The Characterization of Ice Hydrometeor Gamma Size Distributions as Volumes in $N_0–\lambda–\mu$ Phase Space: Implications for Microphysical Process Modeling

GREG M. MCFARQUHAR, TSUNG-LIN HSIEH,* MATT FREER,† JEANA MASCIO,# AND BRIAN F. JEWETT

Department of Atmospheric Sciences, University of Illinois at Urbana–Champaign, Urbana, Illinois

(Manuscript received 16 January 2014, in final form 27 June 2014)

ABSTRACT

Gamma distributions represent particle size distributions (SDs) in mesoscale and cloud-resolving models that predict one, two, or three moments of hydrometeor species. They are characterized by intercept ($N_0$), slope ($\lambda$), and shape ($\mu$) parameters prognosed by such schemes or diagnosed based on fits to SDs measured in situ in clouds. Here, ice crystal SDs acquired in arctic cirrus during the Indirect and Semi-Direct Aerosol Campaign (ISDAC) and in hurricanes during the National Aeronautic and Space Administration (NASA) African Monsoon Multidisciplinary Analyses (NAMMA) are fit to gamma distributions using multiple algorithms. It is shown that $N_0$, $\lambda$, and $\mu$ are not independent parameters but rather exhibit mutual dependence. Although $N_0$, $\lambda$, and $\mu$ are not highly dependent on choice of fitting routine, they are sensitive to the tolerance permitted by fitting algorithms, meaning a three-dimensional volume in $N_0–\lambda–\mu$ phase space is required to represent a single SD. Depending on the uncertainty in the measured SD and on how well a gamma distribution matches the SD, parameters within this volume of equally realizable solutions can vary substantially, with $N_0$, in particular, spanning several orders of magnitude. A method to characterize a family of SDs as an ellipsoid in $N_0–\lambda–\mu$ phase space is described, with the associated scatter in $N_0$, $\lambda$, and $\mu$ observed in prior field campaigns conducted in different conditions. Ramifications for the development of cloud parameterization schemes and associated calculations of microphysical process rates are discussed.

1. Introduction

In many parameterization schemes (e.g., Dudhia 1989; Rotstayn 1997; Reisner et al. 1998; Gilmore et al. 2004; Ferrier 1994; Walko et al. 1995; Meyers et al. 1997; Straka and Mansell 2005; Milbrandt and Yau 2005) developed for models with a myriad of spatial and temporal scales, size distributions (SDs) are represented by either exponential or gamma functions. The number distribution function $N(D)$ for a gamma function is typically given by

$$N(D) = N_0 D^\mu e^{-\lambda D},$$

where $D$ is the maximum dimension of the particle, $N_0$ is the intercept parameter, $\lambda$ is the slope, and $\mu$ is the shape parameter, which is equal to zero for an exponential distribution.

Bulk properties predicted by models, such as mass mixing ratio $q$ and total particle concentration $N_T$ can be written in terms of $N_0$, $\lambda$, and $\mu$, as well as other parameters that describe the characteristics of the cloud particles. Consequently, gamma parameters can equivalently be written in terms of predicted moments. But because many models predict only one or two moments of SDs, additional assumptions about the gamma parameters typically need to be made. For example, in a double-moment parameterization scheme, two parameters can be determined from the predicted moments, with other assumptions needed for the third parameter.

Such assumptions are frequently based on how $N_0$, $\lambda$, and $\mu$, determined from SDs measured in situ during aircraft penetrations through cloud, depend on cloud properties (e.g., $q$) or environmental conditions (e.g., temperature $T$ or vertical air velocity $w$). Knowledge

* Current affiliation: Princeton University, Princeton, New Jersey.
# Current affiliation: University of Utah, Salt Lake City, Utah.

Corresponding author address: Prof. Greg McFarquhar, Department of Atmospheric Sciences, University of Illinois at Urbana–Champaign, 105 S. Gregory Street, MC 223, Urbana, IL 61801. E-mail: mcfarq@atmos.uiuc.edu

DOI: 10.1175/JAS-D-14-0011.1

© 2015 American Meteorological Society
about the dependence of \( N_0, \lambda, \) and \( \mu \) on environmental conditions is important, because the choice of such parameters can have big impacts on simulated cloud properties (e.g., McCumber et al. 1991; McFarquhar et al. 2006). Further, it is now possible to prescribe the relation of \( N_0, \lambda, \) and \( \mu \) on quantities like cloud-scale \( w \) and \( T \) (e.g., Thompson et al. 2008), given that the scales used in models are much finer than those used when bulk moment schemes (e.g., Lin et al. 1983; Rutledge and Hobbs 1983, 1984) were first developed.

Prior studies have fit measured SDs to gamma or exponential functions for both liquid (e.g., Marshall and Palmer 1948; Haddad et al. 1996; Tokay et al. 2008) and ice particles (Gunn and Marshall 1958; Wong et al. 1988; Heymsfield et al. 2002a; McFarquhar and Black 2004; McFarquhar et al. 2007a; Heymsfield et al. 2009). The techniques used to determine fit parameters have ranged from a least squares method minimizing the difference between the observed particle number distribution function and that determined from the fit parameters (e.g., McFarquhar and Heymsfield 1997) to the method of moments (e.g., Smith and Kliche 2005; Smith et al. 2009) to the maximum-likelihood approach (e.g., Haddad et al. 1996).

Knowledge of both the gamma fit parameters and their uncertainties is required. A plethora of studies have investigated uncertainties in the most likely fit parameters derived by specific algorithms. For example, Ulbrich (1985) examined the effect of SD truncation on rainfall integral parameters and on relationships between those variables. Wong and Chidambaram (1985), Chandrasekar and Bringi (1987), Moisseev and Chandrasekar (2007), and others have examined how statistical or stochastic sampling errors or data filtering techniques affect the estimated gamma parameters. Smith et al. (2009) and Handwerker and Straub (2011) further showed a bias in the method of moments used to estimate the gamma parameters, showing that estimators involving higher-order moments could yield physically implausible parameters.

Prior studies have also noted that the \( N_0, \lambda, \) and \( \mu \) characterizing SDs exhibit mutual dependence and hence are not truly independent parameters. For example, Chu and Su (2008) showed that the relation between \( \lambda \) and \( \mu \) for raindrop SDs depended on bulk properties of the measured SDs, and Heymsfield et al. (2002a) showed relations between gamma fit parameters for ice SDs measured in Florida, Brazil, and Kwajalein. However, to the best of our knowledge, there has yet to be a comprehensive investigation on how the selection of fitting algorithm, the various sampling biases and errors, and tolerance allowed on fit parameters affect the most likely estimate of the gamma fit parameters, their interrelation, and, perhaps most importantly, their characterization as a volume of equally plausible values (e.g., McFarquhar and Heymsfield 1997) for ice crystal SDs. These issues are investigated here using in situ observations of ice SDs obtained during recent field campaigns in the Arctic (McFarquhar et al. 2011, 2013; Jackson et al. 2012) and tropics (Heymsfield et al. 2009; Zipser et al. 2009).

The remainder of this paper is organized as follows. Section 2 briefly describes the measurements of the in situ ice particle hydrometeor SDs that are used to investigate how choice of fitting technique and allowed tolerance affect the derived fit parameters. Section 3 describes the different techniques used to determine \( N_0, \lambda, \) and \( \mu \) from the in situ SDs. Section 4 shows the necessity of representing the fit parameters characterizing a single SD as a volume of equally realizable solutions in \( N_0-\lambda-\mu \) phase space and gives examples of such volumes. Section 5 describes an approach for characterizing families of SDs as ellipsoids in \( N_0-\lambda-\mu \) phase space, and section 6 discusses the significance of the study and offers concluding remarks.

