

Bayesian upscaling of aircraft ice measurements to two-dimensional domains for large-scale applications

A. R. Jameson · A. J. Heymsfield

Received: 18 August 2013 / Accepted: 20 November 2013 / Published online: 1 December 2013
© The Author(s) 2013. This article is published with open access at Springerlink.com

Abstract What is new in this manuscript is a method of using aircraft observations from a long horizontal path through an ice cloud to produce properly correlated 2D fields of particle counts consistent with the observations, including all null values, at several different sizes for use in algorithm development and in scientific studies. A Bayesian approach is used to define the distributions of average counts, $P(C)$, at every size. These are, in turn, used to expand the number of values at each particle size by a factor of 50. These data, then fill a square 2D grid of 20×20 km at 100-m resolution. At each grid point, the number concentrations corresponding to each particle size define the particle size distributions (PSD). A method for assuring the proper correlation of counts at each size over the entire grid is devised and discussed. These PSD can then be integrated to yield a number of different quantities over the entire grid such as radar reflectivities and ice water contents. From this perspective, one can then consider the set of observations as just one realization from a much larger ensemble of possible realizations by giving fuller expression to all of the information contained within the observed correlation functions and $P(C)$ s. The in situ observations, however, remain crucial since this method

does not ‘make-up’ new meteorology but simply gives wider expression to the meteorology contained within the observations.

1 Introduction

In previous work (Jameson and Heymsfield 2013), a method is presented for upscaling an ensemble of ice particle measurements along a 79-km flight path in a cirrus cloud produced by the upper level outflow from a tropical thunderstorm. That work focused mainly on the development of a technique to use Bayesian analyses of count statistics and correlation functions corresponding to different size bins in order to generate upscaled and downscaled estimates along a line. At the end of that work a method was suggested for generating 2D upscaled spatial realizations of the particle size distributions (PSD) at each grid point. However, in that work this idea could neither be fully developed nor explored. The purpose, then, of this note is simply to complete that thought by providing an explicit approach for achieving the objective of producing a PSD at every grid point within a 2D region in a manner consistent with the observed statistics (counts and correlations). This approach not only provides detailed size distributions at each grid point, useful for probing remote sensing physics, but these distributions can also be integrated to yield ‘point’ estimates of upscaled quantities such as the total particle concentrations (C_{tot}), radar reflectivity factor at 94 GHz (Z_{94}), the ice water content (IWC), the mean particle size, \bar{D} , and the variances, σ_D^2 of particle sizes.

A key ingredient in this entire process is the use of Bayesian statistical analyses of the observed particle counts in each size bin, that is the number of particles of specified

Responsible editor: C. Simmer.

A. R. Jameson
RJH Scientific, Inc., El Cajon, CA, USA

A. R. Jameson (✉)
5625N. 32nd St., Arlington, VA 22207-1560, USA
e-mail: arjatrjhsci@verizon.net

A. J. Heymsfield
National Center for Atmospheric Research Boulder, Boulder,
CO, USA

sizes counted per unit length of flight path. What this analysis does is to give an estimate of the probability distribution, P , of the average number of particle counts, i.e., $P(C)$. These are the numbers one needs to specify the particle size distribution, PSD.

However, the correct temporal/spatial description of the ice particles requires more than just $P(C)$. There is also the correlation among counts. Calculations for the data described below show that on 10-m spatial scales, the occurrences of different-sized particles are uncorrelated, and even those of the same size are only very weakly correlated. Hence, for these data at 10 m resolution, the counts can be considered to be statistically independent. At coarser size bin resolutions, however, cross-correlation among sizes and auto-correlation at the same sizes begin to appear as a consequence of larger scale structure. These latter auto-correlations appear exponential with variable size dependent $1/e$ correlation lengths (where e is Euler's number). This correlation can be interpreted as a description of the 'structure' which in turn is a reflection of whatever meteorological processes produced it. A knowledge of the meteorology, however, is not required to use the correlation information and for this technique to work. In fact, the technique may well shed light on the meteorology.

In the method described below, the cross-correlation among different sizes is not explicitly incorporated. It turns out that this is not necessary (as illustrated later), because it arises naturally in the methodology (as 2D cross-correlation calculations and, indeed, the existence of patterns reveal). It is accomplished by using the same input field of random numbers (the base field) for all the different size bins each with their own unique correlation length. This will become more apparent below.

Aside from correlation, $P(C)$ only provides a measure of means when the counts exist. Yet, in general there may be extended periods during an interval when particles of a particular size are not observed. This is especially true for the larger particles. Hence, one must also know the fractional time for which the counts are zero, and this varies for each size bin. The method for combining all of these considerations is discussed below.

2 Methodology

Perhaps it is easiest first to present briefly the set of measurements discussed in greater detail in Jameson and Heymsfield (2013). The ice particle measurements were obtained from a Cloud Imaging Probe, (CIP, sizing from about 50–1,000 μm) and a Precipitation Imaging Probe, (PIP, sizing from 200 to >5,000 μm), manufactured by Droplet Measurement Technologies. They have fast

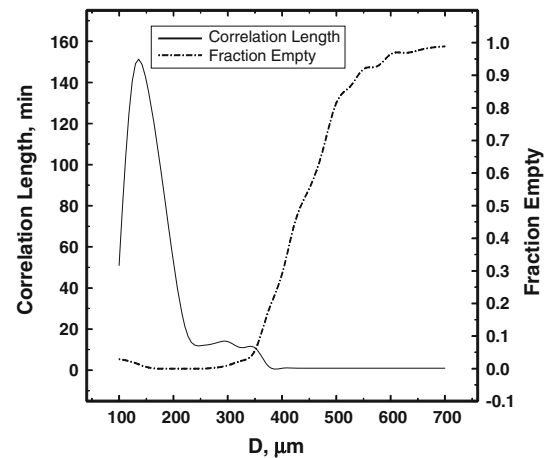


Fig. 1 The $1/e$ correlation length and fraction of null observations per 100 m as measured by the aircraft as a function of particle size

electronics compared to the earlier 2D probes. Every 2D particle imaged by the probes along the transit is sized and, because particle inter-arrival times are recorded, their spatial distribution in the horizontal is known. The sample volumes are derived from the so-called “reconstruction method” (Heymsfield and Parrish 1978).

