P8A.15 L-Moment Estimators as Applied to Gamma Drop Size Distributions

Donna V. Kliche, Paul L. Smith, and Roger W. Johnson
South Dakota School of Mines and Technology, Rapid City, SD, USA

Abstract

The traditional approach with experimental raindrop size data is to use the method of moments in the fitting procedure to estimate the parameters for the raindrop size distribution (RSD) function. However, the moment method is known to be biased. Therefore, we investigated the L-moment method, which is widely used by hydrologists, as an alternative. We applied the L-moment method, along with the moment and maximum likelihood (ML) methods, to simulated samples taken from gamma raindrop populations. A comparison of the bias and the errors involved in the fitting procedures of moments, ML, and L-moments shows that with samples covering the full range of drop sizes, ML and L-moments outperform the method of moments, and for small sample sizes L-moments outperforms ML.

The effects of the absence of small drops in the samples (typical disdrometer minimum size thresholds are 0.3-0.5 mm) on the fitting procedures are also analyzed. Our results show that missing small drops, due to the instrumental constraint, results in a large bias in the case of the L-moment and ML fitting methods; this bias did not decrease with increasing sample size. The very small drops have a negligible contribution to moments of order two or higher, and the bias in the moment methods seems to be about the same as in the case of full samples.

1. INTRODUCTION

Knowledge of the raindrop size distribution (RSD) is essential in the retrieval of rainfall properties utilizing radar remote sensing techniques and in the understanding of the microphysics involved in formation of precipitation. The RSD is usually expressed mathematically in terms of a distribution function, which expresses the number of drops per unit size interval per unit volume of space.

The most widely used description for the raindrop spectrum in space is the size distribution of Marshall and Palmer (1948), which is of exponential form and has two parameters:

$$n(D) = n_0 \exp(-\Lambda D), \quad (D \geq 0) \qquad (1)$$

where $n(D)$ represents the number of raindrops per unit diameter interval and per unit volume of air, D is the drop diameter, and n_0 is the value of $n(D)$ for D = 0. In a semi-logarithmic plot, equation (1) becomes the graph of a straight line with the size (scale) parameter Λ as slope, and n_0 as the y-intercept.

Ulbrich (1983) and Willis (1984), among others, proposed the use of the gamma distribution, since it can give a more appropriate description of the natural variations of the observed RSDs; in addition, the exponential distribution is a special case. In general, a gamma RSD can be expressed by

$$n(D) = n_1 D^{\mu} \exp(-\lambda D), \quad (D \geq 0) \qquad (2)$$

where n_1 is related to the raindrop concentration, μ is the dimensionless shape parameter, and λ is the size (scale) parameter. For a gamma distribution, positive (negative) values of the shape parameter μ indicate concave down (up) shape of the drop spectrum. As can be seen from (2), the units for n_1 are different for a gamma distribution than the units for n_0 in the exponential distribution (1). This can be somewhat confusing, since the same symbol n_0 is widely used in both expressions. The gamma distribution has been widely accepted by the radar meteorology and cloud physics communities (e.g. Wong and Chidambaram, 1985; Chandrasekar and Bringi, 1987; Kozu and Nakamura, 1991; Haddad *et al*., 1996; Tokay and Short, 1996; Ulbrich and Atlas, 1998; Zhang *et al*., 2003), although measurements of RSDs show that even the gamma distribution is not general enough to adequately represent the full range of observed RSDs.

Measurements of raindrop distributions have suggested that the fitted gamma distributions can have a wide range of shape parameter values. Goddard and Cherry (1984) suggested shape parameter μ = 5 to be a better representation than μ = 0, while Ulbrich and Atlas (1984) found that μ = 2 is an appropriate value for the shape parameter of their observed distributions. Later on, Ulbrich and Atlas (1998) found that the shape parameter for the measured RSD can vary from $-2 \leq \mu \leq 5$

Corresponding author address: Dr. Donna V. Kliche, SDSM&T, 501 East Saint Joseph St., Rapid City, SD 57701-3995; E-mail: donna.kliche@sdsmt.edu

with a median value of zero and a mean of 1.67. Kozu and Nakamura (1991) and Tokay and Short (1996) presented distributions of estimated μ values covering a wide range of values.

One important limitation of disdrometer instruments is the effect of censoring the observed raindrop size distributions at lower (e.g. some disdrometers cannot respond to drops < 0.3 mm) and upper (the upper threshold for one electromechanical disdrometer is 5 mm and for one optical disdrometer is 10 mm) drop diameters. The thresholds imposed by the disdrometers generate important questions: Should we be concerned about the missing small drops in the samples? Are important larger drops missing from the observations?

The goal of the present work is to investigate how the L-moment estimators (Hosking and Wallis, 1997) compare with the moment estimators that are known to be biased (e.g. Robertson and Fryer, 1970; Wallis, 1974; Smith and Kliche, 2005; Smith *et al.*, 2007) and with maximum likelihood (ML) estimators that are asymptotically unbiased. We also investigate how the moment, ML, and LM estimators are affected when small drops are missing from the samples.

2. SIMULATION PROCEDURE

The disdrometer measurements of raindrop sizes are numerous. From the statistical point of view, these samples provide an approximate description of the populations from which they are taken. The methods of fitting analytical expressions on the sample data provide the mathematical approach needed to describe the underlying populations of raindrops.

The process has to begin with assuming some form for the raindrop size distribution function. For example, given a measured sample of raindrops, one may assume that a gamma distribution should describe analytically the raindrop spectrum. The next step is to estimate the parameters for the assumed distribution using the sample data. The traditional approach with experimental raindrop size data has been the use of the method of moments to estimate the parameters for the RSD; because of its mathematical simplicity this method is widely used by cloud modelers and radar meteorologists. However, the moment method is biased (Robertson and Fryer, 1970; Wallis, 1974; Haddad *et al*., 1996, 1997; Smith and Kliche, 2005). The outcome of this bias is that the estimated parameters obtained using the moments method often tend to differ significantly from the true values of the population. The great concern is that the biased values can lead to wrong conclusions about the features of the RSD being sampled. Consequently, better parameter-fitting methods such as maximum likelihood (ML) or L-moments may be better suited to this problem.

The bias of the method of moments has been demonstrated by testing it on known RSD populations from which repetitive samples are taken (Smith and Kliche, 2005; Smith *et al.*, 2007). This experiment can be done only through a computer simulation of known raindrop populations. Therefore, we generated simulated gamma raindrop populations using a Monte Carlo technique.

In our computer simulation, we defined a dimensionless size variable y as the ratio of drop diameter over the mass-weighted mean diameter, D_m,

$$y = \frac{D}{D_m} \tag{3}$$

The mass-weighted mean diameter represents the mean volume diameter for a raindrop distribution and is defined as the ratio between the 4th moment and the 3rd moment of the RSD. D_m is a valuable quantity for describing RSD, slightly larger than the median volume diameter D_0, and it also has practical uses. For example, Seliga and Bringi (1976) showed that differential reflectivity, Z_{dr}, can be a function of D_m. This relationship is useful for linking an important cloud microphysics quantity, D_m, and a measurable radar variable, Z_{dr}.

The normalization (3) is used in the present study in order to not be required to keep track of the actual raindrop diameters. We classified the drop sizes into intervals of Δy = 0.02, which represents the size classification procedure similar to drop-measuring instruments.

Another important step in our computer simulation was to express the gamma RSDs as the product between the total raindrop number concentration, N_T, and the corresponding probability density function (PDF) of drop size. In our simulation we also designated N_T to be the mean number of drops in the sample, so that the results can be organized by the value of N_T. The advantage of this approach is that it can be interpreted as representing an instrument with a sample volume of 1 m^3 (independent of the drop size), and it also applies for a sample volume of α m^3 and a mean drop concentration of N_T/α.

To simulate raindrop sampling for purposes of estimating the sampling statistics, we started with a Monte Carlo procedure as described for the exponential case in Smith and Kliche (2005). The simulation starts with a selected value of the population total number concentration, N_T, and draws from a Poisson distribution, with that mean value, to determine the actual number of drops C in a given sample. Then C values of y (normalized raindrop diameters) are randomly drawn from the population PDF of drop sizes. The sampling from any given population and any given sample size was repeated for enough trials to obtain useful estimates of the sampling distributions of all quantities of interest. The intent is to provide indications of the bias and the uncertainty in the estimates of the population parameters, based on a single sample. A Monte Carlo simulation allows such repetitive sampling, and yields the sampling distributions of the parameter estimates, from which the bias and uncertainties can be determined. The mean sample sizes we used were 10, 20, 50, 100, 200, 500 and 1,000 drops.

The size range for the computer-generated gamma raindrop populations is $0 < y \le 3.0$. We used about 1,000,000 drops for the simulated samples; for example, there were 20,000 samples with $N_T = 50$ and 100,000 samples with $N_T = 10$. As the probability of a drop in a gamma PDF with shape parameter $\mu = 2$ being larger than $y = 3.0$ is 2.76×10^{-6} (and it is even smaller for higher values of μ), we are lacking only a few larger drops from a full gamma RSD. Two distinct gamma populations were generated: one that had the shape parameter equal 2 and another with the shape parameter equal 5. Distributions with these two values are mentioned in the literature, and it was recommended that the latter one could be even a better choice for raindrop samples measured at the ground using the disdrometer instruments.

Regardless of the function chosen to represent the RSD, some means of determining the parameters appropriate for any given set of observations is needed. Apart from fitting by eye, as Marshall and Palmer, possibilities include the method of moments, the method of maximum likelihood (ML), and the L-moment method.

