Reduction of Bias from Parameter Variance in Geophysical Data Estimation: Method and Application to Ice Water Content and Sedimentation Flux Estimated from Lidar

MAXIMILIEN BOLOT AND STEPHAN FUEGLISTALER

Princeton University, Princeton, New Jersey

(Manuscript received 24 April 2019, in final form 16 December 2019)

ABSTRACT

This paper addresses issues of statistical misrepresentation of the a priori parameters (henceforth called ancillary parameters) used in geophysical data estimation. Parameterizations using ancillary data are frequently needed to derive geophysical data of interest from remote measurements. Empirical fits to the ancillary data that do not preserve the distribution of such data may induce substantial bias. A semi-analytical averaging approach based on Taylor expansion is presented to improve estimated cirrus ice water content and sedimentation flux for a range of volume extinction coefficients retrieved from spaceborne lidar observations by CALIOP combined with the estimated distribution of ancillary data from in situ aircraft measurements of ice particle microphysical parameters and temperature. It is shown that, given an idealized distribution of input parameters, the approach performs well against Monte Carlo benchmark predictions. Using examples with idealized distributions at the mean temperature for the tropics at 15 km, it is estimated that the commonly neglected variance observed in in situ measurements of effective diameters may produce a worst-case estimation bias spanning up to a factor of 2. For ice sedimentation flux, a similar variance in particle size distributions and extinctions produces a worst-case estimation bias of a factor of 9. The value of the bias is found to be mostly set by the correlation coefficient between extinction and ice effective diameter, which in this test ranged between all possible values. Systematic reporting of variances and covariances in the ancillary data and between data and observed quantities would allow for more accurate observational estimates.

1. Introduction

A common problem in many fields of atmospheric sciences, here discussed for the parameterization of ice water content and sedimentation flux from lidar observations, is that remote measurements often have to be supplemented with a priori information (henceforth also referred to as ancillary data) to estimate geophysical quantities of interest. This can create a challenge for operations involving statistics on measurements—such as estimating bulk properties for a whole range of measurements (e.g., a regional mean)—since such operations require knowledge of the joint distribution of measurements and parameters representing the a priori data. The main difficulty lies in matching the distribution of measurements—typically globally available and obtained from satellite—and the distribution of ancillary parameters—usually coming from limited in situ measurements. Another difficulty consists in computing average quantities of interest by integrating parameterizations over joint distributions of measurements and ancillary parameters, since such integrations are often not analytically tractable and parameterizations are often nonlinear. This paper addresses those difficulties in the specific cases of ice water content and sedimentation flux parameterized from space-based extinction coefficients retrievals, although the methods employed may be applied to other remotely sensed observations that rely on ancillary data for parameterization.

Ice water content and sedimentation flux in upper-tropospheric cirrus clouds are important to characterize at global scale in order to inform climate models of the impact of convection and in situ cloud formation on the flux of water vapor to the stratosphere (Holton and Gettelman 2001; Dessler et al. 2007; Fueglistaler et al. 2009) and in order to compute the radiative impact of such clouds on global climate (Lee et al. 2009; Yang et al. 2010). To ensure sufficient resolution and global coverage, those quantities can be estimated from volume extinction coefficients at 532 nm retrieved from spaceborne

Corresponding author: Maximilien Bolot, mbolot@princeton.edu

DOI: 10.1175/JAS-D-19-0106.1

lidar measurements, but the estimation must be informed on particle microphysical properties via a finite number of extra parameters, such as ice effective diameter. These parameters are not acquired with the main measurement system (at least for optically thin clouds that can only be detected by lidar) and are instead obtained from limited in situ measurements. Often [e.g., in Cloud–Aerosol Lidar with Orthogonal Polarization (CALIOP), version 4.1 (V4.1), IWC data product, described in Heymsfield et al. (2014)], the parameters are specified using fits to in situ measurements, which leads to considering only the center of observed parameter distributions and discarding parameter variance. In the case of ice effective diameter, for instance, the parameterization used in CALIOP V4.1 effectively predicts a distribution that reduces to a single number at any given temperature. The consequences of neglecting distributional aspects of parameters on statistics of ice water content and sedimentation flux must be examined, as well as rectification strategies to address potential biases. In essence, the error coming from neglecting to report the variations derived from ancillary parameters—such as effective diameter—may be greater than the error attached to the choice of parameterization itself.

More generally, there is evidence that unresolved variance addressed with mean value parameterizations can be a problem for the estimation of cloud properties. In situ measurements on board research aircrafts routinely resolve variability on scales as fine as 250 m. Multiple studies pertaining to statistical cloud modeling have examined the distributions of bulk cloud properties and produced modified microphysics formula accounting for subgrid-scale cloud variability (Pincus and Klein 2000; Golaz et al. 2002; Tompkins 2002; Larson et al. 2005; Larson and Griffin 2013). Several recent studies are moving beyond bulk properties by characterizing variability at the particle size distribution level (McFarquhar et al. 2015; Jackson et al. 2015). Given a parametric description of size distributions in terms of gamma functions, they propose that families of particle size distributions be characterized by ellipsoids of feasible parameters. This approach can be used in conjunction with parameterizations to assess impacts on retrieved mean cloud properties since it provides a distribution of parameters that can be sampled from. For instance, McFarquhar et al. (2003) shows that computing average cloud radiative forcing from a single simulation using the most likely parameters instead of a series of simulation randomly sampling the parameter space can produce a bias of several watts per square meter.

To assess the error from ignoring parameter variability, we need to average properties over a distribution representing the correct sampling between observations and ancillary parameters. The problem is akin to finding the average of $f(\mathbf{x})$, where f is a parameterization depending on the distributed variables $\mathbf{x}$, over a distribution function $p(\mathbf{x})$ representing cloud variability:

$$\langle f(\mathbf{x})\rangle = \int f(\mathbf{x})p(\mathbf{x})\,d\mathbf{x}.$$

Note that the problem is quite generic and that $\langle f(\mathbf{x})\rangle$ can have various meanings in this context. For instance, $p(\mathbf{x})$ may represent a joint distribution between observable and ancillary parameters and $\langle f(\mathbf{x})\rangle$ estimates an average value for a range of observations, such as the tropical mean value of ice water content and sedimentation flux taken as example in this paper. In other cases, the sampling may be stratified by observations and $p(\mathbf{x})$ represents the probability of ancillary parameters conditional upon the values of observations. In all cases, the method of averaging over cloud variability is formally the same.

Similar problems of averaging have been discussed in the literature, particularly in the context of statistical upscaling of local microphysics parameterizations. In global simulations of the atmosphere, spatial and temporal resolutions are limited by available computational power and stability considerations of numerical schemes. Parameterizations are thus needed to relate the effects of subgrid-scale variability on the evolution of grid-mean quantities such as microphysical process rates (Zhang 2002; Larson et al. 2005; Larson and Griffin 2013; Chowdhary et al. 2015). Much work has been devoted in particular to systematic errors from neglect of subgrid-scale variability on autoconversion parameterization (Pincus and Klein 2000; Rotstayn 2000; Larson et al. 2001). The errors resulting from ignoring subgrid variability are larger when the processes are nonlinear and the variability is large. For instance, cloud reflectance is systematically underestimated when small-scale variability in cloud optical thickness is neglected (Cahalan et al. 1994) because albedo is a convex function of cloud optical thickness.

A number of strategies have been devised to compute the kind of integral appearing in the definition of $\langle f(\mathbf{x})\rangle$. In some cases, the integral is analytically tractable and $\langle f(\mathbf{x})\rangle$ can be exactly computed (see, e.g., Zhang 2002; Morrison and Gettelman 2008; Larson and Griffin 2013). Unfortunately, the analytic approach is often not feasible except when the structure of the function f and the parameter distribution $p(\mathbf{x})$ are relatively simple. Methods based on random sampling such as Monte Carlo integration are more universally applicable (Pincus et al. 2003; Räisänen and Barker 2004). In that approach, f is evaluated over sample points randomly selected from $p(\mathbf{x})$ and the results are averaged. This method

has the advantage of being widely applicable and is straightforward to implement even when f is a complex computer procedure. However, it exhibits slow convergence and produces statistical noise due to finite size sampling. Some shortcomings of the Monte Carlo method can be mitigated by variance reduction techniques using a more efficient sampling strategy (Larson et al. 2005). Another sampling method, albeit not random, is that of deterministic numerical quadrature. In that approach, random sampling points are replaced by tailored quadrature points and weights. This technique typically achieves greater accuracy with fewer samples than random sampling approaches (Chowdhary et al. 2015).

Here we propose to estimate $\langle f(\mathbf{x})\rangle$ using a semianalytic approach that widens the domain of applicability of analytic integration techniques when the variations of f or $\ln(f)$ can be captured by a low-order Taylor expansion [a similar approach was also used by Griffin and Larson (2013)]. We arrive at an estimator of $\langle f(\mathbf{x})\rangle$ that is analytic and bears a clear relationship to the structure of the function via its gradient and Hessian as well as to the structure of the parameter distribution through its covariance matrix. This is in contrast to sampling methods where $\langle f(\mathbf{x})\rangle$ is an often opaque result of a numerical procedure.

This paper is organized as follows. Section 2 presents the estimation of ice water content and sedimentation flux from volume extinction coefficients retrievals and formulates the need for a finite number of ancillary parameters informing on particle size distribution. The problem of estimating cloud properties for a range of extinction coefficients is also introduced there. Section 3 presents a generic solution (when applicable) to the problem of averaging over parameter distributions representing cloud variability. The solution involves a semianalytic estimator based on Taylor expansion of the parameterizations. Section 4 discusses the application of the generic averaging method to the parameterizations of ice water content and sedimentation flux presented in section 2. Section 5 validates the estimators for ice water content and sedimentation flux against values computed using the Monte Carlo approach. Section 6 discusses estimation biases from ignoring the variance in particle size distributions and associated rectification strategies.

2. Estimating ice water content and sedimentation flux from lidar information in the upper troposphere

In this section, we describe how extinction coefficients retrieved from space lidar may be used to estimate ice water content and sedimentation flux in ice clouds in the upper troposphere. For that purpose, the information contained in extinction coefficients (henceforth referred to as "observations") must be completed by additional information in the form of extra parameters (henceforth referred to as ancillary parameters). We show that the ancillary parameters can be chosen as the ice effective diameter and—for sedimentation flux—the slope parameter of gamma functions used to describe particle size distribution. For ice water content, effective diameter parameterizes the ratio between bulk particulate volume and area, irrespective of particle size distributions and shapes. For the sedimentation flux, however, a consistent estimation requires additional information on particle size distributions and shapes, which warrants the use of an extra parameter and the explicit assumption of gamma-shaped particle sizes distributions. We then discuss the problem of estimating ice water content and sedimentation flux for an ensemble of clouds characterized by a range of extinction coefficients, which is the specific topic of this paper. We show in particular that one must correctly match the distribution of cloud microphysics properties to that of extinction coefficients in order to produce unbiased estimation of cloud properties.

a. *From extinction coefficients to ice water content and sedimentation flux*

The vertical structure and properties of ice clouds in the upper troposphere can be characterized on a global scale by measurements from space lidars. The CALIOP instrument, on board the *Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations* satellite (*CALIPSO*) is a dual-polarization elastic backscatter lidar whose signal penetrates clouds up to 3–5 visible optical depths (Winker et al. 2009; Hunt et al. 2009). CALIOP measures profiles of attenuated backscatter from which cloud features are identified by applying a threshold technique. Profiles of extinction coefficients at 532 nm are then retrieved from attenuated backscatter using constraints on layer transmittance when possible (Young et al. 2018) and corrections for multiple scattering (Garnier et al. 2015). The retrieved extinction coefficients used in this paper are from version 4.1 retrieval algorithm and are taken from the 5-km cloud profile product, where data from multiple laser shots are averaged together to enhance signal, yielding an effective horizontal resolution of 5 km along track.

The primary "observations" used in this work are the extinction coefficients retrieved from attenuated backscatter. For an ensemble of ice particles, the volume extinction coefficient measures the extinction cross section integrated over the particle size distribution:

$$\mathrm{EXT} = \int_0^\infty N(D)\pi\frac{D^2}{4}A_r(D)Q_{\mathrm{ext}}(D)\,dD, \qquad (1)$$

where D is particle size, N is the size distribution function, A_r is the ratio of projected particle area to that of a

disk of same diameter, and Q_{ext} is extinction efficiency. In the geometric optics approximation, which is approximately valid for 532-nm monochromatic radiation interacting with micron-sized and larger ice particles, extinction efficiency asymptotically approaches 2 (Van de Hulst 1957). The extinction coefficient at 532 nm therefore measures twice the cross-sectional area of the particle distribution.

The geophysical quantities of interest investigated in this work are ice water content and sedimentation flux, which are defined respectively by integrating mass and downward momentum over the particle size distributions:

$$\text{IWC} = \int_0^\infty N(D)m(D)\,dD, \tag{2}$$

$$F_{\text{sed}} = \int_0^\infty N(D)m(D)v_t(D)\,dD, \tag{3}$$

where m is particle mass and v_t is particle terminal velocity.

To solve the integrals in Eqs. (1)–(3), one needs to parameterize particle number, area ratio, mass, and terminal velocity according to particle size. Following previous studies on high-altitude cirrus clouds (Schmitt and Heymsfield 2009; Heymsfield and Westbrook 2010; Heymsfield et al. 2013), we assume that the ice particle size distribution follows a gamma function and write

$$N(D) = N_0 D^\mu \exp(-\lambda D), \tag{4}$$

where N_0 is the intercept of the distribution, μ the dispersion parameter, and λ the slope parameter. Note that μ and λ are not independent and exhibit a tight relationship at any altitude (Schmitt and Heymsfield 2009; Heymsfield et al. 2013; McFarquhar et al. 2015). N_0 acts as prefactor in Eq. (4) and scales with total particle number, extinction, ice water content, and sedimentation flux. Particle area ratio, particle mass and terminal velocity are parameterized using power laws of the form

$$A_r(D) = \alpha D^\beta, \tag{5}$$

$$m(D) = aD^b, \tag{6}$$

$$v_t(D) = rD^s, \tag{7}$$

where α, β, a, b, r, and s are obtained through a combination of theory and in situ measurements in cirrus clouds (Schmitt and Heymsfield 2009; Heymsfield et al. 2013; Heymsfield and Westbrook 2010; see appendix B).

Using Eqs. (4)–(7), one can resolve the integrals in Eqs. (1)–(3) in terms of gamma functions. This yields the parameterizations

$$\text{EXT}(\lambda,\mu,N_0) = N_0\frac{\pi}{2}\alpha\lambda^{-(\mu+\beta+3)}\Gamma(\mu+\beta+3), \tag{8}$$

$$\text{IWC}(\lambda,\mu,N_0) = N_0 a\lambda^{-(\mu+b+1)}\Gamma(\mu+b+1), \tag{9}$$

$$F_{\text{sed}}(\lambda,\mu,N_0) = N_0 ar\lambda^{-(\mu+b+s+1)}\Gamma(\mu+b+s+1), \tag{10}$$

where Γ is the gamma function.