Shattered particles are removed objectively using particle inter-arrival times (Field et al. 2006). Particles above several hundred microns are all within the so-called “depth of field”, established by the spacing of the probes sampling arms. The fraction of accepted particles is as follows: 100–200 μm : 0.760368; 200–300 μm : 0.900033; 300–500 μm : 0.904354; 500–1,000 μm : 0.823263. The inter-arrival time threshold is dynamically changing and no specific set time is used. The general features of the data can be found in Fig. 1 in Jameson and Heymsfield (2013).

In this work, we bin particle sizes from 100 to 700 μm in 25-m bins. This accounts for 99.98 % of all observed particles. Setting the origin at the beginning of the flight path, the spatial location of each particle in each size bin is measured from this origin. We can, then, bin count over variable spatial resolutions. For this study, we use a resolution of 100 m which yields 793 observations per size bin. For each of the 24 particle size bins, the $1/e$ correlations lengths and the fractional empty time are calculated. These are illustrated in Fig. 1 and are incorporated into the method as described next.

The distributions $P(C)$ of the mean values, C , are computed for each size bin and at 100-m resolution using the Bayesian technique described briefly in the “Appendix” and in greater detail in Jameson (2007) and in Jameson and Heymsfield (2013). Next, it is necessary to generate a correlated 2D field of uniformly distributed random numbers having the characteristic correlations appropriate to each size bin. To achieve this, we begin with

a 20 km × 20 km square grid having 40,000 points of zero mean, unit variance uniformly distributed random numbers (the base field). This represents a 50-fold increase in the size of the observed data set.

An exponential correlation function having the correct correlation length as given in Fig. 1 is then imposed on this base field for each ice particle size separately using the so-called root matrix method (e.g., Johnson 1994 as discussed in Jameson and Kostinski 1999, 3,924–3,925). The resulting field of numbers must then be renormalized because of distortions to the pdf introduced by this process. This is accomplished by stacking the columns into a vector. That is, each grid consists of columns and rows. We form one long column (vector) by placing one column vertically after another one until all the columns have been used. We then apply the copula transformation method (see Genest and Mackay 1986; Nelsen 1999; Frees and Valdez 1998 and as briefly described in Appendix B in Jameson and Kostinski 2008 with regard to radar signals). Basically, in the copula technique one forms the accumulated pdf of the random variable that one wishes to correlate. One then takes a string of unit variance, zero mean correlated variables having the desired correlation function (often Gaussian or exponential). The accumulated pdf of the desired variable is then used to invert these correlated variables to become a correlated string of the desired variable so that we finally end up with a uniformly distributed vector of properly correlated random numbers lying between 0 and 1, but different for each drop size. That is, the copula method is used twice, once to restore the newly correlated field to be uniformly distributed and a second time to produce the desired correlated field of counts corresponding to the particular $P(C)$.

Returning to the $P(C)$ derived above, it must be remembered that it only applies to non-zero counts. To account for null counts in a 100-m interval, this $P(C)$ is first adjusted so that the number of null counts agrees with the empty fraction in Fig. 1 and the $P(C)$ remains normalized. Using the uniformly distributed vector having values between zero and unity above, the $P(C)$ is then used in the copula method to transform this vector into a vector of C values satisfying $P(C)$ and having the necessary correlation. Finally, this vector is then unstacked to form the 200 × 200 grid of values. This is done for each size bin to yield 24, 20 km × 20 km fields. This provides ‘true’ (in the sense of using mean values) size distributions having the correct correlation and the correct pdf properties including the empty fractions at each size bin and at each of the 40,000 grid points. These can then also be combined to compute integrable variables such as described at the beginning of this introduction and as discussed in the next section.

It is worth mentioning in passing, however, that while seemingly adding unnecessary complexity by going from a

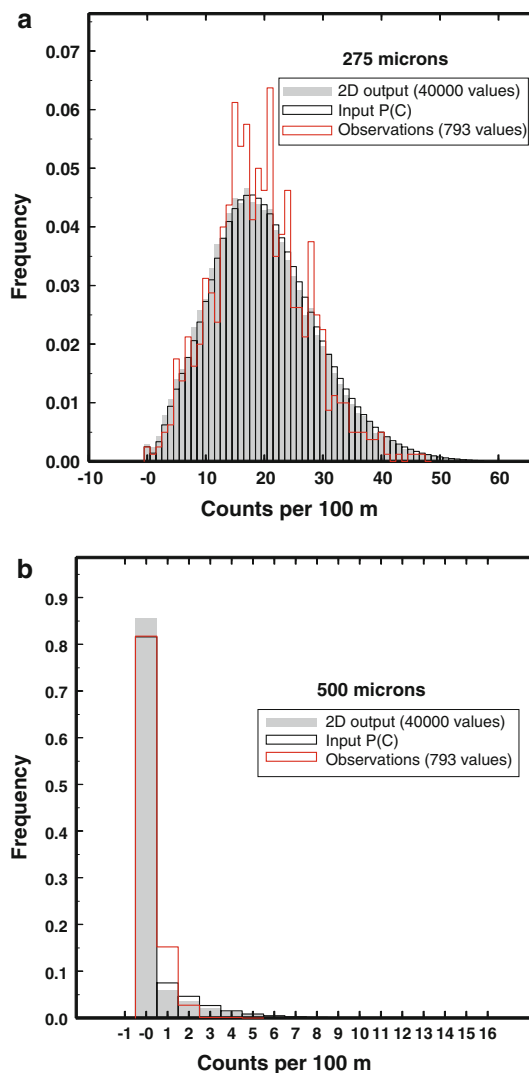


Fig. 2 Histograms of particle counts per 100 m of flight path as observed for 793 observations and as simulated for 40,000 mean values along with plots of the input $P(C)$

correlated 2D field to a vector and then reversing the process after the copula calculations, such a procedure avoids the introduction of periodic artificial decorrelations that occur if one simply takes a 40,000 string of correlated data and tries to convert it directly into a 2D field by chopping it up into 200-long segments and then laying these segments down next to each other. Along the boundaries of the region, the neighboring points will not be properly correlated so that this approach is not correct.

