P2.94

MAXIMUM LIKELIHOOD ESTIMATION OF GAMMA PARAMETERS FOR COARSELY-BINNED AND TRUNCATED RAINDROP SIZE DATA

Roger W. Johnson*
Department of Mathematics and Computer Science
South Dakota School of Mines & Technology

Donna V. Kliche and Paul L. Smith
Institute of Atmospheric Sciences
South Dakota School of Mines & Technology

ABSTRACT

The gamma family of probability densities, which includes the exponential family as a special case, has recently been used to model raindrop size data. The traditional approach of using method of moments to estimate the gamma distribution parameters, however, is known to be biased and can have substantial errors. Methods superior to the method of moments approach include maximum likelihood. In particular, maximum likelihood estimates have been shown to outperform method of moments estimators both in the case in which the full range of drop sizes are observed as well as the case in which small drop sizes fail to be observed because of the inability of disdrometers to record observations below a threshold. The foregoing comments concern the situation in which drop sizes are measured on a continuous scale. In this work we consider drop sizes from gamma distributions which are classified into broad size bins, as would be the case with data obtained from many disdrometers; we do also allow for the possibility of drop sizes below a threshold or above another threshold not being observed. Maximum likelihood performance in this case is investigated through simulation of sampling from gamma distributions with known parameters. In particular, we compare the performance of the maximum likelihood estimates with those of method of moments and a recently developed weighted least squares procedure. The simulation process, which relies in part on numerical optimization as the maximum likelihood estimates are not expressible in closed-form, is conducted using the R statistical package (http://www.r-project.org/). Slight modifications to this code allow parameter estimation with experimental data.

1. INTRODUCTION

When fitting raindrop size data by a gamma distribution the maximum likelihood and L-moments procedures have been shown to outperform commonly used method of moments procedures in terms of both bias and variability (Kliche *et al.* 2008; Johnson *et al.* 2010). In Kliche *et al.* (2008) it was assumed that raindrop size was accurately observed over the entire range of values. In Johnson *et al.* (2010) this was relaxed to allow for values below some threshold to not be observed, as would be the case for many disdrometers. In this paper we likewise allow for data truncation but turn to the situation where raindrop size is only categorized as falling into one of several bins. That is, rather than having precise raindrop sizes we only have counts of the numbers of drops falling into several contiguous bins as would also be the case for many disdrometers. In the two papers referenced both the maximum likelihood and L-moments procedures of parameter estimation were investigated. Here, of the two procedures, we only look at modifying maximum likelihood estimation. Suitable modifications to the L-moments technique have not been identified.

**Corresponding author address*: Dr. Roger Johnson, SDSM&T, 501 E. St. Joseph, Rapid City, SD 57701; e-mail: roger.johnson@sdsmt.edu

Given such coarsely-binned disdrometer data it would seem natural to stick with parameter estimation schemes shown to work with continuous data by simply replacing bin counts with like numbers of bin midpoints. Somewhat surprisingly, the maximum likelihood (and L-moments) estimates of the gamma parameters seriously degrade by doing so. The major technical result of this paper shows how to correctly implement maximum likelihood in the coarsely-binned, truncated data situation.

We compare the performance of our maximum likelihood method to two other estimation procedures through simulation of volume samples from known gamma raindrop distributions. One estimation procedure is a method of moments approach using bin midpoints. The other is the weighted least squares approach of Brawn and Upton (2008).

The R software package (R Development Core Team 2009) was used to perform the simulations. Since the maximum likelihood and method of moments procedures do not give estimates in simple, closed form, some brief implementation details on how R was used are given.

2. MODELING RAINDROP SIZE

We model raindrop size using the gamma drop size distribution (DSD) of, for example, Chandrasekar and Bringi (1987), namely

$$n(D) = N_T \frac{\lambda^{\mu+1}}{\Gamma(\mu+1)} D^{\mu} e^{-\lambda D} \tag{1}$$

Here, the parameters are the total drop number concentration, N_T, the shape parameter μ ($\mu > -1$) and the rate parameter λ ($\lambda > 0$). Also, Γ is the gamma function. This form can be recognized as the product of the mean total number concentration, N_T, and the gamma probability density function (PDF) of drop size. When μ = 0, the gamma DSD reduces to the exponential DSD.

If instrumentation only allows drop sizes above a threshold, call it $D_{\min}$, and below another threshold, call it $D_{\max}$, to be observed, then the truncated drop size distribution for the drops observed is given by

$$\tilde{n}(D) = \tilde{N}_T \frac{\frac{\lambda^{\mu+1}}{\Gamma(\mu+1)} D^{\mu} e^{-\lambda D}}{\left[\frac{\gamma(\mu+1,\lambda D_{\max})}{\Gamma(\mu+1)} - \frac{\gamma(\mu+1,\lambda D_{\min})}{\Gamma(\mu+1)}\right]}, \tag{2}$$

for $D_{\min} < D < D_{\max}$, where $\tilde{N}_T$ is the number concentration for the truncated part of the DSD and

$$\gamma(a,x) \equiv \int_0^x t^{a-1} e^{-t}\, dt$$

is the incomplete gamma function. To see how (2) follows from (1) note, in general, that if $f(x)$ is the untruncated density, then $\tilde{f}(x) = f(x) / [F(D_{\max}) - F(D_{\min})]$ is the truncated density where $F(x) = P(X \le x)$ is the cumulative distribution function.

Note, by the way, that once estimates $\hat{D}_{\min}, \hat{D}_{\max}, \hat{\mu}, \hat{\lambda}$ of the parameters $D_{\min}, D_{\max}, \mu, \lambda$ have been determined,

$$\hat{p} = 1 - \int_{\hat{D}_{\min}}^{\hat{D}_{\max}} \frac{\hat{\lambda}^{\hat{\mu}+1}}{\Gamma(\hat{\mu}+1)} D^{\hat{\mu}} \exp[-\hat{\lambda} D]\, dD \tag{3}$$

is an estimate of the proportion, p, of missing (unobserved) drops. Also note that

$$N_T = \frac{\tilde{N}_T}{1-p} \cong \frac{\tilde{N}_T}{1-\hat{p}} \tag{4}$$

giving us a way of estimating N_T once $\tilde{N}_T$ has been estimated.

3. MAXIMUM LIKELIHOOD FOR COARSE DISDROMETER DATA

Suppose that the disdrometer used is equipped to record counts of drops in the k bins $[a_1,a_2], [a_2,a_3], \ldots, [a_{k-1},a_k], [a_k,a_{k+1}]$ with the a_i specified as known values. We may estimate $D_{\min}$ as the left end point of the smallest bin which contains at least one observation. Likewise, we may estimate $D_{\max}$ as the right end point of the largest bin which contains at least one observation. These estimates, call them $\hat{D}_{\min}$ and $\hat{D}_{\max}$, respectively, are reasonable if the bins are not "too large". Suppose n_i is the number of drops in bin $[a_i, a_{i+1}]$ and let $n = \sum_{i=1}^{k} n_i$.