## 2. In situ observations of SDs

To ensure that the techniques developed for this study apply to a broad range of conditions, data from two separate field projects in geographically diverse regimes are used to characterize the SDs. The first set of data was obtained in tropical waves, storms, and depressions and in hurricanes during the National Aeronautic and Space Administration (NASA) African Monsoon Multidisciplinary Analyses (NAMMA) campaign based out of the Cape Verde Islands (Zipser et al. 2009; Heymsfield et al. 2009; Lawson et al. 2010), and the second set in arctic cirrus sampled during transits between Fairbanks and Barrow, Alaska, during the 2008 Indirect and Semi-Direct Aerosol Campaign (ISDAC) (McFarquhar et al. 2011, 2013; Jackson et al. 2012).

The NAMMA data were collected by probes installed on the NASA DC-8 during August–September 2006. Data were collected off the coast of Africa in Tropical Depression 8 (Hurricane Helene, 12 September), Tropical Storm Debby (23 August), two developing African easterly waves (pre-Ernesto on 19 and 20 August and pre-Gordon on 3 and 4 September), and three nondeveloping waves (25 August, 1 September, and flights on 8 and 9 September). The ISDAC data were collected by probes installed on the National Research Council of Canada (NRC) Convair 580 on six research flights on 4, 5, 13, 19, 25, and 27 April 2008; only data collected during the portions of the flight through cirrus were used in this investigation. McFarquhar et al. (2011) describe those flights and data collected in more detail.
Optical array probes were used to determine $N(D)$ over a broad range of cloud and precipitation particle sizes in both projects. The two-dimensional stereo probe (2DS), cloud imaging probe (CIP), and precipitation imaging probe (PIP), nominally measuring particles with $10 < D < 1500 \, \mu m$, $25 < D < 1550 \, \mu m$, and $100 < D < 6200 \, \mu m$, respectively, were used during NAMMA. However, the probes used in this analysis are the 2DS for $25 < D < 1000 \, \mu m$ and the PIP for $D > 1000 \, \mu m$. Many different probes were installed on the NRC Convair 580 during ISDAC, with redundant measures of SDs over the most important size ranges.

The choice of probes to use in the relevant size ranges and cutoffs between probes was determined based on knowledge of the ranges over which different probes are regarded as reliable, by comparing SDs over common size ranges, and by conducting mass closure studies. For example, although the two-dimensional cloud probe (2DC) and CIP can provide information on $N(D)$ for $D < 150 \, \mu m$, these data were not used because the depth of field is not well defined in these size ranges (Baumgardner and Korolev 1997) and because Korolev et al. (2011) and Jackson et al. (2014) showed shattered remnants can dominate $N(D)$ for $D < 500 \, \mu m$. The probes used in ISDAC had shatter mitigating tips (Korolev et al. 2011), but those used during NAMMA did not. All data were processed using filtering techniques to remove shattered particles based on the interarrival time between particles following Field et al. (2006). Image reconstruction techniques (Heymsfield and Parrish 1978) and diameter correction algorithms were also applied.

Forward-scattering probes used to measure SDs included the cloud and aerosol spectrometer (CAS) during NAMMA, and the forward-scattering spectrometer probe (FSSP) and cloud droplet probe (CDP) during ISDAC. Data from the CAS, CDP, and FSSP are not used, because past studies (Gardiner and Hallett 1985; Gayet et al. 1996; Field et al. 2003; McFarquhar et al. 2007b, 2011) have shown significant overestimates of $N(D)$ from the shattering of large ice crystals on probe shrouds and inlets.

Although SDs measured by the 2DS can be contaminated by shattered artifacts, data collected by the 2DS during NAMMA were used after appropriate algorithms to remove shattered artifacts (Lawson 2011) were applied. The ratio of the concentration measured by the 2DS for crystals with $D$ between 150 and $1000 \, \mu m$ to that measured by the CIP, and by the 2DS with $1000 < D < 3200 \, \mu m$ to that measured by the PIP during NAMMA is plotted as a function of $D$ for different DC-8 true airspeeds in Fig. 1; particles larger than the array width of the 2DS are obtained from reconstruction and hence have larger uncertainties. In the overlap region of $600 < D < 1200 \, \mu m$ between the 2DS and CIP–PIP, the $N(D)$ matched well with an average difference of 27% for NAMMA. There is no strong dependence of the ratios on true airspeed.

As seen in similar comparisons (Lawson et al. 2006; Jackson et al. 2012), the 2DS concentrations are larger than those of the CIP, with the ratio decreasing as $D$ increases. Although this is most likely caused by the slower response time of the CIP, the results of this comparison cannot be generalized to any 2DS–CIP comparison because the response time of CIPs varies. In addition, Lawson (2011) found that shatter-removal algorithms were more effective than shatter mitigating tips for removing artifacts for 2DS data, whereas Korolev et al. (2011) found tips more effective than algorithms for the 2DC probes, which is most similar to the CIP. Thus, the 2DS was used to characterize $N(D)$ for $25 < D < 1000 \, \mu m$ during NAMMA; the size bin of $15 < D < 25 \, \mu m$ was not used, because a prior study (Jackson et al. 2012) suggested large discrepancies in this channel compared to data measured by another probe, the cloud droplet probe. In the absence of any better information, a threshold $D$ of $1000 \, \mu m$ was chosen for switching between the 2DS and PIP for the composite SDs. The PIP is used to represent $N(D)$ for the larger particle sizes, because its larger sample volume allows for a more significant sampling. For ISDAC. Jackson et al. (2012) determined that the SDs could be best represented by a combination of measurements from the 2DS ($25 < D < 300 \, \mu m$), the 2DC ($300 < D < 800 \, \mu m$), and the two-dimensional precipitation probe (2DP) ($D > 800 \, \mu m$) from a combination of probe intercomparisons and mass closure tests: these same SDs are used in this study.
The SDs were averaged for 10 s, or over about 2.1 and 1.0 km for NAMMA and ISDAC, respectively, given the true airspeeds of the DC-8 and the NRC Convair 580. The 10-s average represents a trade-off between greater horizontal resolution and reduced sampling uncertainty. Assuming that the particles are randomly distributed in space, statistical sampling arguments show that the standard deviation in the probability distribution of the number of particles measured in each size bin is proportional to the square root of the number of particles sampled. Error bars in the SDs were thus calculated as 
\[ (p \pm p^{1/2})/(V_s \Delta D), \]
where \( p \) is the actual number of particles counted in each bin for the 10-s-averaged SD, \( V_s \) is the sample volume, and \( \Delta D \) is the bin width (Hallett 2003). There are also other uncertainties in the SDs derived from NAMMA and ISDAC because of the unknown effectiveness of routines designed to remove shattered artifacts, statistical sampling issues, and the choices of different probes to characterize the SDs in different projects. However, as this study is directed mainly at determining the efficacy and uncertainties associated with the fitting techniques themselves and not the characteristics of the SDs, the other uncertainties in the measured SDs are not quantified.

### 3. Fit techniques

Several techniques were used to determine the parameters of gamma distributions best matching observed SDs: 1) a technique based on the moments derived from the observed SDs (Heymsfield et al. 2002a), 2) a newly developed incomplete gamma fitting technique (IGF), 3) a discrete version of the IGF, 4) a standard gamma fit minimizing the \( \chi^2 \) difference between observed and fit SDs, and 5) a normalized gamma fit in which the \( \chi^2 \) difference was weighted by the inverse of the observed SD. These fit techniques and their application to SDs observed during NAMMA and ISDAC are described in this section.

