptions to derive from the microwave radiometer brightness temperatures. N_d and r_e from the surface data need a complicated algorithm to derive, and τ from the surface data is calculated using Eqn. 3. Thus, the surface-derived τ would capture the errors in the surface retrieval of r_e. While there are biases in the comparison, given the substantial differences in the two independent data sets, we conclude that the comparisons demonstrate a reasonably consistent picture of the cloud field during the overpass.

The GV arrived at the ship at approximately 02 UTC on 2/18 and operated in the vicinity of the ship for roughly 2 hours. It conducted ramps, level legs within the cloud layer, and legs above and below the layer for aerosol and remote sensing applications. We compare data collected during this time by gathering the aircraft data within 50 km of the ship. The effective radius is derived from the aircraft CDP data in the upper ½ of the layer (above 1.2 km) and the aircraft N_d is collected from the CDP data in the lowest ½ of the layer. The comparison of N_d and r_e distributions are shown in Fig. 8. The aircraft r_e data are bimodal while the ship retrieved r_e are unimodal and centered on the lower mode of the aircraft r_e distribution. We interpret the lack of bimodality in the ship-based r_e data as being due to the algorithm not converging in regions of heavier drizzle as noted above. The aircraft penetrations of drizzle and non-precipitating clouds results in the bimodality shown in Fig. 8. The N_d distributions are broadly similar, but the ship results are biased to lower values. It is unclear the extent to which there is a bias toward the lower part of the cloud layer in the ship data. Regardless, both distributions are centered just in excess of 100 cm^{-3}. This comparison suggests that the surface-based OE algorithm can reasonably replicate the cloud layer properties.

https://doi.org/10.5194/egusphere-2023-2606
Preprint. Discussion started: 21 December 2023

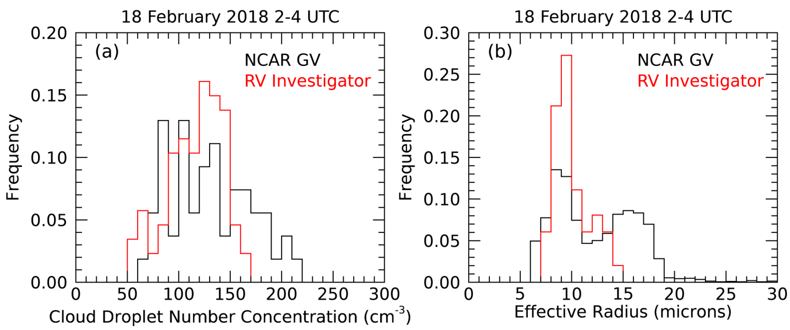

Figure 8. Comparison of N_d (a) and r_e (b) derived from the surface-based data collected on the RV Investigator (red) with data collected from the NCAR GV on 18 February 2018. Cloud properties are compiled over the period from 2-4 UTC.

Finally, we compare with the MODIS-derived cloud properties from overpasses of the ships during the Marcus and Capricorn campaigns. With MODIS instruments on the Terra and Aqua satellites and the ships being at sea over extended periods, we found several events where suitable low-level clouds occurred over the ships during MODIS overpasses. Table 4 lists the information about the 14 overpasses of the ships that we use for the comparison in Fig. 9. Our approach was to examine a 50 km region of MODIS data centered on the ship, and we compiled surface data from 90-minute periods before and after an overpass. We find reasonable agreement in the comparisons. The LWP is an interesting quantity since, as stated above, it is independent of the N_d - r_e retrieval. The LWP from the MODIS data, on the other hand, is derived from the τ and r_e algorithm that uses the Nakajima and King (1990) bi spectral method so that the MODIS LWP would carry forward any uncertainties in τ and r_e. The agreement, however, is reasonable with little bias. Most of the cases have LWP<200 g m^{-2} since we focus on non- to lightly precipitating cloud scenes. The r_e cases range over values that are very small corresponding to cases near the Antarctic continent with high N_d and no precipitation to r_e that exceeds 15 μm. The comparison in r_e is unbiased with a good correlation. While N_d also has a good correlation, there does appear to be a slight bias in the comparison, with the surface data being, on average, 20-30 cm^{-3} higher than MODIS. The optical depth appears unbiased for values less than ~15 but then seems to show a bias for values of more than 15 with MODIS being larger than the surface-based results. More data is highly desirable to establish how well and under what circumstances these data sets agree or don't, but this preliminary comparison is encouraging.