However, it is also important to show that this methodology in fact captures the content in the set of observations. While one could attempt to find a path that reproduces the observations, there are an infinite number of such paths on the 2D plane making such an effort fruitless. Rather it makes more sense to consider the totality of the statistics realizing that in this example the methodology is

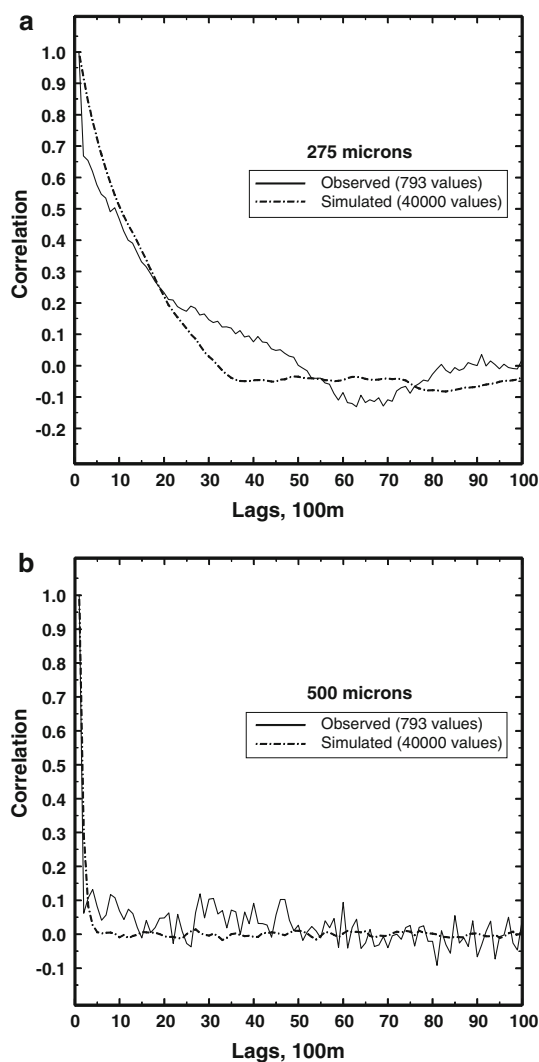


Fig. 3 The observed and simulated auto-correlation functions at the two size bins as in Fig. 2

applied to average particle counts, whereas the observations are but one sample (one realization) from these mean values. With that in mind, Fig. 2 presents two examples, one when there are more numerous particles (averaging 19 per 100 m) and the other when the particles are spatially sparse (averaging 1 per 500 m). In Fig. 2, the histograms of the observations (793 values in red) are plotted along with the output from this method (40,000 values, grey shading) as well as the input distributions of mean values, $P(C)$ (black empty bars). Figure 2a shows that the deviations of the observations is at most 1.5 % from the simulated mean values. The 2D output also faithfully reproduces the input $P(C)$. The results for Fig. 2b are similar, but the departures are somewhat larger due to the sparsity of the particles. Nevertheless, with regard to particle counts, the methodology appears to work.

Figure 3 is a look at the correlation functions of particle counts every 100 m which, in this work, we model using an exponential (consistent with pink spectra). In principle, one could perform parametric fits to the observed correlation functions and use them, but for our purposes here, the exponential was found adequate. Figure 3a shows the observed and simulated correlation functions for particles in the 275- μm -size bin, while Fig. 3b shows them for the sparser 500- μm -size bin. The methodology appears to produce correlation functions sufficiently consistent with the observations. We conclude, therefore, that the methodology is capable of producing 2D fields of values consistent with the observations.

3 Results

The motivation for doing all of this is to expand the size of the data set in a manner consistent with the observations. From this perspective one can consider the set of observations as just one small (but crucial) sample from an ensemble of possible sets of observations all having the same statistical properties [$P(C)$ and correlations]. This then allows one to explore a range of unobserved possible values by giving full expression to all of the information contained within the correlation functions and $P(C)$ s beyond what was seen in the finite set of measurements. This will be given explicit meaning below. It also allows one to explore the 2D implications of the measurements particularly useful, say, in the development of remote sensing algorithms. Of course, it is never possible to explore all possibilities, but each new base field will yield a different 2D realization so that one can explore for as long as one is willing to do the calculations.

The 24 fields of counts corresponding to the different drop sizes are converted into number concentrations per liter based upon the aircraft speed and the sample volume of the particle probes. Superposition of all the fields and summing gives the field of C_{tot} , the total particle concentration over the 20 km \times 20 km domain as illustrated in Fig. 4a. It is then straightforward to calculate the mean size, \bar{D} , and variances, σ_D^2 , of sizes at each grid point as well. That is,

$$\bar{D}(i, j) = \frac{\sum_{d_k=D_1}^{D_{24}} d_k c_{ij}(d_k)}{C_{\text{tot}}} \quad (1)$$

$$\begin{aligned} \sigma_D^2 &= \overline{D^2(i, j)} - [(\bar{D}(i, j))]^2 \\ &= \frac{\sum_{d_k=D_1}^{D_{24}} d_k^2 c_{ij}(d_k)}{C_{\text{tot}}} - \left[\frac{\sum_{d_k=D_1}^{D_{24}} d_k c_{ij}(d_k)}{C_{\text{tot}}} \right]^2 \end{aligned} \quad (2)$$

where i and j refer to the row and column indices, respectively, and c is the different particle concentration at each size. These two fields are illustrated Fig. 4b, c.