Before describing the fitting techniques, a common nomenclature is introduced. The observed number distribution function in the \( i \)th discrete size bin centered at particle maximum dimension \( D_i \) in a bin of width \( \Delta D_i \) is represented as \( N(D_i) \). This corresponds to the composite SD determined from the 2DS and PIP (NAMMA) or from the 2DS, 2DC, and 2DP (ISDAC). To associate the most physical meaning with the fit coefficients, the number distribution function for the fit distributions is represented as

\[ N_f(D) = N_f \left( \frac{D}{D_0} \right)^\mu e^{-\lambda D} \]

for any \( D \) or as \( N_f(D_i) \) when computed at the midpoint of the bins corresponding to the observations, where \( D_0 \) is some constant, defined here as \( 1.0 \times 10^3 \) \( \mu \)m. The use of \( D_0 \) is mandated so that \( N_o \) has identical units to \( N(D) \) (e.g., \( \text{cm}^{-3} \mu \text{m}^{-1} \) or \( \text{m}^{-4} \)). If the distribution function was written in the form of Eq. (1), \( N_o \) would have units involving a length scale raised to the \( -\mu \) power, which is unphysical and more difficult to interpret.

Many fitting techniques require computation of various moments of the SDs. The \( n \)th moment of the observed SD, \( M_{on} \), is determined as

\[ M_{on} = \sum_i D_i^n N(D_i) \Delta D_i, \]

where the summation is over all bins between a minimum dimension \( D_{\text{min}} \) and maximum dimension \( D_{\text{max}} \) where observations are available. Using the error bars that characterize the uncertainties in the measured SDs, a minimum \( [N_{\text{min}}(D_i)] \) and maximum \( [N_{\text{max}}(D_i)] \) number distribution function in each size bin can be determined as

\[ N_{\text{min}}(D_i) = N(D_i) - \frac{\sqrt{p_i}}{V_s \Delta D_i} \]

and

\[ N_{\text{max}}(D_i) = N(D_i) + \frac{\sqrt{p_i}}{V_s \Delta D_i}, \]

where \( p_i \) is the number of particles sampled in the \( i \)th bin.

The second term on the right-hand side of Eqs. (4) and (5) could also be replaced by any other suitable uncertainty estimate in the SD. From \( N_{\text{min}}(D_i) \) and \( N_{\text{max}}(D_i) \), the minimum \( (M_{on,\text{min}}) \) and maximum \( (M_{on,\text{max}}) \) value of the \( n \)th moment can be computed as

\[ M_{on,\text{min}} = \sum_i D_i^n N_{\text{min}}(D_i) \Delta D_i \]

and

\[ M_{on,\text{max}} = \sum_i D_i^n N_{\text{max}}(D_i) \Delta D_i, \]

respectively. The \( n \)th moment of the fit SD \( (M_f) \) is determined as

\[ M_f = \int_{D_{\text{min}}}^{D_{\text{max}}} N_f(D)D^n dD, \]

where the integration of the fit distribution covers the same range \( D_{\text{min}} \) to \( D_{\text{max}} \) on which the original data were based. For some techniques, the \( n \)th moment of the fit SD must be computed using the same bin boundaries and midpoints upon which the original in situ data were based. This is represented as \( M_{fon} \) and is given by
where three parameters uniquely describe a gamma distribution, it is possible to determine them from three moments of a measured SD. Heymsfield et al. (2002a) used such an approach, developing equations by analytically integrating the gamma function between 0 and $D_{\text{max}}$ and determined $\mu$ as the real root of the fourth-order polynomial given by

$$
(1-F)\mu^4 + (8-18F)\mu^3 + (24-119F)\mu^2 + (32-342F)\mu + (16-360F) = 0,
$$

where $F$ is given by

$$
F = \frac{M_{o2}^2}{M_{ob} M_{o1}^4},
$$

and $M_{on}$ obtained from the observed SDs. The other fit parameters were given by

$$
\lambda = \frac{M_{o1}(\mu+2)}{M_{o2}},
$$

and

$$
N_0 = \frac{M_{o1} \lambda^{\mu+2} D_0^\mu}{\Gamma(\mu+2)},
$$

where the extra factor $D_0^\mu$ in Eq. (13) is added because of the difference between Eq. (2) used here and Eq. (1) in Heymsfield et al. (2002a). The Heymsfield et al. (2002a) method has the advantage of ease of use but is inconsistent in that its derivation assumes integration between 0 and $D_{\text{max}}$, but the observed $M_{on}$ are only based on crystals with dimensions between $D_{\text{min}}$ and $D_{\text{max}}$. Further, no information about the possible range of the derived fit parameters is provided.

In reality, moments derived by integrating the gamma distribution should match the observed moments only when the integration is performed from $D_{\text{min}}$ to $D_{\text{max}}$. Thus, Freer and McFarquhar (2008) developed an incomplete gamma fitting method, where fit parameters were constrained so that moments derived using the incomplete gamma function to integrate the gamma function matched those obtained from measured SDs as closely as possible. By computing three moments from the observed SDs and forcing $M_{on} = M_{fn}$, there are three equations with three unknowns ($\lambda$, $\mu$, and $N_0$) for which a unique solution should exist. However, because no analytic solution to the resulting system of equations exists, $N_0$, $\mu$, and $\lambda$ are solved by minimizing

$$
\chi^2 = \sum_{j=1}^{3} \frac{3}{\sqrt{M_{on}(j) M_{fn}(j)}} \left( \frac{M_{on}(j) - M_{fn}(j)}{M_{on}(j) M_{fn}(j)} \right)^2,
$$

where $n(j)$ for $j = 1, 2, 3$ and 3 indicates the three moments used in the minimization procedure [e.g., $n(j) = 0$ if the zeroth moment or total number concentration is computed]. Handwerker and Straub (2011) also describe an iterative solution for determining gamma size distribution parameters based on moments with arbitrary accuracy, which is essentially equivalent to the IGF method.

Because observed SDs are defined only at specific discrete bin sizes, there can be large differences in integrated and observed moments simply because the integrated function is not computed at the same $D$. This can cause differences between $M_{on}$ and $M_{fn}$, especially when there is a large change in $N(D)$ occurring between the boundaries of some bins. For example, for a gamma distribution with $N_0 = 1.0 \times 10^7 \text{ m}^{-3}$, $\mu = 1$, and $\lambda = 1.0 \times 10^4 \text{ m}^{-1}$ with $D_{\text{min}} = 25 \mu\text{m}$ and $D_{\text{max}} = 30 000 \mu\text{m}$ as plausible values for an ice crystal SD, the first, second, third, and sixth moments computed from analytic integration are, respectively, $2.00 \times 10^{-3} \text{ m}^{-2}$, $6.00 \times 10^{-2} \text{ m}^{-1}$, $2.40 \times 10^{-10} \text{ m}^3$, and $5.04 \times 10^{-20} \text{ m}^3$; $1.62 \times 10^{-3} \text{ m}^{-2}$, $5.58 \times 10^{-7} \text{ m}^{-1}$, $2.32 \times 10^{-10} \text{ m}^3$, and $4.32 \times 10^{-20} \text{ m}^3$ when computed using an incomplete gamma function; and $1.62 \times 10^{-3} \text{ m}^{-2}$, $5.57 \times 10^{-7} \text{ m}^{-1}$, $2.32 \times 10^{-10} \text{ m}^3$, and $4.31 \times 10^{-20} \text{ m}^3$ when computed using the discrete bin sizes used to characterize SDs measured during NAMMA. To determine the potential impact of differences in moments on derived fit parameters, a discrete incomplete gamma fit (DIGF) was developed as an alternate fitting procedure. The DIGF is identical to the IGF, except that the $M_{on}$ were used instead of the $M_{fn}$ when minimizing the difference between the observed and fit moments. Thus, the DIGF minimized

$$
\chi^2 = \sum_{j=1}^{3} \frac{3}{\sqrt{M_{on}(j) M_{fn}(j)}} \left( \frac{M_{on}(j) - M_{fn}(j)}{M_{on}(j) M_{fn}(j)} \right)^2
$$

to derive the most likely $N_0$, $\mu$, and $\lambda$, as well as the corresponding minimum $\chi^2$.