Table 4. List of the MODIS overpasses shown in Fig. 9.

Date/Time	Location	Satellite	Campaign
2018/02/04, 0415 UTC	65.6°S, 150.0°E	Aqua	Capricorn 2
2018/02/05, 0415 UTC	63.9°S, 150.0°E	Aqua	Capricorn 2
2018/02/07, 2350 UTC	62.8°S, 143.6°E	Terra	Capricorn 2
2018/02/13, 0545 UTC	63.9°S, 132.1°E	Aqua	Capricorn 2

https://doi.org/10.5194/egusphere-2023-2606
Preprint. Discussion started: 21 December 2023

2018/02/20, 0010 UTC	50.2°S, 143.7°E	Terra	Capricorn 2
2018/01/02, 0110 UTC	66.3°S, 110.5°E	Terra	Marcus
2018/01/05 0140 UTC	66.2°S, 110.2°E	Terra	Marcus
2018/01/06 0720 UTC	66.5°S, 108.8°E	Aqua	Marcus
2018/01/06 0225 UTC	64.0°S 111.3°E	Terra	Marcus
2018/01/10 0425 UTC	47.0°S 142.6°E	Terra	Marcus
2018/02/23 0805 UTC	59.3°S, 89.3°E	Aqua	Marcus
2018/02/24 0305 UTC	56.9°S, 95.4°E	Terra	Marcus

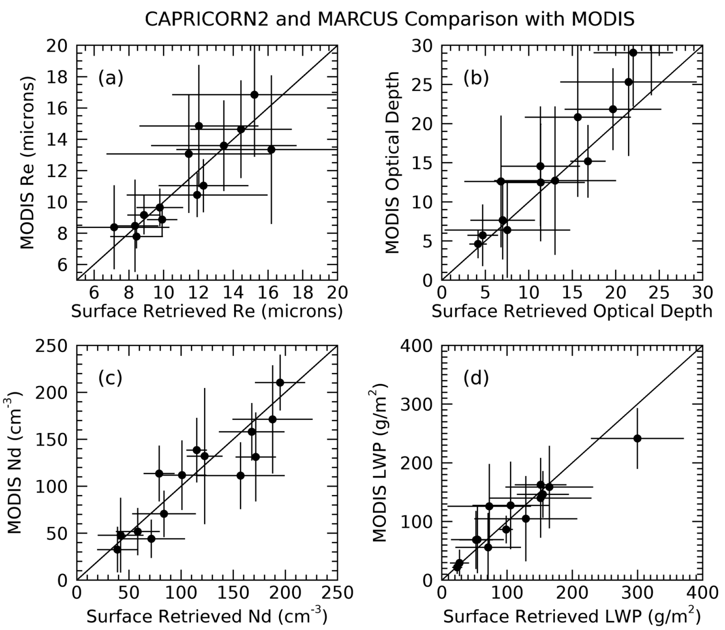

Figure 9. Comparison of MODIS derived cloud properties with cloud properties derived from data collected during the Marcus and Capricorn 2 campaigns in the Southern Ocean during Austral Summer 2018. Error bars are 1 standard deviation of the retrieved cloud properties during the time and over the spatial extent of the two data sets.

4. Discussion

https://doi.org/10.5194/egusphere-2023-2606
Preprint. Discussion started: 21 December 2023

Since we focused our analysis on the 2/18 case study, it seems desirable to explore this case a bit more and illustrate what can be learned from the surface-based remote sensing of cloud – especially when combined with aircraft and satellite data. We find that the aircraft, satellite, and surface-based data sources provide similar and very interesting characterizations of the cloud and CCN on 2/18. Twohy et al. (2021) in their supplemental information show that the airmass above the marine boundary layer on 2/18 had one of the highest sulfur-based concentrations of CCN recorded during Socrates at 224 cm^{-3}. The air mass observed on 2/18 followed a trajectory from the deep south from over the Antarctic continent and the biologically productive waters of the Southern Ocean. The high concentrations of sulfate CCN in the free troposphere imply new particle formation along the trajectory was likely responsible for the high CCN (McCoy et al., 2021). The CCN at the surface measured on the R/V Investigator was near 210 cm^{-3} – slightly lower than that measured on the aircraft.

