AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE
10.1002/2016JD025491

# Sampling variability effects in drop-resolving disdrometer observations

**Michael L. Larsen[1] and Katelyn A. O'Dell[1,2]**

[1]Department of Physics and Astronomy, College of Charleston, Charleston, South Carolina, USA, [2]Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado, USA

**Key Points:**
- Previous simulation studies of disdrometer sampling variability may have been too optimistic
- Small disdrometric samples tend to underestimate all moments of the drop size distribution
- Sampling can have interesting effects on disdrometer-derived Z-R relationships

**Correspondence to:**
M. L. Larsen,
LarsenML@cofc.edu

**Citation:**
Larsen, M. L., and K. A. O'Dell (2016), Sampling variability effects in drop-resolving disdrometer observations, *J. Geophys. Res. Atmos.*, *121*, 11,777–11,791, doi:10.1002/2016JD025491.

Received 9 JUN 2016
Accepted 19 SEP 2016
Accepted article online 27 SEP 2016
Published online 10 OCT 2016

**Abstract** Rain events are often studied with raindrop disdrometers—instruments designed to measure and record individual drop properties. Data from these instruments reveal that the instantaneous raindrop arrival rate is highly variable, with drop counts in consecutive time intervals occasionally changing by an order of magnitude or more. These drop count fluctuations result from some unknown combination of real physical changes in the underlying rainfall in conjunction with statistical fluctuations due to instrumental imperfections, limited sampling area, and the correlated nature of drop arrivals. Here empirical observations are used to drive a simulation that explores the degree to which observed variability in bulk rainfall observations might be produced by statistical fluctuations. It is shown that meaningful conclusions about true underlying rain properties require more drops than previously thought. Revisions to the interpretation of disdrometer data are discussed.

## 1. Introduction

Rain is highly variable in time and space. Observations suggest that there is some degree of statistical structure all the way down to temporal scales shorter than 1 s [*Jameson et al.*, 1999] and spatial scales of meters or less [*Gires et al.*, 2015].

Because of statistical fluctuations, accurately measuring and reporting the most basic of rain variables at even a single location is complicated. Multiple identical detectors can be placed right next to each other and see notably different results [see, e.g., *Thurai et al.*, 2011; *Tokay et al.*, 2005, 2013]. Quantification of any variable's spatial or temporal variability (which requires measurements at multiple locations or times) is even more complicated.

Despite the challenges associated with characterizing the spatial and temporal variability of rainfall on small scales, significant attention has been given to the issue in recent years. The spatial and temporal scales of physical variability have been determined to be smaller than a radar pixel [see, e.g., *Jameson and Kostinski*, 2000; *Miriovsky et al.*, 2004; *Jaffrain et al.*, 2011]. There is even increasing evidence that the scales of physical variability may be closer to "point-detector" sampling volumes than previously thought [see, e.g., *Jameson et al.*, 2015a, 2015b].

The differences in measured rain properties in very close ($<\sim$100 m) locations are due to a large combination of factors including, though not necessarily limited to, (i) physical variability, (ii) detector imperfections, and (iii) finite-sampling effects. Normally, all three of these effects are present, but the investigator trying to understand the natural world is typically interested in the physical variability. Even if the physical variability during a detector sampling interval is negligible, estimation of the "physical state" of the system at a point requires a thorough understanding of both detector imperfections and finite-sampling effects.

Many experiments have attempted to explore instrumental limitations by deploying several detectors (sometimes identical, sometimes different) in close proximity and examining the differences in measured results [see, e.g., *Donnadieu*, 1980; *Campos and Zawadzki*, 2000; *Habib et al.*, 2001; *Ciach*, 2003; *Salles and Creutin*, 2003; *Miriovsky et al.*, 2004; *Tokay et al.*, 2005; *Krajewski et al.*, 2006; *Jaffrain and Berne*, 2011; *Thurai et al.*, 2011; *Tokay et al.*, 2013]. The details of these studies vary significantly, but the central problem becomes one of attribution; if two adjacent detectors do not see the same rain, the differences are typically assumed to be due to instrumental error and/or bias. Even if physical variability is not relevant on the spatiotemporal scales of the study, however, the differences between two detectors result from a combination of instrumental

limitations/biases/imperfections and natural sampling variability. Assigning all differences in observations by adjacent detectors to instrument differences is simply not true; due to the correlated structure of rain and the statistical sampling variability, the adjacent detectors very well could have had a different number of drops—or drops of different sizes—fall through their sensing region during the time interval in question.