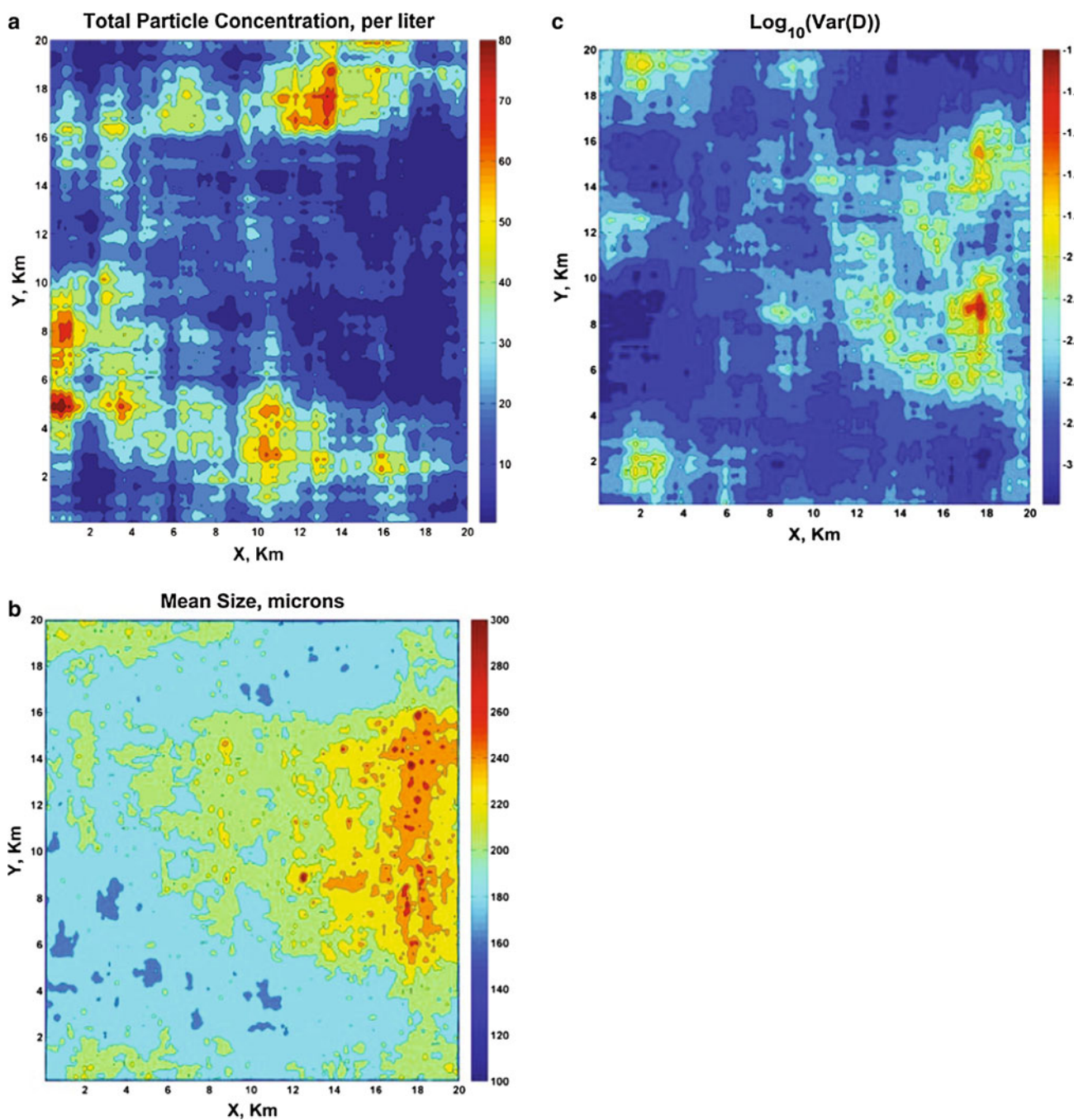


Fig. 4 Contour plots of a the total particle concentration, b mean particle size and c variance of particle sizes over the 20 × 20 km grid

These are the basic observable field from which other quantities can be computed. In particular, using the relation for these data (Heymsfield et al. 2009)

$$\begin{aligned}
 m &= 0.4765D^3 \quad D < 0.0115 \text{ cm} \\
 m &= 6.21 \times 10^{-3}D^{2.05} \quad D \geq 0.0115 \text{ cm}
 \end{aligned}
 \tag{3}$$

where m is the mass in g and D is the particle dimension in cm, the ice water content (IWC) is calculated from

$$\text{IWC}(i,j) = \sum_{d_k=D_1}^{D_{24}} m_k c_{i,j}(d_k)
 \tag{4}$$

and is plotted in Fig. 5a.

There were only 19 particles of size larger than 725 μm out to 1,050 μm along the entire 79 km path. When they did occur in a 100 m distance bin, there were usually fewer than two and on only one occasion were there four.

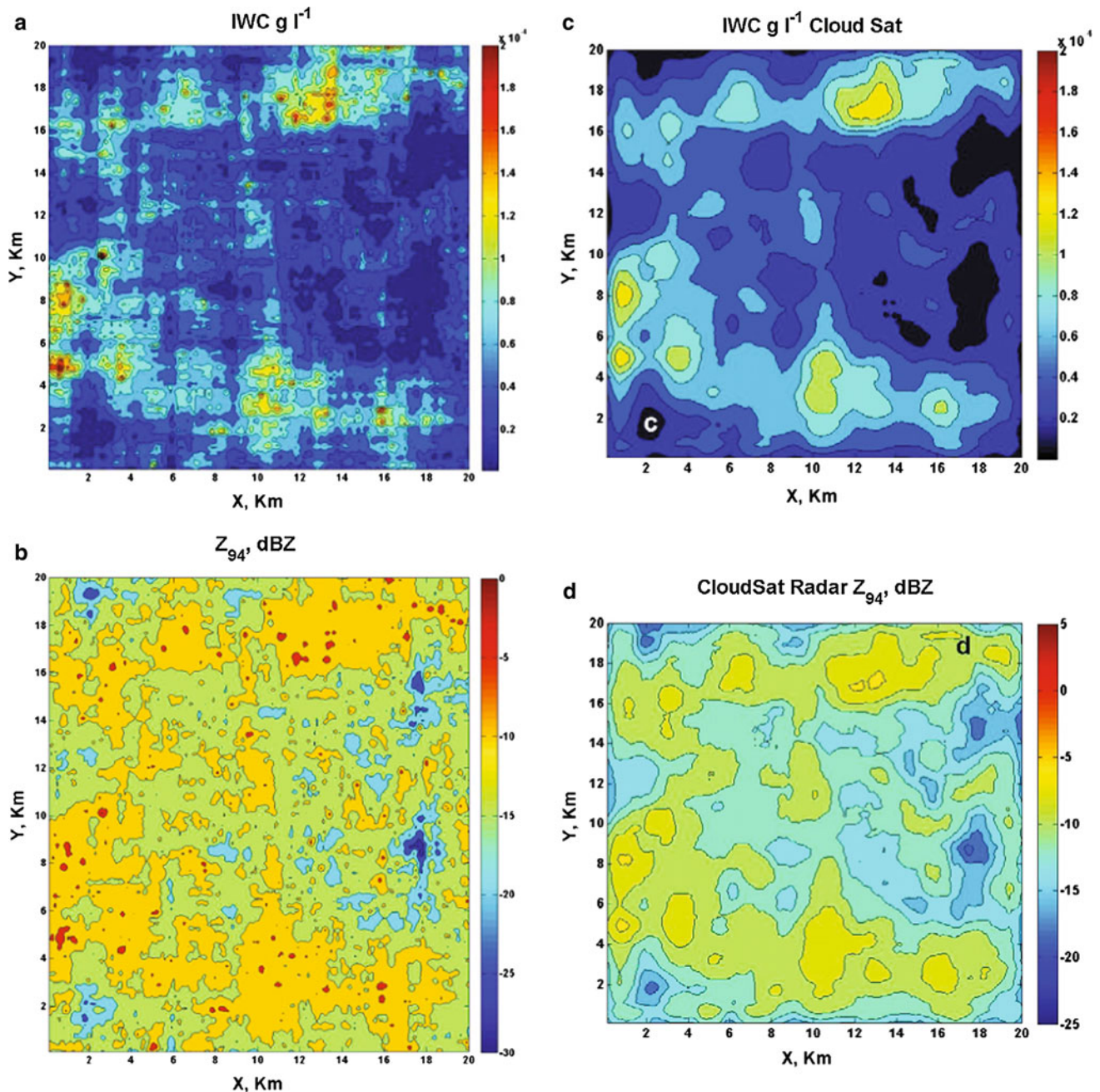


Fig. 5 Contours of **a** the intrinsic ice water content and **b** the intrinsic radar reflectivity factor at 94 GHz expressed as dBZ derived from the PSD over the entire grid compared with the water content, **c** averaged over the CloudSat beam and **d** the observed CloudSat radar