On the other hand, the N_d seems to be consistently in the 100 cm^{-3} range from the surface, ship, and MODIS except for the near-cloud top maxima in Nd observed by the GV in the ramp demonstrated in Fig. 4. The other ramps (not shown) also had values of N_d near the CCN values of 200-250 cm^{-3}. We speculate that the difference between CCN and N_d is mostly likely due to precipitation droplet scavenging and coalescence process that is actively generating drizzle. The high CCN from the free troposphere transported to this location from the south is likely mixing into the marine boundary layer through entrainment (the cloud top spike in N_d in Fig. 4) and being processed through clouds explaining the lower surface CCN. The cloud properties (N_d in the 100 cm^{-3} range) are a drizzle and coalescence damped response to the high free tropospheric CCN.

This brief case study illustrates what is possible using surface-based remote sensing with instrumentation that has become common – a microwave radiometer, w-band radar, and depolarization lidar. Combined with CCN and other ancillary data sources, we can directly probe the processes that govern the properties of clouds that, in turn, modulate the Earth's albedo and control the sensitivity of the Earth's climate to changes in atmospheric composition.

5. Summary and Conclusions

Given the importance of knowing cloud droplet number concentrations (N_d) in low-level clouds for understanding how these clouds interact with aerosol and precipitation-producing processes to influence the earth's albedo, we have explored two techniques that allow us to derive N_d and layer effective radius (r_e) using surface-based remote sensing techniques with an emphasis on the information brought to this problem by lidar data. The depth a laser pulse penetrates a cloud layer is a function of the amount of water droplet cross-sectional area presented to the laser pulse, and that cross-sectional area is dependent upon the N_d and the liquid water content (q). This observable is quantified by the lidar attenuated backscatter, β_{obs}, (Eqn. 4) that is modulated by the directionality of the scattering as represented by the multiple scattering factor. As the lidar beam penetrates a cloud layer, the signal initially increases until two-way attenuation causes the signal to reach a maximum, after which it decays exponentially depending upon multiple scattering. The rate of increase in β_{obs} is easily quantified if N_d and q are known, or turning the problem around, one can calculate N_d if $\beta_{obs}$is observed, and q is known. The math becomes more tractable where the lidar signal is at a maximum (a distance we term r_{max}) since the rate of change of $\beta_{obs}$is zero (Eqn 9) there. The liquid water content, q, can be expressed in terms of the rate of increase of q with height for an adiabatic cloud which can be made more realistic by scaling the q profile by an adiabaticity factor that can be derived from LWP and cloud layer depth. This simple model (Eqn. 9) can be implemented with an estimated cloud depth, LWP, and a lidar. The effective radius near cloud top can then be derived easily. This simple method, however, is very sensitive to uncertainty in r_{max} which is, in turn, dependent on the vertical resolution of the lidar. Since r_{max}

https://doi.org/10.5194/egusphere-2023-2606
Preprint. Discussion started: 21 December 2023

typically ranges from a few 10's to maybe 100 m, the uncertainty in derived N_d becomes prohibitively large for range resolutions much above 5 m. Most lidars in operational networks, however, have range bin spacings of 15 or more meters. The uncertainty in r_{max} translates predictably into uncertainties in r_e.

To lessen the effects of uncertainty in r_{max}, we bring more information to bear on the problem by quantifying the cloud layer extinction in terms of the rate of decay of the lidar signal beyond r_{max} using a published methodology (Li et al., 2011). In addition, we use the radar reflectivity near cloud top as a constraint on the liquid water content profile and r_e. This is cast in an optimal estimation (OE) algorithm that seeks to balance the uncertainty in the observations and uses prior information such as CCN concentrations that provide an upper limit on N_d. The OE algorithm is only marginally successful in reducing the uncertainty in N_d and r_e. The uncertainties, especially on N_d remain substantial since r_{max} provides the most significant information on N_d and the other measurements provide minimal constraint on N_d as quantified in the Jacobian (K_x) matrix. What we find interesting is that the use of CCN as a prior constraint allows us to balance the information content in r_{max} and the other observations with what we know as a significant constraint on N_d and, therefore r_e. Overall, the OE uncertainties that are shown to be reasonable through a bootstrapping experiment and through comparison to aircraft data, are in the range of just under a factor of 2 for N_d and 30% for r_e for lidar range bins of 10-15 m. The only way to reduce this uncertainty is to have dedicated lidar measurements that have vertical resolution less than about 5 m. Using comparisons with in-situ aircraft data and with cloud properties derived from MODIS, we show that the OE algorithm provides results consistent with the uncertainty in the data and retrievals.

