The Bias in Moment Estimators for Parameters of Drop Size Distribution Functions: Sampling from Exponential Distributions

PAUL L. SMITH AND DONNA V. KLICHE

Institute of Atmospheric Sciences, South Dakota School of Mines and Technology, Rapid City, South Dakota

(Manuscript received 25 August 2004, in final form 27 December 2004)

ABSTRACT

The moment estimators frequently used to estimate parameters for drop size distribution (DSD) functions being "fitted" to observed raindrop size distributions are biased. Consequently, the fitted functions often do not represent well either the raindrop samples or the underlying populations from which the samples were taken. Monte Carlo simulations of the process of sampling from a known exponential DSD, followed by the application of a variety of moment estimators, demonstrate this bias. Skewness in the sampling distributions of the DSD moments is the root cause of this bias, and this skewness increases with the order of the moment. As a result, the bias is stronger when higher-order moments are used in the procedures. Correlations of the sample moments with the size of the largest drop in a sample ($D_{\max}$) lead to correlations of the estimated parameters with $D_{\max}$, and, in turn, to spurious correlations between the parameters. These things can lead to erroneous inferences about characteristics of the raindrop populations that are being sampled. The bias, and the correlations, diminish as the sample size increases, so that with large samples the moment estimators may become sufficiently accurate for many purposes.

1. Introduction

Investigators frequently acquire observations of raindrop sizes and seek to describe the drop size distributions (DSDs) of the underlying populations from which the samples were taken by analytical expressions, with the exponential or gamma function being most common. Although moment methods to estimate parameters for the DSD functions have become more or less traditional, Haddad et al. (1996, 1997) pointed out that such methods are biased. Hydrologists have long been aware of this bias (e.g., Wallis et al. 1974; Maidment 1993), and Smith and Kliche (2003) provided examples of the bias for the case of raindrop observations. Nevertheless, the intuitive appeal of the moment approach seems almost irresistible, and the associated mathematical manipulations lend a convincing aura. Yet the methods are indeed biased, in the statistical sense that the expected values of the "fitted" parameters differ from the parameters of the underlying raindrop populations, and so can lead to erroneous inferences about the characteristics of the DSDs being sampled.

Corresponding author address: Dr. Paul L. Smith, IAS, South Dakota School of Mines and Technology, Rapid City, SD 57701. E-mail: paul.smith@sdsmt.edu

The bias in the moment methods can be demonstrated by testing their ability to recover parameters of known DSDs from which samples are taken. This must be done by computer simulation, because DSDs in nature are inherently unknown. The simulations herein use a Monte Carlo simulation procedure similar to that described in Smith et al. (1993) and outlined below. This paper gives results for samples taken from a hypothetical exponential DSD fitted with various moment-based procedures; corresponding results for samples taken from gamma DSDs will be presented later.

2. Simulation of raindrop sampling

We seek to determine the sampling distributions of DSD parameters that are fitted to raindrop observations by moment methods. This is done by simulating repetitive sampling from a specified population DSD, which is exponential in form for the present paper. An exponential DSD is usually expressed in the form

$$n(D) = n_0 \exp(-\Lambda D), \tag{1}$$

where $n(D)$ is the number concentration of drops of diameter D per unit size interval, and n_0 and Λ are concentration and size parameters, respectively, of the

DSD. On conventional semilogarithmic plots, as introduced by Marshall and Palmer (1948), $n(D)$ becomes a straight line with intercept n_0 and slope $-\Lambda$.

For the purposes of these simulations, (1) is more conveniently expressed in terms of the total drop number concentration N_T and the mass-weighted mean diameter D_m ($=4/\Lambda$) as

$$n(D) = (4N_T/D_m)\exp(-4D/D_m). \quad (2)$$

Defining a dimensionless size variable $y = D/D_m$, this becomes

$$n(y) = 4N_T\exp(-4y). \quad (3)$$

This can be recognized as the product of the mean number concentration N_T and the probability density function (PDF) of drop size $4\exp(-4y)$. For convenience herein, we also designate the mean sample size (number of drops) as N_T. This could be viewed as representing a volume-sampling instrument with a sample volume of 1 m^3 (independent of the drop size). However, a sample volume of α m^3 and a mean drop concentration N_T/α would lead to the same mean sample size and the same sampling statistics. Thus, with the drop sizes normalized to D_m, the results can be organized simply by the value of N_T.

The simulation proceeds from a selected value of N_T by first drawing from a Poisson distribution with a mean value N_T to determine the actual total number of drops C in a given sample. Then, C values of y drawn from the exponential PDF establish the (normalized) sizes of those drops. Normalized values for the six sample moments of M_{1S} through M_{6S} are next calculated for each sample, and then various moment-based calculations (discussed in section 4 and summarized in the appendix) are applied to estimate the DSD parameters. For purposes of the present work, we classified the drop sizes into intervals of $\Delta y = 0.02$, representing the size classification procedure that is common to drop-measuring instruments, and truncated the exponential PDF at $y = 3.0$. Repetition of the sampling and fitting process yields the desired distributions. We used about 1 000 000 drops (e.g., 50 000 samples with $N_T = 20$ and 5000 samples with $N_T = 200$) in the simulations; because the probability of a drop in an exponential PDF being larger than $y = 3.0$ is 6×10^{-6}, we are lacking a total of about six larger drops from a full exponential DSD.

3. Characteristics of sampling distributions

Before considering the moment-based fits, we first examine some characteristics of the distributions of the sample moments $M_{iS} = \Sigma_c D^i$ [$i = 3$ gives the sample moment related to liquid water concentration (LWC); $i = 6$ gives radar reflectivity factor Z] themselves. Sampling from long-tailed DSDs, like the exponential, exhibits certain general features. Sample values of the moments are unbiased: the expected, or mean, value of M_{iS} corresponds to that moment of the drop population being sampled. However, the sampling distributions are skewed, as indicated by the fractional standard deviations calculated by Gertzman and Atlas (1977) and as shown in Smith et al. (1993). According to Gertzman and Atlas, the fractional standard deviation (FSD) σ_i/M_i for sample values of the DSD moment M_i as determined with a volume-sampling instrument studying an exponential DSD would be

$$\frac{\sigma_i}{M_i} = \frac{1}{\sqrt{V_s}}\frac{\sqrt{M_{2i}}}{M_i}, \quad (4)$$

where V_S is the instrument-sampling volume. The general form for the moments of an exponential DSD can be written

$$M_i = \frac{i!\,N_T D_m^i}{4^i}. \quad (5)$$

Thus,

$$\frac{\sigma_i}{M_i} = \frac{1}{\sqrt{N_T V_s}}\sqrt{\frac{(i+1)(i+2)\ldots(2i)}{i!}}. \quad (6)$$

The product $N_T V_S$ in the denominator of (6) is the mean sample size, which, for a sample volume of 1 m^3, would be numerically equal to N_T. Thus, for $i = 3$ (LWC) and $N_T = 100$, (6) gives $(\sigma_3/M_3) = 0.447$. The distribution of a positive quantity (M_{3S}) with an FSD this large is necessarily skewed.