Other fitting procedures were also used to determine the gamma fit parameters. The $\chi^2$ difference between the observed number distribution function at specific bin sizes $N(D_i)$ and the fit number distribution function at the same sizes $N_f(D_i)$ defined by

$$
\chi^2 = \sum_{i=1}^{N} \left[ N(D_i) - N_f(D_i) \right]^2
$$
was minimized to determine \( N_0, \lambda, \) and \( \mu \). An alternate fitting method gave alternate weights to the terms in the summation in Eq. (14), rather than weighting them all identically. Following McFarquhar and Heymsfield (1997), the \( \chi^2 \) difference defined by

\[
\chi^2 = \frac{N}{N_f} \left[ \frac{N(D) - N_f(D)}{N_f(D)} \right]^2
\]  

(17)

was minimized in a procedure henceforth called the normalized gamma fit.

With the exception of the Heymsfield et al. (2002a) technique, the fit procedures minimized \( \chi^2 \) through an iterative procedure that determined the minimum \( \chi^2 \) (hereafter \( \chi^2_{\text{min}} \)) and corresponding \( N_0, \mu, \) and \( \lambda \) once \( \chi^2 \) changes in a step were less than a predefined threshold. Thereafter, it was assumed that all \( N_0, \mu, \) and \( \lambda \) that had \( \chi^2 < \chi^2_{\text{min}} + \Delta \chi^2 \) were equally plausible solutions. The \( \Delta \chi^2 \) value used to define the confidence region in the three-dimensional phase space of \( N_0, \mu, \) and \( \lambda \) ideally should contain a large percentage of the probability distribution function (Press et al. 1992) but not such a large percentage that implausible fit parameters are included.

In this study, the uncertainty in measured SDs was used to define \( \Delta \chi^2 \). For the IGF fitting technique \( \Delta \chi^2 = \text{max}(\Delta \chi^2, \chi^2_{\text{min}}) \), where

\[
\Delta \chi^2 = \frac{1}{2} \left\{ \sum_{j=1}^{3} \left[ \frac{M_{\text{on},\text{min}(j)} - M_{\text{on}(j)}}{\sqrt{M_{\text{on},\text{min}(j)} M_{\text{on}(j)}}} \right]^2 + \sum_{j=1}^{3} \left[ \frac{M_{\text{on},\text{max}(j)} - M_{\text{on}(j)}}{\sqrt{M_{\text{on},\text{max}(j)} M_{\text{on}(j)}}} \right]^2 \right\}.
\]  

(18)

with \( \chi^2_{\text{min}} \) equal to the minimum value of \( \chi^2 \) in Eq. (14), which corresponds to the occurrence of the most likely fit parameters. The \( \Delta \chi^2 \) value represents the average difference in \( \chi^2 \) between moments used to determine the fit parameters and moments determined from SDs using the minimum and maximum values of \( N(D) \) obtained from statistical sampling uncertainties. When \( \Delta \chi^2 < \chi^2_{\text{min}} \), the gamma function fit moments are further from those of the observed SD than are the moments derived from the minimum or maximum SDs. For these cases, \( \Delta \chi^2 \) is set equal to \( \chi^2_{\text{min}} \) to permit larger variance in fit parameters because the gamma model does not match the observed SD very well. While SDs that did not fit the gamma model well could have been excluded from subsequent analysis, such a step was not taken because parameterizations for numerical models assume that all SDs are represented by gamma functions. If poorly fit SDs were excluded from analysis, then the variance in naturally occurring SDs would be reduced and hence not well represented in model schemes. Thus, the approach above allows for increased uncertainties when gamma functions do not well represent measured SDs. The tolerance in \( \Delta \chi^2 \) allowed for other fitting methods is similarly defined.

The implementation of the fitting algorithms provides both the most likely values of the gamma fit parameters as well as a range of possible values that defines a volume of equally realizable solutions. To represent this volume in \( N_0-\mu-\lambda \) phase space in a framework that can be easily incorporated into Monte Carlo modeling simulations to determine how uncertainties in fit parameters affect generated model parameters, an ellipsoid is used to represent the volume. If a vector \( \mathbf{x}_0 = (N_0, \mu, \lambda) \) corresponds to the minimum value of \( \chi^2 \) defined by Eq. (14), (15), (16), or (17), the Taylor expansion to second order of \( \chi^2(\mathbf{x}) \) at \( \mathbf{x}_0 \) is given by

\[
\chi^2(\mathbf{x}) = \chi^2(\mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0)^T \mathbf{J}(\mathbf{x}_0) (\mathbf{x} - \mathbf{x}_0),
\]  

(19)

where \( \mathbf{J} \) is the Jacobian of \( \chi^2 \), which is 0 at \( \mathbf{x}_0 \) (the minimum), and \( \mathbf{H} \) is the Hessian matrix of \( \chi^2 \), which depends only on \( \mathbf{x}_0 \). Therefore,

\[
\Delta \chi^2 = \chi^2(\mathbf{x}) - \chi^2(\mathbf{x}_0) = \frac{1}{2} (\mathbf{x} - \mathbf{x}_0)^T \mathbf{H}(\mathbf{x}_0) (\mathbf{x} - \mathbf{x}_0),
\]  

(20)

describes the equation of an ellipsoid once \( \Delta \chi^2 \) is specified.

If the eigenvalues of the inverse of \( \mathbf{A} = \mathbf{H}(\mathbf{x}_0)/(2\Delta \chi^2) \) are given by \( a^2, b^2, \) and \( c^2 \), and the eigenvectors are \( \hat{a}, \hat{b}, \) and \( \hat{c} \), the equation of the ellipsoid can be written as

\[
\left( \frac{r_1}{a} \right)^2 + \left( \frac{r_2}{b} \right)^2 + \left( \frac{r_3}{c} \right)^2 = 1,
\]  

(21)

where

\[
\mathbf{x} - \mathbf{x}_0 = r_1 \hat{a} + r_2 \hat{b} + r_3 \hat{c},
\]  

(22)

and the volume of the ellipsoid is given by \( 4/3 \pi abc = 4/3 \pi \sqrt{\det(\mathbf{A}^{-1})} \). Thus, knowledge of the eigenvectors and eigenvalues of the Hessian matrix gives a complete characterization of the volume of equally realizable solutions. The Hessian matrix is determined using an
algorithm from Moshtagh (2006), which effectively finds the minimum-volume ellipsoid enclosing all \((N_0, \mu, \lambda)\), whose corresponding \(\chi^2 < \Delta \chi^2\), by converting the problem into a concave optimization problem, which is solved by a gradient ascent method.