The likelihood function in this case is multinomial. In particular, using the above notation the likelihood function, L, is

$$L(\mu,\lambda) = \binom{n}{n_1 \quad n_2 \quad \cdots \quad n_k} p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k} \propto p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k}$$

where

$$p_i = p_i(\mu,\lambda) = \frac{\left[\int_{a_i}^{a_{i+1}} \frac{\lambda^{\mu+1} D^{\mu} e^{-\lambda D}}{\Gamma(\mu+1)} dD\right]}{\left[\frac{\gamma(\mu+1,\lambda\hat{D}_{\max})}{\Gamma(\mu+1)} - \frac{\gamma(\mu+1,\lambda\hat{D}_{\min})}{\Gamma(\mu+1)}\right]} = \frac{\left[\frac{\gamma(\mu+1,\lambda a_{i+1})}{\Gamma(\mu+1)} - \frac{\gamma(\mu+1,\lambda a_i)}{\Gamma(\mu+1)}\right]}{\left[\frac{\gamma(\mu+1,\lambda\hat{D}_{\max})}{\Gamma(\mu+1)} - \frac{\gamma(\mu+1,\lambda\hat{D}_{\min})}{\Gamma(\mu+1)}\right]}$$

is an estimate of the chance of an observation falling in the bin $[a_i, a_{i+1}]$. To maximize L over μ and λ it is equivalent to maximize

$$\ln L(\mu,\lambda) = \ln\left[p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k}\right] = \sum_{\{i: n_i > 0\}} n_i \ln p_i \tag{5}$$

Note that the above summation is performed only over those indices i for which $n_i > 0$.

To numerically optimize (5) we use the R statistical package (R Development Core Team 2009, http://www.r-project.org/). In this package the desired maximization is performed using the *optim* function. This numerical nonlinear optimization routine, based on the algorithm of Nelder and Mead (1965), uses ordinary method of moments estimates (see (7)) as the starting point for estimating μ and λ. Further details may be found in the Appendix.

By way of reminder, once we have the estimates $\hat{D}_{\min}, \hat{D}_{\max}, \hat{\mu}, \hat{\lambda}$ we may estimate N_T from an estimate of $\tilde{N}_T$ using (3) and (4). A natural estimate of $\tilde{N}_T$ here is simply the number of drops observed.

4. A METHOD OF MOMENTS PROCEDURE

Various combinations of moments based on samples from the DSDs have commonly been used by atmospheric scientists to estimate the parameters of the underlying population distributions. For example, in the case of the gamma distribution these include the zeroth, 3rd, and 6th moments (Szyrmer *et al*. 2005); the 2nd, 3rd, and 4th moments (Smith 2003; Kliche *et al.* 2008); the 2nd, 4th, and 6th moments (Ulbrich and Atlas 1998; Vivekanandan *et al.* 2004); and the 3rd, 4th, and 6th moments (Ulbrich 1983; Kozu and Nakamura 1991; and Tokay and Short 1996).

The bias is stronger and the errors greater when higher order moments are used in calculating the parameters (Kliche *et al.* 2008; Smith *et al.* 2009). Consequently, we consider the use of the 1st, 2nd and 3rd moments in this section.

4.1 *Continuous Measurements, Observations Not Truncated*

The general form for the moments M_i of the untruncated gamma DSD function (1) can be written as

$$E(M_i) = \int_0^\infty D^i n(D)\,dD = N_T \frac{(\mu+1)(\mu+2)\cdots(\mu+i)}{\lambda^i} \qquad (6)$$

where i is a non-negative integer and, for a sample $D_1, D_2, ..., D_C$ of size C, the sample moments are

$$M_i \equiv \sum_{k=1}^{C} D_k^i$$

Setting $E(M_i) = M_i$ for i equal to 1,2,3 gives three equations in the three unknowns μ, λ, N_T. In particular, we obtain

$$N_T \frac{(\mu+1)}{\lambda} = M_1$$
$$N_T \frac{(\mu+1)(\mu+2)}{\lambda^2} = M_2$$
$$N_T \frac{(\mu+1)(\mu+2)(\mu+3)}{\lambda^3} = M_3$$

Solving these gives the method of moments estimates

$$\hat{\mu} = \frac{3M_2^2 - 2M_1M_3}{M_1M_3 - M_2^2}, \quad \hat{\lambda} = \frac{M_1M_2}{M_1M_3 - M_2^2},$$
$$\hat{N}_T = \frac{M_1^2 M_2}{2M_2^2 - M_1M_3} \qquad (7)$$

4.2 *Continuous Measurements, Observations Truncated*

For the truncated gamma DSD (2) we have the following generalization of (6)

$$E(M_i) = \int_{D_{\min}}^{D_{\max}} D^i \tilde{n}(D)\,dD = \tilde{N}_T \frac{(\mu+1)(\mu+2)\cdots(\mu+i)}{\lambda^i} \frac{\left(\frac{\gamma(i+\mu+1,\lambda D_{\max})}{\Gamma(i+\mu+1)} - \frac{\gamma(i+\mu+1,\lambda D_{\min})}{\Gamma(i+\mu+1)}\right)}{\left(\frac{\gamma(\mu+1,\lambda D_{\max})}{\Gamma(\mu+1)} - \frac{\gamma(\mu+1,\lambda D_{\min})}{\Gamma(\mu+1)}\right)}$$