The exponential RSD and the corresponding equations for the moment and ML fitting methods are discussed in detail in Smith and Kliche (2005), Smith *et al.* (2005), and Kliche *et al.* (2006). With the present paper, we provide the L-moment equations and the results for the gamma case.

We applied the moment, ML, and L-moment methods for all samples from the computer-generated gamma raindrop spectra. Since disdrometers have instrumental limitations at small drop sizes (typically cannot respond to drop sizes < 0.3 mm or so), we decided to investigate the effect of this problem by withdrawing small drops from the simulated samples. We chose to impose a threshold and to eliminate from each generated sample the drops that have normalized sizes satisfying the condition

$$y_i = \frac{D_i}{D_m} \le 0.2 \qquad (4)$$

In other words, we eliminated drops smaller than 0.2 mm if $D_m = 1$ mm or 0.6 mm if $D_m = 3$ mm. For the “censored” samples we applied the same fitting procedure as in the case of full samples.

In the case of the gamma RSD with a shape parameter $\mu = 2$, 12% of the drops in the population have $y \le 0.2$, so that on average with $N_T = 50$ six of the drops will be removed from each sample by imposing this threshold. For the case of a gamma distribution having $\mu = 5$, about 1% of drops in the population have $y \le 0.2$, so that on average with $N_T = 50$ fewer than one drop is removed from each sample; the samples practically remain intact after imposing this threshold.

The gamma distribution has a more convenient representation in terms of the total drop number concentration N_T (Chandrasekar and Bringi, 1987). This form was adopted by Smith *et al.* (2005) and Kliche *et al.* (2006), who also included the mass-weighted mean diameter $D_m = (\mu + 4)/\lambda$, as shown below:

$$n(D) = N_T \frac{(\mu+4)^{\mu+1}}{\Gamma(\mu+1)} \frac{D^{\mu}}{D_m^{\mu+1}} \exp\left[-(\mu+4)D/D_M\right] \qquad (5)$$

where the parameters are N_T, μ ($\mu > -1$) and $D_m > 0$, and $\Gamma(x)$ 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. Equation (5) is similar to the one recommended by Chandrasekar and Bringi (1987), in which we used D_m instead of their use of D_0. When $\mu = 0$, the gamma RSD reduces to the exponential RSD.

Another formulation of the gamma RSD involving N_W, which is the concentration parameter normalized with respect to LWC (Bringi and Chandrasekar, 2001), is also used by the radar community:

$$n(D) = \frac{3}{128} N_W \frac{(\mu+4)^{\mu+1}}{\Gamma(\mu+4)} \left(\frac{D}{D_m}\right)^{\mu} \tag{6}$$
$$x \exp[-(\mu+4)D/D_m]$$

and N_W is an "intercept" or concentration parameter defined by

$$N_W = \frac{256}{\pi \rho_w} \left(\frac{LWC}{D_m^4}\right) \tag{7}$$

and ρ_w is the water density. N_W is considered to have the same meaning as the intercept n_0 of an equivalent exponential distribution with the same water content and same D_m as the gamma RSD with form (5) (Bringi and Chandrasekar, 2001). Equation (6) changes to the exponential form for $\mu = 0$, and $N_W = n_0$.

3. PARAMETER-FITTING PROCEDURES

To test the effectiveness of the moment, ML and L-moment methods in recovering the parameters of a known gamma RSD, we changed equation (5) to reflect the normalized drop diameter as described in equation (3). The equation used in our simulations for gamma RSD is given by

$$n(y) = N_T \frac{(\mu+4)^{\mu+1}}{\Gamma(\mu+1)} y^{\mu} \exp[-(\mu+4)y] \tag{8}$$

3.1 The Moment Method Applied to Gamma RSD

The traditional approach with experimental RSD data has been to use the *method of moments* to estimate the parameters for the RSDs. We treat this method first, to provide a basis of comparison for the L-moment and ML methods. Various combinations of moments calculated based on samples from the RSD can be used to estimate the parameters of the underlying population distributions. For example, in the case of the gamma distribution, Szyrmer *et al*. (2005) used zero moment, 3rd moment, and 6th moment in their fitting procedure; Smith (1998) suggested the combination of 2nd moment, 3rd moment, and 4th moment; Ulbrich (1983), Kozu and Nakamura (1991), and Tokay and Short (1996) used higher moments 3rd, 4th, and 6th moment, while Ulbrich and Atlas (1998) used 2nd, 4th, and 6th moments in their study.

However, the moment method is known to be biased (Robertson and Fryer, 1970; Wallis, 1974; Haddad *et al*., 1996, 1997; Smith and Kliche, 2005; Uijlenhoet *et al*., 2006), which means that the fitted functions often do not correctly represent the raindrop populations, and sometimes not even the samples. The bias is stronger when higher-order moments are considered in calculating the parameters of the "fitted" functions, and the combination of 2nd, 3rd, and 4th moments typically gives the smallest bias for three-parameter distributions. (Although lower-order moments would be desirable in such estimations, they can be poorly determined because of instrumental deficiencies.)

The general form for the moments of a gamma RSD (5) can be written as

$$M_i = N_T(\mu+1)(\mu+2)...(\mu+i)\left[\frac{D_m}{\mu+4}\right]^i \tag{9}$$

In our simulation we used the normalized form (3) for the raindrop diameters, rather than specific drop diameters. Therefore, the six sample moments M_{1S} through M_{6S} are calculated for each sample as normalized sample moments, m_i, which are defined by

$$M_{iS} = m_i D_m^i \qquad \text{where}$$
$$m_i = \sum_C \left(\frac{D}{D_m}\right)^i = \sum_C y^i \tag{10}$$

This format has the advantage of allowing dimensionless expressions for the parameter estimates of the RSD, where the actual drop size does not have to be included. The sample values of the moments are expected to be unbiased, therefore the mean values of the sample moments, M_{iS}, should represent the moments of the drop population from which we are sampling.

The estimated parameters $\hat{\mu}$, $\frac{\hat{\lambda}}{\lambda}$ (or $\frac{\hat{D}_m}{D_m}$) and $\hat{N}_T$ (or $\frac{\hat{N}_W}{N_W}$) - values of dimensional parameters normalized by dividing by corresponding population value - for the gamma RSD are calculated using the expressions given in Appendix A.

The sampling distributions of sample moments from long-tailed RSDs like the gamma are skewed; Uijlenhoet *et al.* (2006) provide details on the mathematical framework to quantify such skewness. Figure 1 indicates how this skewness varies with the sample size in the case of the median sample moment M_3 and the median sample moment M_6. The general tendency is for the sample moments to be lower than the corresponding population values, more so for the case μ = 2 than the case μ = 5; this behavior is the ultimate cause of the bias in the moment methods for estimating RSD parameters. The skewness in the sampling distributions for the moments increases with the order of moment M_i, and decreases with increasing sample size N_T and with increasing values of μ. With samples of hundreds or thousands of drops the skewness may become small enough to be negligible.

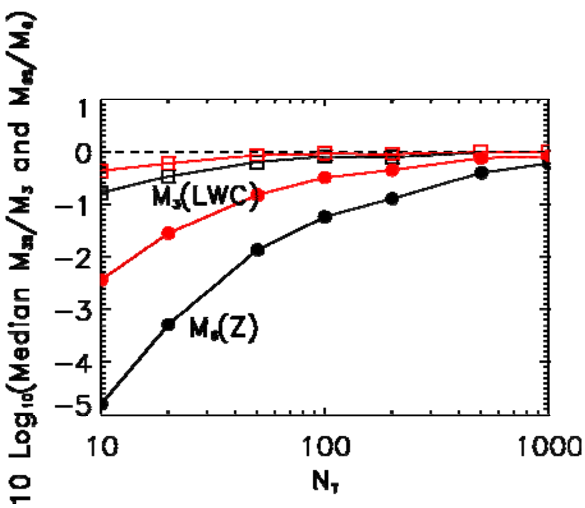

Figure 1: Plot of median values of 3rd sample moment M_{3S} (proportional to LWC) and 6th sample moment M_{6S} (proportional to Z) versus mean sample size; sample moments are normalized with respect to the corresponding population value. Gamma μ = 2: black dots correspond to M_{6S} values and black squares represent M_{3S} values. Gamma μ = 5: red dots for M_{6S} and red squares for M_{3S}. Horizontal dashed line indicates the population value.

Figures 2, 3, and 4 show examples of histograms for the estimated gamma parameters, $\hat{D}_m / D_m$, $\hat{\mu}$, and $\hat{N}_W / N_W$ in the case of a gamma distribution having μ = 2. Each figure includes results for both full and censored samples. From Figure 2 one can see that D_m tends to be underestimated, and the censored case gives essentially the same results as the full samples. Figure 3 shows that the gamma shape parameter (μ) tends to be overestimated, and the overestimation for the case of censored samples is somewhat larger than the case of original samples.

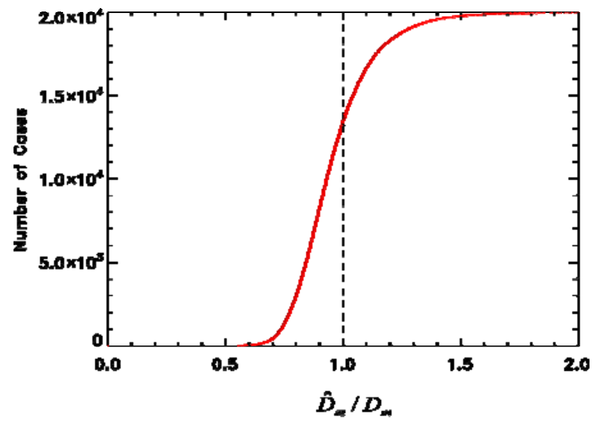

Figure 2: Cumulative histograms of normalized estimated mass-weighted mean diameter using the $M_2M_3M_4$ set; the censored-sample (red) curve is indistinguishable from the full-sample curve (black). Population RSD: gamma, μ = 2, N_T = 50. Vertical dashed line indicates the population value.