Inspection of (6) shows that the skewness increases with the order i of the moment M_i, and decreases with increasing sample size N_T. Figure 1 illustrates the former property for the sample moments M_{3S} (LWC) and M_{6S} (Z), for the case $N_T = 100$. The general tendency is for the sample moments to be lower than the corresponding population values; this behavior is the ultimate cause of the bias in the moment methods for estimating DSD parameters. As shown in section 4, the increase of the skewness with the moment order leads to greater biases when higher moments are employed. Figure 2 illustrates how the median sample moments approach the population values as the sample size increases. The skewness of the sampling distributions, and the resulting biases, decrease in a similar manner.

Sampling the numerous small drops can be a major instrumental problem for exponential DSDs, and adequately sampling the relatively rare large drops is also

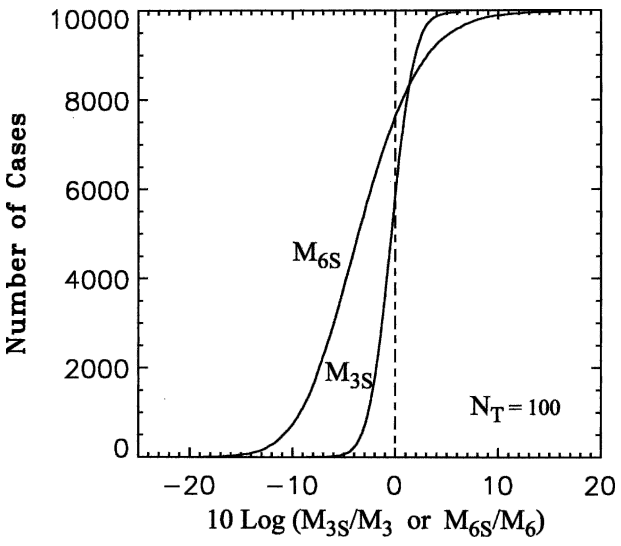

FIG. 1. Cumulative sampling distributions of the third sample moment M_{3S} (proportional to LWC) and sixth moment M_{6S} (proportional to Z); sample values are normalized with respect to the corresponding population value. Vertical dashed line denotes population value. Population DSD: exponential; mean sample size: 100 drops.

a concern. Fewer than one drop in 400 in an exponential DSD is larger than $D = 1.5D_m$, and only about one in 3000 is larger than $D = 2D_m$. However, the drops larger than $1.5D_m$ contribute more than 15% of the LWC and more than 60% of the reflectivity factor. Consequently, the relatively large but relatively rare drops tend to be important in determining the moments of physical significance. The sample values of these moments are, therefore, correlated with the size of the largest drop in each sample (e.g., Fig. 3). Table 1 demonstrates that these correlations are stronger for higher-order moments, and remain appreciable even for fairly large sample sizes. The sample moments are, in turn, correlated with each other—an artifact of the sampling variability as discussed in Smith et al. (1993).

Figure 5 of Smith et al. (1993) showed that the maximum drop size in an exponential DSD is rarely approached in samples of even hundreds of drops. The distribution of values along the abscissa in Fig. 3 here demonstrates the same thing. There is clearly no basis for assuming truncation of the underlying DSD at the maximum observed drop diameter, with samples of such sizes.

4. The bias in moment estimators

The essence of the moment approach for estimating parameters for DSD functions is to use the same number of moments calculated from observed raindrop size distributions as there are parameters in the function to be fitted. Analytical expressions for the selected moments of that function are solved algebraically for the needed parameters, and observed values of the sample moments are then entered into the resulting equations

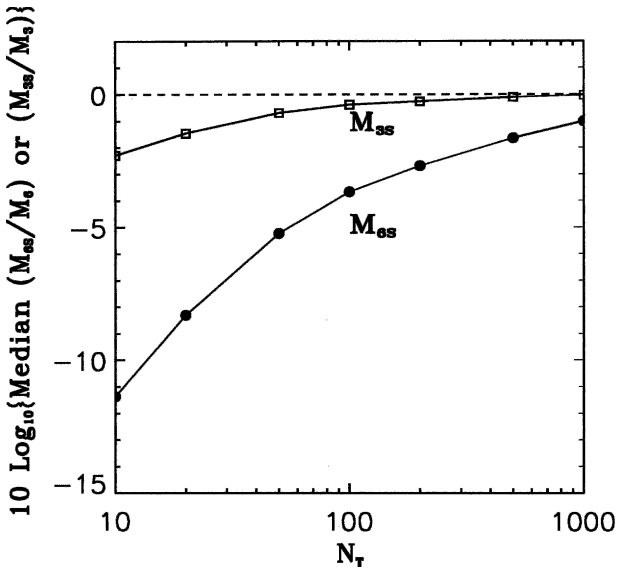

FIG. 2. Plot of median values of third sample moment M_{3S} (proportional to LWC) and sixth moment M_{6S} (proportional to Z) vs mean sample size; sample values are normalized with respect to the corresponding population value. Horizontal dashed line denotes population value. Population DSD: exponential.

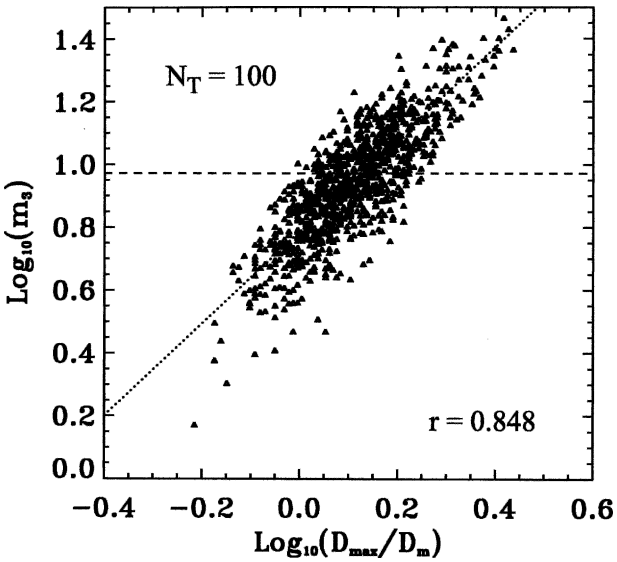

FIG. 3. Sampling distribution of normalized third moment (m_3, proportional to LWC) plotted against (normalized) maximum drop size in the sample. Population DSD: exponential; mean sample size: 100 drops. Dashed line indicates population value of m_3; dotted line shows regression relationship.