Given we have estimates of $D_{\min}$ and $D_{\max}$ (e.g. the smallest and largest observations seen, respectively, in this case) the three parameters μ, λ and $\tilde{N}_T$ remain to be estimated. As before we set $E(M_i) = M_i$ for i equal to 1,2,3 to estimate μ, λ and $\tilde{N}_T$. In particular, we can solve these three equations by successfully minimizing the function

$$\begin{aligned}
&\mathrm{MM123}(\mu,\lambda,\tilde{N}_T) = \\
&\left[1 - \frac{\tilde{N}_T}{M_1}\frac{(\mu+1)}{\lambda}\frac{\left(\frac{\gamma(\mu+2,\lambda\hat{D}_{\max})}{\Gamma(\mu+2)} - \frac{\gamma(\mu+2,\lambda\hat{D}_{\min})}{\Gamma(\mu+2)}\right)}{\left(\frac{\gamma(\mu+1,\lambda\hat{D}_{\max})}{\Gamma(\mu+1)} - \frac{\gamma(\mu+1,\lambda\hat{D}_{\min})}{\Gamma(\mu+1)}\right)}\right]^2 + \\
&\left[1 - \frac{\tilde{N}_T}{M_2}\frac{(\mu+1)(\mu+2)}{\lambda^2}\frac{\left(\frac{\gamma(\mu+3,\lambda\hat{D}_{\max})}{\Gamma(\mu+3)} - \frac{\gamma(\mu+3,\lambda\hat{D}_{\min})}{\Gamma(\mu+3)}\right)}{\left(\frac{\gamma(\mu+1,\lambda\hat{D}_{\max})}{\Gamma(\mu+1)} - \frac{\gamma(\mu+1,\lambda\hat{D}_{\min})}{\Gamma(\mu+1)}\right)}\right]^2 + \\
&\left[1 - \frac{\tilde{N}_T}{M_3}\frac{(\mu+1)(\mu+2)(\mu+3)}{\lambda^3}\frac{\left(\frac{\gamma(\mu+4,\lambda\hat{D}_{\max})}{\Gamma(\mu+4)} - \frac{\gamma(\mu+4,\lambda\hat{D}_{\min})}{\Gamma(\mu+4)}\right)}{\left(\frac{\gamma(\mu+1,\lambda\hat{D}_{\max})}{\Gamma(\mu+1)} - \frac{\gamma(\mu+1,\lambda\hat{D}_{\min})}{\Gamma(\mu+1)}\right)}\right]^2
\end{aligned} \qquad (8)$$

Once again, this optimization problem may be accomplished by using the *optim* function in the R software package (R Development Core Team 2009) with further details in the Appendix. As before, method of moments estimates for the untruncated case (recall (7)) were used as a starting point.

For further details on method of moments for truncated observations from a gamma, see Vivekanandan *et al.* (2004).

4.3 *Discrete Measurements, Observations Truncated*

A natural approach here – and the one used in the simulation below, is to use the method immediately above with the diameters in the moment calculations replaced by the bin midpoints. That is, when computing $M_i \equiv \sum_{k=1}^{C} D_k^i$ use bin midpoints for the values of D_k. As discussed in the section on maximum likelihood estimation we may estimate $D_{\min}$ as the left end point of the smallest bin which contains at least one observation and $D_{\max}$ as the right end point of the largest bin which contains at least one observation.

5. BRAWN AND UPTON METHOD FOR COARSE DISDROMETER DATA

In this section the parameter estimation method of Brawn and Upton (2008) is briefly outlined. We are content, however, to present their method for volume raindrop samples rather than surface samples. See Brawn and Upton (2008) for the modifications to the presentation below for surface samples.

Rather than observing individual drop sizes, they assume that frequency counts of drops within known bins are observed. As before let n_j be the observed count or frequency of drops in bin *j.* Additionally, let D_j be the midpoint of bin *j*, and w_j be the width of bin *j*. Consider the untruncated gamma DSD given by (1), namely

$$n(D) = N_T \frac{\lambda^{\mu+1}}{\Gamma(\mu+1)} D^{\mu} \, e^{-\lambda D} \equiv N_0 D^{\mu} \, e^{-\lambda D}$$

The frequency count n_j of drops in bin *j* should be approximately equal to the DSD at the midpoint of that bin multiplied by the width of that bin. That is, it should be the case that

$$n_j \approx n(D_j) w_j = N_0 w_j D_j^{\mu} \, e^{-\lambda D_j}$$

or

$$\ln(n_j) \approx \ln(N_0) + \ln(w_j) + \mu \ln(D_j) - \lambda D_j \qquad (9)$$

Loosely speaking, we wish to choose parameters to make the left- and right-hand sides of (9) agree as well as possible. Brawn and Upton (2008) first rewrite the right-hand side of (9) in terms of just N_0 and λ. To write μ on the right-hand side of (9) in terms of λ use

$$\frac{E(D^k)}{E(D^{k-1})} = \frac{\mu+k}{\lambda}$$

(for any *k* such that $\mu + k > 0$). It then follows that

$$\mu = \frac{E(D^k)}{E(D^{k-1})} \lambda - k$$

and

$$\mu \approx \frac{S_k}{S_{k-1}} \lambda - k \qquad (10)$$

where $S_k \equiv \sum_j n_j D_j^k$ and $S_{k-1} \equiv \sum_j n_j D_j^{k-1}$. Brawn and Upton (2008) recommend *k* be taken to be 3.5 (the results in Smith *et al.* (2009) suggest that the ratio of neighboring moments of a gamma distribution can be estimated with relatively small bias and error). Replacing the approximation for μ in (10) into (9) and performing some algebra then gives

$$c_j \approx \phi + b_j \lambda \qquad (11)$$

where

$$\phi \equiv \ln(N_0)$$
$$b_j \equiv \ln(D_j) \frac{S_k}{S_{k-1}} - D_j$$
$$c_j \equiv \ln(n_j) + k \ln(D_j) - \ln(w_j)$$

Note in equation (11) that, in practice, b_j and c_j are readily calculated from the data (once *k* is selected) while ϕ and λ are unknown.

The parameters ϕ and λ are then chosen to solve the weighted least squares problem

$$\underset{\phi,\lambda}{\text{minimize}} \sum_j n_j [c_j - (\phi + b_j \lambda)]^2$$

(One subtlety here: Note that $\phi \equiv \ln(N_0)$ depends on both μ and λ so that ϕ and λ are functionally related. The minimization performed here, however, allows ϕ and λ to be independently and freely varied.) From the least squares literature (see, for example, Draper and Smith 1981, equation (2.11.10)) it follows that

$$\hat{\lambda} = \frac{n\left(\sum_j n_j b_j c_j\right) - \left(\sum_j n_j c_j\right)\left(\sum_j n_j b_j\right)}{n\left(\sum_j n_j b_j^2\right) - \left(\sum_j n_j b_j\right)^2}$$

and

$$\hat{\phi} = \frac{\left(\sum_j n_j c_j\right)\left(\sum_j n_j b_j^2\right) - \left(\sum_j n_j b_j c_j\right)\left(\sum_j n_j b_j\right)}{n\left(\sum_j n_j b_j^2\right) - \left(\sum_j n_j b_j\right)^2}$$

where

$$n \equiv \sum_j n_j$$

(in the summations only sum out over those *j* for which n_j is strictly greater than zero). From this N_0 may be estimated by using $\phi \equiv \ln(N_0)$, i.e. $\hat{N}_0 = \exp[\hat{\phi}]$. Using (7) note that μ may then be estimated as

$$\hat{\mu} = \frac{S_k}{S_{k-1}}\hat{\lambda} - k$$

6. SIMULATION PROCEDURE

Comparisons of parameter-fitting procedures to evaluate biases and errors are readily done through computer simulation of repetitive sampling from known raindrop populations. The simulated gamma DSDs are represented as the product between the total drop number concentration, N_T, assumed to follow a Poisson distribution, and the corresponding probability density function (PDF) of drop size, as in (1). (We used the *rpois* and *rgamma* functions in the R package to generate these random observations.) In our simulations we took the mean number of drops in the samples to be the numerical value of $\tilde{N}_T = (1-p)N_T$ (recall (3) and (4)), and organized the results by the value of $\tilde{N}_T$. This approach can be interpreted as representing an instrument with a sample volume of 1 m^3 (independent of the drop size), or a sample volume of α m^3 with a mean drop concentration of N_T / α.