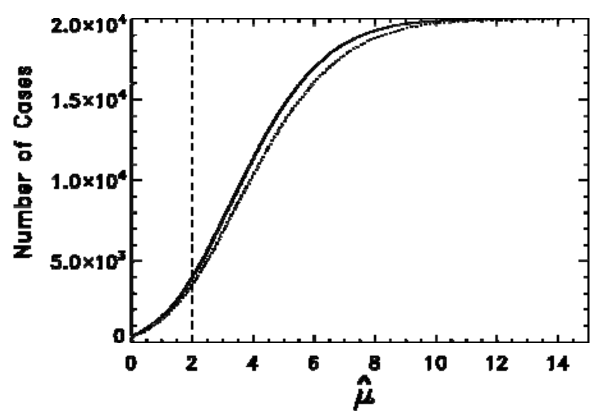

Figure 3: Cumulative histograms of estimated gamma shape parameter using the $M_2M_3M_4$ set; original sample values are shown as a continuous black line, and the censored sample values are shown as a black dotted line. Population RSD: gamma, μ = 2, N_T = 50. Vertical dashed line indicates the population value.

Figure 4 shows the cumulative histogram for the normalized values of the concentration parameter (normalized with respect to LWC) $\hat{N}_W/N_W$, and demonstrates that it tends to be overestimated. The overestimation is essentially the same whether full or censored samples are considered.

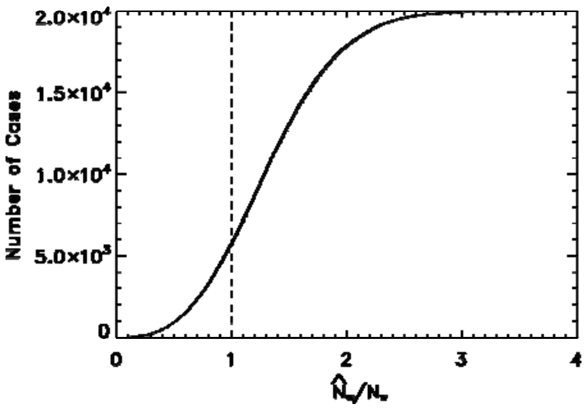

Figure 4: Cumulative histograms of normalized estimated gamma concentration parameter $\hat{N}_W/N_W$ using the $M_2M_3M_4$ set; the censored-sample curve (red) is indistinguishable from the full-sample curve (black). Population RSD: gamma, μ = 2, N_T = 50. Vertical dashed line indicates the population value.

Table 1 gives the mean, median, and root mean square (RMS) error values for each gamma estimator using the $M_2M_3M_4$ set. The values in parentheses correspond to the case of censored samples. In this table, values for the normalized estimated total number concentration, $\hat{N}_T$, are also included, and demonstrate that it tends to be underestimated, more so in the case of censored samples than in the case of full samples. Normalized mean values different from 1 show that the estimated gamma parameters are biased. The greatest (relative) bias, and also the largest error, occurs for the shape parameter.

For each sample we calculated the gamma RSD parameters using the three moment combinations listed in Appendix A. Figure 5 shows the median estimates of μ for the three-moment combinations: $M_2M_3M_4$, $M_2M_4M_6$, and $M_3M_4M_6$, for a gamma RSD with shape parameter μ = 2, versus sample size. This figure demonstrates again that the gamma RSDs "fitted" this way tend to have overestimated shape parameters. Figure 5 also illustrates how the increase of the skewness with the moment order translates into greater biases for the estimated parameters when higher-order moments are used. In the case of the $M_2M_3M_4$ combination, the bias is the smallest; consequently, we use only the $M_2M_3M_4$ moments from here on. As further illustrated in Figure 5, in the case of a gamma population having μ = 5, the estimated shape parameter exhibits a smaller bias than in the case μ = 2; with μ = 5 the drop-size spectra are narrower, the skewness in the distributions of m_i is reduced, and the bias is less.

Table 2 includes the mean and RMS error values of the estimated shape parameter vs. sample size for two different population shape parameters. These results show that as the sample size increases, the bias and the errors decrease. Thus for samples having hundreds or thousands of drops from a gamma RSD, the moment estimators' bias can be small or negligible; therefore, the moment estimators may be sufficiently close to the true population values. The values in parentheses correspond to the censored samples case; the bias is again larger, as are the errors, with the wider distribution (μ = 2), and less so with the narrower one (μ = 5).

Table 1. Normalized Moment estimators for gamma distribution (μ = 2), in the case of the $M_2M_3M_4$ set. Mean sample size N_T = 50. Values in parentheses correspond to censored samples.

Parameter	Mean	Median	RMS Error
$\hat{\mu}/\mu$	1.90 (2.05)	1.81 (1.94)	1.39 (1.56)
$\hat{D}_m/D_m$	0.96 (0.96)	0.93 (0.93)	0.17 (0.17)
$\hat{N}_W/N_W$	1.33 (1.32)	1.29 (1.28)	0.62 (0.62)
$\hat{N}_T/N_T$	0.93 (0.90)	0.88 (0.85)	0.60 (0.67)
$\hat{\lambda}/\lambda$	1.45 (1.50)	1.37 (1.42)	0.73 (0.80)

Table 2: Mean value and RMS error of the estimated shape parameter vs. sample size for two gamma distributions. Values in parentheses correspond to censored samples.

N_T	$\mu = 2$		$\mu = 5$	
	Mean	RMS Error	Mean	RMS Error
10	8.99 (10.59)	9.85 (19.21)	12.93 (13.08)	12.68 (12.98)
20	5.79 (6.37)	5.21 (6.03)	8.99 (9.07)	6.30 (6.39)
50	3.80 (4.11)	2.78 (3.12)	6.77 (6.82)	3.37 (3.42)
100	3.08 (3.32)	1.91 (2.14)	5.97 (6.00)	2.28 (2.30)
200	2.64 (2.85)	1.37 (1.54)	5.56 (5.59)	1.61 (1.63)
500	2.27 (2.44)	0.86 (0.97)	5.22 (5.25)	1.02 (1.04)
1000	2.17 (2.33)	0.64 (0.73)	5.14 (5.17)	0.79 (0.80)

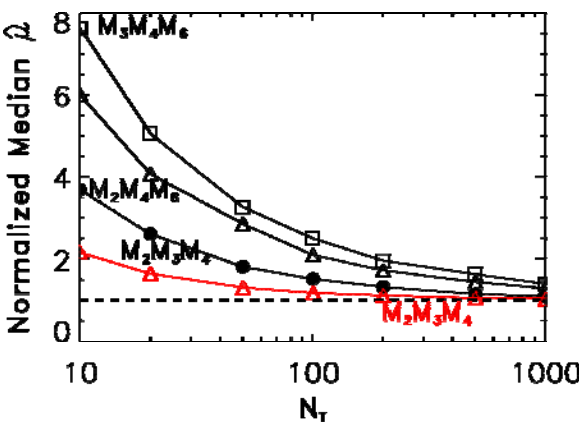

Figure 5: Variation of normalized estimated median value of gamma RSD shape parameter (population μ = 2, in black), as estimated from the indicated sets of three sample moments (dots-$M_2M_3M_4$, triangles-$M_2M_4M_6$, squares-$M_3M_4M_6$) with mean sample size N_T. Corresponding values for population shape parameter μ = 5 in red, for the case $M_2M_3M_4$. Population RSD: gamma; horizontal dashed line indicates the normalized population value.

Based on the results shown for the moment estimators for gamma RSD, we can conclude that:

(a) the moment estimators are biased:
- the shape parameter μ tends to be <u>overestimated</u>,
- the scale parameter λ tends to be <u>overestimated</u>,
- the mass weighted mean diameter D_m tends to be <u>underestimated</u>,
- the concentration normalized with respect to LWC, N_W , tends to be <u>overestimated</u>,
- the total number concentration N_T tends to be <u>underestimated</u>,

(b) the bias in the moment estimators decreases with increasing shape parameter of the population gamma distribution,
(c) the bias in moment estimators decreases with increasing sample size,
(d) the case of missing small drops seems to have little effect on these estimators.

3.2 The L-moment Method Applied to Gamma RSD

The L-moment method has apparently not been used by radar meteorologists and cloud modelers until recently, when a first attempt to use this method was done by Kliche *et al*. (2006). We also applied this method for all the computer-generated raindrop spectra. In a nutshell, the L-moments can be explained as follows: consider estimating m parameters p_1, p_2,…,p_m of a probability density function. Using the method of moments, we set

$$\frac{1}{C}\sum_{k=1}^{C} D_k^i = E\left(\frac{1}{C}\sum_{k=1}^{C} D_k^i\right) = h_i\left(p_1,\ldots,\ p_m\right)$$

for m values of i, say i = 1, 2, …, m to get m equations in m unknowns (C is the number of drops in the sample). By way of contrast, L-moments procedures set for the L-moments l_i

$$\hat{l}_i = E\left(\hat{l}_i\right) = g_i\left(p_1,\ldots,p_m\right)$$

for i = 1, 2, …, m to get m equations in m unknowns. *What makes the L-moments procedure attractive is that the (sample) L-moments* $\hat{l}_1,\hat{l}_2,\hat{l}_3,\ldots$ *are always* <u>*linear*</u> *in the observations.*

For example, for two parameters we need only the first two values,

$$\hat{l}_1 = b_0$$
$$\hat{l}_2 = 2b_1 - b_0 \tag{11}$$

where

$$b_0 = \frac{1}{C}\sum_{k=1}^{C} D_k = \overline{D}, \quad b_1 = \frac{1}{C(C-1)}\sum_{k=1}^{C-1} kD_{(k+1)} \tag{12}$$

are the probability-weighted sample moments: b_0 is the sample mean; b_1 is a measure of the dispersion of the data values about their mean. Before calculating b_1, we first order the observations as $D_{(1)} < D_{(2)} < \ldots < D_{(C)}$. Then, the ratio of the two estimated L-moments is given by

$$\hat{l}_2 / \hat{l}_1 = (2b_1 - b_0)/b_0 \tag{13}$$

from which an iterative procedure (Hosking 1990) yields estimates for the parameters.