Here an attempt is made to use a simulation to explore an upper bound for possible effects of sampling variability via the use of a data-driven Monte Carlo simulation. Such an approach has been undertaken before in at least four well-respected and thought-provoking publications [see, e.g., *Smith et al.*, 1993; *Jameson and Kostinski*, 2002a; *Smith and Kliche*, 2005; *Smith et al.*, 2009]. However, all of these studies used models of drop arrival statistics that do not always adhere particularly well to recent observations of rain drop arrival statistics. By constructing a simulation with more realistic drop arrival statistics, the work presented here clearly demonstrates that sampling variability effects are quite possibly even more pronounced than previously reported.

## 2. Simulation

### 2.1. General Scope and Structure of the Simulation

The Monte Carlo simulation developed here is very similar to those described in *Smith et al.* [1993] and *Jameson and Kostinski* [2002a]. The simulations are designed to replicate the process of measuring real rainfall with a disdrometer but with the (artificial) constraint that the underlying system that is being "measured" is unchanging in time.

More explicitly, in these simulations we proceed as if there are "true" (and fixed) distributions that govern the relative abundance of drops of different size. Differences in variables measured from sample to sample depend on incomplete sampling of the underlying drop size distribution and drop number fluctuations caused by sampling and natural variability. The detector is assumed "perfect" and the only reason that the "true" drop size distribution is not measured in each sample is because a finite number of drops are observed.

### 2.2. Simulation Procedure

The basic simulation procedure is as follows.

1. For each sample, draw the number $k_i$ of drops that are observed. $k_i$ is drawn from a specified distribution with mean number of drops per sample $\bar{N}$. (This distribution will be called the sample-size distribution. The probability of each sample containing $k$ raindrops when the mean number of drops per measurement is $\bar{N}$ will be written as $p_s(k|\bar{N})$.)
2. For each of the $k_i$ drops in the sample, assign a drop diameter based on some preset distribution. (This distribution will be called the size-assignment distribution. The probability that a particular raindrop is assigned a diameter $D$ will be written as $p_d(D)$.)
3. After establishing the number of drops and the sizes for each drop in each sample, calculate any desired quantities for each sample. Following typical rain science convention, these quantities include $D_m$ (mass-weighted mean diameter), LWC (liquid water content), $R$ (rain rate), $R^*$ (the empirical fourth moment of the drop size distribution, often used as a proxy for rain rate) [see, e.g., *Joss and Gori*, 1978], and/or $Z$ (radar reflectivity factor).

This procedure is repeated for a fixed number of samples at each value of $\bar{N}$ of interest. Note that, by construction, both $p_s(k|\bar{N})$ and $p_d(D)$ are treated as constant; all samples draw from the same underlying distributions. Each sample in the simulation draws, as needed, from these fixed distributions—thus rendering each measurement statistically independent of all other measurements. The resulting distributions associated with each bulk rainfall variable are the quantities of primary interest in this study.

### 2.3. Driving the Simulation With Real Data

The overall model/structure described above is designed to replicate the intuitive approaches taken in *Smith et al.* [1993], *Jameson and Kostinski* [2002a], *Smith and Kliche* [2005], and *Smith et al.* [2009]. Details follow below, but overall similarities and differences between this study and those completed in prior studies are presented in the appendix.