We used about 1,000,000 drops for each simulation run. For example, we drew 20,000 samples with $\tilde{N}_T$ = 50 and 1,000 samples with $\tilde{N}_T$ = 1,000.

Four distinct gamma populations were generated. Reasoning that the truncation issue would be immaterial unless a significant portion of the drops were truncated, we used shape parameters μ = 2 and μ = 5 along with rate parameters λ chosen to give both a 10% chance and a 25% chance of drop sizes being smaller than 0.313 mm (the lower threshold for the Joss-Waldvogel disdrometer, chosen for purposes of illustration). A summary, listing the particular rate parameter values, is given in Table 1. The right-hand column of Table 1 shows the rainfall rate that would be associated with a mean sample size of $\tilde{N}_T$ = 1,000 drops.

For each selection of population parameters and sample size we recorded the following numerical summary measures of estimates: the mean, the median and the normalized root mean squared error. The normalized root mean squared error of the estimates $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_n$ of θ is defined to be the square root of $\sum_{i=1}^{n}[(\hat{\theta}_i / \theta - 1]^2 / n$. Boxplots were also generated during the simulation to compare estimates graphically; only selected boxplots are included in this paper.

In our simulation we use the Joss-Waldvogel disdrometer (JWD) bins stated in the Appendix of Brawn and Upton (2008): $[a_1, a_2] = [0.313 \text{ mm}, 0.405 \text{ mm}]$, $[a_2, a_3] = [0.405 \text{ mm}, 0.505 \text{ mm}], \ldots,$ $[a_{20}, a_{21}] =$ = [5.148 mm, 5.60 mm]. From this, note the midpoints D_j and widths w_j easily follow (e.g. D_1 = (0.313 mm + 0.405 mm)/2 = 0.359 mm, w_1 = 0.405 mm – 0.313 mm = 0.092 mm, etc.). While 5.60 mm is used as the upper truncation point in our simulations only rarely, if ever, were observations above this cutoff generated and, consequently, then truncated. For the parameter combination μ = 2 and λ = 3.52 mm^{-1} the chance of an observation above 5.60 mm is about 2.5 x 10^{-6}. Likewise, for μ = 2 and λ = 5.52 mm^{-1} this chance is about 2.0 x 10^{-10}, for μ = 5 and λ = 10.07 mm^{-1} this chance is about 1.1 x 10^{-16}, and for μ = 5 and λ = 13.48 mm^{-1} this chance is essentially zero.

Chance of a drop below 0.313 mm	Shape Parameter, μ	Rate Parameter, λ (mm^{-1})	Mean $((\mu+1)/\lambda)$ (mm)	Standard Deviation $(\sqrt{\mu+1}/\lambda)$ (mm)	Rainfall Rate (mm/h) $\tilde{N}_T = 1{,}000$
10%	2	3.52	0.85	0.49	16.30
	5	10.07	0.60	0.24	2.45
25%	2	5.52	0.54	0.31	3.06
	5	13.48	0.45	0.18	0.94

Table 1. Parameter values used in the simulations.
The mean and standard deviation values refer to the PDF given in (1).

While we use the size bins associated with a surface-based disdrometer for purposes of illustration, the simulated samples are volume samples. This is done for simplicity as well as to facilitate comparisons with our earlier work. The results of simulations of surface-sampling instruments will differ in detail, but that will have no effect on the major conclusions.

7. SIMULATION RESULTS

We begin by comparing estimates obtained using the maximum likelihood method accounting for the coarse binning of data (MLB) with those obtained from the method of moments method previously described using bin midpoints (MM123).

We generally find less bias using the MLB parameter estimates. Along these lines, note in Tables 2-5 that the mean and median MLB estimates, as a general rule, are closer to the true parameter values than those for the MM123 estimates. Based on this limited experimentation, the two methods behave most similarly in terms of average performance when dealing with large sample sizes from narrow distributions. Note, in particular, the simulations associated with Tables 4 and 5 for which the population standard deviations are smallest (0.24 mm and 0.18 mm, respectively) among the four parameter combinations examined. In these situations, when the sample size is 500 or more the mean and median estimates of the MLB and MM123 methods are quite similar (see also Figure 3). With smaller sample sizes (see, as an example, Figure 1) or wider distributions (see, as an example, Figure 2), however, the average performance of the MLB estimates is substantially better than that of MM123.

Turning to the variability of the MLB and MM123 estimates we again note fairly similar performance in the case of larger sample sizes from narrow gamma distributions. The distributions of the MLB and MM123 estimates in Figure 3, for example, show nearly identical interquartile ranges; the only apparent difference is a greater possibility of substantially underestimating both parameters when using MM123 estimates, as evidenced by the larger number of small-valued outliers. Similar performance in the case of large sample sizes from narrow distributions aside, it is notable that for all parameter combinations and sample sizes used in the simulation study we always obtained a smaller normalized root mean squared error using MLB estimates.

We now compare MLB estimates with the estimates of Brawn and Upton (BU). Putting aside average performance for a moment, the variability of these two methods is quite similar across the various parameter combinations and sample sizes used in the simulation. Note, for example, that the interquartile ranges and overall spread of the two estimates are quite similar in Figures 1-3. Unfortunately, the simulations indicate that the BU estimates do not always converge to the true population values as the sample size increases. In particular, in Tables 2-5 examine the values of the estimate means and medians for a mean sample size of 1,000. In all but the case of the broadest distribution considered (μ = 2 and λ = 3.52 mm^{-1}) we see that the BU estimates tend to seriously underestimate the true parameter values. From Figure 3 we see that over 75% of both estimates are below the true population values. While it is not entirely clear when this bias for large samples occurs, comparing Table 1 to Tables 2, 3 and 4 perhaps suggests that the magnitude of the bias increases with the amount of truncation.