Intuitively, one can see that the L-moments should outperform the method of moments procedure when two or more parameters are being estimated. When using the method of moments, observations, and in particular outliers, are being raised to powers greater than one, magnifying their importance in the obtained sample. There is no such “inflation” of observations when estimating using L-moments.

The L-moment method was applied to the gamma RSDs. In the case of the gamma distribution, to estimate the shape parameter $\hat{\mu}$ using this method we solve an equation by iteration as described by Hosking (1990) and Bowman and Shenton (1988), and as shown in Kliche *et al.* (2006). In the case of the gamma distribution given in (5), the work of Hosking (1990) establishes that

$$\frac{l_2}{l_1}\frac{1}{\sqrt{\pi}}\frac{\Gamma(\mu + 3/2)}{\Gamma(\mu + 2)} \tag{14}$$

where Γ is the gamma function and l_1, l_2 are the first two L-moments as defined above. Equations (11) give the estimated first and second L-moments (which are unbiased estimates of l_1, l_2) expressed in terms of the probability-weighted sample moments b_0, b_1. This suggests estimating μ by $\hat{\mu}_L$ (the L-moments estimate of μ) which satisfies

$$\frac{\hat{l}_2}{\hat{l}_1} = \frac{(2b_1 - b_0)}{b_0} = \frac{1}{\sqrt{\pi}}\frac{\Gamma(\hat{\mu}_L + 3/2)}{\Gamma(\hat{\mu}_L + 2)} \tag{15}$$

Using (12),

$$\frac{2b_1 - b_0}{b_0} = \frac{2}{C(C-1)\overline{D}}\sum_{k=1}^{C-1} kD_{k+1} - 1 =$$
$$\frac{2}{C(C-1)\overline{y}}\sum_{k=1}^{C-1} ky_{k+1} - 1 \tag{16}$$

The left side of (15) can be calculated using equation (16); the middle form is appropriate for experimental data, while the last form is used in our simulations. The shape parameter estimate is then calculated from (15) by iteration using recursion. Once the shape parameter is determined, the estimator for the scale parameter is calculated from

$$\hat{\lambda}_L = \frac{\hat{\mu}_L + 1}{\hat{\ell}_1} = \frac{\hat{\mu}_L + 1}{\overline{D}} \tag{17}$$

The estimated gamma PDF must obey the relationship $\hat{\lambda}_L = (\hat{\mu}_L + 4)/(\hat{D}_m)_L$, so

$$(\hat{D}_m)_L = (\hat{\mu}_L + 4)\overline{D}/(\hat{\mu}_L + 1) \tag{18}$$

In the simulations, we obtain normalized estimates of these dimensional parameters by dividing by the respective population values and using $\lambda = (\mu + 4)/D_m$:

$$\frac{\hat{\lambda}_L}{\lambda} = \frac{\hat{\mu}_L + 1}{\mu + 4}\frac{D_m}{\overline{D}} = \frac{\hat{\mu}_L + 1}{(\mu + 4)\overline{y}} \tag{19}$$

$$\frac{(\hat{D}_m)_L}{D_m} = \frac{\hat{\mu}_L + 4}{\hat{\mu}_L + 1}\overline{y} \tag{20}$$

Since there is no L-moment estimator for the mean total number concentration parameter, we used the ML estimator for $\hat{N}_T$, namely the sample size C.

Figure 6 (left) shows the moment estimator for shape parameter μ and (right) the L-moment shape parameter estimator for a gamma population with μ = 2 and the mean sample size of N_T = 50. The original sample case is shown in blue and the censored sample case is shown in red. Figure 6 demonstrates that while the moment estimator tends to overestimate μ, the L-moment estimator gives results close to the true population value in the case of original samples. The situation, however, changes dramatically when the case of censored samples is considered: the L-moment estimator is substantially overestimating the population value.

Figure 7 shows the corresponding histograms for the moments (left) and the L-moment (right) shape parameter estimator for the gamma population having μ = 5. The censored samples have less effect when the population RSD has a shape parameter μ = 5 than in the case of μ = 2. This is due to the fact that for μ = 5 the distribution is narrower and shifted more to the right, which implies fewer small drops in this RSD.

Table 3 gives the normalized estimated mean values for the cases shown in Figures 6 and 7, and the corresponding values for the normalized mass-weighted mean diameter and scale parameter. The values for censored samples are shown in parentheses.

Table 4 shows the variation of L-moment estimated shape parameter for the gamma RSDs studied as a function of sample size. The bias decreases with increasing size of the sample, and becomes negligible for samples having $N_T \geq 100$. The L-moments biases are smaller, and less sensitive to missing small drops.

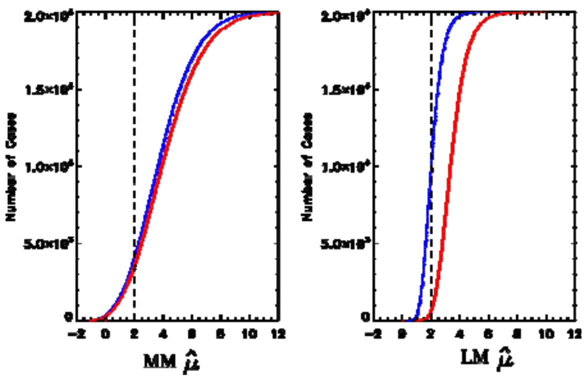

Figure 6: (left) Cumulative histograms of estimated shape parameter using the $M_2M_3M_4$ set: original sample values are shown in blue (mean = 3.8), and the censored sample values are shown in red (mean = 4.1). (right) Cumulative histograms of estimated shape parameter using L-moments: in blue (mean = 2.10) for original samples, and in red for censored samples (mean = 3.45). Population RSD: gamma, μ = 2, N_T = 50. Vertical dashed line indicates the population value.

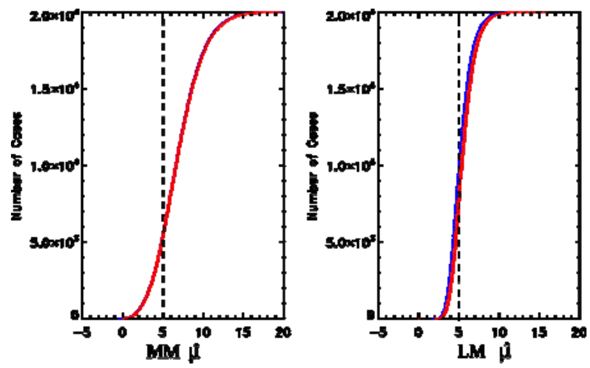

Figure 7: (left) Cumulative histograms of estimated shape parameter using the $M_2M_3M_4$ set: original sample values are shown in blue (mean = 6.77), and the censored sample values are shown in red (mean = 6.81). (right) Cumulative histograms of estimated shape parameter using L-moments: in blue (mean = 5.19) for original samples, and in red for censored samples (mean = 5.48). Population RSD: gamma, μ = 5, N_T = 50. Vertical dashed line indicates the population value.

Table 3. Comparison of moment and L-moment normalized mean estimator values for two gamma RSDs (N_T = 50).Values in parentheses correspond to censored samples.

Parameter	$\mu = 2$		$\mu = 5$	
	$M_2M_3M_4$	*L-moment*	$M_2M_3M_4$	*L-moment*
$\hat{\mu} / \mu$	1.9 (2.05)	1.05 (1.72)	1.35 (1.36)	1.04 (1.1)
$\hat{D}_m / D_m$	0.96 (0.96)	1.01 (0.94)	0.98 (0.98)	1.01 (1.0)
$\hat{\lambda} / \lambda$	1.45 (1.50)	1.04 (1.36)	1.26 (1.26)	1.0 (1.08)

Table 4. L-moment estimated shape parameter mean, median and RMS error values for the two gamma RSDs as a function of the sample size. Values in parentheses correspond to censored samples.

N_T	$\mu = 2$			$\mu = 5$		
	Mean	Median	RMS Error	Mean	Median	RMS error
10	2.26 (3.76)	2.01 (3.35)	1.32 (2.63)	6.51 (6.84)	5.11 (6.35)	6.18 (5.71)
20	2.26 (3.76)	2.01 (3.35)	1.32 (2.63)	5.52 (5.84)	5.01 (5.32)	2.57 (2.73)
50	2.10 (3.45)	2.01 (3.32)	0.68 (1.76)	5.20 (5.49)	5.01 (5.30)	1.36 (1.47)
100	2.04 (3.36)	2.00 (3.29)	0.45 (1.51)	5.09 (5.38)	5.00 (5.30)	0.91 (1.01)
200	2.03 (3.34)	2.01 (3.32)	0.31 (1.42)	5.05 (5.33)	5.03 (5.30)	0.61 (0.71)
500	2.01 (3.30)	2.00 (3.28)	0.19 (1.33)	5.01 (5.29)	5.00 (5.28)	0.39 (0.49)
1000	2.00 (3.29)	2.00 (3.29)	0.14 (1.31)	5.01 (5.29)	5.00 (5.28)	0.27 (0.40)

These results demonstrate that

a) the L-moment estimators are biased, but their bias is smaller than the bias of the moment estimators (and also, from comparison with Table 5, smaller than that of the ML estimators):
   - the L-moment shape parameter is slightly overestimated;
   - the L-moment mass-weighted mean diameter and scale parameter are also slightly overestimated;

b) the L-moment estimator's bias decreases with increasing shape parameter for the population RSD;

c) the L-moment estimator's bias decreases with increasing sample size;

d) the case of missing small drops affects the L-moment estimators by showing an increase in the bias of the estimators, but to a lesser extent than the ML estimators.