Because different techniques are used to fit observed SDs, it is important to determine if discrepancies in gamma fit parameters derived in past studies are caused by natural variations in SDs or by use of varying fit techniques. It is also important to determine how the interrelationship between gamma fit parameters depends on the fitting technique and variance in allowed parameters. These issues are addressed in the next section.

4. \(N_0, \lambda, \text{ and } \mu\) parameters derived from fits to single SDs

Before examining the variation of gamma fit parameters with environmental and cloud conditions, the variance of parameters due to choice of fitting technique and the allowed confidence region around the minimum \(\chi^2\) for the fit must be understood. These issues are addressed in this section using one SD measured during NAMMA at 1335:07 UTC 4 September 2006 and another measured during ISDAC at 1808:05 UTC 4 April 2008.

All five fitting techniques described in section 3 were used to determine the most likely gamma fit parameters describing these two SDs. The IGF and DIGF techniques were implemented using the first, second, and sixth moments (i.e., \(M_{01}, M_{02}, \text{ and } M_{06}\)) of the observed SD in order to match the choice of moments in Heymsfield et al. (2002a) and to choose moments representative of those appropriate for model parameterization and remote sensing retrieval schemes. For example, the second moment is closely related to extinction and the sixth moment to radar reflectivity; however, there is not exact correspondence because ice crystals are nonspherical.

Figure 2 shows an SD observed during NAMMA with the best fits obtained from the five different techniques. Although the results of the fits appear similar to each other and the observed SD, there is actually a spread in the parameters describing the SD. For example, the IGF gives \(N_0 = 9.92 \times 10^{10} \text{ m}^{-4}\), \(\mu = 1.62\), and \(\lambda = 1.05 \times 10^4 \text{ m}^{-1}\); the Heymsfield et al. (2002a) equations give \(N_0 = 2.66 \times 10^{11} \text{ m}^{-4}\), \(\mu = 1.84\), and \(\lambda = 1.12 \times 10^4 \text{ m}^{-1}\); and the standard gamma fit gives \(N_0 = 7.81 \times 10^{12} \text{ m}^{-4}\), \(\mu = 2.54\), and \(\lambda = 1.40 \times 10^5 \text{ m}^{-1}\). The moments derived from the IGF and DIGF parameters match those of the observed SD well, with differences typically less than 1%. For example, the first, second, and sixth moments differ by 0.1%, 0.3% and, 10.4% from those observed.

On the other hand, the same moments from the standard gamma fit differ by 4.1%, 1.6%, and 15.9% from those observed. But, because of the visual similarity of the fits, it is difficult to identify which fitting technique provides the optimal fit. In fact, the optimal result and fitting technique may depend on the application for which the fit parameters are required. However, because of the large variation in gamma fit parameters, such as of a factor of 80 in the case of \(N_0\), the cause of the large variation needed to be explored.

To investigate the performance of the different fitting techniques, the moments derived from the observed SDs were compared against those determined from the fit SDs. The moments from the different fit SDs are calculated by integrating between \(D_{\text{min}}\) and \(D_{\text{max}}\) in a way that is consistent with the manner in which the fit scheme operates. For example, for the IGF, the fit moments are calculated analytically using the incomplete gamma function, but, for the DIGF, they are calculated by numerically integrating using the number distribution function calculated at the midpoint of each of the observed size bins.

Figure 3 shows the relationship between the first moment \(M_1\) and the second moment \(M_2\) for several SDs observed during NAMMA and for various fits to these SDs. The length of the line connecting the moments derived from the fits to those from the observed SDs is proportional to the difference of the moments. When the moments for the fits are equal to those observed, they are not visible, indicating there was very close agreement between moments. Not surprisingly, the moments for the IGF and DIGF fits are closest to those
of the observed SDs, because those two techniques were specifically designed to force the moments between $D_{\text{min}}$ and $D_{\text{max}}$ to closely match those observed. For instance, $M_1$ is 0.45% (0.80%) different on average than $M_{\text{fit}}$ for the IGF (DIGF) fits, but 5.7% (10%, 9.8%) for the standard (normalized, Heymsfield et al. 2002a) fits. Similarly, $M_2$ is 0.69% (1.2%) different than $M_{\text{fit}}$ for the IGF (DIGF) fits, but 10% (11%, 5.0%) for the standard (normalized, Heymsfield et al. 2002a) fits. Thus, for applications requiring the best estimate of observed moments, the IGF and DIGF techniques would provide the best fit. Hence, the IGF is used in subsequent investigations to examine the sensitivity of the fit parameters to the allowed tolerance of the fit.

With the suggestion from Fig. 2 that a single $N_{0\gamma-\mu-\lambda}$ triplet cannot characterize a single SD, it was necessary to determine if volumes in $N_{0\gamma-\mu-\lambda}$ phase space characterizing the SD within an allowed confidence region could explain the variation seen in Fig. 2. To do this, a $\Delta \chi^2$ that estimates how close $\chi^2$ must be to $\chi^2_{\text{min}}$ for deriving acceptable fit parameters was selected based on the statistical uncertainty in the measured SD, as explained in section 2. However, it is possible that a different choice of $\Delta \chi^2$ may be more appropriate for other specific problems. For instance, for calculation of cloud radiative effects, selecting $\Delta \chi^2$ providing a derived surface radiation budget within 5 W m$^{-2}$ might be appropriate; or, for a model parameterization scheme, $\Delta \chi^2$ providing derived mass-weighted fall speeds within 5% might be reasonable. Alternatively, $\Delta \chi^2$ could be determined as a function of confidence level and the number of degrees of freedom, with a 68.3% confidence level and 3 degrees of freedom corresponding to $\Delta \chi^2$ of 3.53 (Press et al. 1992). Nevertheless, for this study, the statistical uncertainty in the measured SD is used to define $\Delta \chi^2$ that determines the tolerance of the fit parameters.

Using the results of the IGF technique, Fig. 4 shows the three-dimensional ellipsoid that characterizes the range of $N_{0\gamma-\mu-\lambda}$ that provide $\chi^2 < \chi^2_{\text{min}} + \Delta \chi^2$ for the SD depicted in Fig. 2. In $N_{0\gamma-\mu-\lambda}$ phase space, the eigenvalues characterizing the ellipsoid are $2.6 \times 10^{-2}$, 7.7, and $1.1 \times 10^{6}$, with corresponding eigenvectors of $(0.59, -0.80, -1.3 \times 10^{-4})$, $(-0.80, -0.59, 8.1 \times 10^{-4})$, and $(7.3 \times 10^{-4}, 3.8 \times 10^{-4}, 1.3 \times 10^{-7})$. Only values with $-1 < \mu < 5$ and with $\lambda > 0$ are used in the creation of the ellipsoid; values with $\mu$ up to 100 or 200 can be realized, but such values would cause problems with numerical stability in any sort of implementation of a parameterization scheme. Further, although reasonable agreement between the moments of the observed and fit SDs are obtained with such large $\mu$, the fit SDs do not visually resemble those observed. Values of $\mu < -1$ are also problematic for integration of a complete gamma function and thus are not included when finding the points that meet the $\chi^2$ criterion. However, the ellipsoid enclosing these points inevitably extends somewhat to $\mu < -1$ and $\mu > 5$, and those values should not be used in the application of an associated parameterization scheme.