TABLE 1. Correlations between sample moments and maximum drop size in a sample. Population DSD: exponential.

Mean sample size (N_T)	Sample moment			
	M_{2S}	M_{3S}	M_{4S}	M_{6S}
10	0.890	0.968	0.985	0.995
20	0.839	0.939	0.970	0.990
50	0.740	0.888	0.944	0.980
100	0.634	0.828	0.912	0.968
200	0.558	0.775	0.880	0.956
500	0.435	0.676	0.820	0.933
1000	0.341	0.610	0.781	0.918

to estimate the parameters. The appendix summarizes the relevant mathematical expressions used here. The use of moment methods for rain DSDs evidently began with Waldvogel's (1974) paper on the "N_0 jump" of DSDs. He used observed values of moments M_{3S} and M_{6S} (i.e., LWC and Z) along with (5), or its equivalent, to determine pairs of parameters for exponential functions like (1) that purportedly represented the observed DSDs. However, most functions that are fitted in this way do not represent well either the samples upon which they are based or the underlying populations from which the samples were taken.

The introductory section pointed out that such moment estimators are biased, and the fact that estimates of DSD parameters obtained in this way are biased was actually indicated in Smith et al. (1993). Figure 4, reproduced from that paper, compares sample estimates of $\hat{D}_m$ with the value of D_m for the exponentially distributed population from which the (simulated) samples were drawn. In this example, more than 80% of the values of $\hat{D}_m$ are underestimates, and the mean (the expected value) is about 78% of the population value. In terms of the more familiar exponential slope parameter Λ ($=4/D_m$), this means that $\hat{\Lambda}$ is generally overestimated.

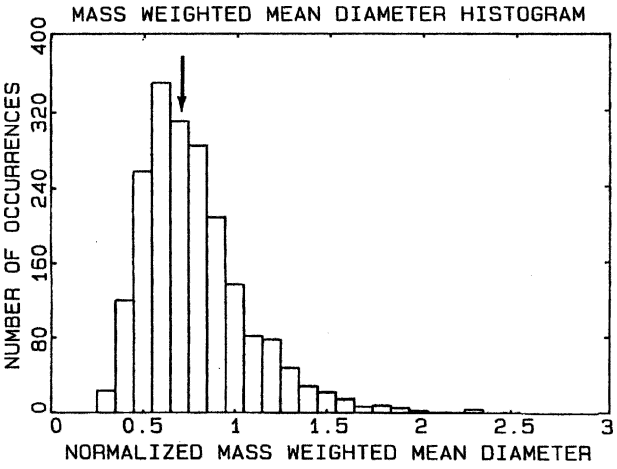

FIG. 4. Example of sampling distribution for (normalized) mass-weighted mean diameter $\hat{D}_m/D_m$. Population DSD: exponential; mean sample size: 20 drops. The population value falls at 1.0 on the abscissa; arrow indicates median sample value. [From Smith et al. 1993.]

a. Estimators for exponential functions

Figure 5, based on an "ideal sample" of 100 drops from an exponential DSD, illustrates the bias in the moment-method fits. This ideal sample, unlike the randomly drawn samples used elsewhere in this paper, was constructed from the cumulative PDF of the drop size using the systematic procedure described below. It provides as close a representation of the PDF as could be achieved with a sample of 100 drops. Construction of the sample began with the "inverse cumulative PDF" N_L, showing the number of drops of diameter D or larger:

$$N_L(D) = N_T \exp(-4D/D_m), \tag{7}$$

where N_T is the total number of drops in the sample. All 100 drops have $D > 0$, while (with the drop sizes quantized in intervals of $\Delta D/D_m = 0.02$) only 92 have $D/D_m \geq 0.02$. Thus, we assigned $D/D_m = 0.01$ to eight drops in the sample. Similarly, only 85 drops have $D/D_m \geq 0.04$, so $D/D_m = 0.03$ was assigned to the next

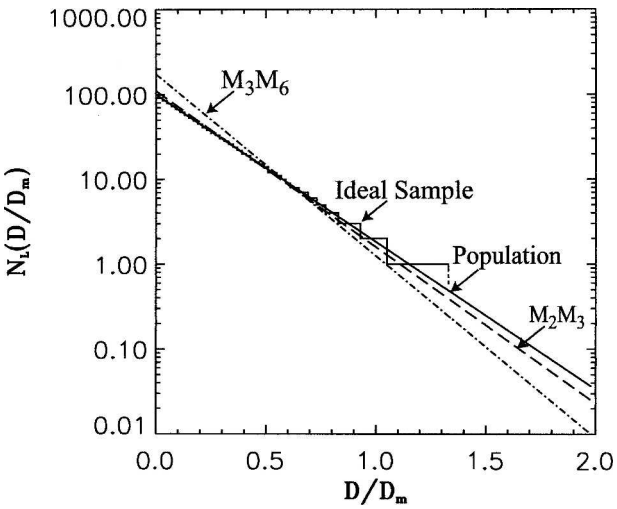

FIG. 5. Illustration of an "ideal sample" of 100 drops from an exponential DSD, along with two moment-based exponential fits to the sample; plot uses the "inverse cumulative" N_L format advocated in Smith (1982). Solid line represents underlying drop population, while stair-step plot represents the ideal sample constructed as described in the text. Dash–dot line represents Waldvogel fit based on moments M_3 and M_6; dashed line shows a similar fit based on M_2 and M_3.

7 drops in the sample, and so on. The stair-step plot in Fig. 5 represents the resulting sample.

The figure includes two moment-method fits to the sample for exponential DSD functions. The two are the "Waldvogel fit," based on moments M_3 and M_6, as employed in Waldvogel (1974), and one based on M_2 and M_3. The one using parameters based on M_3 and M_6 does not represent either the "observed" sample or the original population DSD very well; the smaller discrepancy resulting when the lower moments (M_2 and M_3) are used in the calculation is evident.

The foregoing discussion and the specific example in Fig. 5 suggest the general nature of the bias in moment estimators for parameters of exponential DSD functions: they tend to overestimate both the concentration (intercept) parameter n_o and the size (slope) parameter Λ. In terms of the parameters of (2), they tend to underestimate D_m and overestimate N_T, yielding fits having too many drops that are too small when compared with the original raindrop population. The biases are greater when higher moments are used in the procedure; Fig. 6 illustrates this behavior for n_0. Consequently, procedures that use sample values of reflectivity in the moment calculations lead to greater biases than ones employing only lower moments.