3.3 The Maximum Likelihood (ML) Method

Another approach with experimental RSD data would be to use the *maximum likelihood* (ML) method to estimate the parameters for the RSDs. The likelihood function represents a fundamental concept in statistical inference, and it indicates how likely a particular population is to produce an observed sample. Mathematically speaking, ML estimators are expected to be asymptotically unbiased. Thom (1958) included the approximate solutions of the maximum likelihood equations for the gamma distribution using an asymptotic approach. Choi and Wette (1969) published a numerical technique using the maximum likelihood method to estimate the parameters of the gamma distribution and did the first published test of the bias of the ML estimates. Mielke (1976) introduced a rapidly converging iterative procedure to determine the ML parameter estimates for the gamma distribution. The ML method advocated by Haddad *et al.* (1996, 1997) should provide more accurate estimates of the RSD parameters than the moment estimators, even though the ML estimators have some bias (Choi and Wette, 1969). Nevertheless, its use by the radar community and cloud modelers seems to be quite limited (Haddad *et al.*, 1996, 1997; Cho *et al.*, 2004; Smith *et al.*, 2005; Kliche *et al.*, 2006).

The ML method was applied for the two gamma populations of μ = 2 and μ = 5, for various sample sizes. Appendix B gives the steps and the equations used in estimating the gamma parameters. Table 5 shows the ML estimated shape parameter mean, median and RMS error values for the two gamma RSDs studied as a function of sample size. The ML estimator shows some bias for small samples, but the bias diminishes with increasing sample size to become essentially negligible for large samples (N_T > 100). In the case of censored samples, however, the ML estimator values have large biases which decrease little with increasing sample size. This effect is less serious for the narrower distribution (μ = 5) containing fewer very small drops.

Based on the results shown in Table 5 and others not shown here, we can conclude that:

a) the ML estimators have some biases, with complete samples
   - the shape parameter is slightly overestimated;
   - the scale parameter is slightly overestimated.
   - the mass-weighted mean diameter is slightly underestimated;

b) the ML estimator's bias decreases with increasing population RSD shape parameter, and with increasing sample size;

c) in the case of full samples, the bias of the ML estimators is smaller than the bias of the moment estimators, but a bit greater than that of the L-moment estimators.

d) the case of missing small drops seems to strongly affect the ML estimators, showing a marked increase in the bias of the estimators.

Table 5. ML estimated shape parameter mean, median and RMS error values for the two gamma RSDs as a function of the sample size. Values in parentheses correspond to censored samples.

N_T	$\mu = 2$			$\mu = 5$		
	Mean	Median	RMS Error	Mean	Median	RMS Error
10	3.43 (6.13)	2.56 (4.69)	4.25 (6.88)	7.94 (8.26)	6.19 (6.62)	7.93 (7.01)
20	2.53 (4.57)	2.26 (4.13)	1.45 (3.34)	6.04 (6.52)	5.48 (5.95)	2.87 (3.12)
50	2.18 (3.97)	2.09 (3.84)	0.67 (2.20)	5.37 (5.82)	5.19 (5.63)	1.40 (1.59)
100	2.08 (3.80)	2.04 (3.74)	0.43 (1.91)	5.17 (5.62)	5.08 (5.54)	0.90 (1.09)
200	2.04 (3.75)	2.02 (3.73)	0.30 (1.81)	5.09 (5.54)	5.07 (5.50)	0.60 (0.81)
500	2.01 (3.69)	2.01 (3.67)	0.18 (1.71)	5.02 (5.47)	5.01 (5.45)	0.37 (0.60)
1000	2.00 (3.67)	2.00 (3.67)	0.13 (1.68)	5.01 (5.50)	5.01 (5.45)	0.26 (0.56)

4. COMPARISON OF ESTIMATORS

4.1 Complete Samples (samples have the full range of drop sizes)

Cumulative histograms of the estimated shape parameter for original samples using the moment, ML and L-moment fitting methods for the case of μ = 2 and small sample sizes (N_T = 10) are shown in Figure 8. The moment estimator gives the greatest bias (overestimation); the ML estimator is also biased, but not as much as the moment estimator; the L-moment estimator gives the smallest bias of the three.

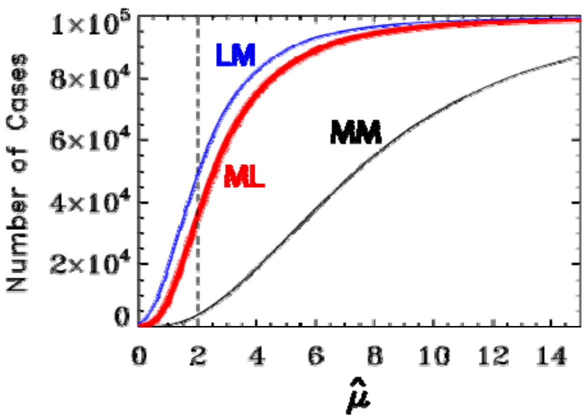

Figure 8: Cumulative histograms of the estimated gamma shape parameter using the moment, ML and L-moment methods in the case of small sample sizes: moments in black (mean = 8.98); ML in red (mean = 3.43); L-moment in blue (mean = 2.73). Population RSD: gamma, μ = 2, N_T = 10. Vertical dashed line indicates the population value.

Figure 9 shows the cumulative histograms of estimated shape parameter for complete samples using the three fitting methods in the case of medium sample sizes, N_T = 50. The moment method still gives the greatest bias (overestimation); ML and L-moment estimators give the closest results to the population value, both slightly overestimating. The L-moment method again gives the best estimation.

Figure 10 shows the same graph as in Figure 9, but in the case of a gamma distribution having μ = 5. From comparing Figures 9 and 10, it is obvious that the results for the two gamma distributions studied are comparable. Table 6 gives examples of the mean estimated shape parameter values for the three fitting procedures in the case of the two gamma distributions studied. Normalized values different from 1 indicate biased estimates.

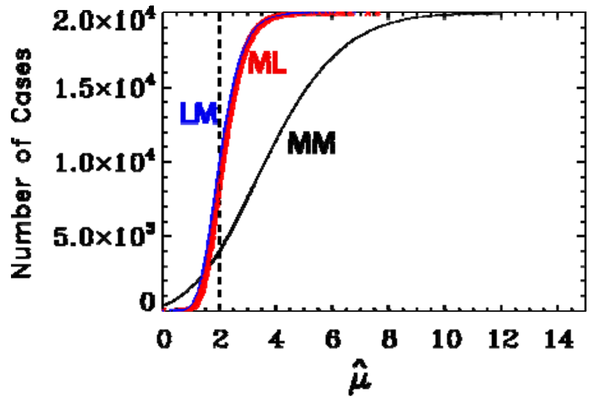

Figure 9: Cumulative histograms of the estimated gamma shape parameter using the moment, ML and L-moment methods in the case of medium sample sizes: moments in black (mean = 3.80); ML in red (mean = 2.18); L-moment in blue (mean = 2.09). Population RSD: gamma, μ = 2, N_T = 50. Vertical dashed line indicates the population value.

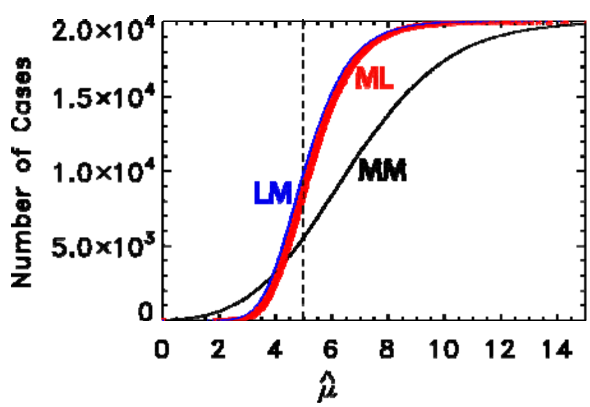

Figure 10: Cumulative histograms of the estimated gamma shape parameter using the moment, ML and L-moment methods in the case of medium sample sizes: moments in black (mean = 6.77); ML in red (mean = 5.37); L-moment in blue (mean = 5.19). Population RSD: gamma, μ = 5, N_T = 50. Vertical dashed line indicates the population value.

Table 6. Mean estimated shape parameter values for the gamma RSDs studied (N_T = 50).

Estimated parameter	$\mu = 2$		$\mu = 5$	
	Mean Value	Normalized Mean	Mean Value	Normalized Mean
Moment $\hat{\mu}$	3.8	**1.9**	6.77	**1.35**
ML $\hat{\mu}$	2.18	**1.09**	5.37	**1.07**
L-moment $\hat{\mu}$	2.09	**1.05**	5.19	**1.04**

Comparing the normalized values of the means for the three estimators as shown in Table 6, one can see that L-moment and ML methods give superior results to moment estimators, with slightly better results given by the L-moment method in the case of medium sample sizes.