To allow for easier identification of the range of acceptable $N_0$, $\mu$, and $\lambda$, Fig. 5 shows projections of the three-dimensional ellipsoid in $N_{0\gamma-\mu}$, $N_{0\gamma-\lambda}$, and $\mu-\lambda$ phase space. It is apparent from both Figs. 4 and 5 that
a wide range of \( N_0, \mu, \) and \( \lambda \) can characterize the SD, with possible values of \( N_0 \) covering more than 10 orders of magnitude even with the restriction of \(-1 < \mu < 5\) and \( \lambda > 0 \). However, the spread of \( N_0, \mu, \) and \( \lambda \) associated with the different fitting techniques in Fig. 2 is included in the surfaces shown in Figs. 4 and 5. For example, the \( N_0 \) of \( 2.66 \times 10^{11} \text{ m}^{-4} \) and \( \mu \) of 1.84 derived using the Heymsfield et al. (2002a) fitting approach are within the surface depicted in Fig. 5, which is the range of possible solutions derived from the IGF technique. This suggests that the different fitting techniques all provide reasonable fits to the SD, as there is a wide range of fit parameters that almost equally characterize the minimum \( \chi^2 \) allowed in the fit parameters. It is also important to note that \( N_0, \mu, \) and \( \lambda \) are mutually dependent parameters in these volumes: the codependence of parameters must thus be taken into account when characterizing how the fit parameters vary with cloud and environmental conditions.

To translate the uncertainty in fit parameters characterized by the ellipsoid in Fig. 4 into uncertainties in the original SDs, Fig. 6 shows the NAMMA SD plotted in Fig. 2 together with several alternate representations of gamma distributions characterizing it, determined by uniformly choosing values of \( N_0, \lambda, \) and \( \mu \) from the ellipsoid in Fig. 4. At first glance, the variation of \( N(D) \) seems strikingly large, with \( N(D) \) varying by orders of magnitude for \( D < 150 \mu \text{m} \). However, the variability is less for larger \( D \), which are more important for deriving fit parameters determined by forcing the first, second, and sixth moments of the gamma distribution to match the observed moments. In addition, the solutions from

---

**Fig. 5.** Projection of volume of equally realizable solutions of \( N_0, \mu, \) and \( \lambda \) for fit to 10-s-averaged SD measured during NAMMA at 1335:07 UTC 4 Sep 2006 in (a) a \( \mu-N_0 \) plane, (b) a \( \lambda-\mu \) plane, and (c) a \( \lambda-N_0 \) plane.

**Fig. 6.** As in Fig. 2, except the fits to the NAMMA SD shown in green are obtained by randomly selecting values of \( N_0, \mu, \) and \( \lambda \) from the ellipsoid characterizing the volume of equally realizable solutions.
the different fit techniques depicted in Fig. 2 are all within the range of these alternate representations of the gamma function. Analysis of the simulated moments from the gamma functions plotted in Fig. 6, discussed below, also show reasonable agreement with the moments derived from the observed SD.

This analysis was repeated for an SD measured in a very different meteorological situation, synoptically generated cirrus encountered during a flight of the NRC Convair 580 between Fairbanks and Barrow. Figure 7 shows the SD derived from the 2DS, 2DC, and 2DP at 1808:05 UTC 4 April 2008 together with the best fit derived from all five techniques. As for the distribution in Fig. 2, all fits appear to match the SD closely, but there is significant variation in the fit parameters. For example, the IGF gives $N_0$ of $4.93 \times 10^{12}$ m$^{-4}$, $\mu$ of 2.01, and $\lambda$ of $2.25 \times 10^4$ m$^{-1}$, whereas the Heymsfield et al. (2002a) equations give $N_0$ of $1.11 \times 10^{15}$ m$^{-4}$, $\mu$ of 3.07, and $\lambda$ of $2.84 \times 10^4$ m$^{-1}$, and the standard gamma fit gives $N_0$ of $2.60 \times 10^{13}$ m$^{-4}$, $\mu$ of 2.35, and $\lambda$ of $2.34 \times 10^4$ m$^{-1}$.

A three-dimensional ellipsoid characterizing the $N_0$-$\mu$-$\lambda$ providing $\chi^2 < \chi^2_{\min} + \Delta \chi^2$ for this SD was again generated as for the NAMMA SD. In $N_0$-$\mu$-$\lambda$ phase space, the eigenvalues characterizing the ellipsoid are $1.9 \times 10^{-2}$, 18, and $5.2 \times 10^8$, with corresponding eigenvectors of $(0.51, -0.86, -5.4 \times 10^{-5})$, $(0.86, 0.51, -4.3 \times 10^{-4})$, and $(4.0 \times 10^{-4}, 1.8 \times 10^{-4}, 1-9.5 \times 10^{-8})$. Figure 8 shows projections in $N_0$-$\mu$, $\lambda$-$N_0$, and $\mu$-$N_0$ phase space of parameters that give $\chi^2 < \chi^2_{\min} + \Delta \chi^2$ for the IGF fit, based on a restriction of the fit parameters to $\mu > -1$, $\mu < 5$, and $\lambda > 0$. But the ellipsoid enclosing these points does extend to $\mu < -1$ and $\mu > 5$, as before; the values with $\mu < -1$ are not displayed in the projections. As in
Fig. 5, there is a large range of $N_0$ and $\mu$ that can characterize this single SD, with $N_0$ varying by 12 orders of magnitude, $\mu$ varying between $-1$ and 5, and $\Lambda$ varying from less than $5 \times 10^3$ m$^{-1}$ to greater than $4 \times 10^4$ m$^{-1}$. Again, the values of the fit parameters from the different fitting techniques are all contained in the volume of equally realizable solutions. However, given the mutual dependence of fit parameters, it is important that their covariance be taken into account in the implementation of any parameterization or remote sensing retrieval scheme.

It is also important to know the degree to which this variation in fit parameters affects quantities derived from gamma functions used in parameterization or retrieval schemes. One important parameter used in cloud-resolving models is the mass-weighted fall speed $V_q$, which is defined following McFarquhar and Black (2004) and others as

$$V_q = \frac{\int_{D_{\min}}^{D_{\max}} V(D) m(D) N(D) dD}{\int_{D_{\min}}^{D_{\max}} m(D) N(D) dD},$$  \hspace{1cm} (23)

where $m(D)$ characterizes how the ice crystal mass varies with $D$, and $V(D)$ characterizes how the fall velocity of an ice crystal varies with $D$. Typically, it is assumed that $m(D) = \alpha D^\beta$ and $V(D) = \alpha D^\beta$, where $\alpha$, $\beta$, $a$, and $b$ are estimated depending on dominant crystal habit from results of previous studies (e.g., Locatelli and Hobbs 1974; Mitchell 1996; Mitchell and Heymsfield 2005; Heymsfield and Westbrook 2010). In this study, it is assumed that $\alpha = 0.017$ kg m$^{-1}$s$^{-1}$ and $\beta = 1.9$, because it provided the best match for the mass estimated from the in situ size and that measured by counterflow virtual impactor probe (Mascio 2013) distributions on 4 September 2006 during NAMMA. Further, it is assumed that $a = 11.72$ m$^{1.5}$s$^{-1}$ and $b = 0.41$, corresponding to aggregates of side planes, columns, and bullets (Locatelli and Hobbs 1974).