In *Smith et al.* [1993], *Smith and Kliche* [2005], and *Smith et al.* [2009] the sample-size distributions were set to be a Poisson distribution:

$$p_s(k|\bar{N}) = \frac{\bar{N}^k \exp\left(-\bar{N}\right)}{k!}. \quad (1)$$

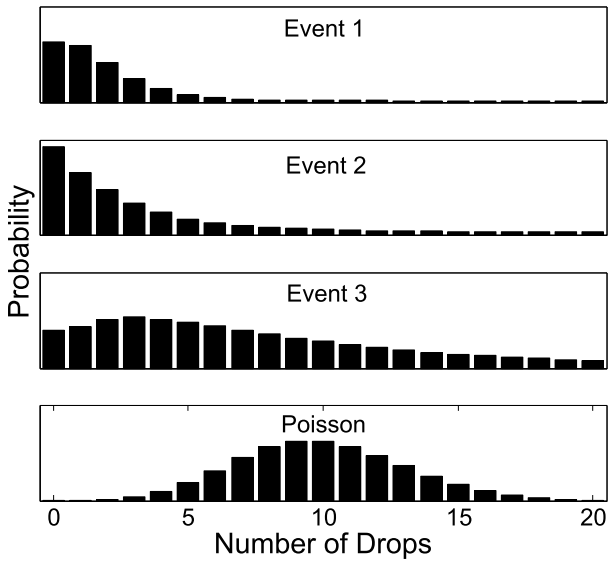


**Figure 1.** The distribution of sample sizes observed by the 2DVD for each of the three events used in the simulations presented here, as well as a Poisson distribution. Each panel forms a probability distribution of sample sizes and thus integrates to unity (though all panels are truncated to values between 0 and 20 drops to show more detail). Here the time per sample for each event was fixed so that an average of 10 drops/sample are expected. Notice the substantial breadth and skewness of these distributions when compared to a Poisson distribution.

Even at the time of the study, the investigators of *Smith et al.* [1993] acknowledged that observations of fractal behavior in drop arrivals "might mean that a distribution broader than the Poisson should be used for simulating the sample numbers." The authors further asserted "....We do not expect the substantive findings of this paper to change significantly with other, broader-type distributions." The veracity of that claim was tested in *Jameson and Kostinski* [2002a], where the sample-size distribution was set to be a discrete uniform random distribution, running from (approximately) 0 to $2\bar{N}$. Some of the conclusions drawn from these studies agreed with each other (e.g., both studies saw that convergence to the underlying intrinsic values improves with increasing mean sample sizes), but at times the studies differed (e.g., *Smith et al.* [1993] saw a definite bias to underestimation of all quantities for small sample sizes, whereas no deterministic underestimation is evident in the *Jameson and Kostinski*, 2002a's [2002a] work. Also, inferences about the general behavior of statistical *Z-R* relations differed in these two studies).

In this study, a slightly different approach is taken. Rather than prescribing a particular distribution for $p_s(k|\bar{N})$, this simulation uses measured data to develop a distribution that retains some elements of real physical and sampling variability. More formally, let a sequence of drop observations (where drop arrival times and diameters are recorded) include a total of $N_t$ drops observed in total time $T$. Thus, the average measured interdrop time can be set to $\tau \equiv T/N_t$. If samples are desired having an average of $\bar{N}$ drops, then one can partition the observed drop arrival times into $N_t/\bar{N}$ disjoint intervals, each of duration $\bar{N}\tau$. By looking at the frequency distribution of the numbers of drops observed in these time intervals, an empirical distribution of sample sizes with known mean $\bar{N}$ is generated. This empirical distribution then serves as the sample-size distribution $p_s(k, \bar{N})$ to drive the simulation. Examples of such empirical distribution functions are shown in Figures 1–3 (first to third panels).