For large sample sizes (N_T = 1000), Figure 11 shows the cumulative histograms of the estimated shape parameter for the original samples. All three fitting methods give essentially correct results, with the moment method slightly overestimating. The L-moment and ML estimators give the best estimation with smaller scatter.

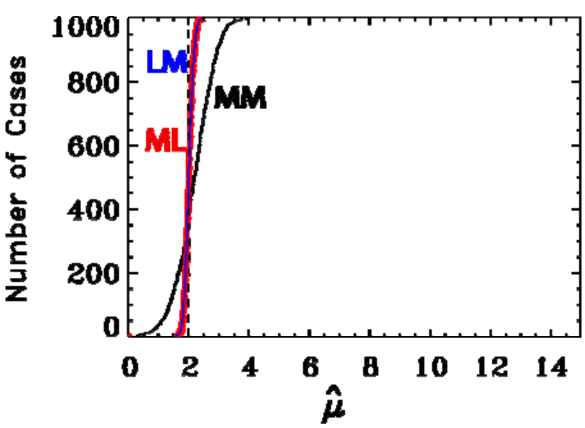

Figure 11: Cumulative histograms of the estimated gamma shape parameter in the case of large sample size: moments in black (mean = 2.17); ML in red (mean = 2.00); L-moments in blue (mean = 2.00). Population RSD: gamma, μ = 2, N_T = 1000. Vertical dashed line indicates the population value.

Figure 12, taken from Kliche *et al.* (2006), shows the variation of median value of the gamma RSD shape parameter with mean sample size for the three fitting methods analyzed. Superiority of the L-moments and ML methods is evident. One important feature shown is that the L-moment method gives superior results to the ML method for small sample sizes, and it is not much influenced by the sample sizes.

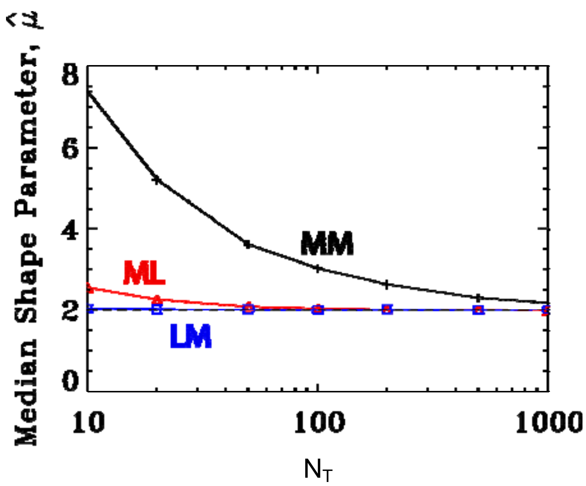

Figure 12: Variation of median value of gamma RSD estimated shape parameter using the method of moments (in black) for the case of $M_2M_3M_4$ combinations, the method of maximum likelihood (in red) and the method of L-moments (in blue). Population RSD: gamma, μ = 2, case of original samples. Horizontal dash line indicates population value.

In addition, Figure 13 shows the variation of the median value of the second parameter for the gamma distribution, i.e., mass-weighted mean diameter $\hat{D}_m$, with N_T, for the case of the three fitting methods analyzed. Several features are significant in this figure:

- The moment estimated mass-weighted mean diameter $\hat{D}_m$ has a strong tendency to underestimate; this tendency decreases with the increase of the sample size.

- L-moment and ML estimators also tend to underestimate for small sample sizes, but the bias decreases rapidly for medium and large sample sizes ($N_T \geq 50$).
- The L-moment estimator gives the best results and seems to not be affected as much as the ML and moment estimators by the sample size.

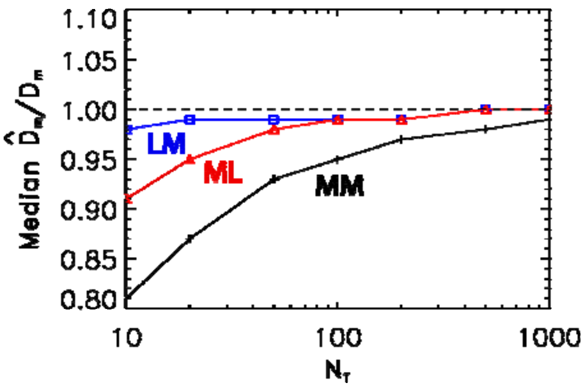

Figure 13: Variation of median value of gamma RSD normalized estimated mass-weighted mean diameter using the method of moments (in black) for the case of $M_2M_3M_4$ combinations, the method of maximum likelihood (in red) and the method of L-moments (in blue). Population RSD: gamma, μ = 2, case of complete samples. Horizontal dash line indicates population value.

4.2 Censored samples (small drops missing in the samples)

The case of missing small drops from the samples was also investigated. The same minimum drop-size threshold (Equation 4) was applied to all original samples and smaller drops were discarded. In this case, a comparison of the gamma shape parameter estimators for medium sample sizes (N_T = 50) is shown in Figure 14. As noted earlier, the L-moment and ML estimators are strongly affected by the absence of small drops. All the estimators now tend to overestimate the shape parameter, but the moment method is still inferior. For the gamma distribution with μ = 5, a similar comparison is shown in Figure 15. Here the effect on the L-moment and ML estimators is less pronounced, because there are fewer small drops in the narrower distribution.

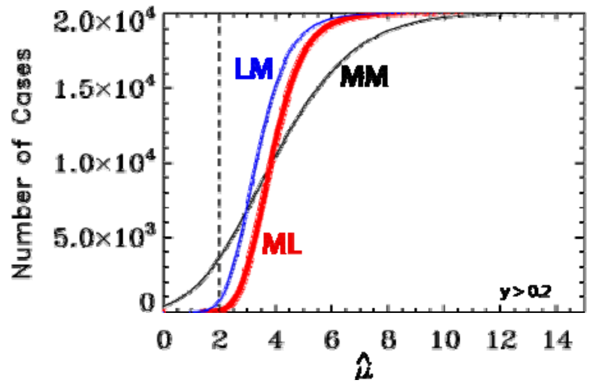

Figure 14: Cumulative histograms of the gamma shape parameter estimators in the case of censored samples for medium sample size; moments in black (mean = 4.11); ML in red (mean = 3.97); L-moments in blue (mean = 3.45). Population RSD: gamma, μ = 2, N_T = 50. Vertical dash line indicates the population value.

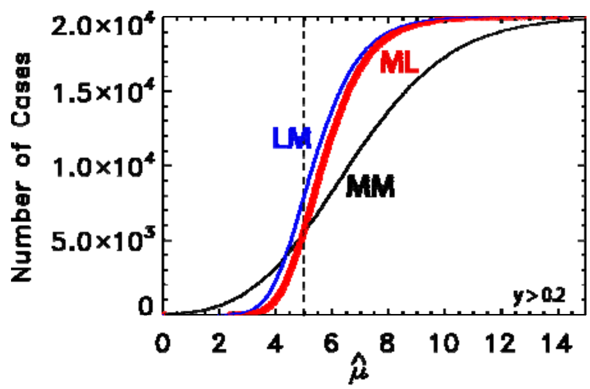

Figure 15: Cumulative histograms of the estimated gamma shape parameter in the case of censored samples: moments in black (mean = 6.81); ML in red (mean = 5.82); L-moments in blue (mean = 5.48). Population RSD: gamma, μ = 5, N_T = 50. Vertical dash line indicates the population value.

The mean shape parameter values for the three fitting methods are provided in Table 7 below for the sample case having N_T = 50. As shown by the normalized values in this table, all the estimators are biased in the case of censored samples. However, the largest bias is still given by the moment method, followed by ML and L-moments respectively. This bias decreases with increased value of the shape parameter of the raindrop population.

From these results one can conclude:

a) the L-moment and ML methods are very sensitive to small drops missing from the samples for the gamma μ = 2 distribution; the bias in this case is comparable to the bias in the moment method.

b) For the gamma μ = 5 distribution, which is narrower, none of the estimators are as much affected by the small drops missing; in this case both L-moments and ML give superior results to the moment method, with L-moments giving slightly better results than ML.

Figure 16 illustrates the changes for the median estimated shape parameter $\hat{\mu}$ with increasing sample size, while Figure 17 shows the variation of median estimated mass-weighted mean diameter $\hat{D}_m$ with increasing sample size. The mean values for the two parameters follow similar tendencies.

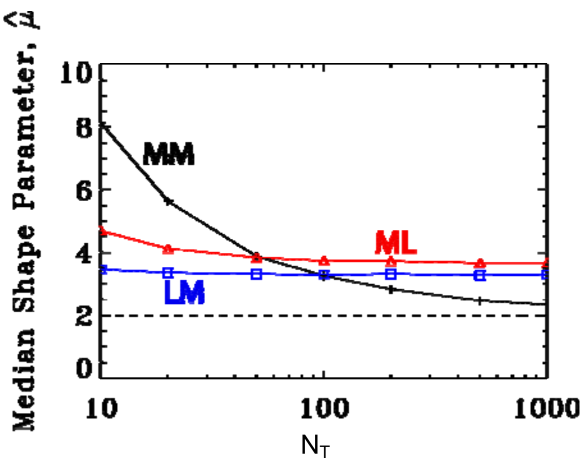

Figure 16: Variation of the median estimated shape parameter values using the method of moments (in black) for the case of $M_2M_3M_4$ combinations, the method of maximum likelihood (in red) and the method of L-moments (in blue). Population RSD: gamma, μ = 2, case of censored samples. Horizontal dash line indicates population value.