reflectivity factor at 94 GHz averaged over the beam and expressed as dBZ. The IWC and Z_{94} for CloudSat are larger than the intrinsic values because of the effects from averaging over the beam

Consequently, such scarce particles are ignored in this work. The radar reflectivity factor at a frequency of 94 GHz (Z_{94} , used by radars for studying clouds) can be estimated using the methods outlined in Matrosov and Heymsfield (2008) and Heymsfield et al. (2008) as modified by using a ‘soft sphere’ Mie scattering approximation (see Liu 2004). These values are plotted in Fig. 5b. Of course, the field will look somewhat different when viewed

from space, say, by CloudSat. To get a feeling for what happens, Fig. 5c, d illustrate when the nominal 1.3 km resolution is applied to each grid point for the IWC and Z_{94} . While some of the same features are evident between the intrinsic and radar values, the radar obviously smooths the intrinsic field as one would expect. To see how the relations among variables might be affected, the ice water content (IWC) is plotted versus Z_{94} in Fig. 6. Obviously,

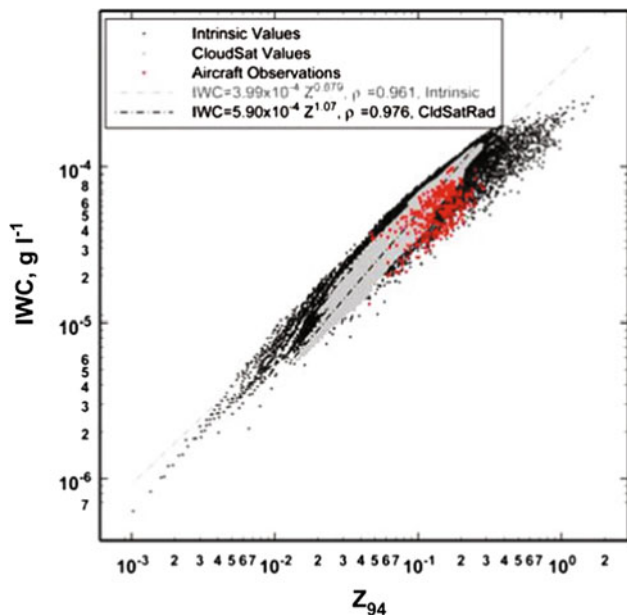


Fig. 6 A comparison between the intrinsic ice water contents (IWC) and theoretical Z_{94} to what the CloudSat radar would ‘see’ from space. The distribution of both Z_{94} and IWC ‘seen’ by the radar are obviously somewhat different from the intrinsic field because of the beam size. The relations between the two variables are not identical. The red symbols represent the data derived directly from aircraft measurements

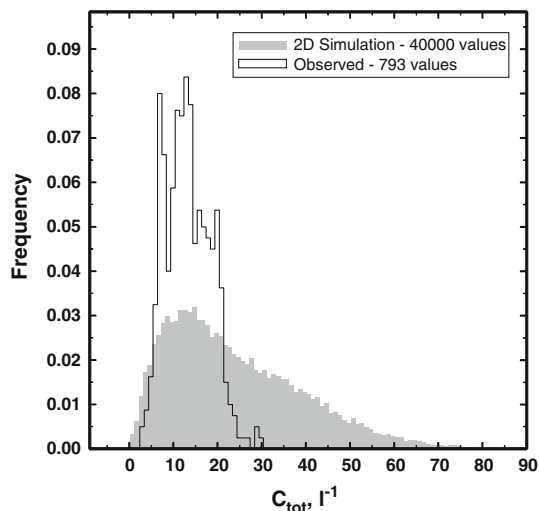


Fig. 7 A comparison between the observed and 2D grid histograms of the total particle concentration

the two relations differ between the intrinsic and radar values suggesting caution when attempting to apply relations strictly based upon the physics to actual radar measurements. The aircraft-derived values fall nicely within the domain of simulated values as well. This demonstrates the significant multiplicative effect of the 2D expansion to

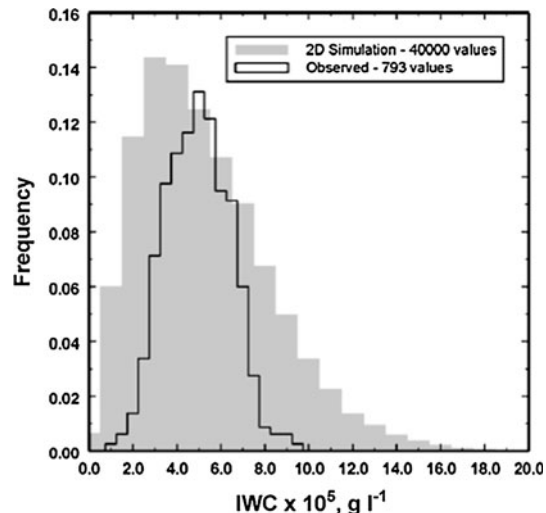


Fig. 8 Similar to Fig. 7 except for the ice water content (IWC)

the extent of even providing a range of values potentially useful for deducing relations among variables.

It is clear from both Figs. 4 and 5 that different flight paths of an airplane through these fields would yield a wide variety of different sets of observations, likely distinct from the actual observations. The reason for this is that these data fields are a fuller expression of all of the information contained within the correlation functions and $P(C)$ s than is provided by the single realization observed. One way of seeing this is by comparing the histograms of C_{tot} (Fig. 7) and the ice water content (IWC) (Fig. 8). While contained within the 2D fields, the observations are but a subset within the much larger distribution evident in the 2D fields. Another way of showing this is also to look at the relation between C_{tot} and IWC as illustrated in Fig. 9 where the observations lie within the set of possibilities expressed in the 2D fields, but they are clearly a much more limited set as one would expect for any single realization (flight path) through these fields. However, it must be remembered that these results are only from using one base field of random numbers. Different base fields will produce different data fields so that the results above are neither unique nor necessarily complete.