Extension of this argument would appear to suggest that using the lowest moments—M_{0S}, the sample size, and M_{1S}, related to the mean drop diameter—would yield the smallest biases of all. Such would indeed be the case in a simple mathematical exercise like that involved here. However, as noted in Testud et al. (2001), Smith (2003), and elsewhere, instrument responses to very small drops are highly variable and often suspect. Thus, trying to use moments that are lower than M_{2S} in the analysis would introduce another kind of uncertainty into the moment procedures. There may even be problems with using M_{2S}; instruments that do not adequately sense drops as small as $0.2–0.3D_m$ may significantly underestimate the value of M_2 (as illustrated in Table 2). If, say, $D_m = 1$ mm, this could be a challenging instrumentation problem. For that reason no approach using the lowest-order moments is considered here.

The skewness in the sampling distributions for the moments diminishes as the sample size increases (e.g., Fig. 2), so the bias in the moment estimators should also decrease with increasing sample size. Figure 7 shows that to be the case; with samples of hundreds or thousands of drops, the bias may become small enough to be negligible for many purposes.

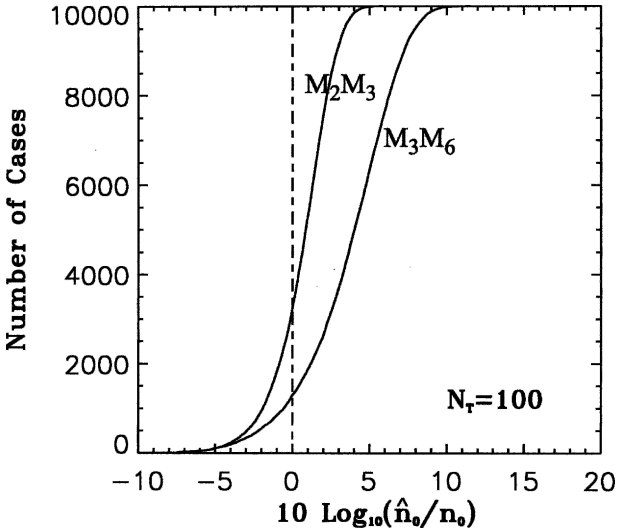

FIG. 6. Cumulative distributions of values of exponential concentration (intercept) parameter $\hat{n}_0$ as fitted to simulated samples using indicated pairs of sample moments; fitted values are normalized with respect to the population value of n_0. Vertical dashed line denotes population value. Population DSD: exponential; mean sample size: 100 drops.

b. *Estimators for gamma functions*

More recently, estimators involving three moments have been used in attempts to fit gamma distributions to DSD observations. Examples include Ulbrich (1983), Kozu and Nakamura (1991), Smith (1993), Tokay and Short (1996), and Ulbrich and Atlas (1998). While these procedures can produce closer fits to the observations, at least for part of the size spectrum, the bias that arises when moment methods are used with an assumed gamma form for the DSD is of even greater concern. A gamma DSD is often expressed in the form

$$n(D) = n_1 D^{\mu} \exp(-\lambda D), \tag{8}$$

with concentration parameter n_1, distribution shape parameter μ, and size (scale) parameter λ. An alternative form using the same parameters as (2) is

TABLE 2. Contributions of small drops to some population moments (exponential DSD).

Drops smaller than	Contribute this percentage To M_2	To M_3
$0.2D/D_m$	4.74	0.91
$0.25D/D_m$	8.03	1.90
$0.3D/D_m$	12.05	3.38
$0.35D/D_m$	16.65	5.37
$0.4D/D_m$	21.66	7.88
$0.45D/D_m$	26.94	10.87
$0.5D/D_m$	32.33	14.29

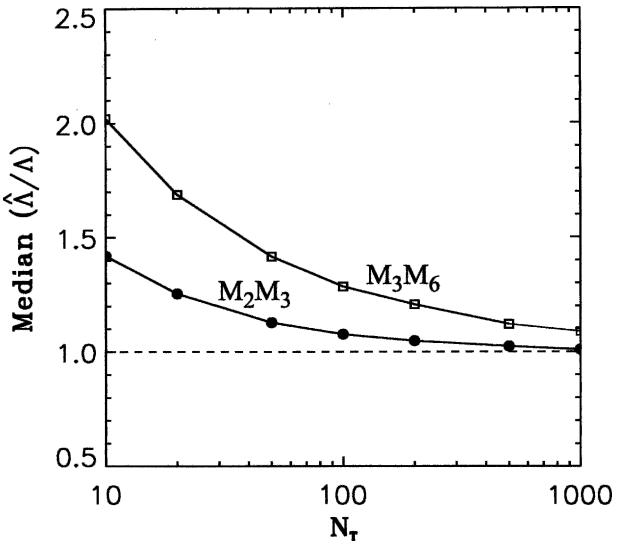

FIG. 7. Variation of the median value of exponential DSD size (slope) parameter $\hat{\Lambda}$, as estimated from the two sample moments indicated on the respective curves, with mean sample size. Ordinate indicates ratio of median sample estimate of $\hat{\Lambda}$ to population value; horizontal dashed line denotes population value. Population DSD: exponential.

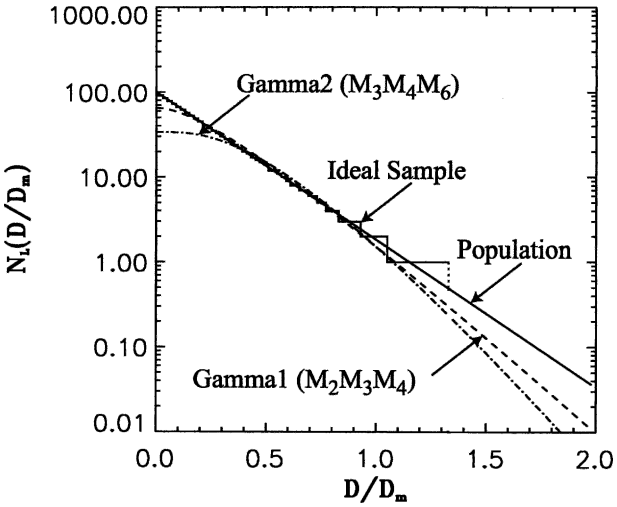

FIG. 8. As in Fig. 5, but showing two gamma-function fits, based on the indicated sets of moments, to the ideal sample of 100 drops from an exponential DSD. The gamma curves have shape parameters closely approximating those based on the indicated moments.

$$n(D) = N_T \frac{(\mu + 4)^{\mu+1}}{\mu!} \frac{D^\mu}{D_m^{\mu+1}} \exp[-(\mu + 4)D/D_m], \quad (9)$$

while a third form using the normalized concentration parameter N_W (Bringi and Chandrasekar 2001) is

$$n(D) = \frac{3}{128} N_W \frac{(\mu + 4)^{\mu+1}}{(\mu + 3)!} \left(\frac{D}{D_m}\right)^\mu \exp[-(\mu + 4)D/D_m]. \quad (10)$$

An exponential DSD corresponds to a gamma DSD with shape parameter $\mu = 0$; thus, in the present simulations, gamma functions fitted to the observations should have estimates of μ that are close to zero.