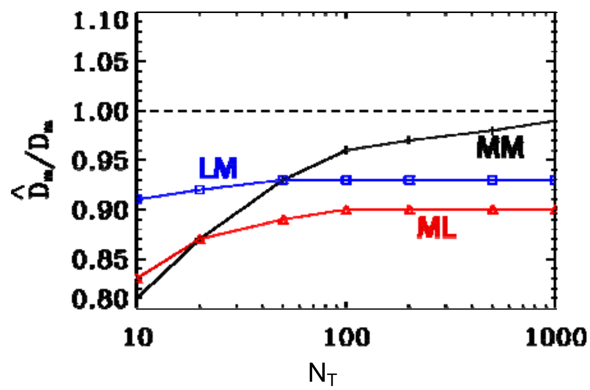

Figure 17: Variation of normalized median estimated mass-weighted mean diameter using the method of moments (in black) for the case of $M_2M_3M_4$ combinations, the method of maximum likelihood (in red) and the method of L-moments (in blue). Population RSD: gamma, μ = 2, case of censored samples. Horizontal dash line indicates population value.

Table 7. Mean estimated shape parameter values for the gamma RSDs studied in the case of censored samples and N_T = 50.

Estimated parameter	$\mu = 2$ ($y > 0.2$)		$\mu = 5$ ($y > 0.2$)	
	Mean Value	Normalized Mean	Mean Value	Normalized Mean
Moment $\hat{\mu}$	4.11	**2.05**	6.81	**1.36**
ML $\hat{\mu}$	3.96	**1.98**	5.82	**1.16**
L-moment $\hat{\mu}$	3.45	**1.73**	5.48	**1.10**

From both figures and Table 7 one can conclude that, in the case of censored samples,

- the bias in the L-moment and ML estimators is large and does not decrease much with increasing sample size;
- the moment method, although it has much stronger bias at small and medium sample sizes than the L-moment and ML, actually gives superior results as the sample size increases into hundreds of drops. This behavior for the case of the moment method is not surprising, since the moments used in the estimation of the gamma parameters ($M_2M_3M_4$) are not sensitive to very small drops. Also, as the sample size increases, the probability of having larger drops in the sample increases, and these larger drops are the ones that contribute the most to the higher moments used in such estimations ($M_2M_3M_4$).

5. CONCLUSIONS FOR THE GAMMA DISTRIBUTIONS

The main goal for the present work was to evaluate the biases and uncertainties in estimating the parameters of population raindrop size distribution functions from individual samples drawn from those populations using the L-moment method. The populations of raindrops from which samples are collected are unknown; therefore, it is hard to estimate which fitting method gives the best results based on direct disdrometer measurements. Therefore, the present attempt used the approach of building computer-simulated raindrop populations and then randomly extracting samples from these populations. Various fitting procedures were applied to estimate parameters for the two gamma distributions studied (μ = 2) and (μ = 5). The L-moments method was applied for the first time, and the results were compared to the ones obtained with the moment method and the maximum likelihood method. The resulting sampling statistics provided the basis for our comparison and estimation of biases and uncertainties.

The L-moment parameters for gamma distributions have the smallest bias of the three fitting methods studied (moments, ML, L-moments). In the case of the gamma RSDs, the L-moments estimator's bias is small and decreases with increasing shape parameter (narrower) for the population RSDs; it does not seem to be much affected by the sample size, outperforming the ML estimators for all cases studied.

In the case of censored samples, the L-moment method has strong sensitivity (comparable to the ML method) for all raindrop populations studied. The bias in the L-moment estimated parameters remains small for small samples, and is practically zero for larger sample sizes, in the case of the narrower RSDs, but the bias is substantial for all sample sizes in the case of the wider RSDs.

When samples covering the complete range of drop sizes are available, the best fitting method is the L-moment method, followed closely by the ML method, with moment methods being the last one on the performance scale. The bias is practically zero using the L-moment method for the original samples, and it is not dependent on the sample size. The moment approach is acceptable only when large sample sizes (order of 1000 drops or more) are available.

A different picture for the L-moment and ML estimators emerges when small drops are not included in the samples (censored samples). Our results show that both methods are very sensitive to the missing small drops, and that their results are not reliable in such applications, especially for the case of wider RSDs. This behavior does not change even with increasing size of the samples. However, in the case of some narrower RSDs, the L-moment and ML methods can still give results sufficiently close to the population values. Overall, in the case of censored samples, the moment method gives the best results for all RSDs studied, especially when large sample sizes are available.

These results are applied to real data collected with instrumentation on board the T-28 research aircraft and discussed in Kliche *et al.* (2007b).

Acknowledgement: This research work was partially supported by NSF Grant ATM-0531690.

APPENDIX A – Moment Method Estimators

a. Moments: M_2, M_3, M_4

$$\hat{\mu} = \frac{3m_4 m_2 - 4m_3^2}{m_3^2 - m_4 m_2}$$

$$\frac{\hat{\lambda}}{\lambda} = \frac{(\hat{\mu}+4)m_3}{(\mu+4)m_4}$$

or, in terms of D_m, $\frac{\hat{D}_m}{D_m} = \frac{m_4}{m_3}$ and

$$\hat{N}_T = \left(\frac{m_2^2}{m_4}\right)\frac{\hat{\alpha}}{(2-3\hat{\alpha})(1-2\hat{\alpha})} \quad ; \quad \hat{\alpha} = \frac{m_3^2}{m_4 m_2}.$$

or, in terms of N_W

$$\frac{\hat{N}_W}{N_W} = \frac{(\mu+4)^3}{N_T(\mu+1)(\mu+2)(\mu+3)}\frac{m_3^5}{m_4^4}$$

b. Moments: M_2, M_4, M_6

$$\hat{\mu} = \frac{7 - 11\hat{\eta} - (\hat{\eta}^2 + 14\hat{\eta} + 1)^{1/2}}{2(\hat{\eta}-1)} \quad \text{with}$$

$$\hat{\eta} = \frac{m_4^2}{m_2 m_6}$$

$$\frac{\hat{\lambda}}{\lambda} = \frac{1}{\mu+4}\left[\frac{(\hat{\mu}+3)(\hat{\mu}+4)m_2}{m_4}\right]^{1/2}$$

$$\frac{\hat{D}_m}{D_m} = \left[\frac{(\hat{\mu}+4)}{(\hat{\mu}+3)}\frac{m_4}{m_2}\right]^{1/2}$$

$$\hat{N}_T = \frac{(\hat{\mu}+3)(\hat{\mu}+4)}{(\hat{\mu}+1)(\hat{\mu}+2)}\frac{m_2^2}{m_4}$$

$$\frac{\hat{N}_w}{N_w} = \frac{(\mu+4)^3}{N_T(\mu+1)(\mu+2)(\mu+3)}\left\{\left[\frac{(\hat{\mu}+3)m_2}{(\hat{\mu}+4)}\right]^5 \frac{1}{m_4^3}\right\}^{1/2}$$

c. Moments: M_3, M_4, M_6

$$\hat{\mu} = \frac{11\hat{G} - 8 + \left[\hat{G}\left(\hat{G}+8\right)\right]^{1/2}}{2(1-\hat{G})} \quad \text{with}$$

$$\hat{G} = \frac{m_4^3}{m_3^2 m_6}$$

$$\frac{\hat{\lambda}}{\lambda} = \frac{(\hat{\mu}+4)}{(\mu+4)}\frac{m_3}{m_4}$$

$$\frac{\hat{D}_m}{D_m} = \frac{m_4}{m_3}$$

$$\hat{N}_T = \frac{(\hat{\mu}+4)^3}{(\hat{\mu}+1)(\hat{\mu}+2)(\hat{\mu}+3)}\frac{m_3^4}{m_4^3}$$

$$\frac{\hat{N}_w}{N_w} = \frac{(\mu+4)^3}{N_T(\mu+1)(\mu+2)(\mu+3)}\frac{m_3^5}{m_4^4}$$

APPENDIX B – Maximum Likelihood Method

With the general form of the gamma distribution given in equation (5), the gamma PDF can be represented as a two-parameter density function:

$$f(D;\mu,D_m) = \frac{(\mu+4)^{\mu+1}}{\Gamma(\mu+1)}\frac{D^{\mu}}{D_m^{\mu+1}} \times \exp\left[-(\mu+4)\frac{D}{D_m}\right] \quad (D>0) \qquad (21)$$

The ML method is described below as it is shown in Smith *et al.* (2005) and Kliche *et al.* (2006). If we consider the above gamma PDF and recall that $\lambda = (\mu+4)/D_m$, then the likelihood function is given by:

$$L(D_1, D_2, ..., D_C; \lambda, \mu) = \prod_{i=1}^{C} f(D_i) = \frac{\lambda^{C(\mu+1)}}{[\Gamma(\mu+1)]^C}\left[\prod_{i=1}^{C} D_i\right]^{\mu} \exp\left(-\lambda\sum_{i=1}^{C} D_i\right) \qquad (22)$$

where C is the number of drops in the sample.

First we calculate the natural log of the likelihood function:

$$\ln L(D;\lambda,\mu) = C(\mu+1)\ln\lambda - C\ln\Gamma(\mu+1) + \mu\ln\left[\prod_{i=1}^{C} D_i\right] - \lambda\sum_{i=1}^{C} D_i \qquad (23)$$

The scale parameter λ for the gamma RSD is estimated by taking the derivative $\frac{\partial \ln L}{\partial \lambda} = 0$, from which we obtain $\hat{\lambda} = \frac{C(\hat{\mu}_{ML}+1)}{\sum_{i=1}^{C} D_i} = \frac{\hat{\mu}_{ML}+1}{\overline{D}}$ where $\overline{D}$ is the arithmetic mean of the drop diameters.