Replacing Eq. (8) into Eqs. (9) and (10), one can therefore rewrite ice water content and ice sedimentation flux as a function of extinction:

$$\text{IWC}(\lambda,\mu,\text{EXT}) = \text{EXT}\frac{2a}{\pi\alpha}\lambda^{(\beta-b+2)}\frac{\Gamma(\mu+b+1)}{\Gamma(\mu+\beta+3)}, \tag{11}$$

$$F_{\text{sed}}(\lambda,\mu,\text{EXT}) = \text{EXT}\frac{2ar}{\pi\alpha}\lambda^{(\beta-b-s+2)}\frac{\Gamma(\mu+b+s+1)}{\Gamma(\mu+\beta+3)}. \tag{12}$$

Equations (11) and (12) allow to formally estimate ice water content and sedimentation flux from retrievals of extinction coefficients supplemented with the knowledge of ancillary parameters λ and μ. These parameters inform on the normalized form of the particle size distribution [cf. Eq. (4)].

b. *Transformation of parameters*

As long as information on normalized particle size distribution remains conveyed over two degrees of freedom, one can always change variables inside the set of ancillary parameters. For reasons that will be explained, we perform the change of independent variable $\mu \rightarrow D_e$, where D_e is the effective diameter of the ice particle size distribution (Foot 1988; McFarquhar and Heymsfield 1998):

$$D_e = \frac{3}{\rho_i}\frac{\text{IWC}}{\text{EXT}}. \tag{13}$$

Using the definitions of ice water content and extinction, one can write the relationship between variables as

$$D_e(\lambda,\mu) = \frac{6a}{\rho_i\pi\alpha}\lambda^{(\beta-b+2)}\frac{\Gamma(\mu+b+1)}{\Gamma(\mu+\beta+3)}. \tag{14}$$

Equation (14) must be numerically inverted to yield λ or μ as function of the remaining variables. In the new coordinate system, ice water content and sedimentation flux are reparameterized as follows:

$$\text{IWC}(D_e,\text{EXT}) = \text{EXT}\frac{\rho_i}{3}D_e, \tag{15}$$

$$F_{\text{sed}}(\lambda,D_e,\text{EXT}) = \text{EXT}\frac{2ar}{\pi\alpha}\lambda^{(\beta-b-s+2)} \times \frac{\Gamma[\mu(\lambda,D_e)+b+s+1]}{\Gamma[\mu(\lambda,D_e)+\beta+3]}, \tag{16}$$

where $\mu(\lambda, D_e)$ designates the value of μ inverted from Eq. (14).

The reparameterization leading to Eqs. (15) and (16) is done on physical ground. Ice effective diameter is one of the most important properties of cirrus clouds, both from a mass distribution and radiative perspective, and has been sufficiently measured to warrant its use as a modeling parameter. The use of effective diameter in Eq. (15) essentially converts from area to volume information for an ensemble of ice particles whose characteristics correspond to those observed in situ in the upper troposphere. Equation (15) is used as a basis for the retrieval of ice water content from spaceborne lidar extinction retrievals in conjunction with temperature-dependent parameterizations of D_e (Heymsfield et al. 2014). One goal of this paper is thus to quantify the inherent bias of such algorithms in conditions where D_e (ancillary information) and EXT (measured by lidar) may be correlated.

The parameterization of sedimentation flux at Eq. (16) is consistent with that of ice water content at Eq. (15). That is, Eqs. (15) and (16) are formally equivalent to Eqs. (11), (12), and (14). Modeling sedimentation flux as a function of ice effective diameter also makes physical sense. In situ measurements show a compact relationship between mass-weighted terminal velocity v_m and D_e, therefore a compact relationship must also exist between F_{sed} and D_e (Schmitt and Heymsfield 2009). This suggests that the variations of sedimentation flux are essentially controlled by D_e and EXT in the formulation of Eq. (16), with a lesser dependency upon λ (cf. section 5 for further discussion). The reparameterization of sedimentation flux may thus allow for a simpler modeling.

c. *The problem of estimating cloud properties for a range of extinction values*

We now consider the problem of estimating ice water content and sedimentation flux for an ensemble of cloud particles with a distribution of extinction coefficients established from spaceborne lidar observations. Cloud variability must manifest itself not only through a range of extinction coefficients, but also through a range of microphysical properties and a corresponding distribution of parameters such as the effective diameter. The mean value associated with a range of extinction coefficients must therefore be computed by integration over the joint distribution of extinction and effective diameter $p(\text{EXT}, D_e)$, which, in the case of ice water content, can be written

$$\langle \text{IWC} \rangle = \int \text{EXT} \frac{\rho_i}{3} D_e p(\text{EXT}, D_e)\, d\text{EXT}\, dD_e, \qquad (17)$$

where the angle brackets represent the averaging operator. An important difficulty in resolving Eq. (17) is the specification of the joint distribution $p(\text{EXT}, D_e)$ since extinction retrievals come from satellite measurements and parameters such as D_e typically come from (limited) in situ measurements.

In the current CALIOP V4.1 approach to parameterizing ice water content, the joint distribution of extinction and effective diameter is such that the range of effective diameters reduces to a single number at each temperature. This is because D_e is parameterized using an empirical relationship obtained from fitting of in situ measurements (Heymsfield et al. 2014). The parameterization, which uses Eq. (9e) of Heymsfield et al. (2014), yields $\ln(D_e)$ as a linear function of temperature and the variability of D_e at each temperature, shown in Fig. 11a of that paper, is ignored. Therefore, the joint distribution of extinction and effective diameter at a given temperature in the V4.1 parameterization is of the form

$$p(\text{EXT}, D_e) = p(\text{EXT})\delta[D_e - D_e^{\text{param}}(T)]. \qquad (18)$$

In the above expression, $D_e^{\text{param}}(T)$ represents the parameterized value as a function of temperature, $p(\text{EXT})$ represents the distribution of volume extinction coefficients retrieved from CALIOP measurements at that temperature and $\delta[D_e - D_e^{\text{param}}(T)]$ (where δ is the Dirac function) expresses the fact that all effective diameters collapse to a single number. This situation is illustrated in Fig. 1a and one can see that the joint distribution at Eq. (18) would predict a mean ice water content value of $1.28 \times 10^{-6}\,\text{kg m}^{-3}$ (for a mean temperature such as found at 15 km, see caption for more details).

In reality, one would expect ice effective diameter to fluctuate for a given sampling temperature. Those fluctuations would result in an observed distribution of effective diameters with nonzero variance that can be estimated at each temperature. In those conditions, and assuming lognormality for $p(D_e)$ (see section 4), one would expect the logarithms of effective diameters to be normally distributed with a mean value parameterized by $\ln[D_e^{\text{param}}(T)]$ and some nonzero amount of variance. Given that extinction values are observed to be approximately lognormally distributed at any given altitude (see section 4), the joint distribution of the logarithms of extinction and effective diameter can be reasonably expected as a bivariate normal distribution with unknown correlation. This situation is illustrated in Fig. 1b for a plausible variance of $\ln(D_e)$ of 0.25 and a hypothetical correlation coefficient of 0.5 for illustrative purpose. The corresponding joint distribution $p(\text{EXT}, D_e)$ would result in a mean ice water content value of $1.61 \times 10^{-6}\,\text{kg m}^{-3}$, significantly higher than predicted with the current parameterization used in CALIOP V4.1. The exact value depends of course on the unknown correlation between extinction and effective diameter. Note also that this analysis assumes a simple

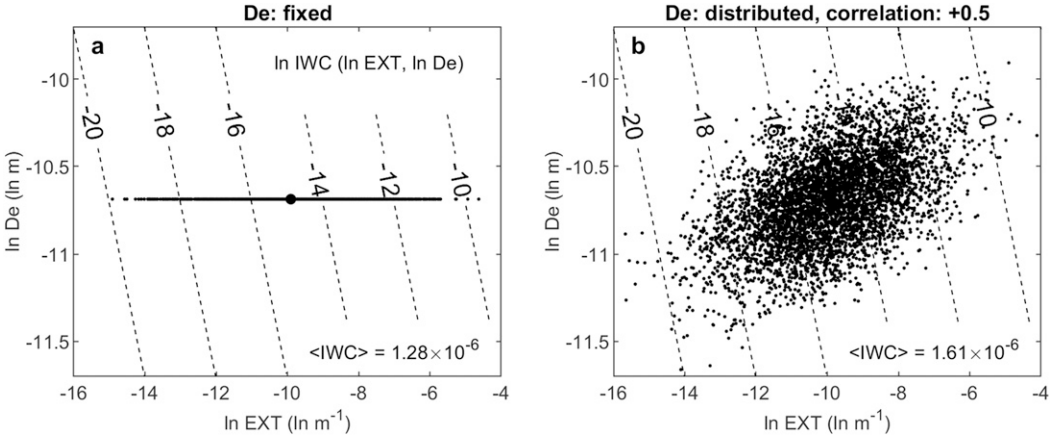

FIG. 1. (a) Scatterplot from the joint distribution of extinction coefficients and effective diameter at −71°C as calculated using Eq. (9e) of Heymsfield et al. (2014). The mean temperature was calculated from interpolated MERRA-2 data at 15 km in the tropics. A normal distribution is chosen for ln(EXT) based on the mean and variance of extinction measured by CALIOP at 15 km (see Fig. 3). Because ln(D_e) is a simple linear function of temperature in the CALIOP V4 parameterization, the distribution of ln(D_e) reduces to a single number. Contour lines of calculated IWC are represented in the background as dashed lines. Integration over the distribution yields a mean IWC of 1.28×10^{-6} kg m^{-3}. (b) As in (a), but for a ln(D_e) distribution estimated from Fig. 11a of Heymsfield et al. (2014) at −71°C. The standard deviation is estimated at 0.25. A bivariate normal distribution featuring a hypothetical correlation coefficient of 0.5 is chosen for the sake of illustration. Integration over the distribution yields a mean IWC of 1.61×10^{-6} kg m^{-3}.

mapping between altitude and mean temperature, which is sufficient to illustrate the importance of variability in D_e. A rigorous estimator of IWC also needs to take into account the histogram of temperature at each level since the distributions of D_e and extinction coefficients at any given level are matched by temperature.

Distributional aspects of ancillary parameters such as the ice effective diameter may therefore warrant consideration when trying to project geophysical variables of interest such as ice water content from a distribution of extinction coefficients. This issue is actually general and will also affect the estimation of other variables, such as the sedimentation flux in cirrus clouds. In that latter case, other microphysical parameters have to be considered as well, such as the slope of the particle size distribution [see Eq. (16)]. The fundamentals of estimation for an ensemble of clouds with a range of extinction coefficients are the same for any geophysical variable of interest though. They always involve integrating a parameterization over a joint distribution of extinction coefficients and ancillary parameters, such as in Eq. (17). The mathematical aspects of integration over distributions of parameters to estimate average values will be studied in the next section from a generic point of view, with emphasis on a new technique to quickly estimate such integrals.

3. A new generic method for averaging over parameter distributions

In this section, we lay out the mathematical foundation of the general problem of integrating over distributed parameters. Geophysical quantities of interest are represented by functions depending on a number of distributed parameters. Those distributed parameters are used to convey variability of the state of the atmosphere. Let us call $f(\mathbf{x})$ a function of a distributed vector $\mathbf{x} = [x_1, x_2, \ldots, x_N]^{\mathrm{T}}$ containing N distributed parameters. We are interested in estimating the mean (or expected value) and variance of f:

$$\langle f(\mathbf{x}) \rangle = E[f(\mathbf{x})] = \int f(\mathbf{x}) p(\mathbf{x})\, d\mathbf{x}, \tag{19}$$

$$\mathrm{Var}[f(\mathbf{x})] = \int [f(\mathbf{x}) - \langle f(\mathbf{x}) \rangle]^2 p(\mathbf{x})\, d\mathbf{x}, \tag{20}$$

where $p(\mathbf{x})$ is the distribution of $\mathbf{x}$, that is, the joint distribution of the set of parameters $x_1, \ldots, x_N$. For simple analytical function, Eqs. (19) and (20) may be solved analytically. In general, however, $f(\mathbf{x})$ is either too complex to solve the kind of integrals in Eqs. (19) and (20) or $f(\mathbf{x})$ is a numerical procedure of the input parameters. A more suitable approach to solving those integrals would be the Monte Carlo method, where samples $(\mathbf{x}_i)_{i=1}^N$ are generated from $p(\mathbf{x})$ and the mean and variance of f are empirically computed. The main drawback of the Monte Carlo approach is its slow convergence and its tendency to produce noisy estimators at prohibitive cost when the number of samples is large. Variance reduction techniques can improve convergence over standard Monte Carlo (McKay et al. 2000; Larson et al. 2005). Numerical

quadrature approximations to the integral involved in Eq. (19) is another option (Chowdhary et al. 2015). In this work, we present an approach to computing the mean and variance of $f(\mathbf{x})$ based on Taylor expansions around the mean of the distributed vector of parameters $\langle \mathbf{x} \rangle$.

We assume that f can be Taylor expanded around $\langle \mathbf{x} \rangle$ over the typical range of fluctuations of $\mathbf{x}$ and is sufficiently well approximated to second order so that one can write

$$f(\mathbf{x}) \sim f(\langle \mathbf{x} \rangle) + \nabla f(\langle \mathbf{x} \rangle)^{\mathrm{T}} (\mathbf{x} - \langle \mathbf{x} \rangle) + \frac{1}{2} (\mathbf{x} - \langle \mathbf{x} \rangle)^{\mathrm{T}} \mathbf{H}(\langle \mathbf{x} \rangle)(\mathbf{x} - \langle \mathbf{x} \rangle), \tag{21}$$

where $\nabla f(\langle \mathbf{x} \rangle)$ and $\mathbf{H}(\langle \mathbf{x} \rangle)$ are the gradient and Hessian matrix of f evaluated at $\langle \mathbf{x} \rangle$ [$\mathbf{H}(\langle \mathbf{x} \rangle)$ is the matrix whose entries are the second-order derivatives of f: $\mathbf{H}_{ij}(\langle \mathbf{x} \rangle) = (\partial^2 f / \partial x_i \partial x_j)(\langle \mathbf{x} \rangle)$].

Applying the expectation operator to Eq. (21) yields

$$\langle f(\mathbf{x}) \rangle \sim f(\langle \mathbf{x} \rangle) + \frac{1}{2} E[(\mathbf{x} - \langle \mathbf{x} \rangle)^{\mathrm{T}} \mathbf{H}(\langle \mathbf{x} \rangle)(\mathbf{x} - \langle \mathbf{x} \rangle)] \tag{22}$$

after identifying $\langle \cdots \rangle = E[\cdots]$. Note that the expectation of the linear form vanishes. Since the quadratic form evaluates to a scalar, one can rewrite Eq. (22) as

$$\langle f(\mathbf{x}) \rangle \sim f(\langle \mathbf{x} \rangle) + \frac{1}{2} \mathrm{tr}\{E[(\mathbf{x} - \langle \mathbf{x} \rangle)^{\mathrm{T}} \mathbf{H}(\langle \mathbf{x} \rangle)(\mathbf{x} - \langle \mathbf{x} \rangle)]\}, \tag{23}$$

where the trace operator (tr) has been introduced. Using the linearity of the trace and its invariance under cyclic permutation (Mathai and Provost 1992, p. 50), one can transform Eq. (23) further and write

$$\langle f(\mathbf{x}) \rangle \sim f(\langle \mathbf{x} \rangle) + \frac{1}{2} \mathrm{tr}\{\mathbf{H}(\langle \mathbf{x} \rangle) E[(\mathbf{x} - \langle \mathbf{x} \rangle)(\mathbf{x} - \langle \mathbf{x} \rangle)^{\mathrm{T}}]\}. \tag{24}$$

One recognizes the definition of the variance–covariance matrix on the RHS of Eq. (24):

$$\mathbf{\Sigma} = E[(\mathbf{x} - \langle \mathbf{x} \rangle)(\mathbf{x} - \langle \mathbf{x} \rangle)^{\mathrm{T}}]. \tag{25}$$

Thus, the approximate expression for the mean of $f(\mathbf{x})$ is given by

$$\langle f(\mathbf{x}) \rangle \sim f(\langle \mathbf{x} \rangle) + \frac{1}{2} \mathrm{tr}[\mathbf{H}(\langle \mathbf{x} \rangle)\mathbf{\Sigma}] \tag{26}$$

for $\mathbf{x}$ distributed with mean $\langle \mathbf{x} \rangle$ and covariance matrix $\mathbf{\Sigma}$. Because Eq. (26) derives from a direct Taylor expansion of f, we shall call this approach the direct approach to estimating $\langle f(\mathbf{x}) \rangle$.