Figure 9 shows the normalized frequency distribution of $V_q$ determined for the SD depicted in Fig. 2. It is generated by selecting $N_0$, $\mu$, and $\Lambda$ that characterize this SD from a discrete grid of all possible values in the ellipsoid characterizing the volume of equally realizable solutions depicted in Fig. 4. The integration in Eq. (23) used to generate Fig. 9 was performed between $D_{\min}$ and $D_{\max}$ to be consistent with observations, even though typical integrations for bulk microphysical schemes are performed between 0 and $\infty$. For consistency, such schemes should ideally use incomplete gamma functions to integrate only over size ranges upon which SDs are obtained. However, increased computational time makes such an approach infeasible, and the severity of the truncation error varies with the component SDs.

Even though there is a large variation in the fit parameters shown in Fig. 4, there is a comparatively smaller variation in $V_q$, with $V_q$ ranging from 0.4 to 0.54 m s$^{-1}$, corresponding to an approximately 20% variation about the mean $V_q$ of approximately 0.48 m s$^{-1}$. There is a smaller variation in $V_q$ than in individual fit parameters, because the computation of $V_q$ takes into account the mutual dependence of the gamma fit parameters.

In addition to $V_q$, several other important parameters in bulk cloud schemes must be known for determining microphysical process rates. All these process rates are proportional to different moments of the SDs. Milbrandt and McTaggart-Cowan (2010) show that moments $M_{2+b}$, $M_{1+h}$, and $M_{h}$ are important for terms involving collection, $M_{1}$ and $M_{1.5+0.5b}$ are important for diffusion and melting, and $M_{b+h}$ and $M_{b}$ are important for sedimentation. Figure 10 shows the variation of these moments and of $M_a$, assuming $b = 0.41$ and $\beta = 1.9$. As with the variation of $V_q$, there is smaller variation of these model-relevant moments than in the variation of fit parameters shown in Fig. 4. Further, the range of the calculated moments matches well with the indicated mean and range of the moments estimated from the observed SDs and uncertainties given by Eqs. (6) and (7). Thus, when representing gamma fit parameters in terms of environmental or cloud conditions for models or parameterization schemes, it is crucial to take into account the mutual dependence of the parameters.
FIG. 10. Normalized frequency distribution of (a) \(M_{2.41}\), (b) \(M_{1.41}\), (c) \(M_{0.41}\), (d) \(M_{1}\), (e) \(M_{1.71}\), (f) \(M_{2.31}\), (g) \(M_{1.9}\), and (h) \(M_6\). All frequency distributions obtained by selecting \(N_x, \lambda\), and \(\mu\) from the volume of equally realizable solutions depicted in Fig. 4 for the SD measured during NAMMA at 1335:07 UTC 4 Sep 2006. Green line represents best estimate of moment from observed SD, and red bars represent minimum and maximum estimates of moment from observed SD.
5. Characterization of families of SDs

When developing parameterization schemes, multiple SDs are used to characterize how \( N_0, \mu, \) and \( \lambda \) depend on cloud and environmental conditions. Thus, families of SDs obtained in similar conditions rather than single SDs are used to characterize the interrelationships of gamma parameters and their dependence on external environmental and cloud conditions. Traditionally, curves defining the relationship between \( N_0, \mu, \) and \( \lambda \) have been developed using databases where each SD is represented by a specific \( N_0, \mu, \) and \( \lambda \) value, rather than by a surface of equally realizable solutions, as in section 4. In this section, a methodology for constructing volumes in \( N_0-\mu-\lambda \) phase space to characterize families of SDs sampled in similar cloud and environmental conditions is developed using a composite of the surfaces derived to describe each of the SDs.

For any family of SDs, ellipsoids describing the volume of equally realizable solutions for the fit to each individual SD were found by the method described in section 4. To determine the single ellipsoid characterizing the family, the optimal region of overlap of these individual ellipsoids was determined. Denoting the set of the individual ellipsoids by \( \mathbf{S} \) and defining an equally spaced map of grid points in \( N_0-\mu-\lambda \) phase space, it is noted that an arbitrary point \( \mathbf{p} \) is inside an individual ellipsoid centered at \( \mathbf{x}_0 \) when

\[
(p - x_0)^T \mathbf{A}(p - x_0) \leq 1,
\]

where the eigenvalues and eigenvectors of \( \mathbf{A}^{-1} \) describe the principal axes of the ellipsoid. By assuming that points inside some fraction of the ellipsoids in \( \mathbf{S} \) were within the volume characterizing the family of SDs, the minimum-volume ellipsoid enclosing these points was then found using the algorithm from Moshtagh (2006).

To illustrate how such a volume is calculated, data obtained on all NAMMA flights are used to construct a three-dimensional plot of the ellipsoid that characterizes the volume of equally realizable solutions in \( N_0-\mu-\lambda \) phase space in Fig. 11. The ellipsoid was drawn to represent all points contained within 1% of the individual ellipsoids. In \( N_0-\mu-\lambda \) phase space, the eigenvalues characterizing the ellipsoid are 0.94, 32, and \( 8.2 \times 10^8 \), with corresponding eigenvectors of \((0.52, -0.86, -6.6 \times 10^{-5}), (0.86, 0.52, -2.9 \times 10^{-7})\), and \( (2.8 \times 10^{-4}, 9.5 \times 10^{-5}, 1.45 \times 10^{-5}) \). Two-dimensional projections of this ellipsoid in \( N_0-\mu, N_0-\lambda, \) and \( \mu-\lambda \) phase space are shown in Fig. 12 to better illustrate the fraction of phase space included in the volume. In \( N_0-\mu-\lambda \) phase space, the eigenvalues characterizing the ellipsoid are 1.2, 54, and \( 1.3 \times 10^9 \), with corresponding eigenvectors of \((0.53, -0.85, -6.1 \times 10^{-5}), (0.85, 0.53, -2.1 \times 10^{-4})\), and \( (2.1 \times 10^{-4}, 5.7 \times 10^{-5}, 1.23 \times 10^{-8}) \). One main difference between the ellipsoid for the ISDAC SDs depicted in Fig. 13 compared to that for the NAMMA SDs depicted in Fig. 11 is that it covers a larger volume of phase space, as clearly seen from the projections in Figs. 14 and 12. Further, the values of \( N_0, \mu, \) and \( \lambda \) for ISDAC cover a much larger fraction of the SDs in the family, with some values corresponding to 20% of the SDs, whereas few of the \( N_0, \mu, \) and \( \lambda \) values for NAMMA occur in more than 4% of the SDs.
The sizes of the ellipsoids describing the family of NAMMA and ISDAC SD parameters in Figs. 11 and 13 depend significantly on the choice of percentage threshold over which the individual ellipsoids must overlap. There is uncertainty in the best choice for this threshold. For cases with large natural variability of SDs within a family, there would be minimal overlap between individual ellipses, and a low threshold would be needed to construct a family ellipsoid. On the other hand, if many individual ellipsoids were similar, then a larger threshold could be used. The fact that a 1% threshold was necessary to get adequate overlap for NAMMA suggests that there is large natural variability in the SDs within the family; the greater amount of overlap for the ISDAC data suggests there may have been less variability in the shape of these SDs. The greater overlap for ISDAC also suggests the uncertainty volumes of individual SDs are typically larger for ISDAC than for NAMMA, which is consistent with fewer ice crystals being measured in the optically thinner arctic cirrus in ISDAC than in the NAMMA hurricane clouds.