The example in Fig. 8, for the same ideal sample as that in Fig. 5, suggests that not to be the case. The curve for the gamma fit based on moments M_{2S}, M_{3S}, and M_{4S}, as suggested in Smith (2003), has a shape parameter $\mu = 1.0$; the actual calculated value of $\hat{\mu}$ was 1.013, but cumulants are not available for noninteger values of μ. Using the higher moments M_{3S}, M_{4S}, and M_{6S}, as done by several of the aforementioned authors, would yield $\hat{\mu} = 3.365$—a more strongly biased estimate. The second gamma curve in Fig. 8, for $\mu = 3.0$, provides a close approximation. The gamma fits match part of the sample distribution reasonably well, but in no way do they correspond to the population from which the sample was taken. As noted, the fitted value of the gamma M_2, M_3, M_4 shape parameter $\hat{\mu}$ is about 1.0; the expected value of $\hat{\mu}$ for samples of this size, drawn at random from an exponential DSD ($\mu = 0$), is $\hat{\mu} = 1.50$. For gamma fits based on moments M_3, M_4, M_6, the median value of $\hat{\mu}$ is about 4.5. Thus, sampling the exponential DSD with samples of this size and attempting to fit a gamma DSD function by moment methods will not reveal that the underlying DSD is exponential.

Figure 9 illustrates the bias in the moment estimators for μ, for different choices of the three moments used in the procedure. The general tendency is to overestimate μ. Here again, the lower moments yield smaller biases; use of the sample reflectivity values (M_{6S}) leads to the strongest bias among the examples illustrated. This bias can be quite misleading. Here, one has sampled from what is actually an exponential DSD, used moment methods in attempts to fit parameters of gamma distributions to the observations, found (biased) high values of $\hat{\mu}$, and probably concluded, quite erroneously, that the population DSD was gamma after all.

The hybrid approach used by Testud et al. (2001) does not employ a moment-based calculation to determine $\hat{\mu}$, but the estimators for their other gamma parameters (D_m and N_W) are biased. Figure 10 illustrates this bias for values of $\hat{D}_m$, which as noted earlier, tend to be underestimates. Figure 11 shows corresponding results for values of $\hat{N}_W$, which tend to be overestimates. (In both instances, when M_{3S} and M_{4S} are among the estimating moments, the third moment of the group has no effect on these estimates.)

As the sample size increases, the sampling variability

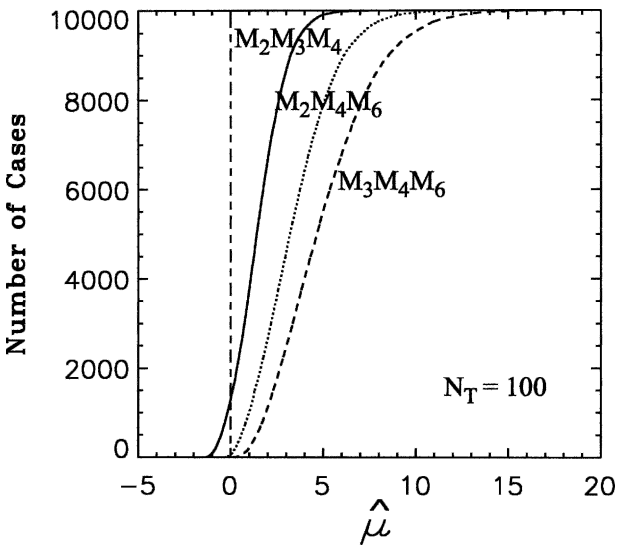

FIG. 9. Cumulative distributions of values of the gamma DSD shape parameter $\hat{\mu}$, as estimated from the three sample moments indicated for the respective curves. Population DSD: exponential, where the vertical dashed line indicates the population value $\mu = 0$. Mean sample size: 100 drops.

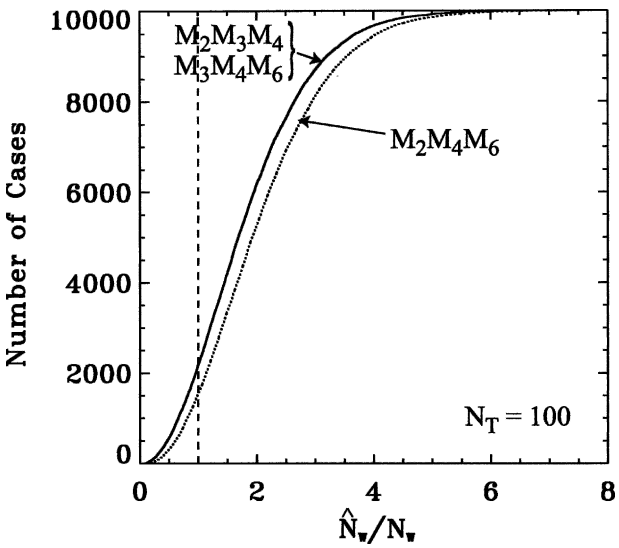

FIG. 11. As Fig. 10, but for estimated values of the normalized gamma DSD concentration parameter $\hat{N}_W$. Abscissa indicates the ratio of the sample estimate of $\hat{N}_W$ to the population value (which, for exponential DSD, is same as n_0); vertical dashed line denotes population value.

and the skewness in the sample moments decrease, and the correlations between the various sample moments weaken (Smith et al. 1993). Consequently, the biases diminish. Figure 12 illustrates this behavior for the shape parameter $\hat{\mu}$. Thus, with very large samples, the moment methods may give approximations of the population parameters that become sufficiently accurate for practical purposes, even when the wrong functional form is assumed.