Several comments can be made on Eq. (26). Perhaps the most obvious observation is that this equation reduces to the following equality if f is linear:

$$\langle f(\mathbf{x}) \rangle = f(\langle \mathbf{x} \rangle). \tag{27}$$

Running a model with mean parameters as inputs produces unbiased mean physical quantities only if the model is linear in its parameters. If the model is nonlinear in its parameters, a systematic bias appears that can be estimated by the term $(1/2)\mathrm{tr}[\mathbf{H}(\langle \mathbf{x} \rangle)\mathbf{\Sigma}]$. Note that Eq. (26) holds exactly if f is quadratic, irrespective of the distribution of $\mathbf{x}$.

In one dimension, Eq. (26) reduces to the expression

$$\langle f(x) \rangle \sim f(\langle x \rangle) + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \sigma^2, \tag{28}$$

where σ is the standard deviation of x. Therefore, if f is convex, one has

$$\langle f(x) \rangle \geq f(\langle x \rangle), \tag{29}$$

and the opposite is true if f is concave. Equation (29) is known as Jensen's inequality (Jensen 1906) and the quantity $\langle f(\mathbf{x}) \rangle - f(\langle \mathbf{x} \rangle) \sim (1/2)(\partial^2 f / \partial x^2)\sigma^2$ is sometimes referred to as the Jensen gap in the literature (Abramovich and Persson 2016). Equation (26) can be seen as a generalization of Jensen's inequality in multiple dimensions and the term $(1/2)\mathrm{tr}[\mathbf{H}(\langle \mathbf{x} \rangle)\mathbf{\Sigma}]$ estimates the Jensen gap in those conditions. The term $(1/2)\mathrm{tr}[\mathbf{H}(\langle \mathbf{x} \rangle)\mathbf{\Sigma}]$ can be viewed as a correction to model outputs computed from mean parameter values in order to reduce the systematic evaluation bias and rectify the outputs toward their expected values.

The variance of $f(\mathbf{x})$ can also be estimated by recognizing that

$$\mathrm{Var}[f(\mathbf{x})] = E[(f(\mathbf{x}) - E[f(\mathbf{x})])^2], \tag{30}$$

and using Eqs. (21) and (26), which gives

$$\mathrm{Var}[f(\mathbf{x})] \sim E\left(\left\{ \nabla f(\langle \mathbf{x} \rangle)^{\mathrm{T}} (\mathbf{x} - \langle \mathbf{x} \rangle) + \frac{1}{2} (\mathbf{x} - \langle \mathbf{x} \rangle)^{\mathrm{T}} \mathbf{H}(\langle \mathbf{x} \rangle)(\mathbf{x} - \langle \mathbf{x} \rangle) - \frac{1}{2} \mathrm{tr}[\mathbf{H}(\langle \mathbf{x} \rangle)\mathbf{\Sigma}] \right\}^2 \right). \tag{31}$$

On expanding Eq. (31) one can write

$$\begin{aligned}\mathrm{Var}[f(\mathbf{x})] \sim\ & E\{[\nabla f(\langle\mathbf{x}\rangle)^{\mathrm{T}}(\mathbf{x}-\langle\mathbf{x}\rangle)][\nabla f(\langle\mathbf{x}\rangle)^{\mathrm{T}}(\mathbf{x}-\langle\mathbf{x}\rangle)]\}\\ &+E\left\{\frac{1}{4}[(\mathbf{x}-\langle\mathbf{x}\rangle)^{\mathrm{T}}\mathbf{H}(\langle\mathbf{x}\rangle)(\mathbf{x}-\langle\mathbf{x}\rangle)][(\mathbf{x}-\langle\mathbf{x}\rangle)^{\mathrm{T}}\mathbf{H}(\langle\mathbf{x}\rangle)(\mathbf{x}-\langle\mathbf{x}\rangle)]\right\}\\ &+E\{[\nabla f(\langle\mathbf{x}\rangle)^{\mathrm{T}}(\mathbf{x}-\langle\mathbf{x}\rangle)][(\mathbf{x}-\langle\mathbf{x}\rangle)^{\mathrm{T}}\mathbf{H}(\langle\mathbf{x}\rangle)(\mathbf{x}-\langle\mathbf{x}\rangle)]\}\\ &-E\{[\nabla f(\langle\mathbf{x}\rangle)^{\mathrm{T}}(\mathbf{x}-\langle\mathbf{x}\rangle)]\mathrm{tr}[\mathbf{H}(\langle\mathbf{x}\rangle)\boldsymbol{\Sigma}]\}\\ &-E\left\{\frac{1}{2}(\mathbf{x}-\langle\mathbf{x}\rangle)^{\mathrm{T}}\mathbf{H}(\langle\mathbf{x}\rangle)(\mathbf{x}-\langle\mathbf{x}\rangle)\mathrm{tr}[\mathbf{H}(\langle\mathbf{x}\rangle)\boldsymbol{\Sigma}]\right\}\\ &+\frac{1}{4}\mathrm{tr}[\mathbf{H}(\langle\mathbf{x}\rangle)\boldsymbol{\Sigma}]^2. \end{aligned} \tag{32}$$

Assuming a normal distribution for $\mathbf{x}$, one can use the following identities on the variance of quadratic forms (see, e.g., Mathai and Provost 1992, p. 53):

$$\begin{aligned}&E\{[(\mathbf{x}-\langle\mathbf{x}\rangle)^{\mathrm{T}}\mathbf{H}(\langle\mathbf{x}\rangle)(\mathbf{x}-\langle\mathbf{x}\rangle)][(\mathbf{x}-\langle\mathbf{x}\rangle)^{\mathrm{T}}\mathbf{H}(\langle\mathbf{x}\rangle)(\mathbf{x}-\langle\mathbf{x}\rangle)]\}\\ &\quad=2\mathrm{tr}\{[\mathbf{H}(\langle\mathbf{x}\rangle)\boldsymbol{\Sigma}]^2\}+\mathrm{tr}[\mathbf{H}(\langle\mathbf{x}\rangle)\boldsymbol{\Sigma}]^2, \end{aligned} \tag{33}$$

and on the covariance between linear and quadratic forms

$$E\{[\nabla f(\langle\mathbf{x}\rangle)^{\mathrm{T}}(\mathbf{x}-\langle\mathbf{x}\rangle)][(\mathbf{x}-\langle\mathbf{x}\rangle)^{\mathrm{T}}\mathbf{H}(\langle\mathbf{x}\rangle)(\mathbf{x}-\langle\mathbf{x}\rangle)]\}=0 \tag{34}$$

(since the assumption of Gaussianity ensures that third-order centered moments of $\mathbf{x}$ are 0). One can then show that the expression at Eq. (32) eventually simplifies to

$$\mathrm{Var}[f(\mathbf{x})]\sim\nabla f(\langle\mathbf{x}\rangle)^{\mathrm{T}}\boldsymbol{\Sigma}\nabla f(\langle\mathbf{x}\rangle)+\frac{1}{2}\mathrm{tr}\{[\mathbf{H}(\langle\mathbf{x}\rangle)\boldsymbol{\Sigma}]^2\}, \tag{35}$$

when $\mathbf{x}$ is normally distributed.

The term $\nabla f(\langle\mathbf{x}\rangle)^{\mathrm{T}}\boldsymbol{\Sigma}\nabla f(\langle\mathbf{x}\rangle)$ on the RHS of Eq. (35) means that to a good approximation the variance of $f(\mathbf{x})$ is given by the variance of the parameters projected onto the gradient of the model. For a model that is linear in its parameters, the variance of $f(\mathbf{x})$ is exactly given by $\nabla f(\langle\mathbf{x}\rangle)^{\mathrm{T}}\boldsymbol{\Sigma}\nabla f(\langle\mathbf{x}\rangle)$ and thus depends on the mean, the covariance matrix of the parameters and the gradient of the model.

In some situations, a function of f may be better approximated by the first few terms of its Taylor series. For instance, one may seek an expansion of $\ln(f)$ and write

$$\begin{aligned}\ln f(\mathbf{x})\sim\ & \ln f(\langle\mathbf{x}\rangle)+\nabla\ln f(\langle\mathbf{x}\rangle)^{\mathrm{T}}(\mathbf{x}-\langle\mathbf{x}\rangle)\\ &+\frac{1}{2}(\mathbf{x}-\langle\mathbf{x}\rangle)^{\mathrm{T}}\mathbf{H}^{\ln}(\langle\mathbf{x}\rangle)(\mathbf{x}-\langle\mathbf{x}\rangle), \end{aligned} \tag{36}$$

where $\mathbf{H}^{\ln}$ is the Hessian matrix of $\ln(f)$. Equation (36) may be better suited to capture the low-order variations of the function than Eq. (21) depending on the functional form of f. One can then reuse the results at Eqs. (26) and (35) to build estimators of $\langle\ln f(\mathbf{x})\rangle$ and $\mathrm{Var}[\ln f(\mathbf{x})]$:

$$\langle\ln f(\mathbf{x})\rangle\sim\ln f(\langle\mathbf{x}\rangle)+\frac{1}{2}\mathrm{tr}[\mathbf{H}^{\ln}(\langle\mathbf{x}\rangle)\boldsymbol{\Sigma}], \tag{37}$$

$$\mathrm{Var}[\ln f(\mathbf{x})]\sim\nabla\ln f(\langle\mathbf{x}\rangle)^{\mathrm{T}}\boldsymbol{\Sigma}\nabla\ln f(\langle\mathbf{x}\rangle)+\frac{1}{2}\mathrm{tr}\{[\mathbf{H}^{\ln}(\langle\mathbf{x}\rangle)\boldsymbol{\Sigma}]^2\}. \tag{38}$$

In this situation, it is possible to build an estimator of $\langle f(\mathbf{x})\rangle$ from the estimators at Eqs. (37) and (38), as we explain in appendix A. Making use of Eq. (A5), we arrive at the following estimator for $\langle f(\mathbf{x})\rangle$:

$$\begin{aligned}\langle f(\mathbf{x})\rangle\sim\ & f(\langle\mathbf{x}\rangle)\\ &\times\exp\left(\frac{1}{2}\nabla\ln f(\langle\mathbf{x}\rangle)^{\mathrm{T}}\boldsymbol{\Sigma}\nabla\ln f(\langle\mathbf{x}\rangle)+\frac{1}{2}\mathrm{tr}[\mathbf{H}^{\ln}(\langle\mathbf{x}\rangle)\boldsymbol{\Sigma}]\right.\\ &\left.+\frac{1}{4}\mathrm{tr}\{[\mathbf{H}^{\ln}(\langle\mathbf{x}\rangle)\boldsymbol{\Sigma}]^2\}\right). \end{aligned} \tag{39}$$

Because Eq. (39) derives from a Taylor expansion of $\ln(f)$, we shall say that $\langle f(\mathbf{x})\rangle$ was estimated using a logarithmic approach, as opposed to the direct approach of Eq. (26). Note, however, that the logarithmic approach relies itself on the direct approach for estimation of $\langle\ln f(\mathbf{x})\rangle$ and $\mathrm{Var}[\ln f(\mathbf{x})]$.

Note that while the direct approach yields exact results for quadratic functions, the logarithmic approach yields exact results for power laws and lognormal distributions, up to a transformation of parameter.

Whether a Taylor expansion of f or $\ln(f)$ is chosen, the quality of the resulting estimator depends on convergence considerations of the Taylor series [Eqs. (21) and (36)]. In particular, for the estimators to make sense when the domain of convergence is finite, the variance of the parameter distribution (i.e., the diagonal of $\boldsymbol{\Sigma}$) must be low enough to ensure that the parameters take the bulk of their values within the domain of convergence.

The main results of this section are summarized in Fig. 2 where the view is assumed that the parameter space is partitioned into observable and ancillary subspaces. The estimators of $\langle f(\mathbf{x})\rangle$ at Eqs. (26) and (39) can

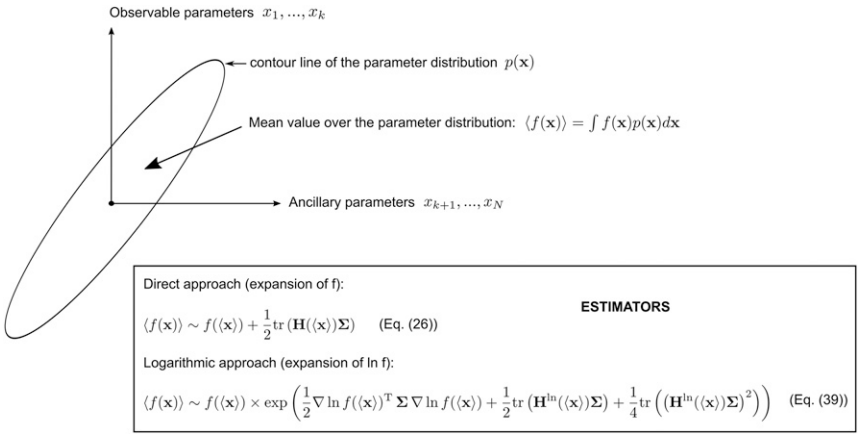

FIG. 2. Summary of the results from section 3: N parameters are assumed, where k dimensions correspond to observations; $N - k$ dimensions correspond to ancillary (a priori) parameters; $p(\mathbf{x})$ is a joint distribution of the observable and the ancillary parameters; and $\langle f(\mathbf{x})\rangle$ is the mean value of f over the parameter distribution, where f is a geophysical quantity of interest evaluated from the parameters $\mathbf{x}$. We construct estimators of $\langle f(\mathbf{x})\rangle$ by Taylor expanding to second order either f (direct approach) or $\ln(f)$ (logarithmic approach) around the mean parameter values, depending on which expansion makes more sense. In the above formula, $\mathbf{\Sigma}$ represents the covariance matrix of the parameters while $\mathbf{H}$ and $\mathbf{H}^{\ln}$ represent the Hessian matrices of f and $\ln(f)$.

then be used to estimate mean values over an ensemble of observations and are applied in the next sections to the problem of estimating mean cirrus cloud ice water content and sedimentation flux from lidar observations. One should remember that these estimators are very generic in nature and applicable beyond the specific problem treated in this article. In particular, they constitute suitable options for the upscaling of microphysical process rates.