Moreover, it was mentioned above that cross-correlation among counts at different sizes was not explicitly introduced into the data generation process. Yet, it does appear naturally as suggested above. Indeed, Fig. 10 illustrates that the procedure for generating the 2D fields leads to ‘natural’ cross-correlation in a manner quite similar to the observations. Two-dimensional cross-correlations also confirm this conclusion. This likely occurs because we are using the same correlated base map which then naturally produces cross-correlation among the different sizes.

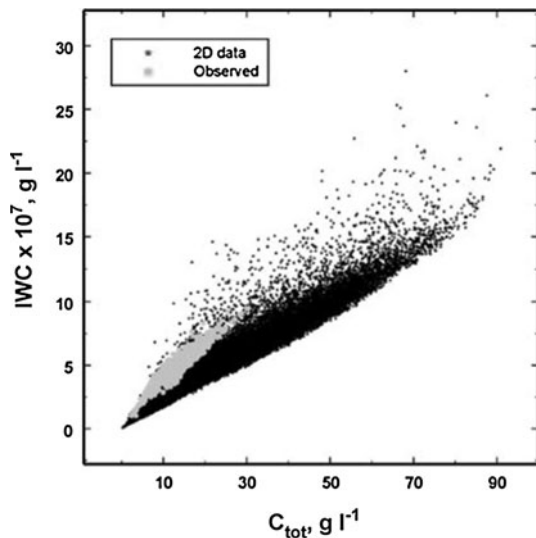


Fig. 9 The IWC vs the total concentration of particles, C_{tot} , as observed and as expressed over the 2D grid demonstrating that the observations are a subset of the total field of possibilities

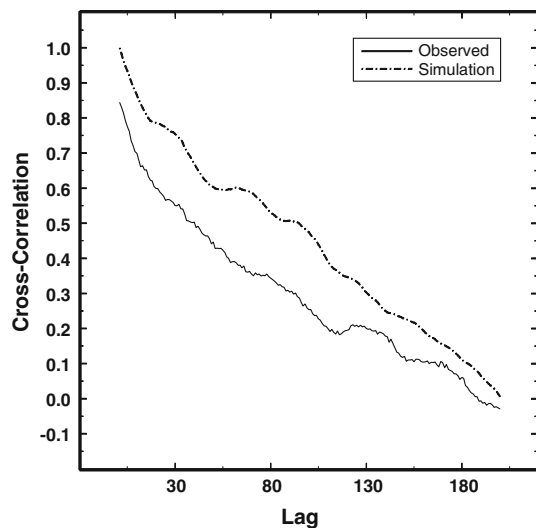


Fig. 10 The cross-correlation functions between the concentrations of 287 and 312 μm -sized particles as observed in the aircraft data and as it appears for the 2D grid at $Y = 50$ km and $1 \leq X \leq 200$ km demonstrating the presence of reasonable cross-correlation in the 2D values

These various outputs can be used to produce maps of other variables as well such as plots of different parameters for different fits. For example, \bar{D} and σ_D^2 can be considered as the parameters of Gaussian fits to the PSD. Alternatively, the same variables can be combined to estimate the parameters of a simple (2-variable) gamma or log-normal distribution as found in several studies of ice cloud particle size distributions (Heymsfield et al. 2002; Tian et al. 2009), namely the shape parameter, $(\mu = \bar{D}^2/\sigma_D^2 - 1)$, and the scale parameter or slope, $\Lambda (= \bar{D}/\sigma_D^2)$. In addition, C_{tot} is

turned into the gamma distribution drop concentration parameter, $N_o = C_{\text{tot}}/[\Gamma(\mu + 1)\bar{D}^\mu]$, where Γ is the gamma function as discussed in Heymsfield et al. (2002). For the fields in this study these variables are shown in Fig. 11. Moreover, these values can be used to inspect relations among the different variables. For example, a plot of N_o vs. Λ (Fig. 12) suggests that this relation differs somewhat from that in Heymsfield et al. (2008) which is not surprising given the different meteorological conditions.

4 Summary

In this work, we present one approach for addressing the upscaling of measurements of ice particle size distributions (PSD) and the associated integrable properties. An extended set of aircraft observations in cloud over 79 km allowed us to evaluate the potential of Bayesian statistics for use in such upscaling. Using the approach described in detail in Jameson (2007) and Jameson and Heymsfield (2013), the Bayesian analysis of particle counts in 24 different size bins from 100–725 μm yielded the probability distribution, $P(C)$, of mean counts (C) over the observation interval. (While 19 particles larger than 725 μm were observed along the entire 79 km path, their sparseness reduced their appearance to infrequent shot noise that could be safely ignored.) These $P(C)$ were adjusted to account for the frequency of null counts for each size bin.

A square 2D grid of $200 \times 200 \times 100$ m ‘points’ was filled with uniformly distributed random numbers having zero mean and unit variance. For each size bin, this field was correlated using the different observed correlation distance for each size and using the root matrix correlation method. As explained in the text, these fields were then adjusted and the copula method was used to generate 2D fields of mean particle concentrations at each size bin over the 40,000 grid points. Combining these fields over all particle sizes produced the particle size distribution of mean concentrations at each grid point having the correct correlations and distributions of mean values of the particles as actually observed. These PSD can then be integrated appropriately to yield a number of different quantities over the entire grid.