Finally, a case study is explored that shows how synergistic remote sensing data from the surface, especially when combined with aircraft and satellite data, can be exploited. The February 18, 2018 case study that took place in the Southern Ocean near 56 S and 141 E shows how aerosol transport and likely new particle formation from the biologically productive waters of the deep south modulated the cloud properties that existed on this day. The CCN measured at the surface and from the GV aircraft was about a factor of two larger than the ~100 cm^{-3} N_d inferred from the ship and MODIS data and observed by the GV. This difference between N_d and CCN was likely a response to the widespread precipitation processes that were occurring on this day.

6. Competing Interests

The author declares no conflict of interest.

7. Code and Data Availability

All data used in this study are available in public archives. MODIS cloud products can be found for Terra and Aqua at https:/doi.org/10.5067/TERRA/MODIS/L3M/CHL/2018 and http://dx.doi.org/10.5067/MODIS/MYD06_L2.006. ARM data can be obtained at https://www.arm.gov/data/. SOCRATES data are available at https://data.eol.ucar.edu/project/SOCRATES, CAPRICORN 2 data are available at https://doi.org/10.25919/5f688fcc97166. Computer code for this study including all analysis code and graphic generation code is written in the IDL language. Code is available upon request to the corresponding author.

8. Acknowledgments

This work was supported by NASA Grant 80NSSC21k1969, DOE ASR Grants DE-SC00222001 and NSF Grant 2246488. The author would like to thank Sally Benson for her expertise in adapting code and generating figures. Alain Protat enabled the author's participation in the CAPRICORN 2 campaign and provided the cloud radar and lidar data from the R/V Investigator. Ruhi Humphries provided the CCN data filtered for ship exhaust. This Research was supported, and data were obtained from the

Atmospheric Radiation Measurement (ARM) User Facility, a U.S. Department of Energy (DOE) Office of Science user facility managed by the Biological and Environmental Research Program. Technical, logistical, and ship support for MARCUS were provided by the Australian Antarctic Division through Australia Antarctic Science projects 4292 and 4387 and we thank Steven Whiteside, Lloyd Symonds, Rick van den Enden, Peter de Vries, Chris Young and Chris Richards for assistance. The author would like to thank the staff of the Australian Marine National Facility for providing the infrastructure and logistical and financial support for the voyages of the RV Investigator. Funding for these voyages was provided by the Australian Government and the U.S. Department of Energy. All data used in this study are available in public archives.

9. References

Albrecht, B. A. (1989). Aerosols, cloud microphysics, and fractional cloudiness. *Science*, *245*(4923), 1227–1230. https://doi.org/10.1126/science.245.4923.1227

Austin, R. T., & Stephens, G. L. (2001). Retrieval of stratus cloud microphysical parameters using millimeter-wave radar and visible optical depth in preparation for CloudSat: 1. algorithm formulation. *Journal of Geophysical Research: Atmospheres*, *106*(D22), 28233–28242. https://doi.org/10.1029/2000jd000293

Bai, H., Gong, C., Wang, M., Zhang, Z., & L'Ecuyer, T. (2018). Estimating precipitation susceptibility in warm marine clouds using multi-sensor aerosol and cloud products from A-Train Satellites. *Atmospheric Chemistry and Physics*, *18*(3), 1763–1783. https://doi.org/10.5194/acp-18-1763-2018

Cooper, S. J., L'Ecuyer, T. S., Gabriel, P., Baran, A. J., & Stephens, G. L. (2006). Objective assessment of the information content of visible and infrared radiance measurements for cloud microphysical property retrievals over the global oceans. part II: Ice clouds. *Journal of Applied Meteorology and Climatology*, *45*(1), 42–62. https://doi.org/10.1175/jam2327.1

Foskinis, R., Nenes, A., Papayannis, A., Georgakaki, P., Eleftheriadis, K., Vratolis, S., Gini, M. I., Komppula, M., Vakkari, V., & Kokkalis, P. (2022). Towards reliable retrievals of cloud droplet number for non-precipitating planetary boundary layer clouds and their susceptibility to aerosol. *Frontiers in Remote Sensing*, *3*. https://doi.org/10.3389/frsen.2022.958207