Somewhat surprisingly, the maximum likelihood method developed for accurately recorded, but possibly truncated, drop size data (described in Johnson *et al.* 2010) using bin midpoints (ML in the figures) performed poorly. While we do not give numerical summaries of performance for these ML estimates, note from Figures 1-3 that this “continuous data” maximum likelihood procedure does not fare nearly as well as the MLB procedure. The continuous procedure on bin midpoints tends to seriously underestimate both parameter values.

	MLB	BU	MM123
$\tilde{N}_T$	Mean; Median; Normalized RMSE	Mean; Median; Normalized RMSE	Mean; Median; Normalized RMSE
50	2.26; 2.17; 0.57	2.54; 2.45; 0.60	2.84; 2.75; 0.86
100	2.15; 2.09; 0.39	2.28; 2.24; 0.42	2.51; 2.45; 0.61
500	2.02; 2.02; 0.16	2.02; 2.03; 0.18	2.11; 2.10; 0.25
1000	2.02; 2.01; 0.12	2.01; 2.01; 0.13	2.07; 2.07; 0.17

Table 2a. Performance of μ estimates as a function of sample size for MLB, BU and MM123 estimation procedures in the case μ = 2, λ = 3.52 mm^{-1} (10% chance of missing drops with truncation = 0.313 mm).

	MLB	**BU**	**MM123**
$\tilde{N}_T$	Mean; Median; Normalized RMSE	Mean; Median; Normalized RMSE	Mean; Median; Normalized RMSE
50	3.83; 3.68; 0.34	4.10; 3.97; 0.39	4.33; 4.18; 0.49
100	3.69; 3.61; 0.24	3.82; 3.76; 0.27	4.00; 3.90; 0.34
500	3.55; 3.53; 0.10	3.54; 3.53; 0.11	3.61; 3.60; 0.13
1000	3.54; 3.53; 0.07	3.52; 3.52; 0.08	3.57; 3.56; 0.09

Table 2b. Performance of λ estimates as a function of sample size for MLB, BU and MM123 estimation procedures in the case μ = 2, λ = 3.52 mm^{-1} (10% chance of missing drops with truncation = 0.313 mm).

	MLB	**BU**	**MM123**
$\tilde{N}_T$	Mean; Median; Normalized RMSE	Mean; Median; Normalized RMSE	Mean; Median; Normalized RMSE
50	2.44; 2.24; 0.86	2.73; 2.55; 0.88	3.20; 3.03; 1.16
100	2.21; 2.14; 0.57	2.30; 2.21; 0.58	2.87; 2.70; 0.86
500	2.06; 2.05; 0.25	1.96; 1.96; 0.26	2.35; 2.36; 0.34
1000	2.02; 2.02; 0.17	1.90; 1.91; 0.19	2.12; 2.13; 0.23

Table 3a. Performance of μ estimates as a function of sample size for MLB, BU and MM123 estimation procedures in the case μ = 2, λ = 5.52 mm^{-1} (25% chance of missing drops with truncation = 0.313 mm).

	MLB	**BU**	**MM123**
$\tilde{N}_T$	Mean; Median; Normalized RMSE	Mean; Median; Normalized RMSE	Mean; Median; Normalized RMSE
50	6.23; 5.85; 0.46	6.54; 6.22; 0.47	7.14; 6.81; 0.59
100	5.86; 5.69; 0.29	5.91; 5.75; 0.30	6.64; 6.36; 0.42
500	5.61; 5.59; 0.12	5.42; 5.42; 0.13	5.92; 5.92; 0.16
1000	5.56; 5.55; 0.08	5.34; 5.36; 0.10	5.63; 5.66; 0.11

Table 3b. Performance of λ estimates as a function of sample size for MLB, BU and MM123 estimation procedures in the case μ = 2, λ = 5.52 mm^{-1} (25% chance of missing drops with truncation = 0.313 mm).

	MLB	**BU**	**MM123**
$\tilde{N}_T$	Mean; Median; Normalized RMSE	Mean; Median; Normalized RMSE	Mean; Median; Normalized RMSE
50	5.58; 5.34; 0.47	5.53; 5.33; 0.44	6.08; 5.93; 0.54
100	5.31; 5.21; 0.32	5.16; 5.07; 0.31	5.78; 5.79; 0.40
500	5.05; 5.03; 0.14	4.83; 4.82; 0.14	5.04; 5.03; 0.16
1000	5.03; 5.01; 0.10	4.80; 4.80; 0.10	5.02; 4.98; 0.13

Table 4a. Performance of μ estimates as a function of sample size for MLB, BU and MM123 estimation procedures in the case μ = 5, λ = 10.07 mm^{-1} (10% chance of missing drops with truncation = 0.313 mm).

	MLB	**BU**	**MM123**
$\tilde{N}_T$	Mean; Median; Normalized RMSE	Mean; Median; Normalized RMSE	Mean; Median; Normalized RMSE
50	11.03; 10.55; 0.36	10.87; 10.48; 0.34	11.62; 11.26; 0.40
100	10.58; 10.36; 0.24	10.28; 10.12; 0.23	11.12; 11.05; 0.29
500	10.16; 10.12; 0.10	9.75; 9.72; 0.11	10.06; 10.01; 0.12
1000	10.10; 10.08; 0.07	9.68; 9.66; 0.08	10.01; 9.96; 0.09

Table 4b. Performance of λ estimates as a function of sample size for MLB, BU and MM123 estimation procedures in the case μ = 5, λ = 10.07 mm^{-1} (10% chance of missing drops with truncation = 0.313 mm).

	MLB	**BU**	**MM123**
$\tilde{N}_T$	Mean; Median; Normalized RMSE	Mean; Median; Normalized RMSE	Mean; Median; Normalized RMSE
50	5.89; 5.46; 0.70	5.31; 4.91; 0.64	6.64; 6.09; 0.81
100	5.40; 5.20; 0.46	4.67; 4.52; 0.45	5.93; 5.50; 0.59
500	5.10; 5.04; 0.19	4.33; 4.31; 0.24	5.22; 5.19; 0.23
1000	5.05; 5.03; 0.14	4.30; 4.28; 0.20	5.05; 5.03; 0.16

Table 5a. Performance of μ estimates as a function of sample size for MLB, BU and MM123 estimation procedures in the case μ = 5, λ = 13.48 mm^{-1} (25% chance of missing drops with truncation = 0.313 mm).