This approach offers a method for greatly expanding the size of any observed data set while retaining all of the important statistical characteristics of the observations. In this example, the set of data was expanded by a factor of 50. From this perspective one can then consider the set of observations as just one sample (but a crucial one) from an ensemble of possible sets of observations all having the same statistical properties [$P(C)$ and

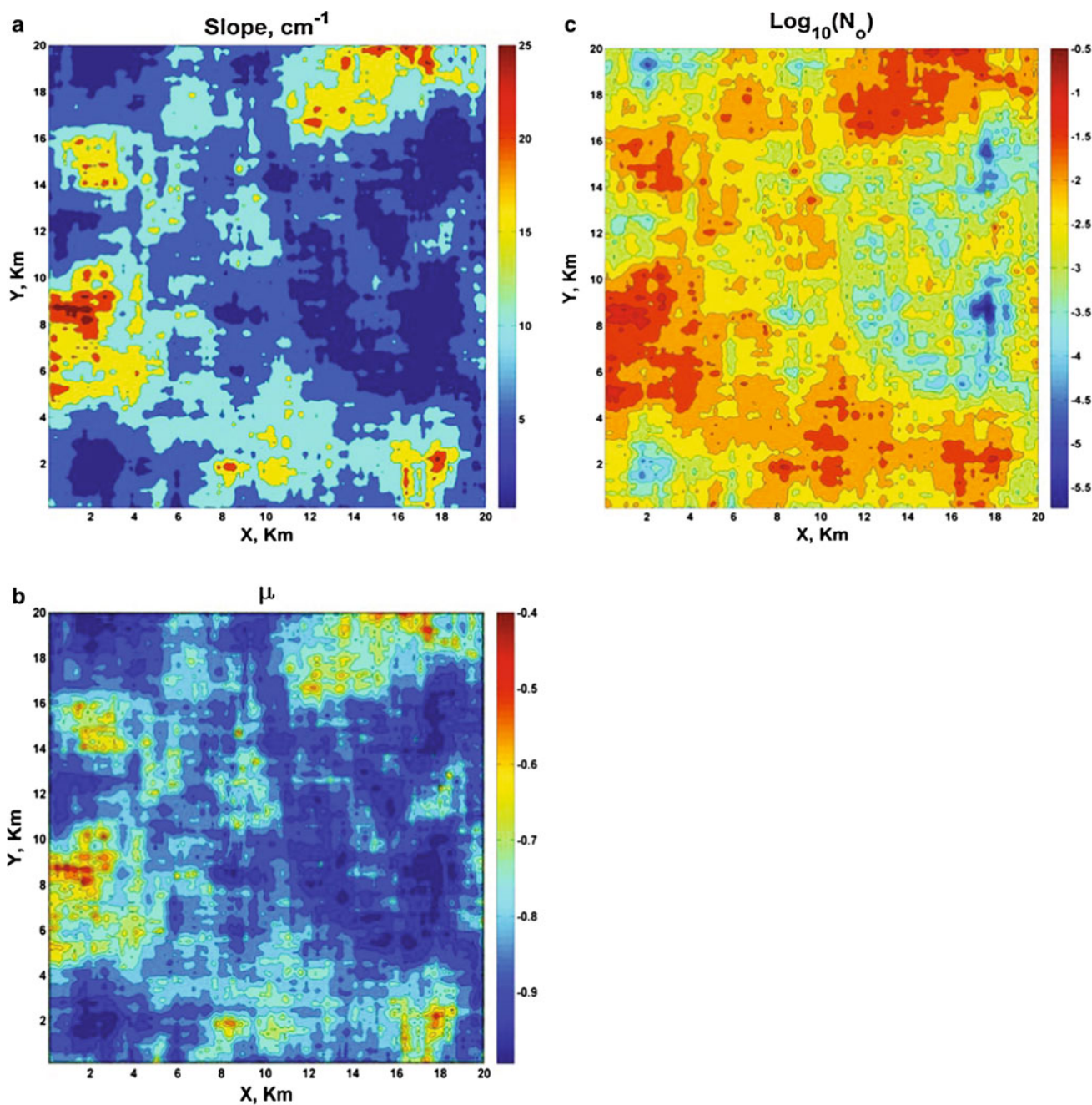


Fig. 11 Contour plots of parameters for a simple gamma (two parameter) fit to the 2D values for **a** the slope or scale factor **b** μ , the shape factor and **c** N_o the particle concentration term

correlations]. This then allows one to explore a range of possible values that may not have been observed in the more finite set of measurements actually observed by giving full expression to all of the information contained within the correlation functions and $P(C)$ s. The observations, however, remain crucial since without them, there is nothing to explore, and it is important to realize that this method does not ‘make-up’ new meteorology, but simply gives a fuller expression to the meteorology that was observed.

Aside from the highlights above, one of the other principal advantages of this approach is that there are no scaling assumptions. Furthermore, after the first realization, it is possible, then, to readily generate an ensemble of realizations using different random number base maps described above.

While the initial setup is somewhat labor intensive (the lead author can provide information), it mostly uses commercial software. Once established, it took the authors about 2 h to complete the first realization for a new set of

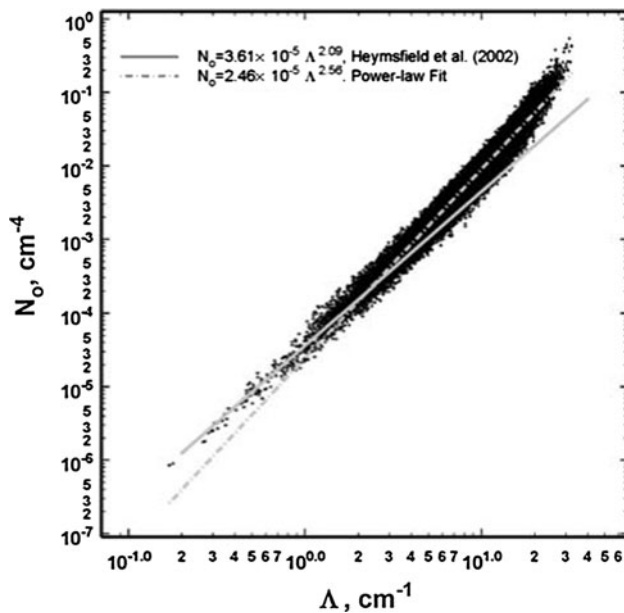


Fig. 12 A comparison of two power-law relations between the slope (Λ) of simple gamma fits to the 2D values and to observations in Heymsfield et al. (2002). Differences are to be expected because of the different meteorological conditions between the two sets of observations

data, while subsequent realizations took less than 0.5 h. Furthermore, if one had multiple aircraft observations at different altitudes, one could, in principle, even generate simulations in three dimensions. Finally, this method can be applied to other counting data such as disdrometer measurements in rain.

Acknowledgments This work was supported by the National Science Foundation (NSF) under grant AGS130087 as well as by the United States Social Security Administration. Support for A.H. came from the National Center for Atmospheric Research, NASA grant NNX10AH67G, and the NSF Science and Technology Center for Multiscale Modeling at CSU CMMAP. The authors are also grateful to Mr. Aaron Bansemer for providing help for computing Z_{94} .