Gettelman, A., & Morrison, H. (2015). Advanced Two-moment bulk microphysics for global models. part I: Off-line tests and comparison with other schemes. *Journal of Climate*, *28*(3), 1268–1287. https://doi.org/10.1175/jcli-d-14-00102.1

Grosvenor, D. P., Sourdeval, O., Zuidema, P., Ackerman, A., Alexandrov, M. D., Bennartz, R., Boers, R., Cairns, B., Chiu, J. C., Christensen, M., Deneke, H., Diamond, M., Feingold, G., Fridlind, A., Hünerbein, A., Knist, C., Kollias, P., Marshak, A., McCoy, D., … Quaas, J. (2018). Remote sensing of droplet number concentration in warm clouds: A review of the current state of knowledge and perspectives. *Reviews of Geophysics*, *56*(2), 409–453. https://doi.org/10.1029/2017rg000593

Heintzenberg, J., & Charlson, R. J. (2009). *Clouds in the perturbed climate system: Their relationship to energy balance, atmospheric dynamics, and precipitation*. MIT Press.

Hu, Y., Vaughan, M., McClain, C., Behrenfeld, M., Maring, H., Anderson, D., Sun-Mack, S., Flittner, D., Huang, J., Wielicki, B., Minnis, P., Weimer, C., Trepte, C., & Kuehn, R. (2007). Global statistics of liquid water content and effective number concentration of water clouds over ocean derived from combined Calipso and Modis measurements. *Atmospheric Chemistry and Physics*, *7*(12), 3353–3359. https://doi.org/10.5194/acp-7-3353-2007

Hu, Yongxiang, Lu, X., Zhai, P.-W., Hostetler, C. A., Hair, J. W., Cairns, B., Sun, W., Stamnes, S., Omar, A., Baize, R., Videen, G., Mace, J., McCoy, D. T., McCoy, I. L., & Wood, R. (2021). Liquid phase cloud microphysical property estimates from Calipso Measurements. *Frontiers in Remote Sensing*, *2*. https://doi.org/10.3389/frsen.2021.724615

Hu, Yongxiang, Winker, D., Vaughan, M., Lin, B., Omar, A., Trepte, C., Flittner, D., Yang, P., Nasiri, S. L., Baum, B., Holz, R., Sun, W., Liu, Z., Wang, Z., Young, S., Stamnes, K., Huang, J., & Kuehn, R. (2009). Calipso/Caliop Cloud Phase

Discrimination algorithm. *Journal of Atmospheric and Oceanic Technology*, *26*(11), 2293–2309. https://doi.org/10.1175/2009jtecha1280.1

Kollias, P., Miller, M. A., Luke, E. P., Johnson, K. L., Clothiaux, E. E., Moran, K. P., Widener, K. B., & Albrecht, B. A. (2007). The Atmospheric Radiation Measurement Program Cloud profiling radars: Second-generation sampling strategies, processing, and Cloud Data Products. *Journal of Atmospheric and Oceanic Technology*, *24*(7), 1199–1214. https://doi.org/10.1175/jtech2033.1

Kollias, P., Puigdomènech Treserras, B., & Protat, A. (2019). *Calibration of the 2007–2017 Record of ARM Cloud Radar Observations Using CloudSat*. https://doi.org/10.5194/amt-2019-34

Lawson, P., Gurganus, C., Woods, S., & Bruintjes, R. (2017). Aircraft observations of Cumulus microphysics ranging from the tropics to midlatitudes: Implications for a "new" Secondary ice process. *Journal of the Atmospheric Sciences*, *74*(9), 2899–2920. https://doi.org/10.1175/jas-d-17-0033.1

Li, J., Hu, Y., Huang, J., Stamnes, K., Yi, Y., & Stamnes, S. (2011). A new method for retrieval of the extinction coefficient of water clouds by using the tail of the CALIOP signal. *Atmospheric Chemistry and Physics*, *11*(6), 2903–2916. https://doi.org/10.5194/acp-11-2903-2011

L'Ecuyer, T. S., Gabriel, P., Leesman, K., Cooper, S. J., & Stephens, G. L. (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. *Journal of Applied Meteorology and Climatology*, *45*(1), 20–41. https://doi.org/10.1175/jam2326.1