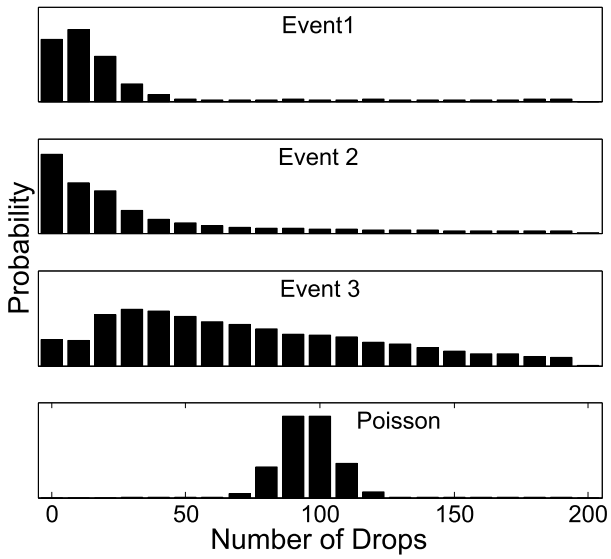


**Figure 2.** Same as Figure 1, except with an average of 100 drops/sample for each panel. (This corresponds to a time interval 10 times longer than Figure 1.) Here bins are made with a bin width of 10 drops. It does not appear that all four panels would have the same mean; this is due to the very long tail of the measured distributions where totals of drops as high as 1193, 1606, and 1617 for the three events (respectively) were obtained.

As can be clearly seen from the figures, the breadth of an empirical sample size distribution greatly exceeds that of a Poisson distribution with the same mean—especially for large $\bar{N}$. This will be seen to have substantial impact on the results in the simulations below.

In *Smith et al.* [1993], *Jameson and Kostinski* [2002a], and *Smith and Kliche* [2005] the drop

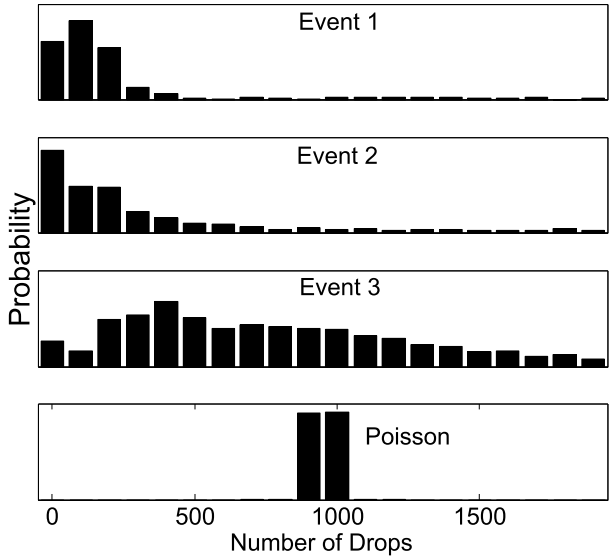


**Figure 3.** Same as Figures 1 and 2, except with an average of 1000 drops/sample. (This corresponds to a time interval 100 times longer than Figure 1.) Substantial differences both between events and when comparing events to a Poisson distribution are clearly evident. Note that though the plots were truncated to only show sample sizes up to 2000 drops, the tails of the distribution functions extend beyond the region shown. (The largest samples for the three events included 7250, 10850, and 10469 drops, respectively.) The bin width in this figure is 100 drops/bin.

size distribution was also established *apriori* as a decaying exponential distribution following long-standing convention [e.g., *Marshall and Palmer*, 1948]. Conversely, *Smith et al.* [2009] used the increasingly popular Gamma distribution. In this study a different methodology was employed by using an actual (observed event averaged) $p_d(D)$, rather than an assumed theoretical size distribution (see Figure 4).

A concrete physical interpretation of using such a data-driven simulation procedure may be of some benefit. Multiple studies [e.g., *Jameson and Kostinski*, 2001; *Jameson et al.*, 2015b] have noted that the drop size distribution can explicitly be written:

$$n(D)\mathrm{d}D = n_\circ \times p(D)\mathrm{d}D, \qquad (2)$$

where $n_\circ$ is the total number of drops in a unit sample volume, $p(D)$ is a probability density function of diameter, and both $n_\circ$ and $p(D)$ have implicit time dependence. The approach utilized here ignores the time dependence of $p(D)$ since (i) measurements of $p(D)$ suggest that it shows far less variability in time than $n_\circ$ and (ii) recent other studies [e.g., *Jameson et al.*, 2015a; *Jameson and Larsen*, 2016] have clearly revealed that the influence of $n_\circ$ dominates over the influence of $p(D)$ for many physical scenarios.