4. Application to the estimation of mean ice water content and sedimentation flux

We now discuss the application of the generic averaging method presented in section 3 to the parameterizations of ice water content and sedimentation flux developed in section 2. To proceed, we first need to specify the joint distribution of the parameters of the problem, that is the joint distribution of volume extinction coefficients and microphysical parameters (effective diameter and, in the case of sedimentation flux, the slope of the particle size distribution). For normal (or lognormal distribution, up to a transformation), the specifications are encoded in the mean parameter values $\langle \mathbf{x}\rangle$ and the covariance matrix $\mathbf{\Sigma}$. To compute the estimators developed in section 3, we also need to produce the gradient and Hessian of the parameterizations for ice water content and sedimentation flux.

a. *Distribution of parameters*

We have shown in section 2 that the distributed parameters to consider for ice water content and sedimentation flux estimation are the volume extinction coefficient, the ice effective diameter and the slope of the particle size distribution (for sedimentation flux only). We discuss plausible distributions of these parameters, based on information acquired from satellite observations or in situ measurements. In the case of ancillary parameters inferred from in situ measurements, the distributions are not always precisely reported in the literature. For the purpose here, namely, to investigate the impact of parameter variance on mean value estimators, the distributions can be reasonably approximated by postulating a plausible amount of variance. A plausible model of the joint parameter distribution in cirrus clouds is then proposed by adding unknown correlations into the picture. The degeneracy of such distributions with multiple correlation coefficients is characterized, which allows us to project the impact on integral cloud properties. Because observations suggest lognormal distributions, the parameters are redefined as their log-transformed counterparts:

$$x_1 = \ln(\lambda), \tag{40}$$

$$x_2 = \ln(D_e), \tag{41}$$

$$x_3 = \ln(\mathrm{EXT}), \tag{42}$$

which are assumed normally distributed; x_1 and x_2 are the redefined ancillary parameters while x_3 is the redefined volume extinction coefficient. (The values of x_1, x_2, and x_3 henceforth are given in SI units.)

1) Distribution of volume extinction coefficients

Measurements from the Cloud–Aerosol Lidar with Orthogonal Polarization (Winker et al. 2003, 2007; Hunt et al. 2009) suggest that volume extinction values are approximately lognormally distributed in the tropical upper troposphere–lower stratosphere (UTLS). The distribution function of measurements is represented in Fig. 3 as a function of altitude in the tropics. One can check that $\ln[p(x_3)]$ can approximately be fitted with a quadratic function, thus proving that x_3 is quasi-normally distributed, or equivalently that EXT is quasi-lognormally distributed. The mean value and standard deviation of x_3 ($\langle x_3 \rangle, \sigma_3$) can be determined at each altitude by direct numerical integration over the distribution function (Fig. 3, solid line).

2) Distribution of ice effective diameters and slope parameters

In Heymsfield et al. (2014), at Eq. (9e), the authors propose to parameterize ice effective diameter in upper-tropospheric cirrus as

$$\ln(D_e^{\text{param}}) = \begin{cases} \ln(8.33 \times 10^{-5}) + 0.0184T, & T < -71^\circ\text{C} \\ \ln(9.1744 \times 10^{-2}) + 0.117T, & T > -71^\circ\text{C} \end{cases}, \tag{43}$$

where T is air temperature. The parameterization is built by fitting a compilation of in situ measurements acquired at temperatures down to −86°C. One consequence of using Eq. (43) is that the distribution of effective diameters reduces to a single number at a given temperature (see Fig. 1a for illustration), whereas the observational data show substantial variance [cf. Fig. 11a of Heymsfield et al. (2014)]. Therefore, a distribution with a nonzero amount of variance and centered around $\ln(D_e^{\text{param}})$ should be expected (we effectively assume a normal distribution, this assumption being supported by Thornberry et al. (2017), Fig. 5). This is shown in Fig. 1b where a plausible variance of 0.25 of the logarithm of effective diameters has been added to the parameterization. The value of 0.25 has been estimated based on visual inspection of the data in Fig. 11a of Heymsfield et al. (2014). In summary, a plausible distribution for $x_2 = \ln(D_e)$ is the normal distribution with mean $\langle x_2 \rangle = \ln(D_e^{\text{param}})$ and standard deviation $\sigma_2 \sim 0.25$. Note also that while other investigators have proposed empirical parameterizations similar to Eq. (43) (Thornberry et al. 2017), they always focus on parameterizing the center of the distribution of D_e as a function of temperature, not the full distribution.

Similarly, Heymsfield et al. (2013) propose to parameterize the slope parameter of particle size distributions (fitted as gamma functions) in cirrus clouds as

$$\ln(\lambda^{\text{param}}) = \ln(75) - 0.1057T. \tag{44}$$

Using the same reasoning as for effective diameter, we estimate that the distribution of values of $\ln(\lambda)$ can be approximated by a normal distribution centered around $\ln(\lambda^{\text{param}})$ with a nonzero amount of variance, that we roughly estimate at 0.6 based on visual inspection of the data in Fig. 7 of Heymsfield et al. (2013). That is, $\langle x_1 \rangle = \ln(\lambda^{\text{param}})$ and $\sigma_1 \sim 0.6$.

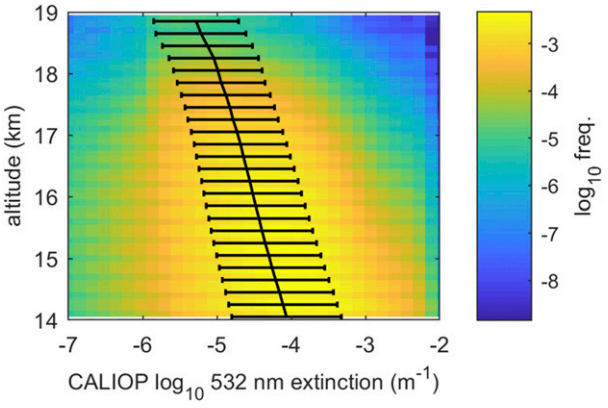

FIG. 3. Histogram of tropical (30°S–30°N) CALIOP 532-nm volume extinction coefficients across altitude and extinction bins, over the period 2007–15. Extinction bins are linearly spaced in $\log_{10}$ scale. Color coding measures fraction of occurrence across all altitudes and extinctions (in $\log_{10}$ units). The solid black line represents the mean vertical profile of $\log_{10}$(EXT) measurements. Error bars correspond to 1-sigma standard deviation. Extinction values are from the CALIOP L2 5-km Cloud Profile Product V4.1 (Young et al. 2018). Data have been quality screened to retain only measurements with quality flags of 0, 1, 2, 16, and 18 and an absolute value of the cloud aerosol discrimination score of at least 50. Both daytime and nighttime profiles are used.

3) Joint distribution of parameters

As stated in subsections 1 and 2, the parameters of the problem (volume extinction coefficient, effective diameter, slope parameter) can be modeled using log-normal distributions, which is equivalent to assuming normal distributions for their log-transformed counterparts (x_1, x_2, x_3):

$$p(x_i) \sim \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left[-\frac{1}{2} \frac{(x_i - \langle x_i \rangle)^2}{\sigma_i^2} \right], \tag{45}$$

where $\langle x_i \rangle$ and σ_I are the mean values and standard deviations of x_i.

The $p(x_i)$ can also be viewed as marginal distributions of a more general joint distribution of parameters $p(\mathbf{x}) = p(x_1, x_2, x_3)$. In the case of independent

parameters, the joint distribution is simply obtained by taking the product of the marginal distributions:

$$p(\mathbf{x}) = p(x_1)p(x_2)p(x_3). \tag{46}$$

In the general case, however, we have to assume that the parameters can be correlated and the joint density is thus given by a generalized trivariate normal distribution:

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}(\det\boldsymbol{\Sigma})^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x} - \langle\mathbf{x}\rangle)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \langle\mathbf{x}\rangle)\right], \tag{47}$$

where $\mathbf{x} = [x_1, x_2, x_3]^{\mathrm{T}}$ and $\boldsymbol{\Sigma}$ is the covariance matrix of the log-transformed parameters. $\boldsymbol{\Sigma}$ is defined as

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{bmatrix}, \tag{48}$$

where the off-diagonal entries σ_{ij} are the covariances measuring the degree of correlation between x_i and x_j:

$$\sigma_{ij} = E[(x_i - \langle x_i\rangle)(x_j - \langle x_j\rangle)]. \tag{49}$$

In the absence of corresponding measurements reported in the literature, the covariances σ_{ij} must be considered as unknown. By virtue of the Cauchy–Schwarz inequality, they can range anywhere between $-\sigma_i\sigma_j$ and $+\sigma_i\sigma_j$. Alternatively, the correlation matrix $\boldsymbol{\Sigma}^{\mathrm{corr}}$ can be defined by the normalization operation:

$$\boldsymbol{\Sigma}^{\mathrm{corr}} = (\mathrm{diag}\boldsymbol{\Sigma})^{-1/2}\boldsymbol{\Sigma}(\mathrm{diag}\boldsymbol{\Sigma})^{-1/2}, \tag{50}$$

where $\mathrm{diag}\boldsymbol{\Sigma}$ is a matrix of diagonal entries containing the variances σ_i^2 and the off-diagonal entries of $\boldsymbol{\Sigma}^{\mathrm{corr}}$ feature the correlation coefficients ρ_{ij}:

$$\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i\sigma_j}, \tag{51}$$

which are normalized between -1 and 1. The decomposition into variances and correlation coefficients separates the information controlling parameter dispersion: the variances control the overall level of variability of the parameters while the correlations coefficients determine how tightly the parameters correlate.

The joint distribution specified at Eq. (47) is degenerate with respect to values of the correlation coefficients. That is, multiple joint distributions differing by their correlation coefficients alone can be generated from the same individual distributions of parameters. Properties resulting from integration over such degenerate distributions, such as ice water content or sedimentation flux, do, however, differ with the configuration of the correlation coefficients, as discussed in sections 5 and 6.

Note that in producing the joint distribution $p(\mathbf{x})$, we assume a simple mapping between altitude and mean temperature to match the marginal distributions $p(x_i)$ together. Since temperature is the real coupling variable between extinction and other parameters, one should also roll temperature variations into this framework. This can be done rigorously by computing joint distributions $p_T(\mathbf{x})$ conditioned on the value of temperature T, integrating over such distributions to produce mean value estimates of ice water content and sedimentation flux conditioned on temperature, and integrating those estimates over the distribution of temperatures at a given level. Our choice of working with mean temperature is motivated by simplicity since the focus of this paper is on the impact of parameter distributional aspects, such as the degeneracy with multiple correlation coefficients, rather than on temperature variations.

b. *Range limitation of the correlations between parameters*

When 3 parameters or more are involved, the values of the unknown covariances, or equivalently the correlation coefficients, have to satisfy some limitations in order to define a valid joint parameter distribution. For a two-dimensional problem, the correlation matrix is simply given by

$$\boldsymbol{\Sigma}^{\mathrm{corr}} = \begin{bmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{bmatrix},$$

and ρ_{12} can take any value between -1 and 1. For a three-dimensional problem however, the situation is complicated by the triangular nature of correlations between variables x_1, x_2 and x_3. That is, the correlations between x_1 and x_2 and between x_2 and x_3 limit the possible correlation between x_1 and x_3.

The condition on the correlation coefficients may be found by requiring that $\boldsymbol{\Sigma}^{\mathrm{corr}}$ be a positive semidefinite matrix. One can show that $\det(\boldsymbol{\Sigma}^{\mathrm{corr}}) > 0$ is a sufficient condition that warrants positivity of the eigenvalues of $\boldsymbol{\Sigma}^{\mathrm{corr}}$.[1] For a three-dimensional problem, the condition on the correlations coefficients is then given by

[1] This can be established from the characteristic polynomial of $\boldsymbol{\Sigma}^{\mathrm{corr}}$, using the fact that $\boldsymbol{\Sigma}^{\mathrm{corr}}$ is symmetric.

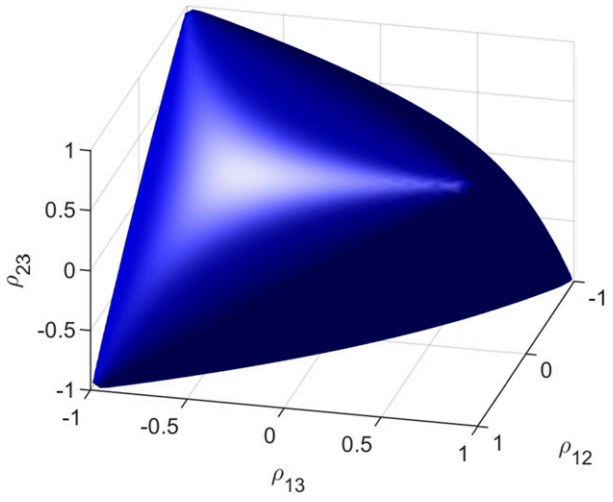

FIG. 4. Joint values of the correlation coefficients ρ_{12}, ρ_{13}, and ρ_{23} for which $p(x_1, x_2, x_3)$ is a valid joint density. The values correspond to $\det(\boldsymbol{\Sigma}^{\text{corr}}) \geq 0$ and are located inside the $\det(\boldsymbol{\Sigma}^{\text{corr}}) = 0$ isosurface ($\boldsymbol{\Sigma}^{\text{corr}}$ being the correlation matrix of the parameters x_1, x_2, and x_3). That isosurface is displayed using a solid representation with artificial lighting.

$$\det(\boldsymbol{\Sigma}^{\text{corr}}) > 0 \Leftrightarrow 1 + 2\rho_{12}\rho_{13}\rho_{23} - (\rho_{12}^2 + \rho_{13}^2 + \rho_{23}^2) > 0\,. \tag{52}$$

As shown in Fig. 4, Eq. (52) defines a volume of $\mathbb{R}^3$ bounded by the isosurface $\det(\boldsymbol{\Sigma}^{\text{corr}}) = 0$. This volume spans the feasible values of the triplet (ρ_{12}, ρ_{13}, ρ_{23}). Therefore, the mean value and variance of $f(\mathbf{x})$ should be evaluated over the volume defined by Eq. (52).

One should keep in mind that, in nature, extreme values of the correlation coefficients between parameters (toward -1 or $+1$) are highly unlikely. The correlations are more likely to remain toward middle values.

c. Gradient and Hessian for model evaluation

We will show in the following section that the parameterizations for ice water content and sedimentation flux have to be estimated from their logarithmic formulation (logarithmic approach described in section 3). This requires one to specify the gradient and Hessian matrix of the log-transformed parameterizations.