	MLB	**BU**	**MM123**
$\tilde{N}_T$	Mean; Median; Normalized RMSE	Mean; Median; Normalized RMSE	Mean; Median; Normalized RMSE
50	15.24; 14.27; 0.48	13.91; 12.99; 0.43	16.40; 15.23; 0.55
100	14.27; 13.82; 0.31	12.69; 12.31; 0.30	15.03; 14.24; 0.39
500	13.68; 13.54; 0.13	12.00; 11.93; 0.17	13.75; 13.67; 0.15
1000	13.57; 13.50; 0.09	11.93; 11.88; 0.15	13.45; 13.44; 0.10

Table 5b. Performance of λ estimates as a function of sample size for MLB, BU and MM123 estimation procedures in the case μ = 5, λ = 13.48 mm^{-1} (25% chance of missing drops with truncation = 0.313 mm).

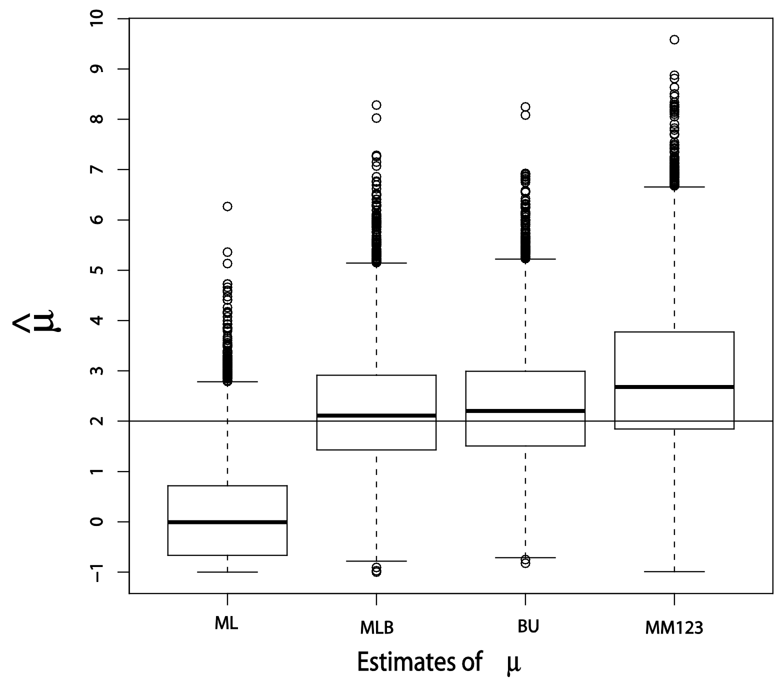

Figure 1a. Comparative boxplots of estimates of μ in the case $\mu = 2, \lambda = 5.52\ \mathrm{mm}^{-1}, \tilde{N}_T = 100$. There is a 25% chance of drop size being below 0.313 in this case.

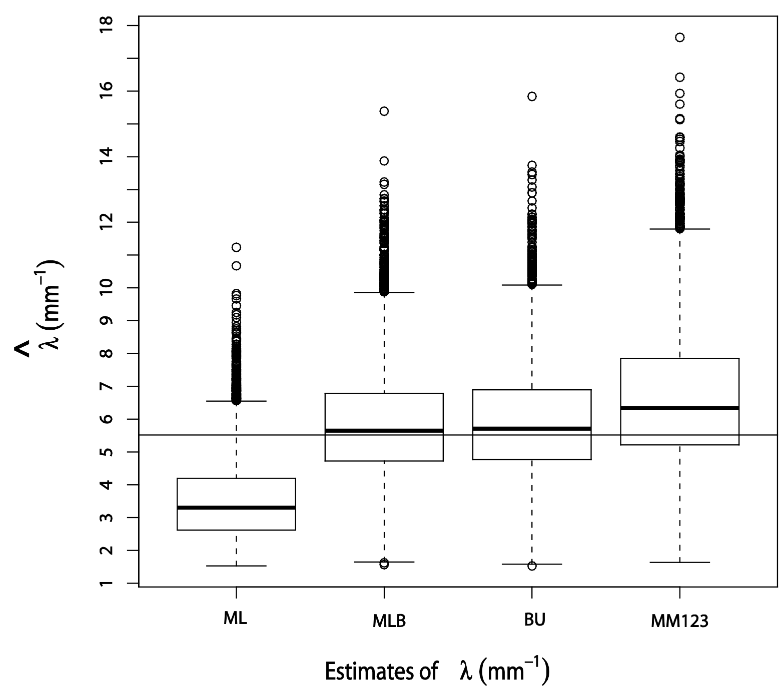

Figure 1b. Comparative boxplots of estimates of λ in the case $\mu = 2, \lambda = 5.52\ \mathrm{mm}^{-1}, \tilde{N}_T = 100$. There is a 25% chance of drop size being below 0.313 in this case.

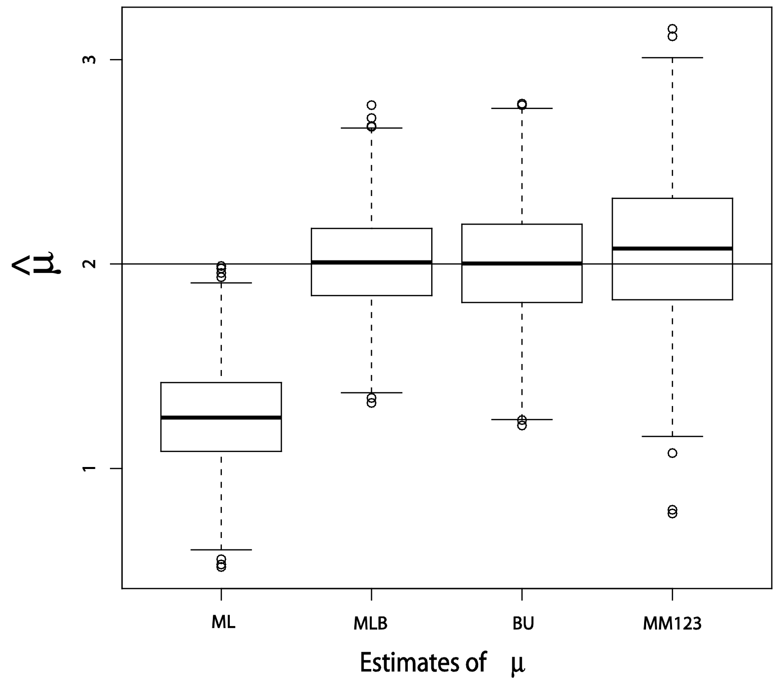

Figure 2a. Comparative boxplots of estimates of μ in the case $\mu = 2, \lambda = 3.52\ \text{mm}^{-1}, \tilde{N}_T = 1000$.
There is a 10% chance of drop size being below 0.313 in this case.