Maahn, M., Turner, D. D., Löhnert, U., Posselt, D. J., Ebell, K., Mace, G. G., & Comstock, J. M. (2020). Optimal estimation retrievals and their uncertainties: What every atmospheric scientist should know. *Bulletin of the American Meteorological Society*, *101*(9). https://doi.org/10.1175/bams-d-19-0027.1

Mace, G. G., Protat, A., Humphries, R. S., Alexander, S. P., McRobert, I. M., Ward, J., Selleck, P., Keywood, M., & McFarquhar, G. M. (2021). Southern ocean cloud properties derived from Capricorn and Marcus Data. *Journal of Geophysical Research: Atmospheres*, *126*(4). https://doi.org/10.1029/2020jd033368

Mather J. 2021. ARM User Facility 2020 Decadal Vision. Ed. by Rolanda Jundt, Robert Stafford, Stacy Larsen, U.S. Department of Energy. DOE/SC-ARM-20-014. 10.2172/1782812.

McFarquhar, G. M., Bretherton, C. S., Marchand, R., Protat, A., DeMott, P. J., Alexander, S. P., Roberts, G. C., Twohy, C. H., Toohey, D., Siems, S., Huang, Y., Wood, R., Rauber, R. M., Lasher-Trapp, S., Jensen, J., Stith, J. L., Mace, J., Um, J., Järvinen, E., … McDonald, A. (2021). Observations of clouds, aerosols, precipitation, and surface radiation over the Southern Ocean: An overview of Capricorn, Marcus, MICRE, and socrates. *Bulletin of the American Meteorological Society*, *102*(4). https://doi.org/10.1175/bams-d-20-0132.1

McCoy, I. L., Bretherton, C. S., Wood, R., Twohy, C. H., Gettelman, A., Bardeen, C. G., & Toohey, D. W. (2021). Influences of recent particle formation on Southern Ocean aerosol variability and low cloud properties. *Journal of Geophysical Research: Atmospheres*, *126*(8). https://doi.org/10.1029/2020jd033529

Miles, N. L., Verlinde, J., & Clothiaux, E. E. (2000). Cloud droplet size distributions in low-level stratiform clouds. *Journal of the Atmospheric Sciences*, *57*(2), 295–311. https://doi.org/10.1175/1520-0469(2000)057<0295:cdsdil>2.0.co;2

Miller, M. A., Jensen, M. P., & Clothiaux, E. E. (1998). Diurnal cloud and thermodynamic variations in the stratocumulus transition regime: A case study using in situ and remote sensors. *Journal of the Atmospheric Sciences*, *55*(13), 2294–2310. https://doi.org/10.1175/1520-0469(1998)055<2294:dcatvi>2.0.co;2

https://doi.org/10.5194/egusphere-2023-2606
Preprint. Discussion started: 21 December 2023

Nakajima, T., & King, M. D. (1990). Determination of the optical thickness and effective particle radius of clouds from reflected solar radiation measurements. part I: Theory. *Journal of the Atmospheric Sciences*, *47*(15), 1878–1893. https://doi.org/10.1175/1520-0469(1990)047<1878:dotota>2.0.co;2

Painemal, D., & Zuidema, P. (2013). The first aerosol indirect effect quantified through airborne remote sensing during vocals-rex. *Atmospheric Chemistry and Physics*, *13*(2), 917–931. https://doi.org/10.5194/acp-13-917-2013

Platnick, S., & Twomey, S. (1994). Determining the susceptibility of cloud albedo to changes in droplet concentration with the advanced very high resolution radiometer. *Journal of Applied Meteorology*, *33*(3), 334–347. https://doi.org/10.1175/1520-0450(1994)033<0334:dtsoca>2.0.co;2

Platnick, S., King, M. D., Ackerman, S. A., Menzel, W. P., Baum, B. A., Riedi, J. C., & Frey, R. A. (2003). The Modis Cloud Products: Algorithms and examples from Terra. *IEEE Transactions on Geoscience and Remote Sensing*, *41*(2), 459–473. https://doi.org/10.1109/tgrs.2002.808301

Platnick, Steven, & Oreopoulos, L. (2008). Radiative susceptibility of cloudy atmospheres to droplet number perturbations: 1. theoretical analysis and examples from Modis. *Journal of Geophysical Research*, *113*(D14). https://doi.org/10.1029/2007jd009654