Taking the logarithm of Eqs. (15) and (16) allows one to easily rewrite the parameterizations as functions of the log-transformed parameters x_i. By doing so, one arrives at

$$\ln\text{IWC}(x_1, x_2, x_3) = x_3 + x_2 + \ln\left(\frac{\rho_i}{3}\right) \tag{53}$$

and

$$\ln F_{\text{sed}}(x_1, x_2, x_3) = x_3 + (\beta - b - s + 2)x_1 + \ln\Gamma[\mu(x_2, x_1) + b + s + 1] - \ln\Gamma[\mu(x_2, x_1) + \beta + 3] + \ln\left(\frac{2ar}{\pi\alpha}\right)\,. \tag{54}$$

For ice water content [Eq. (53)], the components of the gradient are

$$\begin{cases} \dfrac{\partial \ln\text{IWC}}{\partial x_1}(\mathbf{x}) = 0 \\ \dfrac{\partial \ln\text{IWC}}{\partial x_2}(\mathbf{x}) = 1\,, \\ \dfrac{\partial \ln\text{IWC}}{\partial x_3}(\mathbf{x}) = 1 \end{cases} \tag{55}$$

and the Hessian matrix is simply the null matrix: $\mathbf{H}^{\ln}(\mathbf{x}) = 0$. For sedimentation flux [Eq. (54)], the components of the gradient are given by

$$\begin{cases} \dfrac{\partial \ln F_{\text{sed}}}{\partial x_1}(\mathbf{x}) = -s + (2 + \beta - b)\dfrac{\psi_0[\mu(x_2, x_1) + b + 1] - \psi_0[\mu(x_2, x_1) + b + s + 1]}{\psi_0[\mu(x_2, x_1) + b + 1] - \psi_0[\mu(x_2, x_1) + \beta + 3]} \\ \dfrac{\partial \ln F_{\text{sed}}}{\partial x_2}(\mathbf{x}) = \dfrac{\psi_0[\mu(x_2, x_1) + b + s + 1] - \psi_0[\mu(x_2, x_1) + \beta + 3]}{\psi_0[\mu(x_2, x_1) + b + 1] - \psi_0[\mu(x_2, x_1) + \beta + 3]} \\ \dfrac{\partial \ln F_{\text{sed}}}{\partial x_3}(\mathbf{x}) = 1 \end{cases}\,, \tag{56}$$

and the entries of the Hessian matrix are

$$
\begin{cases}
\dfrac{\partial^2 \ln F_{\text{sed}}}{\partial x_1^2}(\mathbf{x}) = -(2+\beta-b)^2 \left(\dfrac{\psi_1[\mu(x_2,x_1)+b+1] - \psi_1[\mu(x_2,x_1)+b+s+1]}{\{\psi_0[\mu(x_2,x_1)+b+1] - \psi_0[\mu(x_2,x_1)+\beta+3]\}^2} \right. \\
\qquad - \dfrac{\{\psi_1[\mu(x_2,x_1)+b+1] - \psi_1[\mu(x_2,x_1)+\beta+3]\}}{\{\psi_0[\mu(x_2,x_1)+b+1] - \psi_0[\mu(x_2,x_1)+\beta+3]\}^3} \\
\qquad \left. \times \dfrac{\{\psi_0[\mu(x_2,x_1)+b+1] - \psi_0[\mu(x_2,x_1)+b+s+1]\}}{\{\psi_0[\mu(x_2,x_1)+b+1] - \psi_0[\mu(x_2,x_1)+\beta+3]\}^3} \right) \\
\dfrac{\partial^2 \ln F_{\text{sed}}}{\partial x_2^2}(\mathbf{x}) = \dfrac{\psi_1[\mu(x_2,x_1)+b+s+1] - \psi_1[\mu(x_2,x_1)+\beta+3]}{\{\psi_0[\mu(x_2,x_1)+b+1] - \psi_0[\mu(x_2,x_1)+\beta+3]\}^2} \\
\qquad - \dfrac{\{\psi_1[\mu(x_2,x_1)+b+1] - \psi_1[\mu(x_2,x_1)+\beta+3]\}}{\{\psi_0[\mu(x_2,x_1)+b+1] - \psi_0[\mu(x_2,x_1)+\beta+3]\}^3} \\
\qquad \times \dfrac{\{\psi_0[\mu(x_2,x_1)+b+s+1] - \psi_0[\mu(x_2,x_1)+\beta+3]\}}{\{\psi_0[\mu(x_2,x_1)+b+1] - \psi_0[\mu(x_2,x_1)+\beta+3]\}^3} \\
\dfrac{\partial^2 \ln F_{\text{sed}}}{\partial x_3^2}(\mathbf{x}) = 0 \\
\dfrac{\partial^2 \ln F_{\text{sed}}}{\partial x_1 x_2}(\mathbf{x}) = -(2+\beta-b) \left(\dfrac{\psi_1[\mu(x_2,x_1)+b+s+1] - \psi_1[\mu(x_2,x_1)+\beta+3]}{\{\psi_0[\mu(x_2,x_1)+b+1] - \psi_0[\mu(x_2,x_1)+\beta+3]\}^2} \right. \\
\qquad - \dfrac{\{\psi_1[\mu(x_2,x_1)+b+1] - \psi_1[\mu(x_2,x_1)+\beta+3]\}}{\{\psi_0[\mu(x_2,x_1)+b+1] - \psi_0[\mu(x_2,x_1)+\beta+3]\}^3} \\
\qquad \left. \times \dfrac{\{\psi_0[\mu(x_2,x_1)+b+s+1] - \psi_0[\mu(x_2,x_1)+\beta+3]\}}{\{\psi_0[\mu(x_2,x_1)+b+1] - \psi_0[\mu(x_2,x_1)+\beta+3]\}^3} \right) = \dfrac{\partial^2 \ln F_{\text{sed}}}{\partial x_2 x_1}(\mathbf{x}) \\
\dfrac{\partial^2 \ln F_{\text{sed}}}{\partial x_1 x_3}(\mathbf{x}) = 0 = \dfrac{\partial^2 \ln F_{\text{sed}}}{\partial x_3 x_1}(\mathbf{x}) \\
\dfrac{\partial^2 \ln F_{\text{sed}}}{\partial x_2 x_3}(\mathbf{x}) = 0 = \dfrac{\partial^2 \ln F_{\text{sed}}}{\partial x_3 x_2}(\mathbf{x}).
\end{cases}
\tag{57}
$$

In Eqs. (56) and (57), ψ_0 and ψ_1 respectively denote the polygamma functions of order 0 and 1 (first and second derivatives of $\ln\Gamma$; see, e.g., Abramowitz and Stegun 1972, p. 260).

5. Validation

In this section, the generic averaging method of section 3 applied to ice water content and sedimentation flux parameterizations is validated against values computed from the Monte Carlo method. Results are computed over the set of possible correlation triplets defined at Eq. (52) and are reported as deviations of Monte Carlo predictions. This way, we can evaluate how the accuracy of the method holds with degenerate joint distributions of parameters (see section 4). We contrast the results obtained by Taylor expanding the parameterizations and their log-transformed counterparts (direct vs logarithmic approach) and show that the logarithmic approach is

the one that makes sense given the formulation of the parameterizations.

a. Ice water content

The results predicted by the logarithmic approach are mathematically exact in the case of ice water content, which trivially validates the approach for this particular parameterization. Indeed, the parameterization for ice water content is quadratic and therefore can be integrated exactly over the parameter distribution. By definition, the mean value of ice water content over the distribution of extinction coefficients and effective diameters is equal to

$$\langle \mathrm{IWC} \rangle = \frac{\rho_i}{3} \langle \mathrm{EXT} D_e \rangle \,. \tag{58}$$

Rewriting the above equation in terms of the log-transformed variables $x_2 = \ln D_e$ and $x_3 = \ln \mathrm{EXT}$, one has

$$\langle \mathrm{IWC} \rangle = \frac{\rho_i}{3} \langle \exp(x_2 + x_3) \rangle \,, \tag{59}$$

which is equivalent to the integral

$$\langle \mathrm{IWC} \rangle = \frac{\rho_i}{3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp(x_2 + x_3) p(x_2, x_3) \, dx_2 \, dx_3, \tag{60}$$

where $p(x_2, x_3)$ is the probability density of a bivariate normal distribution. The double integration in Eq. (60) can be carried out analytically [see, e.g., Larson and Griffin (2013), their Eq. (26)], which yields the result

$$\langle \mathrm{IWC} \rangle = \frac{\rho_i}{3} \exp(\langle x_2 \rangle + \langle x_3 \rangle) \exp\left[\frac{1}{2} (\sigma_2^2 + \sigma_3^2) + \sigma_2 \sigma_3 \rho_{23} \right]. \tag{61}$$

Note that Eq. (61) can also be written as

$$\langle \mathrm{IWC} \rangle = \frac{\rho_i}{3} \langle \mathrm{EXT} \rangle \langle D_e \rangle \exp(\sigma_2 \sigma_3 \rho_{23}) \tag{62}$$

by using Eq. (A5) with the definitions of $\langle \mathrm{EXT} \rangle$ and $\langle D_e \rangle$.

One can check that the logarithmic approach to computing mean ice water content [Eq. (39)] recovers the exact result given in Eqs. (61) and (62). There are two reasons for this fact. First, the logarithm of ice water content is linear in its arguments:

$$\ln(\mathrm{IWC}) = \ln\left(\frac{\rho_i}{3}\right) + x_2 + x_3 \,, \tag{63}$$

which warrants that its Taylor expansion is exact to second order. Therefore, the expressions for $\langle \ln(\mathrm{IWC}) \rangle$ and $\mathrm{Var}[\ln(\mathrm{IWC})]$ developed in section 3 are also exact. Then, it is easy to deduce from the linearity of Eq. (63) that the values of ln(IWC) are normally distributed. In that context, Eq. (A5) becomes exact (see appendix A), hence the exact prediction of $\langle \mathrm{IWC} \rangle$ from $\langle \ln(\mathrm{IWC}) \rangle$ and $\mathrm{Var}[\ln(\mathrm{IWC})]$.

More generally, as noted in section 3, the logarithmic approach will produce exact results for power-law parameterizations of the form

$$f(X_1, X_2, X_3, \ldots) = X_1^{\alpha_1} X_2^{\alpha_2} X_3^{\alpha_3} \ldots, \tag{64}$$

where X_1, X_2, X_3, ... are lognormally distributed. Attempting to compute $\langle f(\mathbf{x}) \rangle$ using a direct approach in this example produces bad results because the corresponding Taylor expansion is too severely truncated at the second order given the variance of the parameter distribution.

b. Sedimentation flux

The parameterization for ice sedimentation flux is nonlinear, both in its direct and logarithmic formulation, and cannot be integrated exactly over the parameter distribution. Therefore, estimators of the mean sedimentation flux produce approximate results that must be validated against values computed from the Monte Carlo method. As for ice water content, we will show that the logarithmic approach performs satisfactorily because the underlying truncated Taylor expansion captures well the variations of the parameterization over the typical range of parameter values.

To begin with, one may contrast the results obtained by Taylor expanding the parameterization versus its logarithmic counterpart following a direct approach. The truncated expansion for F_{sed} and associated mean value estimator are given by

$$F_{\mathrm{sed}}(\langle \mathbf{x} \rangle) + \nabla F_{\mathrm{sed}}(\langle \mathbf{x} \rangle)^{\mathrm{T}} (\mathbf{x} - \langle \mathbf{x} \rangle) + \frac{1}{2} (\mathbf{x} - \langle \mathbf{x} \rangle)^{\mathrm{T}} \mathbf{H}(\langle \mathbf{x} \rangle) (\mathbf{x} - \langle \mathbf{x} \rangle), \tag{65}$$

$$\langle F_{\mathrm{sed}} \rangle \sim F_{\mathrm{sed}}(\langle \mathbf{x} \rangle) + \frac{1}{2} \mathrm{tr}[\mathbf{H}(\langle \mathbf{x} \rangle) \mathbf{\Sigma}], \tag{66}$$

while those for $\ln(F_{\mathrm{sed}})$ are

$$\ln(F_{\mathrm{sed}})(\langle \mathbf{x} \rangle) + \nabla \ln(F_{\mathrm{sed}})(\langle \mathbf{x} \rangle)^{\mathrm{T}} (\mathbf{x} - \langle \mathbf{x} \rangle) + \frac{1}{2} (\mathbf{x} - \langle \mathbf{x} \rangle)^{\mathrm{T}} \mathbf{H}^{\ln}(\langle \mathbf{x} \rangle) (\mathbf{x} - \langle \mathbf{x} \rangle), \tag{67}$$

$$\langle \ln(F_{\mathrm{sed}}) \rangle \sim \ln(F_{\mathrm{sed}})(\langle \mathbf{x} \rangle) + \frac{1}{2} \mathrm{tr}[\mathbf{H}^{\ln}(\langle \mathbf{x} \rangle) \mathbf{\Sigma}]. \tag{68}$$

As one can see in Fig. 5a, $\langle \ln(F_{\mathrm{sed}}) \rangle$ predicted from Eq. (68) tracks very well the Monte Carlo values across all possible configurations of the correlation

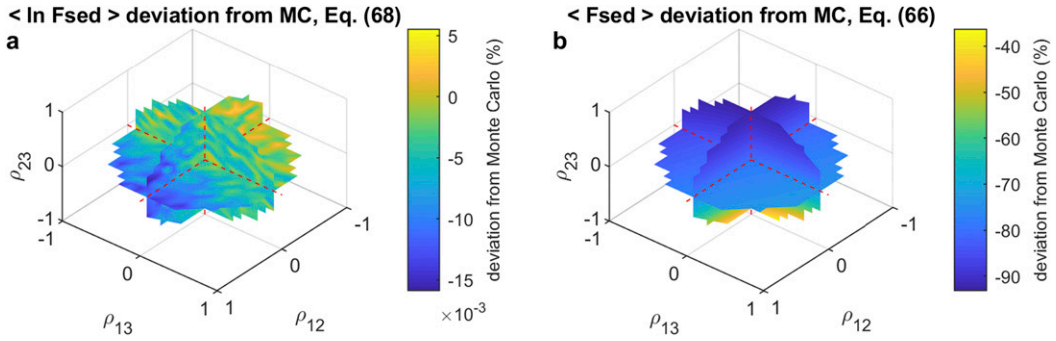

FIG. 5. Deviations of (a) $\langle \ln(F_{\rm sed})\rangle$ and (b) $\langle F_{\rm sed}\rangle$ estimators from corresponding Monte Carlo predictions. The estimators are computed from Eqs. (68) and (66) using a direct Taylor expansion approach. A joint normal distribution of extinction coefficients, effective diameters, and exponential slope parameters with unknown correlation coefficients is used. A normal distribution is chosen for ln(EXT) (parameter 3) based on the mean and variance of extinction measured by CALIOP at 15 km (see Fig. 3). The $\ln(D_e)$ (parameter 2) and $\ln(\lambda)$ (parameter 1) distributions are estimated from Fig. 11a of Heymsfield et al. (2014) and Fig. 7b of Heymsfield et al. (2013) at −71°C, with standard deviations estimated at 0.25 and 0.6, respectively. The mean temperature was calculated from interpolated MERRA-2 data at 15 km in the tropics. The deviations of the estimators from Monte Carlo are functions of the unknown correlation coefficients ρ_{12}, ρ_{13}, and ρ_{23} and are here represented over the planes $\rho_{12} = 0$, $\rho_{13} = 0$, and $\rho_{23} = 0$ for feasible values (cf. Fig. 4).

coefficients. The configurations are those for which the correlation coefficients are within the feasible volume defined by Eq. (52). Note that the absolute deviation from Monte Carlo stays under 15×10^{-3} % while $\langle \ln(F_{\rm sed})\rangle$ varies end to end by 0.4% (not shown). By contrast, $\langle F_{\rm sed}\rangle$ predicted from Eq. (66) does a very poor job tracking the Monte Carlo estimator, as displayed in Fig. 5b.

To understand why Eq. (68) succeeds at being a good estimator whereas Eq. (66) fails, one needs to examine the underlying truncated Taylor expansions [Eqs. (67) and (65)]. Those are plotted at constant extinction in Figs. 6c and 6d. The full function variations are plotted for comparison in Figs. 6a and 6b and the residuals of the truncated expansions in Figs. 6e and 6f. One can see that Eq. (67) is a much better proxy for the full variations of $\ln(F_{\rm sed})$ than Eq. (65) is for $F_{\rm sed}$. The residual of the quadratic expansion of $\ln(F_{\rm sed})$ is small over the domain concentrating most ancillary parameters (6e, domain spanned by the error bars) while that for $F_{\rm sed}$ only vanishes in a tiny neighborhood of the mean parameters (Fig. 6f).