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

Appendix

Here, we briefly describe the application of Bayesian statistics to counting. The interested reader is referred to Jameson (2007, p. 2013–2016) for a more detailed discussion of the derivation of this approach with regard to counting particles. To provide a feel for what is happening, though, we first imagine that we have counted a certain number of particles, n , that represent a random draw from a

Poissonian distribution having a mean value, C , that leads to the so-called likelihood gamma distribution of C (see Jameson 2007, p. 2015). However, we have no idea from which C this n came (i.e., we assume a uniform prior distribution of equal probabilities). The idea, then, is to try to estimate the probability distribution of the various possible C . Given our finite set of observations, this is achieved for a fixed particle size, by superposing all these probabilities of C for all the different observed n and renormalizing the results so that we arrive at our best guess of the so-called posterior distribution of C [i.e., $P(C|D)$ where the vertical bar denotes statistical conditioning, i.e., the probability that C will occur, when n has been observed. Moreover, as discussed in Jameson (2007, p. 2015) and by Sivia (1996, pp. 15–20), the Poissonian assumption is not critical here. Any peaked distribution should work, but the best will emerge the fastest. For example, calculations using a Gaussian distribution give the same results]. That is,

$$P(C|D) = \frac{\sum_{i=1}^T \frac{C^{n_i}}{n_i!} e^{-C} P(n_i|D)}{\sum_{C=0}^{C_{\max}} \left(\sum_{i=1}^T \frac{C^{n_i}}{n_i!} e^{-C} P(n_i|D) \right)} \quad (5)$$

where T is the total number of observations, $P(n_i|D)$ is the probability of observing n particles of diameter D at the i th observation and the denominator is a normalizer of the distribution. The most frequent C then emerge naturally as the most likely expected (MLE) values surrounded by the entire probability density function (pdf) of C as best estimated under the assumptions. This differs significantly from histograms which only show the frequency distribution of observed counts n often at low resolution (coarse bins in order to achieve enough samples) with an outcome that is heavily dependent upon the one realization observed and the total number of samples. $P(C)$, on the other hand, is the observed pdf of the mean values which can emerge quite rapidly from even a limited number of observations.

References

- Field PR, Heymsfield AJ, Bansemer A (2006) Shattering and particle interarrival times measured by optical array probes in ice clouds. *J Atmos Ocean Technol* 23:1357–1371. doi:10.1175/JTECH1922.1
- Frees E, Valdez E (1998) Understanding relationships using copulas. *N Am Actuar J* 2:1–25. doi:10.1080/10920277.1998.10595667
- Genest C, Mackay J (1986) The joy of copulas: bivariate distributions with uniform marginals. *The Am Stat* 40:280–283. doi:10.1080/00031305.1986.10475414
- Heymsfield AJ, Parrish JL (1978) A computational technique for increasing the effective sampling volume of the PMS two-dimensional particle size spectrometer. *J Appl Meteorol* 17:1566–1572. doi:10.1175/1520-0450(1978)017<1566:ACTFIT>2.0.CO;2

- Heymsfield AJ, Bansemmer A, Field PR, Durden SL, Stith JL, Dye JE, Hall W, Grainger CA (2002) Observations and parameterizations of particle size distributions in deep tropical cirrus and stratiform precipitating clouds: results from in situ observations in TRMM field campaigns. *J Atmos Sci* 59:3457–3491. doi:[10.1175/1520-0469\(2002\)059<3457:OAPOPS>2.0.CO](https://doi.org/10.1175/1520-0469(2002)059<3457:OAPOPS>2.0.CO)
- Heymsfield AJ, Matrosov S, Tian L (2008) The 94-GHz radar dim band: relevance to ice cloud properties and CloudSat. *Geophys Res Lett* 35:3. doi:[10.1029/2007GL031361](https://doi.org/10.1029/2007GL031361)
- Heymsfield AJ, Matrosov S, Heymsfield G, Fierro AO (2009) Microphysics of maritime tropical convective updrafts at temperatures from -20 to -60 °C. *J Atmos Sci* 66:3530–3562. doi:[10.1175/2009JAS3107.1](https://doi.org/10.1175/2009JAS3107.1)
- Jameson AR (2007) A new characterization of rain and clouds: results from a statistical inversion of count data. *J Atmos Sci* 64:2012–2028. doi:[10.1175/JAS3950.1](https://doi.org/10.1175/JAS3950.1)
- Jameson AR, Heymsfield AJ (2013) A Bayesian approach to upscaling and downscaling of aircraft measurements of ice particle counts and size distributions. *J Appl Meteor Clim* 52:2075–2088. doi:[10.1175/JAMC-D-12-0301.1](https://doi.org/10.1175/JAMC-D-12-0301.1)
- Jameson AR, Kostinski AB (1999) Fluctuation properties of precipitation Part V: distribution of rain rates—theory and observations in clustered rain. *J Atmos Sci* 56:3920–3932. doi:[10.1175/1520-0469\(1999\)056<3920:FPOPPV>2.0.CO;2](https://doi.org/10.1175/1520-0469(1999)056<3920:FPOPPV>2.0.CO;2)
- Jameson AR, Kostinski AB (2008) The effect of clustering on the uncertainty of differential reflectivity measurements. *J Appl Meteor Clim* 47:2816–2827. doi:[10.1175/2008JAMC1860.1](https://doi.org/10.1175/2008JAMC1860.1)
- Johnson GE (1994) Constructions of particular random processes. *Proc IEEE* 82:270–285. doi:[10.1109/5.265353](https://doi.org/10.1109/5.265353)
- Liu G (2004) Approximation of single scattering properties of ice and snow particles for high microwave frequencies. *J Atmos Sci* 61:2441–2456. doi:[10.1175/1520-0469\(2004\)061<2441:AOSSPO>2.0.CO;2](https://doi.org/10.1175/1520-0469(2004)061<2441:AOSSPO>2.0.CO;2)
- Matrosov SY, Heymsfield AJ (2008) Estimating ice content and extinction in precipitating cloud systems from CloudSat radar measurements. *J Geophys Res* 113:1984–2012. doi:[10.1029/2007JD009633](https://doi.org/10.1029/2007JD009633)
- Nelsen RB (1999) An introduction to copulas. Springer, New York
- Sivia DS (1996) Data analysis a Bayesian tutorial. Oxford University Press, Oxford
- Tian L, Heymsfield GM, Heymsfield AJ, Bansemmer A, Li L, Twohy CH, Srivastava RC (2009) A study of cirrus ice particle size distribution using TC4 observations. *J Atmos Sci* 67:195–216. doi:[10.1175/2009JAS3114.1](https://doi.org/10.1175/2009JAS3114.1)