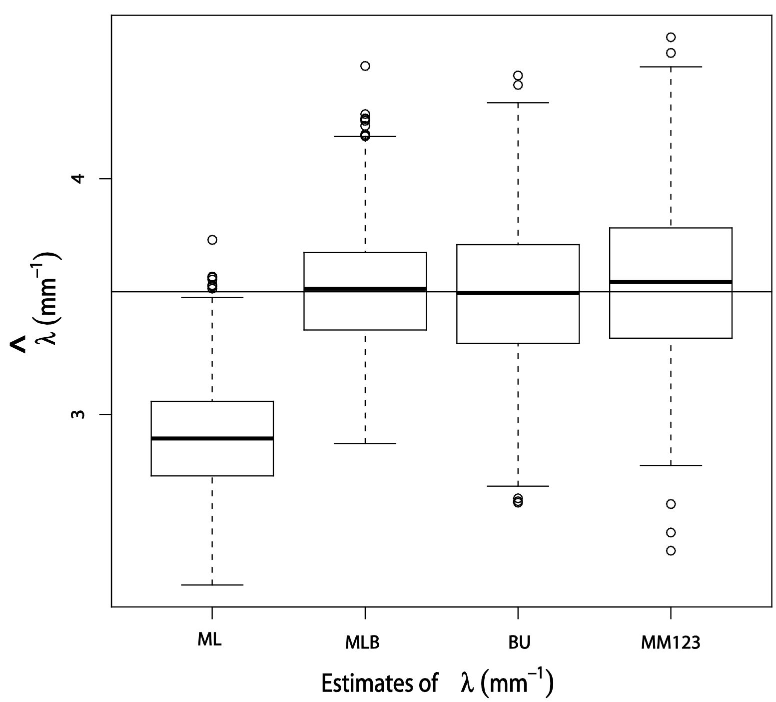

Figure 2b. Comparative boxplots of estimates of λ in the case $\mu = 2, \lambda = 3.52\ \text{mm}^{-1}, \tilde{N}_T = 1000$.
There is a 10% chance of drop size being below 0.313 in this case.

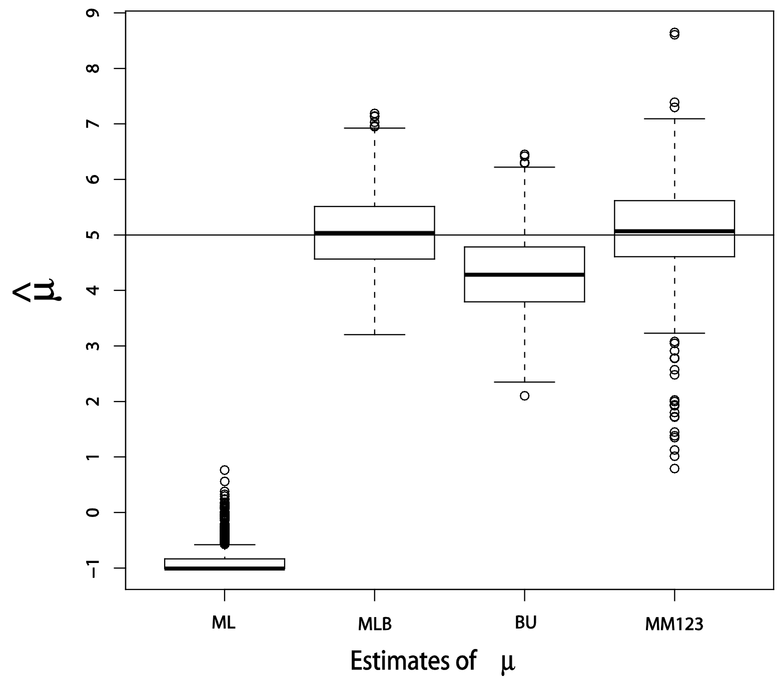

Figure 3a. Comparative boxplots of estimates of μ in the case $\mu = 5, \lambda = 13.48$ mm$^{-1}, \tilde{N}_T = 1000$. There is a 25% chance of drop size being below 0.313 in this case.

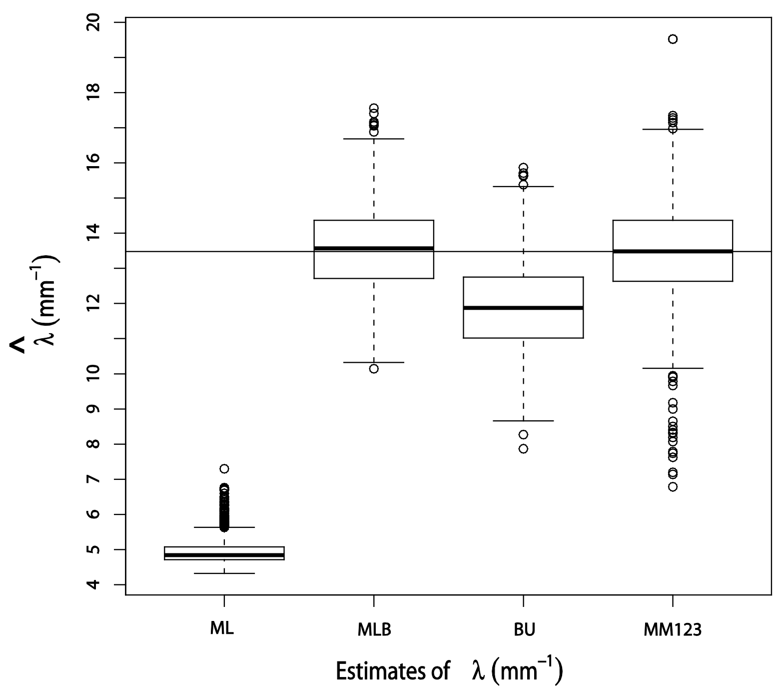

Figure 3b. Comparative boxplots of estimates of λ in the case $\mu = 5, \lambda = 13.48$ mm$^{-1}, \tilde{N}_T = 1000$. There is a 25% chance of drop size being below 0.313 in this case.

8. CONCLUSIONS

In general, the modified maximum likelihood procedure designed to handle coarsely-binned disdrometer data substantially outperforms the method of moments procedure in terms of both bias and variability. For just some parameter combinations with a large enough sample size the method of moments procedure using bin midpoints performs as well as the maximum likelihood procedure. Consequently, we recommend this maximum likelihood procedure be used instead of method of moments.

The maximum likelihood procedure for binned data is also superior to the weighted least squares procedure of Brawn and Upton (2008). A notable problem with the Brawn and Upton (2008) procedure is that parameter estimates, on average, do not always converge to the true parameter values as sample size increases – the shape and rate parameters may be substantially underestimated.