Platt, C. M. (1977). Lidar observation of a mixed-phase altostratus cloud. *Journal of Applied Meteorology*, *16*(4), 339–345. https://doi.org/10.1175/1520-0450(1977)016<0339:looamp>2.0.co;2

Platt, C. M., Winker, D. M., Vaughan, M. A., & Miller, S. D. (1999). Backscatter-to-extinction ratios in the top layers of tropical mesoscale convective systems and in isolated cirrus from Lite Observations. *Journal of Applied Meteorology*, *38*(9), 1330–1345. https://doi.org/10.1175/1520-0450(1999)038<1330:bterit>2.0.co;2

Posselt, D. J., & Mace, G. G. (2014). MCMC-based assessment of the error characteristics of a surface-based combined radar–Passive Microwave Cloud Property Retrieval. *Journal of Applied Meteorology and Climatology*, *53*(8), 2034–2057. https://doi.org/10.1175/jamc-d-13-0237.1

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

Seifert, A., & Beheng, K. D. (2005). A two-moment cloud microphysics parameterization for mixed-phase clouds. part 1: Model description. *Meteorology and Atmospheric Physics*, *92*(1–2), 45–66. https://doi.org/10.1007/s00703-005-0112-4

Simmons, J. B., Humphries, R. S., Wilson, S. R., Chambers, S. D., Williams, A. G., Griffiths, A. D., McRobert, I. M., Ward, J. P., Keywood, M. D., & Gribben, S. (2021). Summer aerosol measurements over the East Antarctic Seasonal Ice Zone. *Atmospheric Chemistry and Physics*, *21*(12), 9497–9513. https://doi.org/10.5194/acp-21-9497-2021

Stephens, G. L. (1978). Radiation profiles in extended water clouds. I: Theory. *Journal of the Atmospheric Sciences*, *35*(11), 2111–2122. https://doi.org/10.1175/1520-0469(1978)035<2111:rpiewc>2.0.co;2

Stevens, B., & Feingold, G. (2009). Untangling aerosol effects on clouds and precipitation in a buffered system. *Nature*, *461*(7264), 607–613. https://doi.org/10.1038/nature08281

Szczap, F., Alkasem, A., Shcherbakov, V., Schmisser, R., Blanc, J., Mioche, G., Gour, Y., Cornet, C., Banson, S., & Bray, E. (2023). Computation of the attenuated backscattering coefficient by the backscattering LIDAR Signal Simulator (bliss) in the framework of the CALIOP/calipso observations. *Atmosphere*, *14*(2), 249. https://doi.org/10.3390/atmos14020249

Tanelli, S., Durden, S. L., Im, E., Pak, K. S., Reinke, D. G., Partain, P., Haynes, J. M., & Marchand, R. T. (2008). CloudSat's cloud profiling radar after two years in orbit: Performance, calibration, and processing. *IEEE Transactions on Geoscience and Remote Sensing*, *46*(11), 3560–3573. https://doi.org/10.1109/tgrs.2008.2002030

Thompson, G., & Eidhammer, T. (2014). A study of aerosol impacts on clouds and precipitation development in a large winter cyclone. *Journal of the Atmospheric Sciences*, *71*(10), 3636–3658. https://doi.org/10.1175/jas-d-13-0305.1

https://doi.org/10.5194/egusphere-2023-2606
Preprint. Discussion started: 21 December 2023

Turner, D. D., Kneifel, S., & Cadeddu, M. P. (2016). An improved liquid water absorption model at microwave frequencies for supercooled liquid water clouds. *Journal of Atmospheric and Oceanic Technology*, *33*(1), 33–44. https://doi.org/10.1175/jtech-d-15-0074.1

Twomey, S. (1974). Pollution and the planetary albedo. *Atmospheric Environment (1967)*, *8*(12), 1251–1256. https://doi.org/10.1016/0004-6981(74)90004-3

Zhao, Y., Mace, G. G., & Comstock, J. M. (2011). The occurrence of particle size distribution bimodality in Midlatitude Cirrus as inferred from ground-based remote sensing data. *Journal of the Atmospheric Sciences*, *68*(6), 1162–1177. https://doi.org/10.1175/2010jas3354.1