The fact that the variations of $\ln(F_{\rm sed})$ are well captured by the lowest orders of a Taylor expansion can be related to the empirical observation of a compact relationship between mass-weighted velocities v_m and effective diameter. Without loss of generality, the expression for $F_{\rm sed}$ at Eq. (16) may indeed be interpreted as

$$F_{\rm sed}(\lambda, D_e, \mathrm{EXT}) = \mathrm{IWC}(D_e, \mathrm{EXT}) v_m(\lambda, D_e). \quad (69)$$

Many studies in microphysics (e.g., Schmitt and Heymsfield 2009) find that v_m is well modeled by power laws of D_e for the dataset they consider. Therefore, one can expect $F_{\rm sed}$ to vary in the following way:

$$F_{\rm sed} \sim \mathrm{EXT} D_e^{\alpha}, \quad (70)$$

where α is a slowly varying function of the parameters λ, D_e, and EXT. The logarithm of $F_{\rm sed}$ is thus expected to behave approximately as

$$\ln(F_{\rm sed}) \sim \ln(\mathrm{EXT}) + \alpha \ln(D_e) \quad (71)$$

over the typical range of variations of the parameters. A quick inspection of Fig. 6a shows that this is the case indeed. At constant extinction, and for $\ln(\lambda)$ and $\ln(D_e)$ spanning the bulk of parameter values, $\ln(F_{\rm sed})$ seems to vary quasi linearly with $\ln(D_e)$ with no other significant dependency. This can be justified more rigorously from linear sensitivity considerations (see appendix C). One can show in particular that the picture put together at Eq. (71) holds because typical measured values of the microphysics parameters are high enough to minimize sensitivity to λ and preclude significant curvature effects.

Note that the accuracy of Eq. (67) as a proxy for full function variations is not uniform over the parameter space. Outlier effects from the quadratic term become significant when $\ln(\lambda)$ and $\ln(D_e)$ are both small or are both large (Fig. 6e). The quality of the associated mean value estimator at Eq. (68) thus depends on the correlation coefficients between parameters, which control

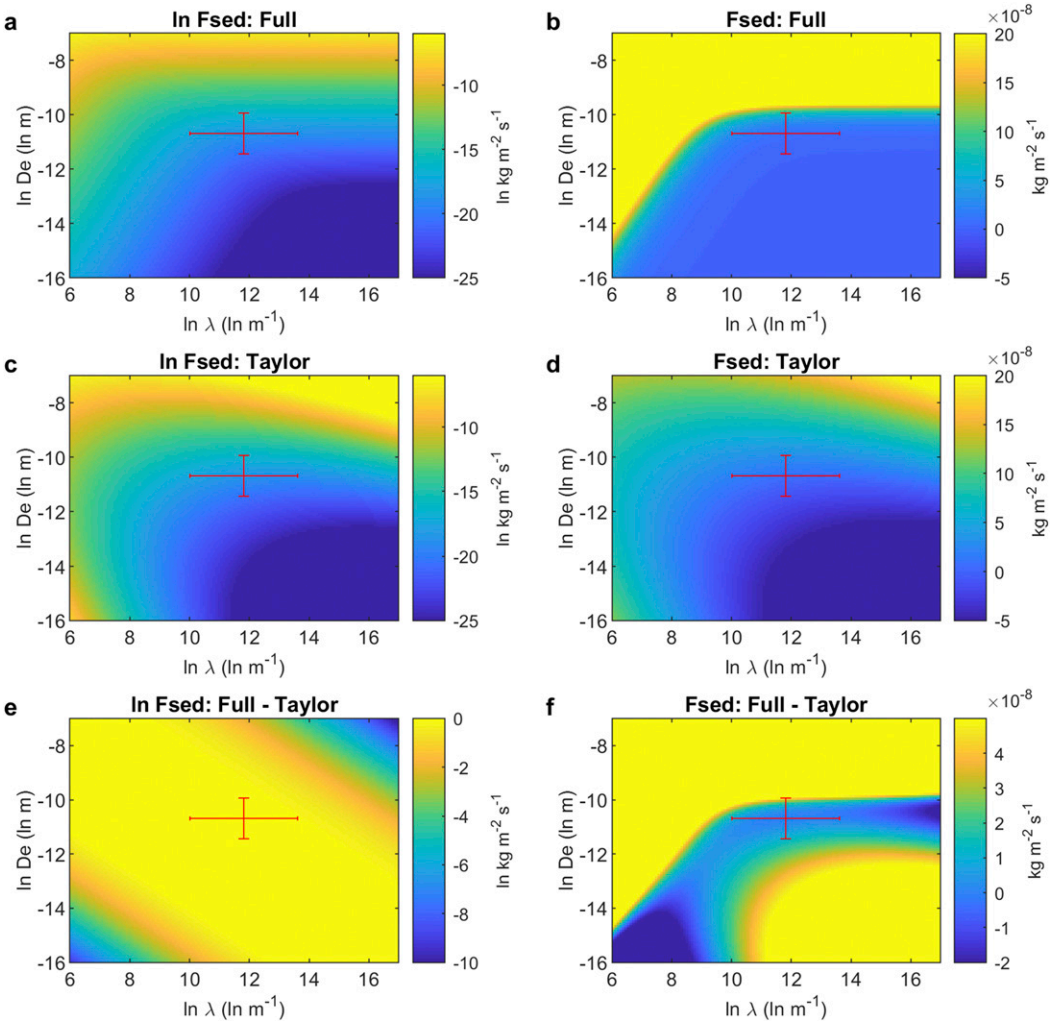

FIG. 6. Variations in the $\ln(\lambda)$–$\ln(D_e)$ plane of (a) $\ln(F_{\rm sed})$ from its full definition, (c) $\ln(F_{\rm sed})$ from its quadratic Taylor expansion, and (e) $\ln(F_{\rm sed})$ from the residuals of the expansion (orders ≥ 3). The variations are computed at constant extinction set to the mean value measured by CALIOP at 15 km (see Fig. 3). The error bars measure a 3-sigma dispersion (i.e., 99.4% of samples are within the area spanned by the error bars) of the distributions of $\ln(D_e)$ and $\ln(\lambda)$ estimated from Fig. 11a of Heymsfield et al. (2014) and Fig. 7b of Heymsfield et al. (2013) at −71°C, with standard deviations estimated at 0.25 and 0.6, respectively. The mean temperature was calculated from interpolated MERRA-2 data at 15 km in the tropics. (b),(d),(f) As in (a), (c), and (e), but for variations of $F_{\rm sed}$ instead of $\ln(F_{\rm sed})$.

the preferential direction of alignment of the parameters. For instance, positive values of the correlation coefficient between microphysics parameters ρ_{12} lead to an accumulation of values in the domain where Eq. (67) diverges from the full function, which degrades the mean value estimator. Negative values of ρ_{12} tend to alleviate this problem. The accuracy of Eq. (68) as an estimator of $\langle \ln(F_{\rm sed}) \rangle$ is thus expected to decrease when ρ_{12} increases, which explains the residual trend perceptible in Fig. 5a.

Having established that Eq. (66) is a bad estimator of $\langle F_{\rm sed} \rangle$, we follow the logarithmic approach of section 3 and construct an estimator from $\langle \ln(F_{\rm sed}) \rangle$ and $\mathrm{Var}[\ln(F_{\rm sed})]$ instead:

$$\langle F_{\rm sed} \rangle \sim F_{\rm sed}(\langle \mathbf{x} \rangle) \times \exp\left(\frac{1}{2} \nabla \ln F_{\rm sed}(\langle \mathbf{x} \rangle)^{\rm T} \mathbf{\Sigma} \nabla \ln F_{\rm sed}(\langle \mathbf{x} \rangle) + \frac{1}{2} \mathrm{tr}[\mathbf{H}^{\ln}(\langle \mathbf{x} \rangle) \mathbf{\Sigma}] + \frac{1}{4} \mathrm{tr}\{[\mathbf{H}^{\ln}(\langle \mathbf{x} \rangle) \mathbf{\Sigma}]^2\} \right). \quad (72)$$

Figure 7a compares $\langle F_{\rm sed} \rangle$ from Eq. (72) to values computed using the Monte Carlo method. As one can

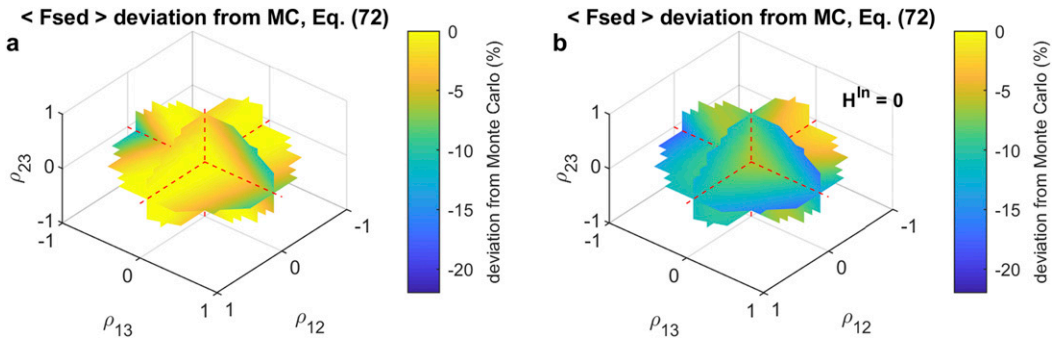

FIG. 7. (a) As in Fig. 5, but for an estimator of $\langle F_{sed} \rangle$ computed from Eq. (72) (logarithmic approach). (b) As in (a), but with curvature terms discarded in Eq. (72) (Hessian matrix $\mathbf{H}^{ln}$ set to 0).

see, Eq. (72) produces results whose absolute value never deviate from the Monte Carlo predictions by more than 15% and are often within 2%–3% of the predictions for most correlation coefficients. Such accuracy compares favorably with the end to end range of variations of the Monte Carlo predictions, which is about a factor of 18. Figure 7b shows the loss of accuracy of predictions from Eq. (72) when the curvature terms in the underlying Taylor expansion are dropped (i.e., $\mathbf{H}^{ln}$ is set to 0). Most configurations of the correlation coefficients show deviations from Monte Carlo values jump from 2%–3% to 5%–10% with maximum deviations reaching 20%. Curvature terms in Eq. (72) are thus important if one wants to correct the estimator to a few percent of the true value and to extend the domain of correction to most correlation coefficients.

More generally, it is the case for nonlinear parameterizations whose variations cannot be exactly captured by a low-order Taylor expansion that the accuracy of the mean value estimator depends on convergence consideration of the Taylor series and on the covariance matrix $\mathbf{\Sigma}$. This implies a dependency on the correlation coefficients between parameters, but also on the variances of the parameters, which differ from problem to problem. Although accuracy should be assessed on a case by case basis, better results are invariably obtained with smaller variances of the parameters. In particular, if the joint distribution of parameters considered throughout this section were to be stratified by values of volume extinction coefficients—such as would be the case for an instantaneous ''scene'' observed by spaceborne lidar—the resulting distribution conditioned on extinction values would exhibit low variance resulting in high estimator accuracy.

6. Discussion

In this section, we interpret the variations in the estimators of $\langle \text{IWC} \rangle$ and $\langle F_{sed} \rangle$ from a cloud physics perspective. We also discuss biases of parameterizations ignoring the variance in particle size distributions and relevant rectification strategies. This point deserves general attention since the information on microphysics parameters often comes from empirical parameterizations derived from in situ measurements (Heymsfield et al. 2014; Thornberry et al. 2017). Because of the way these parameterizations are constructed, the distribution of such parameters reduces to a single number at each temperature (i.e., parameters become deterministic), which biases the estimation of mean ice water content and sedimentation flux when a range of extinction coefficients is involved. The case of the CALIOP V4.1 ice water content parameterization (Heymsfield et al. 2014) using an empirical relationship for ice effective diameter recalled at section 4 [Eq. (43)] is discussed as an example of parameterization where microphysics parameter become deterministic. Results are contrasted with situations where a nonzero amount of variance is added to the parameterization, taking into account the range of uncertainty spanned by ice water content due to degenerate joint parameter distributions (corresponding to the possible range of variations of the correlation coefficients between parameters). Similar results are discussed for the estimated sedimentation flux corresponding to a range of extinction coefficients. Parameterizations featuring deterministic parameters instead of observed distributions shall henceforth be called deterministic parameterizations. Such deterministic parameterizations can be envisioned as a reduction of full parameterizations (using distributed parameters) under the limit of vanishing microphysics parameter variance, that is, $\sigma_1, \sigma_2 \to 0$

a. Ice water content

Figure 8a shows mean ice water content computed from Eq. (61) for a distribution of extinction coefficients and effective diameters as a function of the assumed correlation coefficient between the logarithms of those

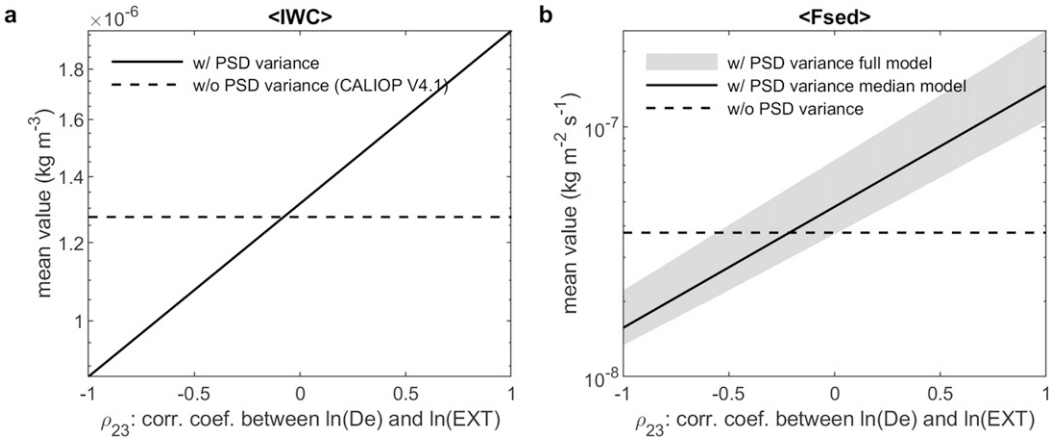

FIG. 8. (a) ⟨IWC⟩ from Eq. (61) as a function of the assumed correlation coefficient between ln(EXT) and ln(D_e) (solid line). A joint lognormal distribution of extinction coefficients and effective diameters is used. A normal distribution is chosen for ln(EXT) based on the mean and variance of extinction measured by CALIOP at 15 km (see Fig. 3). The ln(D_e) distribution is estimated from Fig. 11a of Heymsfield et al. (2014) at −71°C, with an estimated standard deviation of 0.25. The mean temperature was calculated from interpolated MERRA-2 data at 15 km in the tropics. The limit value of ⟨IWC⟩ when the distribution of ln(D_e) reduces to a single number (CALIOP V4.1 parameterization) is symbolically represented as a horizontal dashed line. (b) As in (a), but for $\langle F_{\text{sed}} \rangle$ from Eq. (72). The joint normal distribution is now between extinction coefficients (parameter 3), effective diameters (parameter 2) and exponential slope parameter (parameter 1). The ln(λ) distribution is estimated from Fig. 7b of Heymsfield et al. (2013) at −71°C, with an estimated standard deviation of 0.6. The full solution depends on ρ_{12}, ρ_{13}, and ρ_{23} but is shown as a function of ρ_{23} only, with median model from Eq. (76) overlaid as solid line.

parameters (ρ_{23}). The predicted value of ⟨IWC⟩ grows exponentially with ρ_{23}. From Eq. (62), the minimum and maximum values of ⟨IWC⟩ are given by

$$\langle \text{IWC} \rangle = \frac{\rho_i}{3} \langle \text{EXT} \rangle \langle D_e \rangle \exp(\pm \sigma_2 \sigma_3) \tag{73}$$

and are respectively attained in the totally anticorrelated ($\rho_{23} = -1$) and totally correlated ($\rho_{23} = 1$) cases. The range of predicted values spans a factor

$$\exp(2\sigma_2 \sigma_3) , \tag{74}$$

which amounts to 2.2 for mean parameter values at 15 km (i.e., σ_2 and σ_3 estimated from Heymsfield et al. (2014) and from CALIOP measurements at 15 km, respectively; see caption of Fig. 8a for more details). Note that, in the general case, this factor would depend on the variances of the parameters of the problem.