In any event, the large range of variability of parameters within merged ellipsoids suggests that the variability in fit parameters observed in past studies may be at least partially caused not by variability in how each SD is characterized, but rather by variability in the actual SDs. Thus, the variability in individual SDs must be taken into account when explaining how gamma parameters and their interrelationships vary. Ultimately, in order to place gamma fit parameters in a form conducive to use in model or remote sensing studies, relationships for these ellipsoids describing $N_0$, $\mu$, and $\lambda$ need to be placed in terms of parameters, such as the total water content or temperature. Although such relationships have also been developed in many prior studies (e.g., Heymsfield et al. 2002a; Field et al. 2005; McFarquhar et al. 2007a), it is critical that the mutual dependence of parameters be taken into account and the

![Fig. 12. Projection of ellipsoid in Fig. 9 on (a) $\mu$–$N_0$, (b) $\lambda$–$\mu$, and (c) $\lambda$–$N_0$ phase space.](image1)

![Fig. 13. As in Fig. 11, but for ellipsoid enclosing all SDs measured during ISDAC.](image2)
representation of the $N_0$, $\mu$, and $\lambda$ parameters as volumes of equally realizable solutions is an ideal method of doing this. However, because it is not possible to characterize each SD by a single $N_0-\mu-\lambda$ triplet, the parameterizations should in reality be represented by volumes of equally realizable solutions that may vary with environmental conditions. Mascio (2013) developed relationships that describe how the ellipsoids for the NAMMA SDs vary as functions of temperature, total water content, vertical velocity, and tropical cyclone stage of development. Similar studies need to be performed for the ISDAC data, and for data collected in many other field programs.

The development of volumes of equally realizable parameters has immediate applications for parameterization schemes (e.g., van Lier-Walqui et al. 2014). McFarquhar et al. (2003) used a solution of equally realizable solutions describing the mass contained in small crystal sizes as a function of total ice water content to develop a parameterization of ice crystal effective radius in terms of total ice water content: application of this parameterization by randomly selecting values from the space of equally realizable solutions showed that the average cloud radiative forcing determined from a series of single-column model simulations was not identical to the cloud radiative forcing determined from a single simulation using the most likely parameters. This shows the importance of deriving volumes of equally realizable solutions. Thus, the eigenvectors and eigenvalues of the ellipsoid of equally realizable solutions in $N_0-\lambda-\mu$ phase space are presented here to aid in ultimate application of these parameterizations as probability distributions in future Monte Carlo–type parameterization schemes.

6. Summary

Gamma distributions frequently used to represent size distributions (SDs) in mesoscale and cloud-resolving models and in remote sensing retrieval schemes are characterized by intercept ($N_0$), slope ($\lambda$), and shape ($\mu$) parameters determined by fits to SDs measured in situ within clouds. In this study $N_0$, $\mu$, and $\lambda$ were determined for SDs measured in tropical storms off the western coast of Africa during the NASA African Monsoon Multidisciplinary Analysis (NAMMA) campaign and in cirrus over Alaska during the 2008 Indirect and Semi-Direct Aerosol Campaign (ISDAC) using five different fitting techniques. Although the fit number distribution functions $N(D)$ obtained by the different techniques appear similar for the majority of the observed SDs, there can be large variations, up to orders of magnitude in $N_0$, in the parameters that describe the fit distributions. Assuming that fit parameters within a confidence region $\Delta \chi^2$ about the minimum chi-squared $\chi^2_{\text{min}}$ determined from the statistical uncertainty in the measured size distributions are equally realizable, it was seen that the variation in solutions between

![Fig. 14. Projection of ellipsoid in Fig. 13 on (a) $\mu-N_0$, (b) $\lambda-\mu$, and (c) $\lambda-N_0$ phase space.](image-url)
techniques was less than the variation in $N_\theta$, $\mu$, and $\lambda$ in the confidence region of the fit. Even though there was large variation in the $N_\theta$, $\mu$, and $\lambda$ values characterizing a single SD, the variability in properties derived using these values, such as the mass-weighted terminal velocity, derived using fit parameters in the volume of equally realizable $N_\theta$, $\mu$, and $\lambda$ varied by only 20% when the codependence of the fit parameters was accounted for.

After the volume of equally realizable solutions was generated for each SD within a family, the concentration of such volumes in the $N_\theta-\mu-\lambda$ phase space was determined. A single ellipsoid characterizing the family is found using a numerical routine by determining the minimum-volume ellipsoid that encloses the region where a sufficient number of the individual ellipsoids overlap. Parameters within this volume of equally realizable solutions for a family were shown to vary substantially, with $N_\theta$, in particular, spanning several orders of magnitude. The scatter of $N_\theta$, $\mu$, and $\lambda$ for single SDs is comparable to the scatter of most likely $N_\theta$, $\mu$, and $\lambda$ from multiple SDs, suggesting that previously observed scatter in fit parameters is at least partially caused by uncertainties in fit parameters rather than variations in SDs themselves.

When quantifying the dependence of $N_\theta$, $\mu$, and $\lambda$ on cloud or environmental conditions, it is necessary to determine how the three-dimensional volume in $N_\theta-\mu-\lambda$ phase space depends on cloud or environmental parameters. Future studies are examining this dependence by determining how the fit parameters characterizing SDs obtained in hurricanes during NAMMA vary with cloud and environmental conditions (J. Mascio et al. 2015, unpublished manuscript).

The determination of such volumes has important implications for the development of cloud parameterization schemes. Currently, schemes assume a single relationship for fit parameters as functions of variables like total water content or temperature, which is assumed to hold under all conditions. However, no single relation holds universally, as several different relationships have been proposed to characterize fit parameters. Because the relationship affects microphysical process rates, such as latent heating and cooling, that in turn influence system dynamics, future modeling efforts should use the range of fit parameters described by the ellipsoids: the characterization of the ellipsoids by eigenvectors and eigenvalues allows such relationships to be easily implemented in a stochastic framework. Ensemble simulations can thus determine the range of possible future states taking into account the uncertainty in the microphysical parameterization. Alternatively, the range of solutions can be determined by simulating extremes and medians of particular processes calculated from the range of fit parameters included in the ellipsoids (e.g., parameterizations of maximum, mean, and minimum mass-weighted velocities). Such approaches are plausible, as ensemble convection parameterization schemes (e.g., Bao et al. 2007; Liang et al. 2007; Bouttier et al. 2012) have been used before.

Therefore, the statistical approach introduced here can serve as a basis for an ensemble predictive microphysics scheme. It takes into account uncertainties in the SDs of ice particles but also allows both the mean state and associated uncertainties to be dependent on other variables, such as vertical velocity or temperature. This will be important for future modeling efforts to determine how such uncertainties cascade up which affects the spread in the final state of the system.

Acknowledgments. Data were obtained from the Atmospheric Radiation Measurement Program archive, sponsored by the Department of Energy (DOE), Office of Science, Office of Biological and Environmental Research (BER) Environmental Science Division. This research was supported by the Office of Science (BER), United States DOE under Grants DE-SC0001279 and DE-SC0008500, by the National Science Foundation (NSF) under Grant AGS-1213311, and by the National Aeronautics and Space Administration (NASA) Hurricane Science Program Grant NNX09AB82G and Precipitation Measurement Missions Science Team Grant NNX13AF86G. We are grateful for the efforts of the ISDAC and NAMMA science teams for the collection of high-quality datasets. We are grateful for the comments of J. Milbrandt and two anonymous reviewers, which considerably improved the quality of the manuscript.

REFERENCES


Mascio, J. R., 2013: The use of gamma distributions to quantify the dependence of cloud particle size distributions in hurricanes on cloud and environmental conditions. M.S. thesis, Dept. of Atmospheric Sciences, University of Illinois at Urbana-Champaign, 96 pp. [Available online at http://hdl.handle.net/2142/46624.]