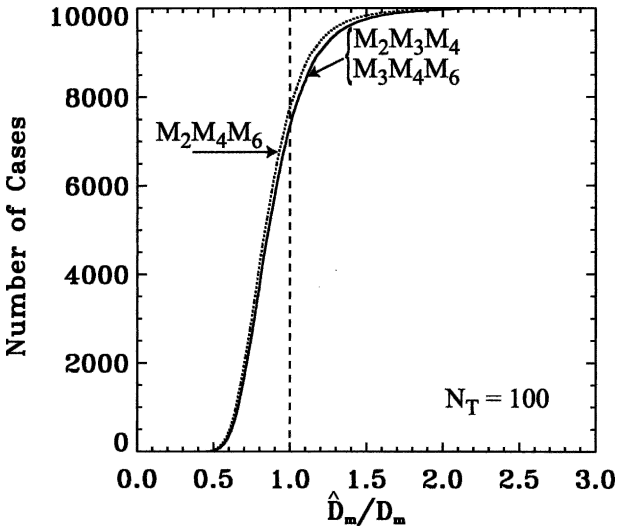

FIG. 10. Cumulative distributions of values of mass-weighted mean diameter $\hat{D}_m$, as estimated from the three sample moments indicated on the respective curves. (When M_3 and M_4 are among the three moments, the estimator is M_{4S}/M_{3S} and does not depend upon the third moment of the group.) Abscissa indicates ratio of estimate $\hat{D}_m$ to population value; vertical dashed line denotes population value. Population DSD: exponential; mean sample size: 100 drops.

5. Related findings

The correlations of the various sample moments with the maximum drop size in a sample (section 3), and the

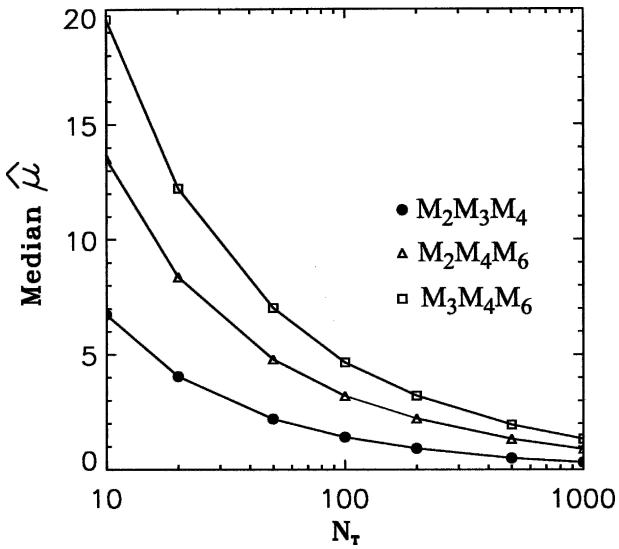

FIG. 12. Variation of the median value of the gamma DSD shape parameter $\hat{\mu}$, as estimated from the three sample moments indicated on the respective curves, with a mean sample size. Population DSD: exponential ($\mu = 0$).

associated correlations between moments, lead to correlation of the fitted parameters with the maximum drop size (e.g., Fig. 13). The parameters are, in turn, correlated with each other (Fig. 14), and such correlations (caused here entirely by sampling variability) could be mistaken for physical relationships. The correlations between estimated parameters are a bit less strong when lower-order moments are used in the fitting process; with $N_T = 100$, the $\hat{n}_0 - \hat{\Lambda}$ correlation is only 0.927 when sample moments M_{2S} and M_{3S} are used as compared with the 0.946 value with moments M_{3S} and M_{6S}, as illustrated in Fig. 14. The figure also illustrates the general tendency for the moment methods to overestimate both the slope and intercept parameters of an exponential function.

The $\hat{n}_0 - \hat{\Lambda}$ correlation actually *increases* as the sample size increases—to values of 0.946 (for moments M_{2S}, M_{3S}) or 0.980 (for moments M_{3S}, M_{6S}) when $N_T = 1000$—though the ranges of the variation of the parameters decrease. In any case, this behavior suggests that one should be wary of inferring physical relationships between such fitted parameters until the effects of the sampling variability have been taken into account.

6. Implications for analysis of experimental data

In trying to relate these simulations to actual raindrop observations, one should keep in mind several factors:

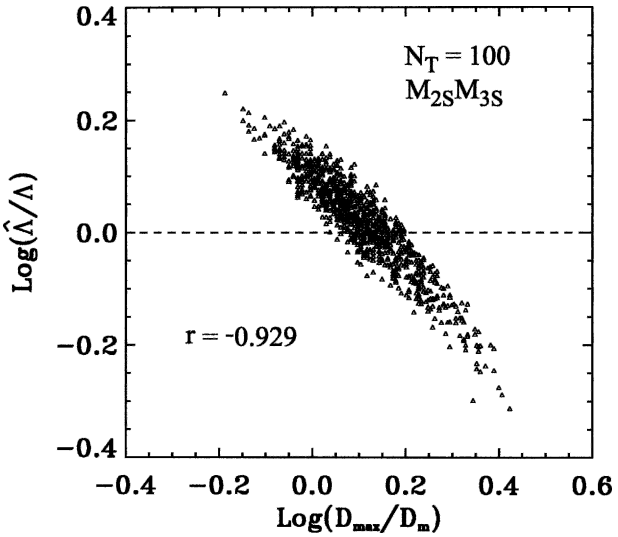

FIG. 13. Scatterplot of normalized values of the DSD slope parameter $\hat{\Lambda}$ for an exponential function, estimated from moments M_{2S} and M_{3S}, vs (normalized) maximum drop size in the sample. Population DSD: exponential; mean sample size: 100 drops. Dashed line indicates population value of Λ.

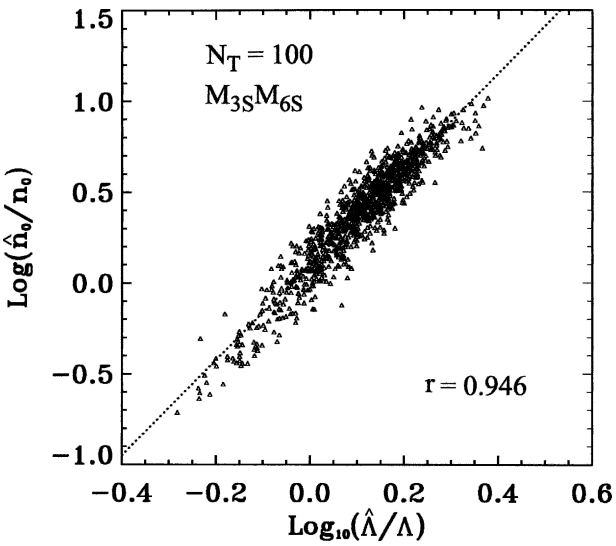

FIG. 14. Scatterplot of normalized values of DSD intercept parameter $\hat{n}_o$ vs slope parameter $\hat{\Lambda}$ for an exponential function, as estimated from moments M_{3S} and M_{6S}. Population DSD: exponential, with a mean sample size of 100 drops. Dotted line indicates the regression relationship.