When the variances of the ancillary parameters vanish (i.e., σ_1, $\sigma_2 \to 0$), as in the CALIOP V4.1 ice water content parameterization, the treatment of particle size distribution becomes deterministic and ⟨IWC⟩ is given by the following deterministic model:

$$\langle \text{IWC} \rangle^{\text{determ}} = \frac{\rho_i}{3} \langle \text{EXT} \rangle D_e^{\text{param}} . \tag{75}$$

The value at Eq. (75) is represented symbolically in Fig. 8a by a horizontal dash line. Note that all dependency upon the correlation coefficient ρ_{23} has vanished since there are no fluctuations of effective radius to consider in this parameterization. Figure 8a shows that a parameterization approach that treats size distributions as deterministic can significantly bias the estimated mean value of ice water content. The estimated value of ⟨IWC⟩ can be biased high by up to a factor of 1.45 or biased low by up to a factor of 1.54 with respect to results from a parameterization taking into account the variance in particle size distribution. Incorporating distributional aspects of the microphysics parameters in parameterizations would require systematic reporting of the joint distributions between extinction coefficients and ice effective diameters from in situ measurements, including the correlation between such parameters.

A physical interpretation of these results can be developed by considering the meaning of the correlation coefficient between the logarithms of extinction coefficients and ice effective diameters (ρ_{23}). A positive correlation coefficient means that positive extinction anomalies are correlated with positive D_e anomalies, meaning that samples with higher projected area have a tendency to occur with heavier particles while samples with lower projected area have a tendency to occur with lighter particles (in a relative sense, i.e., compared to situations with negative correlation coefficients, see below). The predicted value of ⟨IWC⟩ over the distribution of extinction coefficients thus tends to be large in

that case. By contrast, a negative correlation coefficient means that samples with higher projected area have a tendency to occur with lighter particles, and vice versa. This will then produce lower values of ⟨IWC⟩ when averaging over the distribution of extinctions.

b. Sedimentation flux

Figure 8b shows the mean sedimentation flux estimated from Eq. (72) for a distribution of extinction coefficients and microphysics parameters as a function of the correlation between the logarithms of extinction and ice effective diameter (ρ_{23}), by analogy with Fig. 8a. $\langle F_{\text{sed}}\rangle$ is computed over all feasible configurations of the correlation coefficients so the observed dispersion at constant value of the correlation coefficient ρ_{23} is explained by variations in the correlation coefficients between the logarithms of λ and extinction (ρ_{13}) and between the logarithms of λ and effective diameter (ρ_{12}). Most of the variance in $\langle F_{\text{sed}}\rangle$ is controlled by the correlation coefficient between the logarithms of extinction and ice effective diameter (ρ_{23}). ρ_{23} produces a dispersion of $\langle F_{\text{sed}}\rangle$ of about a factor of 10 while ρ_{13} and ρ_{12} produce on their own a dispersion of a factor of 2 only. A median parameterization for $\langle F_{\text{sed}}\rangle$ can thus be derived from Eq. (72) by retaining only the correlation coefficient ρ_{23} and the variances of extinction (σ_3) and effective diameter (σ_2):

$$\langle F_{\text{sed}}\rangle \sim F_{\text{sed}}(\langle \mathbf{x}\rangle)\exp\left(\frac{1}{2}\left\{\left[\frac{\partial \ln(F_{\text{sed}})}{\partial x_2}\right]^2\sigma_2^2+\sigma_3^2\right\} + \frac{\partial \ln(F_{\text{sed}})}{\partial x_2}\sigma_2\sigma_3\rho_{23}\right), \tag{76}$$

where $\partial \ln(F_{\text{sed}})/\partial x_2 = \partial \ln(F_{\text{sed}})/\partial \ln(D_e)$ is computed at Eq. (56) and plotted in Fig. C1. In deriving Eq. (76) from Eq. (72), the curvature terms have been dropped; that is, $\mathbf{H}^{\ln}$ has been set to 0. Note that Eq. (76) implies an underlying model for sedimentation flux of the form

$$F_{\text{sed}} \sim \text{EXT}\, D_e^{\partial \ln(F_{\text{sed}})/\partial x_2}, \tag{77}$$

which exactly corresponds to the empirical scaling law discussed in the previous section [Eq. (70)] upon identifying $\alpha = \partial \ln(F_{\text{sed}})/\partial x_2$. The structure of Eq. (76) for $\langle F_{\text{sed}}\rangle$ is identical to that of Eq. (61) for ⟨IWC⟩ albeit for the coefficient $\partial \ln(F_{\text{sed}})/\partial x_2$, which expresses the greater sensitivity of sedimentation flux to ice effective diameter [$\partial \ln(F_{\text{sed}})/\partial x_2$ effectively acts as a multiplier on σ_2]. This is why the range of values spanned by the median parameterization for sedimentation flux amounts to about a factor of 10, a fivefold increase over ice water content. The minimum and maximum values of $\langle F_{\text{sed}}\rangle$ in the median parameterization are given by

$$\langle F_{\text{sed}}\rangle \sim F_{\text{sed}}(\langle \mathbf{x}\rangle) \times \exp\left[\pm\frac{\partial \ln(F_{\text{sed}})}{\partial x_2}\sigma_2\sigma_3\right], \tag{78}$$

which yields a range of values due to plausible correlations among parameters spanning a factor

$$\exp\left[2\frac{\partial \ln(F_{\text{sed}})}{\partial x_2}\sigma_2\sigma_3\right]. \tag{79}$$

That factor evaluates to 9.32 for mean parameter values at 15 km (i.e., σ_2 and σ_3 estimated from Heymsfield et al. (2014) and from CALIOP measurements at 15 km, respectively; see caption of Fig. 8 for more details) and, in the general case, would depend on the variances of the parameters of the problem.

In the limit of vanishing variance of the ancillary parameters (i.e., $\sigma_1, \sigma_2 \to 0$), $\langle F_{\text{sed}}\rangle$ continuously reaches a limit that does not depend on the correlation coefficient ρ_{23} anymore. The corresponding value is represented symbolically in Fig. 8b by a horizontal dash line. As one can see, the potential estimation bias produced by a parameterization with a deterministic treatment of particle size distribution is much higher than in the case of ice water content. Compared to the median parameterization for $\langle F_{\text{sed}}\rangle$ featuring variance of the microphysics parameter, the deterministic value can be biased high by up to a factor of 3.88 or biased low by up to a factor of 2.40. If one takes into account the additional dependencies of $\langle F_{\text{sed}}\rangle$ upon the correlation coefficients ρ_{13} and ρ_{12} [i.e., full Eq. (72) is used instead of Eq. (76)], the situation gets even worse. Increased sensitivity to effective ice diameter means that the estimation of sedimentation flux for a distribution of extinction coefficients is intrinsically more error prone than that of ice water content.

In summary, the estimator for $\langle F_{\text{sed}}\rangle$ exhibits a structure of variations approximately similar to that for ⟨IWC⟩, albeit with greater sensitivity to D_e. That sensitivity results from a nonlinearity between ice mass and mass-weighted sedimentation velocity v_m. When samples with higher projected area tend to occur with heavier particles, they also happen to fall faster, hence the increase in $\langle F_{\text{sed}}\rangle$ values is proportionally larger than for ⟨IWC⟩ for a same increase of the correlation coefficient ρ_{23}.

7. Conclusions

This paper deals with situations in which observable parameters (directly observed or derived from observations)

have to be complemented with a priori parameters to estimate geophysical quantities of interest.

While the variance in observed parameters is documented by measurements, that in ancillary parameters is often discarded by the use of empirical parameterizations. By mismatching the distributions of observable and ancillary parameters, this can bias the estimation problem.

To resolve such problems, one needs some a priori knowledge of the joint distribution of observable and ancillary parameters and a method to average parameterizations estimating geophysical quantities of interest over the parameter distribution. In contrast to estimators based on sampling strategies or to estimators based on exact integration, we propose a semianalytical approach based on a Taylor expansion that can be applied in nonanalytically tractable cases where a low-order Taylor expansion makes sense.

We have applied the method to estimate ice water content and sedimentation flux in cirrus clouds for a range of volume extinction coefficients retrieved from spaceborne lidar measurements. Extra information on particle size distribution needs to be provided in the form of a finite number of parameters, namely, the effective diameter and the slope of the particle size distribution (when fitted as gamma-type distributions).

We use our semianalytical approach to construct estimators for mean ice water content and sedimentation flux and demonstrate their good performance in benchmarks against Monte Carlo predictions. The case of ice water content is actually trivial as the corresponding estimator recovers exact analytical results.

The value of the estimators depends on the joint distribution between the parameters of the problem, namely, extinction and the ancillary parameters informing on particle size distribution. There is therefore a dependency on the variance of the ancillary parameters and on the correlation coefficients between parameters. The dependency of the estimators on the correlation coefficients appears for finite variances of the ancillary parameters and vanishes when the parameters are treated empirically, as is the case with the empirical ice water content–extinction parameterization of CALIOP V4.1 (Heymsfield et al. 2014). The maximum estimation bias scales with increasing variance of the ancillary parameters and is significantly larger for sedimentation flux than for ice water content. It appears that the correlation coefficient between volume extinction coefficients and ice effective diameters is the most important parameter in setting the value of a regional mean ice water content and sedimentation flux.

One direction to improve estimation procedures is therefore to account for distributional aspects of ancillary parameters. This would require systematic reporting of variances and covariances or other distributional information from in situ measurement programs. Such reporting effort should also cover as many spatial scales as possible and aim at an unbiased sampling strategy, so that the ancillary data might be representative of the remotely sensed data. On the instrumental side, specific requirements need to be placed on high temporal resolution and high accuracy if measuring variability is a mission objective.

The analysis presented here reinforces the importance of variability at the particle size distribution level previously stressed by McFarquhar et al. (2015) and Jackson et al. (2015) and provides a method to efficiently evaluate and understand potential biases in derived geophysical quantities of interest. In particular, it may never be possible to perfectly convert a measured parameter to the geophysical quantity of interest, but to answer relevant scientific questions, it may be sufficient to know the range of possible values accounting for systematic errors in calculating mean quantities, and the approach here is an efficient way to quantify this range.

Acknowledgments. This material is based upon work supported by the National Science Foundation under Grants AGS-1417659 and AGS-1743753.

APPENDIX A

Mean and Variance of F and $\ln(F)$

Let $F(\mathbf{X})$ be a model of a lognormally distributed random vector $\mathbf{X}$, where F here refers to ice water content or sedimentation flux and $\mathbf{X}$ to the vector of input parameters λ, D_e and EXT. The mean and variance of $F(\mathbf{X})$ can be estimated from those of $f(\mathbf{x})$, where $f = \ln(F)$ and $\mathbf{x} = \ln(\mathbf{X})$, and $\mathbf{x}$ is normally distributed.

One can estimate the mean of $f(F)$ by Taylor expanding f in a neighborhood of $\langle F\rangle$ and applying the expectation operator:

$$\langle f(F)\rangle \sim E\left[f(\langle F\rangle) + f'(\langle F\rangle)(F - \langle F\rangle) + \frac{1}{2}f''(\langle F\rangle)(F - \langle F\rangle)^2\right]. \tag{A1}$$

For $f = \ln F$, Eq. (A1) simplifies to

$$\langle f(F)\rangle = \ln\langle F\rangle - \frac{\mathrm{Var}[F]}{2\langle F\rangle^2}. \tag{A2}$$

The variance of $f(F)$ can similarly be estimated from the Taylor expansion. One can write

$$\mathrm{Var}[f(F)] \sim [f'(\langle F\rangle)]^2 \mathrm{Var}[F], \tag{A3}$$

which evaluates to

$$\mathrm{Var}[f(F)] \sim \frac{\mathrm{Var}[F]}{\langle F\rangle^2}. \tag{A4}$$

Therefore, estimators of the mean and variance of $F(\mathbf{X})$ can be obtained from those of $f(\mathbf{x})$ according to

$$\langle F(\mathbf{X})\rangle \sim \exp\left\{\langle f(\mathbf{x})\rangle + \frac{\mathrm{Var}[f(\mathbf{x})]}{2}\right\}, \tag{A5}$$

$$\mathrm{Var}[F(\mathbf{X})] \sim \langle F(\mathbf{X})\rangle^2 \mathrm{Var}[f(\mathbf{x})]. \tag{A6}$$

Equation (A5) means that the mean value $\langle F(\mathbf{X})\rangle$ maps to a value located at a distance $\{\mathrm{Var}[f(\mathbf{x})]\}/2$ from $\langle f(\mathbf{X})\rangle$, in the wing of the distribution of $f(\mathbf{x})$. Coincidentally, Eq. (A5) holds exactly when $F(\mathbf{X})$ is lognormally distributed [i.e., when the values of $f(\mathbf{x})$ are normally distributed].

APPENDIX B

Power-Law Parameterizations for Particle Mass, Area, and Terminal Velocity

Particle mass m and projected area A are parameterized using dimensional power laws of the form

$$m(D) = aD^b, \tag{B1}$$

$$A_r(D) = \frac{A(D)}{\pi D^2/4} = \alpha D^\beta, \tag{B2}$$

where A_r is the ratio of projected area to that of a disc completely enclosing the particle image. It is a more relevant quantity than projected area A for parameterization purposes. We use the values of Schmitt and Heymsfield (2009) and Heymsfield et al. (2013) for the coefficients in Eqs. (B1) and (B2):

$$a = 0.0175 \times 10^{2.02}, \quad b = 2.51,$$
$$\alpha = 0.25 \times 10^{-0.42}, \quad \beta = -0.21.$$

Those values are representative of ice crystals in tropical tropopause layer (TTL) cirrus clouds at temperatures lower than 213.15 K. They have been adjusted for metric units.

To a good approximation, the terminal velocity of falling ice crystals can also be expressed as a dimensional power law of the form

$$v_t(D) = rD^s. \tag{B3}$$

The computation of terminal velocities is mostly a problem of determining the Reynolds number (Re) of falling crystals. We follow the approach of Heymsfield and Westbrook (2010) and evaluate Re from the modified Best number,

$$X^* = \frac{\rho}{\eta^2}\frac{8mg}{\pi A_r^{0.5}}, \tag{B4}$$

using the formula

$$\mathrm{Re} = \frac{\delta_0^2}{4}\left[\left(1 + \frac{4\sqrt{X^*}}{\delta_0^2\sqrt{C_0}}\right)^{1/2} - 1\right]^2. \tag{B5}$$

The coefficients in Eq. (B5) are set to $C_0 = 0.35$ and $\delta_0 = 8.0$, as in Heymsfield and Westbrook (2010), and η is the air dynamic viscosity. Equation (B5) is based on the original formulations of Abraham (1970) and Böhm (1989), modified to reduce the sensitivity of the computed drag force to crystal area ratio A_r. Although originally developed from boundary layer theory, Heymsfield and Westbrook (2010) show that Eq. (B5) may be applied down to the Stokes regime where it yields fall velocities within 30% of the expected values. An expansion of Eq. (B5) at low Reynolds number leads to express v_t as a power law of the form $v_t = rD^s$ with the coefficients:

$$r = \frac{1}{\delta_0^2 C_0}\frac{8ag}{\pi\eta\alpha^{0.5}}, \tag{B6}$$

$$s = -1 + b - 0.5\beta. \tag{B7}$$

Given the mass and shape characteristics assumed at Eqs. (B1) and (B2), ice crystals in the TTL are predicted to fall with $s = 1.615$ (in comparison, perfect spheres would fall with $s = 2$).