Using $\lambda = (\mu + 4)/D_m$ to express the population value of λ, and taking note of (3) we get the normalized ML equation for the scale parameter:

$$\left(\frac{\hat{\lambda}}{\lambda}\right)_{ML} = \frac{(\hat{\mu}_{ML} + 1)}{\overline{y}(\mu + 4)} \quad (24)$$

Using $\hat{\lambda}_{ML} = (\hat{\mu}_{ML} + 4)/(\hat{D}_m)_{ML}$ we obtain the normalized mass-weighted mean diameter's equation:

$$\left(\frac{\hat{D}_m}{D_m}\right)_{ML} = \frac{(\hat{\mu}_{ML} + 4)}{(\hat{\mu}_{ML} + 1)}\overline{y} \quad (25)$$

Maximizing with respect to μ, $\frac{\partial \ln L}{\partial \mu} = 0$, the following expression for the shape parameter is obtained:

$$\ln(\hat{\mu}_{ML} + 1) - \psi(\hat{\mu}_{ML} + 1) = \ln\left[\frac{\overline{D}}{\left(\prod_{i=1}^{C} D_i\right)^{1/C}}\right] \quad (26)$$

where ψ is the "psi" or "digamma" function defined by $\psi(x) = \frac{\Gamma'(x)}{\Gamma(x)}$. The denominator on the right-hand side of (24) can be recognized as the geometric mean diameter. By substituting (3), equation (24) can be rewritten for the simulations as

$$\ln(\hat{\mu}_{ML} + 1) - \psi(\hat{\mu}_{ML} + 1) = \ln\left[\frac{\overline{y}}{\left(\prod_{i=1}^{C} y_i\right)^{1/C}}\right] \quad (27)$$

where y_i is the corresponding dimensionless size for each drop in the simulated sample and $\overline{y}$ is the arithmetic mean. The maximum likelihood estimate for the shape parameter is then obtained by solving (25) by iteration using recursion, as described in Bowman and Shenton (1988):

$$\alpha_{l+1} = \alpha_l \frac{\ln(\alpha_l) - \psi(\alpha_l)}{\ln\left[\overline{y}\Big/\left(\prod_{i=1}^{C} y_i\right)^{1/C}\right]} \quad (28)$$

with starting value $\alpha_1 = \frac{1 + \sqrt{1 + 4z/3}}{4z}$ where $z = \left[\overline{y} / \left(\prod_{i=1}^{C} y_i\right)^{1/C}\right]$ until the sequence converges to α. Then the maximum likelihood value for the shape parameter is given by $\hat{\mu}_{ML} = \alpha - 1$.

However, to implement the iterative procedure it is necessary to evaluate the digamma, or "psi", function in the equation (26). For this task we use a method proposed by Moody (1967), and follow the steps:

a. Apply the identity

$$\psi(x + 1) = \frac{1}{x} + \psi(x)$$

(repeatedly if necessary) to reduce the digamma calculation to computing $\psi(x + 1)$ with x between 0 and 1.

b. Then use

$$\psi(x + 1) \cong \frac{x}{x + 1} - \gamma + \frac{x^7}{2} + \sum_{i=1}^{6} c_i (x^i - x^7)$$

where

i	c_i
1	+0.64493313
2	-0.20203181
3	+0.08209433
4	-0.03591665
5	+0.01485925
6	-0.00472050

and $\gamma = 0.57721566490153286061...$ is Euler's constant. This approximation is good to within 1.3×10^{-8} for $0 \le x \le 1$.

The ML estimator of the third parameter for the gamma distribution, N_T, is given by the total number of drops in the sample, *C*.

REFERENCES

Bowman, K.O. and L.R. Shenton, 1988: *Properties of Estimators for the Gamma Distribution*, Marcel Dekker, New York, NY.

Bringi, V.N., and V. Chandrasekar, 2001: *Polarimetric Doppler Weather Radar: Principles and Applications*, Cambridge University Press, 636p.

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

Cho, H.K., K.P. Bowman and G.R. North, 2004: A comparison of gamma and lognormal distributions for characterizing satellite rain rates from tropical rainfall measuring mission. *J. Appl. Meteor.*, **43**, 1586-1597.

Choi, S.C., and R. Wette, 1969: Maximum likelihood estimation of the parameters of the gamma distribution and their bias. *Technometrics*, **11**, 683-690.

Goddard, J.W.F., and S.M. Cherry, 1984: Quantitative precipitation measurements with dual linear polarization radar. *Preprints, 22nd Conf. on Radar Meteorology*, Zurich, Switzerland, Amer. Meteor. Soc., 352-357.

Haddad, Z.S., S.L. Durden, and E. Im, 1996: Parameterizing the raindrop size distribution. *J. Appl. Meteor.*, **35**, 3-13.

Haddad, Z.S., D.A. Short, S.L. Durden, E. Im, S. Hensley, M.B. Grable, and R.A. Black, 1997: A new parameterization of the rain drop size distribution. *IEEE Transactions of Geosciences and Remote Sensing*, **35**, 532-539.

Hosking, J.R.M., 1990: L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *J. Roy. Statistical Soc.,* **B52**, 105-124.

Hosking, J.R.M. and J.R. Wallis, 1997: *Regional Frequency Analysis: An Approach Based on L-Moments*, Cambridge University Press.

Joss, J. and E. G. Gori, 1978: Shapes of raindrop size distributions. *J. Appl. Meteor.*, **17**, 1054-1061.

Kliche, D.V., P.L. Smith, R.W. Johnson, 2007a: L-Moment estimators as applied to exponential, gamma, and lognormal drop size distributions. *33rd Conf. on Radar Meteorology,* Cairns, Australia, Amer. Meteor. Soc.

Kliche, D.V., A.G. Detwiler, P.L. Smith, and R.W. Johnson, 2007b: L-Moment method applied to observed raindrop size distribution. *33rd Conf. on Radar Meteorology,* Cairns, Australia, Amer. Meteor. Soc.

Kliche, D.V., P.L. Smith and R. W. Johnson, 2006: Estimators for parameters of drop-size distribution functions: sampling from gamma distributions, *12th Conference on Cloud Physics*, Madison, WI, Amer. Meteor. Soc.

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.

Marshall, J. S. and W. McK. Palmer, 1948: The distribution of raindrops with size. *J. Meteor.*, **5**, 165-166.

Mielke, P.W. Jr., 1976: Simple iterative procedures for two-parameter gamma distribution maximum likelihood estimates. *J. Appl. Meteor.*, **15**, 181-183.

Moody, W., 1967: Approximations for the Psi (Digamma) Function. *Mathematics of Computation*, **21**, 112.

Rice, J.A., 1995: *Mathematical Statistics and Data Analysis*, Second Edition, Duxbury Press, 602p.

Robertson, C.A., and J.G. Fryer, 1970: The bias and accuracy of moment estimators. *Biometrika*, **57**, 57-65.

Seliga, T.A., and V.N. Bringi, 1976: Potential use of radar differential reflectivity measurements at orthogonal polarization for measuring precipitation. *J. Appl. Meteor.*, **15**, 69-75.

Smith, P.L., 1998: Raindrop size distributions: exponential or gamma – does it make a difference? *Preprint, Conf. of Cloud Physics and 14th Conf. on Planned & Inadvertent Weather Modification*, Everett, WA, Amer. Meteor. Soc., Boston, MA, 399-402.

Smith, P. L. and D. V. Kliche, 2005: The bias in moment estimators for parameters of drop-size distribution functions: sampling from exponential distributions. *J. Appl. Meteor.*, **44,** 1195-1205.

Smith, P. L., D. V. Kliche and R. W. Johnson, 2005: The bias in moment estimators for parameters of drop-size distribution functions: sampling from gamma distributions. *32nd Conf.*

on Radar Meteorology, Albuquerque, NM, Amer. Meteor. Soc.

Smith, P.L., D.V. Kliche, and R.W. Johnson, 2007: The bias in moment estimators for parameters of drop-size distribution functions: sampling from gamma distributions. Submitted to *J. Appl. Meteor.*

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

Thom, H.C.S., 1958: A note on the gamma distribution. *Monthly Weather Review*, **86**, 117-122.

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.

Uijlenhoet, R., J.M. Porra, D.S. Torres, and J. D. Creutin, 2006: Analytical solutions to sampling effects in drop size distribution measurements during stationary rainfall: estimation of bulk rainfall variables. *J. Hydrology,* **328**, 65-82.

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

Ulbrich, C. W., 1985: The effects of drop size distribution truncation on rainfall integral parameters and empirical relations. *J. Climate Appl. Meteor.,* **24**, 580-590.

Ulbrich, C.W., and D. Atlas, 1984: Assessment of the contribution of differential polarization to improved rainfall measurements. *Radio Sci.*, **19**, 49-57.

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

Wallis, J.R., 1974: Just a Moment! *Water Resources Research*, **10**, 211-219.

Wilks, D.S., 1995: *Statistical Methods in Atmospheric Sciences, An Introduction*. Academic Press, Inc., 465p.

Willis, P.T., 1984: Functional fits to some observed drop size distributions and parameterization of rain. *J. Appl. Meteor.*, **41**, 1648-1661.

Wong, R.K.W., and N. Chidambaram, 1985: Gamma size distribution and stochastic sampling errors. *J. Climate Appl. Meteor.*, **24**, 568-579.

Zhang, G., J. Vivekanandan, E.A. Brandes, R. Meneghini, and T. Kozu, 2003: The shape–slope relation in observed gamma raindrop size distributions: statistical error or useful information? *J. Atmos. Oceanic Technol.*, **20**, 1106-1119.