- The actual population DSDs in nature are not known. There is no assurance that they are exponential, though there are indications that this may be the case (Joss and Gori 1978).
- Observations with surface-sampling instruments, such as impact disdrometers, involve sample volumes that increase with the drop size, which tends to mitigate the skewness in the high-order sample moments, and, consequently, the associated biases.
- The observations include only the actual sample size (C in these simulations). The mean, or expected, sample size (N_T) is not known, though the actual sample size provides a better approximation as it increases.
- Very small drops are generally absent from many such observations, either because such small drops are not present, or because the instruments do not respond to those drops. In a full exponential DSD, 55% of the drops would be smaller than $0.2D_m$, and 86% would be smaller than $0.5D_m$. With typical values of D_m being 1–3 mm, this means the simulations may involve more (and in some cases much more) than twice the total numbers of drops that would be found in corresponding observations. Thus, results given above for $N_T = 100$ might be more applicable to observations with total drop counts of, say, 20–50.

These caveats notwithstanding, certain broad inferences as follows are applicable:

- Values of the parameters for DSD functions as estimated by moment methods will be biased.

- The bias will be stronger when higher moments are employed.
- The bias will diminish as the sample size increases.

Thus, DSD parameters that are estimated using high-order sample moments (e.g., reflectivity) with small sample sizes (say a few tens of drops) are the most highly suspect. Moreover, as suggested in the preceding section, caution should be used in attempting to impute a physical basis to relationships between DSD parameters fitted by moment methods.

Because the actual DSD parameters (even if the DSD were exponential) are unknown, it is difficult to critique in a quantitative manner any given set of results based on moment-method analyses. The most likely indication of the kind of biases discussed here appears in published frequency distributions of the gamma shape parameter $\hat{\mu}$, such as those reported in Kozu and Nakamura (1991) or Tokay and Short (1996). They extend to values as large as 30, which in view of Fig. 9 herein may well be a consequence of the M_3, M_4, M_6 moment approach that is used by those authors. The fact that the maximum $\hat{\mu}$ values reported by Kozu and Nakamura decrease as the rainfall rate increases (which should be accompanied by increases in sample size) lends further weight to this interpretation. A simple way to test this idea would be to stratify the shape parameter estimates by sample size, to look for a trend similar to that shown in Fig. 12.

7. Conclusions

Moment estimators for parameters of DSD functions are inherently biased. They tend to give erroneous values of the DSD parameters unless the drop samples are much larger than those commonly available. In particular, estimates of the gamma shape parameter μ tend to be far larger than the shape parameter of the underlying DSD from which the samples are taken. The bias is strongest for small sample sizes, and are also stronger when higher-order moments of the observed DSDs are used in the "fitting" process.

Moment methods may provide estimates of DSD parameters of sufficient accuracy if very large samples (hundreds, perhaps thousands) of drops are available. Failing that, some alternative approach to fitting the observed DSDs must be used. The maximum likelihood approach suggested by Haddad et al. (1996, 1997) may be satisfactory, though the maximum likelihood estimators are not without bias (Choi and Wette 1969). A variant of the *L*-moment approach used in hydrology (e.g., Maidment 1993) may also be worthy of consideration.

Acknowledgments. This material is based upon work supported by the National Science Foundation under grant ATM-9907812. The authors appreciate the assistance of Prof. S. J. Burges in directing them to references concerning use of moment methods in hydrology.

APPENDIX

Equations for Moment Estimators of DSD Parameters

This appendix presents the expressions that are used to develop the normalized parameter estimates summarized in this paper. To illustrate the moment-method procedures as employed in these simulations, consider first the case of parameters N_T and D_m for an exponential DSD function to be estimated from sample moments M_{2S} and M_{3S}. From (5),

$$M_2 = N_T D_m^2/8; \quad M_3 = 3N_T D_m^3/32. \tag{A1}$$

Algebraic solution yields the relationships

$$N_T = 9M_2^3/2M_3^2; \quad D_m = 4M_3/3M_2. \tag{A2}$$

The sample moments in these simulations are

$$M_{iS} = m_i D_m^i, \tag{A3}$$

where

$$m_i = \sum_C y^i. \tag{A4}$$

Substituting the sample moments into (A2) yields the estimators

$$\hat{N}_T = 9m_2^3/2m_3^2; \quad \hat{D}_m = 4m_3 D_m/3m_2. \tag{A5}$$

From the latter,

$$\frac{\hat{D}_m}{D_m} = \frac{4}{3}\frac{m_3}{m_2}. \tag{A6}$$

For the n_0, Λ parameters of (1), the relationship $\Lambda = 4/D_m$ leads to

$$\hat{\Lambda} = 4/\hat{D}_m = 3m_2/(m_3 D_m), \tag{A7}$$

from which

$$\frac{\hat{\Lambda}}{\Lambda} = \frac{3}{4}\frac{m_2}{m_3}. \tag{A8}$$

To obtain the estimator for n_0, a hybrid expression for the DSD moments is useful:

$$M_i = i! n_0 D_m^{i+1}/4^{i+1}. \tag{A9}$$

Thus,

$$M_2 = n_0 D_m^3/32; \quad M_3 = 3n_0 D_m^4/128. \tag{A10}$$

An algebraic solution of (A10) yields

$$n_0 = 27M_2^4/2M_3^3. \tag{A11}$$

Substituting the sample moments,

$$\hat{n}_0 = 27m_2^4/(2m_3^3D_m). \tag{A12}$$

Making use of the relationship $n_0 = 4N_T/D_m$,

$$\frac{\hat{n}_0}{n_0} = \frac{27}{8N_T}\frac{m_2^4}{m_3^3}. \tag{A13}$$

In this fashion, parameter estimates based on the normalized moments m_i can be compared with the population parameters in dimensionless expressions where the actual particle sizes do not enter. Thus, the only population parameter that appears is N_T (with one exception, discussed in section 1b below). Expressions for the remaining (normalized) moment estimators follow.

a. Exponential functions

1) MOMENTS: M_3, M_6

Parameters n_0, Λ are

$$\frac{\hat{n}_0}{n_0} = \frac{10(15)^{1/3}}{N_T}\frac{m_3^{7/3}}{m_6^{4/3}};$$

$$\frac{\hat{\Lambda}}{\Lambda} = \frac{(15)^{1/3}}{2}\left(\frac{m_3}{m_6}\right)^{1/3}.$$