One can show that the error on F_{sed} caused by considering all particles falling in the Stokes regime and therefore using Eq. (B3) under the integral sign instead of retaining Eq. (B5) and performing a numerical integration of Eq. (3) is on the order of a few percent at most.

APPENDIX C

Local Power-Law Approximation for F_{sed}

Local power-law approximations of the sedimentation flux can be derived from linear sensitivity considerations. Linearizing $\ln(F_{\mathrm{sed}})$ yields an expression of the form

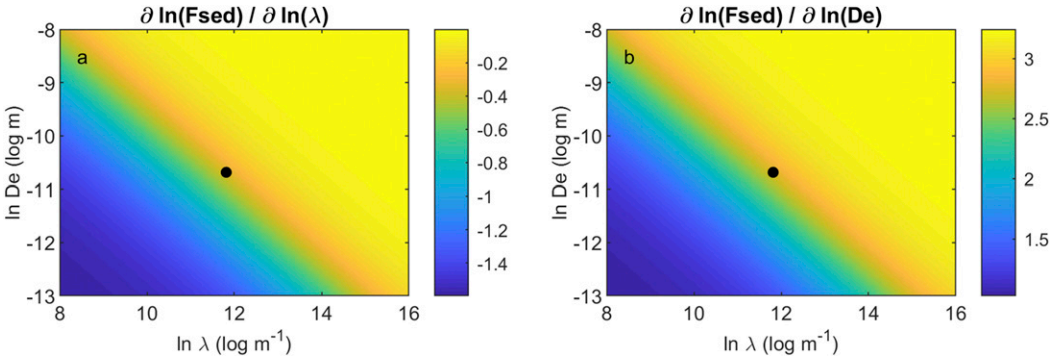

FIG. C1. Variations in the $\ln(\lambda)$–$\ln(D_e)$ plane of (a) $\partial\ln(F_{\rm sed})/\partial\ln(\lambda)$ and (b) $\partial\ln(F_{\rm sed})/\partial\ln(D_e)$. Values are computed from the formula in Eq. (56) for mean parameters at 15 km. The black marker indicates the position of the mean parameters $\langle\ln(\lambda)\rangle$ and $\langle\ln(D_e)\rangle$ at 15 km (for a mean temperature calculated from interpolated MERRA-2 data at 15 km in the tropics).

$$\ln(F_{\rm sed}) \sim \ln({\rm EXT}) + \frac{\partial \ln F_{\rm sed}}{\partial \ln D_e}\ln(D_e) + \frac{\partial \ln F_{\rm sed}}{\partial \ln\lambda}\ln(\lambda) + {\rm constant}, \tag{C1}$$

where the components of the gradient $\partial\ln(F_{\rm sed})/\partial\ln(D_e)$ and $\partial\ln(F_{\rm sed})/\partial\ln(\lambda)$ are to be evaluated at the point of linearization [their full expressions are given at Eq. (56)]. Those components are plotted in Fig. C1 as a function of $\ln(\lambda)$ and $\ln(D_e)$ and from their mean values at 15 km, one concludes that $F_{\rm sed}$ admits at this altitude a local power-law expansion of the form

$$F_{\rm sed} \propto {\rm EXT} D_e^{2.76}\lambda^{-0.35}, \tag{C2}$$

which is equivalent to postulating a mass-weighted velocity of the form

$$\upsilon_m \propto D_e^{1.76}\lambda^{-0.35}. \tag{C3}$$

This tends to confirm the observation that υ_m weakly depends on λ and exhibits a compact relationship with D_e, as noted, for instance, by Schmitt and Heymsfield (2009). This result comes as a direct consequence of the assumption of gamma distributions for particle size distributions. Note also that the result would break down if λ or D_e in ice clouds were lower than the values typically measured. For such values, the sensitivity to λ increases and the local exponents in Eqs. (C2) and (C3) start varying rapidly with the parameters, which is a sign that the linear terms compensate for the missing curvature terms in Eq. (C1).

REFERENCES

Abraham, F. F., 1970: Functional dependence of drag coefficient of a sphere on Reynolds number. *Phys. Fluids*, **13**, 2194–2195, https://doi.org/10.1063/1.1693218.

Abramovich, S., and L. E. Persson, 2016: Some new estimates of the 'Jensen gap.' *J. Inequal. Appl.*, **2016**, 39, https://doi.org/10.1186/S13660-016-0985-4.

Abramowitz, M., and I. A. Stegun, 1972: *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*. Dover, 1046 pp.

Böhm, H. P., 1989: A general equation for the terminal fall speed of solid hydrometeors. *J. Atmos. Sci.*, **46**, 2419–2427, https://doi.org/10.1175/1520-0469(1989)046<2419:AGEFTT>2.0.CO;2.

Cahalan, R. F., W. Ridgway, W. J. Wiscombe, T. L. Bell, and J. B. Snider, 1994: The albedo of fractal stratocumulus clouds. *J. Atmos. Sci.*, **51**, 2434–2455, https://doi.org/10.1175/1520-0469(1994)051<2434:TAOFSC>2.0.CO;2.

Chowdhary, K., M. Salloum, B. Debusschere, and V. E. Larson, 2015: Quadrature methods for the calculation of subgrid microphysics moments. *Mon. Wea. Rev.*, **143**, 2955–2972, https://doi.org/10.1175/MWR-D-14-00168.1.

Dessler, A. E., T. F. Hanisco, and S. Fueglistaler, 2007: Effects of convective ice lofting on H_2O and HDO in the tropical tropopause layer. *J. Geophys. Res.*, **112**, D18309, https://doi.org/10.1029/2007JD008609.

Foot, J. S., 1988: Some observations of the optical properties of clouds. II: Cirrus. *Quart. J. Roy. Meteor. Soc.*, **114**, 145–164, https://doi.org/10.1002/qj.49711447908.

Fueglistaler, S., A. E. Dessler, T. J. Dunkerton, I. Folkins, Q. Fu, and P. W. Mote, 2009: Tropical tropopause layer. *Rev. Geophys.*, **47**, RG1004, https://doi.org/10.1029/2008RG000267.

Garnier, A., J. Pelon, M. A. Vaughan, D. M. Winker, C. R. Trepte, and P. Dubuisson, 2015: Lidar multiple scattering factors inferred from CALIPSO lidar and IIR retrievals of semi-transparent cirrus cloud optical depths over oceans. *Atmos. Meas. Tech.*, **8**, 2759–2774, https://doi.org/10.5194/AMT-8-2759-2015.

Golaz, J.-C., V. E. Larson, and W. R. Cotton, 2002: A PDF-based model for boundary layer clouds. Part II: Model results. *J. Atmos. Sci.*, **59**, 3552–3571, https://doi.org/10.1175/1520-0469(2002)059<3552:APBMFB>2.0.CO;2.

Griffin, B. M., and V. E. Larson, 2013: Analytic upscaling of a local microphysics scheme. Part II: Simulations. *Quart. J. Roy. Meteor. Soc.*, **139**, 58–69, https://doi.org/10.1002/QJ.1966.

Heymsfield, A. J., and C. D. Westbrook, 2010: Advances in the estimation of ice particle fall speeds using laboratory and field

measurements. *J. Atmos. Sci.*, **67**, 2469–2482, https://doi.org/10.1175/2010JAS3379.1.

——, C. Schmitt, and A. Bansemer, 2013: Ice cloud particle size distributions and pressure-dependent terminal velocities from in situ observations at temperatures from 0° to −86°C. *J. Atmos. Sci.*, **70**, 4123–4154, https://doi.org/10.1175/JAS-D-12-0124.1.

——, D. Winker, M. Avery, M. Vaughan, G. Diskin, M. Deng, V. Mitev, and R. Matthey, 2014: Relationships between ice water content and volume extinction coefficient from in situ observations for temperatures from 0° to −86°C: Implications for spaceborne lidar retrievals. *J. Appl. Meteor. Climatol.*, **53**, 479–505, https://doi.org/10.1175/JAMC-D-13-087.1.

Holton, J. R., and A. Gettelman, 2001: Horizontal transport and the dehydration of the stratosphere. *Geophys. Res. Lett.*, **28**, 2799–2802, https://doi.org/10.1029/2001GL013148.

Hunt, W. H., M. A. Vaughan, K. A. Powell, and C. Weimer, 2009: *CALIPSO* lidar description and performance assessment. *J. Atmos. Oceanic Technol.*, **26**, 1214–1228, https://doi.org/10.1175/2009JTECHA1223.1.

Jackson, R. C., G. M. McFarquhar, A. M. Fridlind, and R. Atlas, 2015: The dependence of cirrus gamma size distributions expressed as volumes in N_0-λ-μ phase space and bulk cloud properties on environmental conditions: Results from the Small Ice Particles in Cirrus Experiment (SPARTICUS). *J. Geophys. Res. Atmos.*, **120**, 10 351–10 377, https://doi.org/10.1002/2015JD023492.

Jensen, J. L., 1906: Sur les fonctions convexes et les inégualités entre les valeurs moyennes. *Acta Math.*, **30**, 175–193, https://doi.org/10.1007/BF02418571.

Larson, V. E., and B. M. Griffin, 2013: Analytic upscaling of a local microphysics scheme. Part I: Derivation. *Quart. J. Roy. Meteor. Soc.*, **139**, 46–57, https://doi.org/10.1002/qj.1967.

——, R. Wood, P. R. Field, J. C. Golaz, T. H. Vonder Haar, and W. R. Cotton, 2001: Systematic biases in the microphysics and thermodynamics of numerical models that ignore subgrid-scale variability. *J. Atmos. Sci.*, **58**, 1117–1128, https://doi.org/10.1175/1520-0469(2001)058<1117:SBITMA>2.0.CO;2.

——, J.-C. Golaz, H. Jiang, and W. R. Cotton, 2005: Supplying local microphysics parameterizations with information about subgrid variability: Latin hypercube sampling. *J. Atmos. Sci.*, **62**, 4010–4026, https://doi.org/10.1175/JAS3624.1.

Lee, J., P. Yang, A. E. Dessler, B. C. Gao, and S. Platnick, 2009: Distribution and radiative forcing of tropical thin cirrus clouds. *J. Atmos. Sci.*, **66**, 3721–3731, https://doi.org/10.1175/2009JAS3183.1.

Mathai, A., and S. Provost, 1992: *Quadratic Forms in Random Variables: Theory and Applications*. Vol. 126. Marcel Dekker, 367 pp.

McFarquhar, G. M., and A. J. Heymsfield, 1998: The definition and significance of an effective radius for ice clouds. *J. Atmos. Sci.*, **55**, 2039–2052, https://doi.org/10.1175/1520-0469(1998)055<2039:TDASOA>2.0.CO;2.

——, S. Iacobellis, and R. C. J. Somerville, 2003: SCM simulations of tropical ice clouds using observationally based parameterizations of microphysics. *J. Climate*, **16**, 1643–1664, https://doi.org/10.1175/1520-0442(2003)016<1643:SSOTIC>2.0.CO;2.

——, T.-L. Hsieh, M. Freer, J. Mascio, and B. F. Jewett, 2015: The characterization of ice hydrometeor gamma size distributions as volumes in N0–λ–μ phase space: Implications for microphysical process modeling. *J. Atmos. Sci.*, **72**, 892–909, https://doi.org/10.1175/JAS-D-14-0011.1.

McKay, M. D., R. J. Beckman, and W. J. Conover, 2000: A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, **42**, 55–61, https://doi.org/10.1080/00401706.2000.10485979.

Morrison, H., and A. Gettelman, 2008: A new two-moment bulk stratiform cloud microphysics scheme in the Community Atmosphere Model, version 3 (CAM3). Part I: Description and numerical tests. *J. Climate*, **21**, 3642–3659, https://doi.org/10.1175/2008JCLI2105.1.

Pincus, R., and S. A. Klein, 2000: Unresolved spatial variability and microphysical process rates in large-scale models. *J. Geophys. Res.*, **105**, 27 059–27 065, https://doi.org/10.1029/2000JD900504.

——, H. W. Barker, and J.-J. Morcrette, 2003: A fast, flexible, approximate technique for computing radiative transfer in inhomogeneous cloud fields. *J. Geophys. Res.*, **108**, 4376, https://doi.org/10.1029/2002JD003322.

Räisänen, P., and H. W. Barker, 2004: Evaluation and optimization of sampling errors for the Monte Carlo independent column approximation. *Quart. J. Roy. Meteor. Soc.*, **130**, 2069–2085, https://doi.org/10.1256/QJ.03.215.

Rotstayn, L. D., 2000: On the "tuning" of autoconversion parameterizations in climate models. *J. Geophys. Res.*, **105**, 15 495–15 507, https://doi.org/10.1029/2000JD900129.

Schmitt, C. G., and A. J. Heymsfield, 2009: The size distribution and mass-weighted terminal velocity of low-latitude tropopause cirrus crystal populations. *J. Atmos. Sci.*, **66**, 2013–2028, https://doi.org/10.1175/2009JAS3004.1.

Thornberry, T. D., A. W. Rollins, M. A. Avery, S. Woods, R. P. Lawson, T. V. Bui, and R. S. Gao, 2017: Ice water content-extinction relationships and effective diameter for TTL cirrus derived from in situ measurements during ATTREX 2014. *J. Geophys. Res. Atmos.*, **122**, 4494–4507, https://doi.org/10.1002/2016JD025948.

Tompkins, A. M., 2002: A prognostic parameterization for the subgrid-scale variability of water vapor and clouds in large-scale models and its use to diagnose cloud cover. *J. Atmos. Sci.*, **59**, 1917–1942, https://doi.org/10.1175/1520-0469(2002)059<1917:APPFTS>2.0.CO;2.

Van de Hulst, H. C., 1957: *Light Scattering by Small Particles*. John Wiley and Sons, 470 pp.

Winker, D. M., J. R. Pelon, and M. P. McCormick, 2003: The CALIPSO mission: Spaceborne lidar for observation of aerosols and clouds. *Proc. SPIE*, **4893**, 1–11, https://doi.org/10.1117/12.466539.

——, W. H. Hunt, and M. J. McGill, 2007: Initial performance assessment of CALIOP. *Geophys. Res. Lett.*, **34**, L19803, https://doi.org/10.1029/2007GL030135.

——, M. A. Vaughan, A. Omar, Y. Hu, K. A. Powell, Z. Liu, W. H. Hunt, and S. A. Young, 2009: Overview of the *CALIPSO* mission and CALIOP data processing algorithms. *J. Atmos. Oceanic Technol.*, **26**, 2310–2323, https://doi.org/10.1175/2009JTECHA1281.1.

Yang, Q., Q. Fu, and Y. Hu, 2010: Radiative impacts of clouds in the tropical tropopause layer. *J. Geophys. Res.*, **115**, D00H12, https://doi.org/10.1029/2009JD012393.

Young, S. A., M. A. Vaughan, A. Garnier, J. L. Tackett, J. D. Lambeth, and K. A. Powell, 2018: Extinction and optical depth retrievals for CALIPSO's version 4 data release. *Atmos. Meas. Tech.*, **11**, 5701–5727, https://doi.org/10.5194/amt-11-5701-2018.

Zhang, J., 2002: A new statistically based autoconversion rate parameterization for use in large-scale models. *J. Geophys. Res.*, **107**, 4750, https://doi.org/10.1029/2001JD001484.