When using disdrometers which record (possibly truncated) observations on a continuous scale, the maximum likelihood procedure presented by Mallet and Barthes (2009) and detailed by Johnson *et al*. (2010) is recommended. Simply substituting bin midpoints into this procedure for data from disdrometers that only record bin counts is not recommended – there is a serious degradation in performance by doing so.

When using disdrometers which record (over a truncated range) bin counts, we recommend the maximum likelihood method discussed in this paper be used for the estimation of gamma shape and rate parameters. When dealing with truncated coarsely-binned data from surface samples, by the way, if drop fall speeds are approximated by a power law relationship $v(D) \propto D^d$ the maximum likelihood estimates based on (5) can be used with disdrometer data to estimate $\tilde{\mu} \equiv \mu + d$ and λ. It is of no particular surprise that maximum likelihood should perform well here – such estimates are generally known to perform well in terms of bias and variability, especially for moderate to large sample sizes. See, for example, Norden (1972) and Norden (1973) for further details.

Acknowledgements. This research was supported by NASA under Cooperative Agreement NNX07AL04A.

APPENDIX: R CODE

For the below, when working with experimental data, it is understood that

lower = smallest value of a_i with at least one observation in $[a_i, a_{i+1}]$
upper = largest value of a_{i+1} with at least one observation in $[a_i, a_{i+1}]$

In the Monte Carlo simulation lower = a_1 and upper = a_{k+1}.

Also, the *pgamma* function, namely

$$\text{pgamma}(x,a) = \frac{1}{\Gamma(a)}\int_0^x t^{a-1} e^{-t}\, dt = \frac{\gamma(a,x)}{\Gamma(a)}$$

is a standard function in R.

a. Function *LNLIKELIHOOD* – for use in finding maximum likelihood estimates of μ, λ for binned (discrete) data from a truncated gamma (see (5)):

```
lnlikelihood <- function(x){
 x1 <- x[1]     # mu
 x2 <- x[2]     # lambda
 if ((x1>-1) & (x2>0)) {
    tmp <- 0.0
    for (j in 1:numbins)
       if (n[j]>0)
         {tmp <- tmp - n[j]*log( (pgamma(x2*bins[j+1],x1+1)-pgamma(x2*bins[j],x1+1))/
                                        (pgamma(x2*upper,x1+1)-pgamma(x2*lower,x1+1)) )}
    return(tmp)
    }
 else
   {return(Inf)}
}
```

b. Function *MM123* – for use in finding method of moments estimates of $\mu, \lambda, \tilde{N}_T$ for data from a truncated gamma (see (8)):

```
MM123system <- function(x){
 x1 <- x[1]     # mu
 x2 <- x[2]     # lambda
 x3 <- x[3]     # N
 if ((x1>-1) & (x2>0))
  {return( ( 1.0 - (x3/M1)*((x1+1)/x2)*
                (pgamma(upper*x2,x1+2)-pgamma(lower*x2,x1+2))/
                (pgamma(upper*x2,x1+1)-pgamma(lower*x2,x1+1)) )^2 +
               ( 1.0 - (x3/M2)*((x1+1)/x2)*((x1+2)/x2)*
                (pgamma(upper*x2,x1+3)-pgamma(lower*x2,x1+3))/
                (pgamma(upper*x2,x1+1)-pgamma(lower*x2,x1+1)) )^2 +
               ( 1.0 - (x3/M3)*((x1+1)/x2)*((x1+2)/x2)*((x1+3)/x2)*
                (pgamma(upper*x2,x1+4)-pgamma(lower*x2,x1+4))/
                (pgamma(upper*x2,x1+1)-pgamma(lower*x2,x1+1)) )^2) }
 else
  {return(Inf)}
}
```

REFERENCES

Brawn, D. and G. Upton, 2008: Estimation of an atmospheric gamma drop size distribution using disdrometer data. *Atmos. Res.,* **87**, 66-79.

Chandrasekar, V. and V.N. Bringi, 1987: Simulation of radar reflectivity and surface measurements of rainfall. *J. Atmos. Oceanic Tech.,* **4**, 464-478.

Draper, N. and H. Smith, 1981: *Applied Regression Analysis*, 2nd edition. John Wiley & Sons, New York, NY.

Johnson, R.W, D.V. Kliche, and P.L. Smith, 2010: Comparison of estimators for parameters of gamma distributions with left-truncated samples, submitted to *J. Appl. Meteor*.

Kliche, D.V., P.L. Smith, and R.W. Johnson, 2008: L-moment estimators as applied to gamma drop size distributions. *J. Appl. Meteor.*, **47,** 3117-3130.

Kozu, T. and K. Nakamura, 1991: Rainfall parameter estimation from dual-radar measurements combining reflectivity profile and path-integrated attenuation. *J. Atmos. Oceanic Technol.*, **8**, 259-270.

Mallet, C. and L. Barthes, 2009: Estimation of gamma raindrop size distribution parameters: Statistical fluctuations and estimation errors. *Journal of Atmospheric and Oceanic Technology*, **26**, 1572-1584.

Nelder, J.A. and R. Mead, 1965: A simplex algorithm for function minimization. *Computer Journal,* **7**, 308–313.

Norden, R.H., 1972: A survey of maximum likelihood estimation: Part I, *International Statistical Review*. **40,** 329-354.

Norden, R.H., 1973: A survey of maximum likelihood estimation: Part II. *International Statistical Review*, **41,** 39-58.

R Development Core Team, 2009: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Smith, P.L., 2003: Raindrop size distributions: Exponential or gamma - does the difference matter? *J. Appl. Meteorol.*, **42,** 1031-1034.

Smith, P.L., D.V. Kliche, and R.W. Johnson, 2009: The bias and error in moment estimators for parameters of drop-size distribution functions: Sampling from gamma distributions. *J. Appl. Meteor.*, **48**, 2118-2126.

Szyrmer, W., S. Laroche, and I. Zawadski, 2005: A microphysical bulk formulation based on scaling normalization of the particle size distribution. Part I: Description. *J. Atmos. Sci.*, **62,** 4206-4221.

Tokay, A. and D.A. Short, 1996: Evidence from tropical raindrop spectra of the origin of rain from stratiform versus convective clouds. *J. Appl. Meteor*., **35**, 355-371.

Ulbrich, C.W., 1983: Natural variation in the analytical form of the raindrop size distribution. *J. Climate Appl. Meteor*., **22**, 1764-1775.

Ulbrich, C.W. and D. Atlas, 1998: Rainfall microphysics and radar properties: Analysis methods for drop size spectra. *J. Appl. Meteor.*, **37,** 912-923.

Vivekanandan, J., G. Zhang, and E. Brandes, 2004: Polarimetric radar estimators based on a constrained gamma drop size distribution model. *Journal of Applied Meteorology*, **43**, 217-230.