Parameters N_T, D_m are

$$\hat{N}_T = 20\, m_3^2/m_6;$$

$$\frac{\hat{D}_m}{D_m} = \frac{2}{(15)^{1/3}}\left(\frac{m_6}{m_3}\right)^{1/3}.$$

2) MOMENTS: M_3, M_4 (RESULTS NOT SHOWN)

Parameters n_0, Λ are

$$\frac{\hat{n}_0}{n_0} = \frac{32}{3N_T}\frac{m_3^5}{m_4^4};$$

$$\frac{\hat{\Lambda}}{\Lambda} = \frac{m_3}{m_4}.$$

Parameters N_T, D_m are

$$\hat{N}_T = 32m_3^4/3m_4^3;$$

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

b. Gamma functions

1) MOMENTS: M_2, M_3, M_4

Parameters n_1, μ, and λ are

$$\frac{\hat{n}_1}{(n_1D_m^{\mu-\hat{\mu}})} = \frac{1}{4N_T}\frac{(\hat{\mu}+4)^{\hat{\mu}+4}}{(\hat{\mu}+3)!}\frac{m_3^{\hat{\mu}+5}}{m_4^{\hat{\mu}+4}}$$

(generally $\hat{\mu} \neq \mu$, so the population value of D_m cannot be removed to normalize the expressions involving $\hat{n}_1$),

$$\hat{\mu} = \frac{3m_4m_2 - 4m_3^2}{m_3^2 - m_4m_2}, \quad \text{and}$$

$$\frac{\hat{\lambda}}{\lambda} = \frac{1}{4}\frac{(\hat{\mu}+4)m_3}{m_4}.$$

Parameters N_T, μ, and D_m are

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

with $\alpha = m_3^2/(m_4m_3)$; $\hat{\mu}$ is as above; and

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

Parameters $N_w(n_0^*)$, μ, and D_m are $\hat{\mu}$, $(\hat{D}_m/D_m)$, as above, and

$$\frac{\hat{N}_w}{N_w} = \frac{32}{3N_T}\frac{m_3^5}{m_4^4}.$$

2) MOMENTS: M_2, M_4, M_6

Parameters n_1, μ, and λ are

$$\frac{\hat{n}_1}{(n_1D_m^{\mu-\hat{\mu}})} = \frac{1}{4N_T}\frac{1}{(\hat{\mu}+2)!} \times\left\{[(\hat{\mu}+3)(\hat{\mu}+4)]^{\hat{\mu}+3}\frac{m_2^{\hat{\mu}+5}}{m_4^{\hat{\mu}+3}}\right\}^{1/2};$$

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

with $\eta = m_4^2/(m_2m_6)$; and

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

Parameters N_T, μ, and D_m are

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

$\hat{\mu}$ is as above, and

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

Parameters N_w, μ, and D_m are $\hat{\mu}$, $(\hat{D}_m/D_m)$, as above, and

$$\frac{\hat{N}_w}{N_w} = \frac{32}{3N_T}\left\{\left[\frac{(\hat{\mu}+2)m_2}{(\hat{\mu}+4)}\right]^5 \frac{1}{m_4^3}\right\}^{1/2}.$$

3. MOMENTS: M_3, M_4, M_6

Parameters n_1, μ, and λ are

$$\frac{\hat{n}_1}{(n_1 D_m^{\mu-\hat{\mu}})} = \frac{1}{4N_T}\frac{(\hat{\mu}+4)^{\hat{\mu}+4}}{(\hat{\mu}+3)!}\frac{m_3^{\hat{\mu}+5}}{m_4^{\hat{\mu}+4}};$$

$$\hat{\mu} = \frac{11G - 8 + [G(G+8)]^{1/2}}{2(1-G)},$$

with $G = m_4^3/m_3^2 m_6$; and

$$\frac{\hat{\lambda}}{\lambda} = \frac{1}{4}\frac{(\hat{\mu}+4)m_3}{m_4}.$$

Parameters N_T, μ, and D_m are

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

$\hat{\mu}$ is as above, and

$$\hat{D}_m/D_m = m_4/m_3.$$

Parameters N_w, μ, and D_m are $\hat{\mu}$, $(\hat{D}_m/D_m)$, as above, and

$$\frac{\hat{N}_w}{N_w} = \frac{32}{3N_T}\frac{m_3^5}{m_4^4}.$$

The expressions for $\hat{D}_m$ and $\hat{N}_w$ are identical to those for the M_2, M_3, M_4 moment set; those for $\hat{n}_1$ and $\hat{\lambda}$ are algebraically the same but involve different values for $\hat{\mu}$.

REFERENCES

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

Choi, S. C., and R. Wette, 1969: Maximum likelihood estimation of the parameters of the gamma distribution and their bias. *Technometrics,* **4,** 683–690.

Gertzman, H. S., and D. Atlas, 1977: Sampling errors in the measurement of rain and hail parameters. *J. Geophys. Res.,* **82,** 4955–4966.

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

——, 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 Trans. Geosci. Remote Sens.,* **35,** 532–539.

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

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.

Maidment, D., Ed., 1993: *Handbook of Hydrology.* McGraw-Hill, 1424 pp.

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

Smith, J. A., 1993: Marked point process models of raindrop-size distributions. *J. Appl. Meteor.,* **32,** 284–296.

Smith, P. L., 1982: On the graphical presentation of raindrop size data. *Atmos.–Ocean,* **20,** 4–16.

——, 2003: Raindrop size distributions: Exponential or gamma—Does the difference matter? *J. Appl. Meteor.,* **42,** 1031–1034.

——, and D. V. Kliche, 2003: The bias in moment estimators for parameters of drop size distribution functions. Preprints, *31st Conf. on Radar Meteorology*, Seattle, WA, Amer. Meteor. Soc., 47–50.

——, Z. Liu, and J. Joss, 1993: A study of sampling-variability effects in raindrop size observations. *J. Appl. Meteor.,* **32,** 1259–1269.

Testud, J., S. Oury, R. A. Black, P. Amayenc, and X. Dou, 2001: The concept of "normalized" distribution to describe raindrop spectra: A tool for cloud physics and remote sensing. *J. Appl. Meteor.,* **40,** 1118–1140.

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.

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

Waldvogel, A., 1974: The N_0 jump of raindrop spectra. *J. Atmos. Sci.,* **31,** 1067–1078.

Wallis, J. R., N. C. Matalas, and J. R. Slack, 1974: Just a moment! *Water Resour. Res.,* **10,** 211–219.