Here temporal structures (drop clustering) on certain time scales are still retained via the use of the empirical sample-size distribution function $p_s(k|\bar{N})$. Thus, some elements of both the sampling and physical variability are retained by driving the simulation with this measured data.

Ultimately, the simulation process described here (as well as those systems explored by *Smith et al.* [1993], *Jameson and Kostinski* [2002a], *Smith and Kliche* [2005], and *Smith et al.* [2009]) acts to homogenize the size-assignment distribution; no evolution of this distribution is allowed during the event.

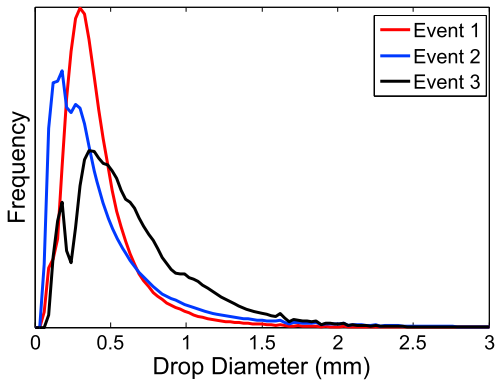


**Figure 4.** The normalized distribution of drop sizes observed by the 2DVD for each of the three events used in the simulations presented here. These probability density functions are used to assign the drop sizes in the simulations.

### 2.4. Intrinsic Event Values

Because the underlying size-assignment distribution $p_d(D)$ is treated as constant, each simulation has deterministic intrinsic values for $D_m$, LWC, $R$, $R^*$, and $Z$. Typically, these bulk quantities are computed in terms of the drop size distribution (which is a bit different from $p_d(D)$) (see discussions in *Kostinski and Jameson* [1999], *Jameson and Kostinski* [2001], *Kruger and Krajewski* [2002], and *Salles and Creutin* [2003]).

The relationship described in *Kruger and Krajewski* [2002] was used to transform the size-assignment distribution $p_d(D)$ into the drop size distribution $n(D)$. In the notation used here, that transformation is given by

$$n(D) = \frac{[N_t p_d(D)]}{AT[v_t(D)]}, \qquad (3)$$

**Table 1.** Basic Information Associated With the Three Events That Were Used to Drive the Simulations Presented in This Manuscript[a]

| Event Number | Date (2014) | Event Duration (Hours) | Total Drops | $\overline{R}$ (mm/h) | $\overline{Z}$ (mm$^6$/m$^3$) | $\overline{D_m}$ (mm) |
|---|---|---|---|---|---|---|
| 1 | Jan 14 | 7.07 | 587766 | 0.8976 | 234.90 | 0.8495 |
| 2 | Jul 6 | 5.42 | 1693811 | 8.9734 | 7991.4 | 1.5032 |
| 3 | Nov 23 | 12.77 | 1383778 | 3.8090 | 1434.5 | 1.1989 |

[a] All data are reported as recorded by the 2DVD at the disdrometer array site.

where there were a total of $N_t$ drops in the event, the cross-sectional area of the detector was $A$, the total duration of the event was $T$, and $v_t(D)$ is the terminal velocity of drops of diameter $D$ (determined using the ninth order power law approximation given in *Foote and du Toit* [1969]).

In terms of this "traditional" drop size distribution $n(D)$ (where $n(D)\mathrm{d}D$ identifies the number of drops *per unit volume* with drop diameter between $D$ and $D + \mathrm{d}D$) the quantities of interest are found via relationships like the following:

$$D_m = \frac{\int_0^\infty n(D)D^4\mathrm{d}D}{\int_0^\infty n(D)D^3\mathrm{d}D} \tag{4}$$

$$R = \frac{6\pi}{10^4}\int_0^\infty n(D)v_t(D)D^3\mathrm{d}D \tag{5}$$

$$Z = \int_0^\infty n(D)D^6\mathrm{d}D \tag{6}$$

[see, e.g., *Smith et al.*, 1993; *Steiner and Smith*, 2000]. In these relationships, it is assumed that $D$ is measured in mm, $v_t(D)$ is measured in m s$^{-1}$, and $n(D)$ is reported in units of (mm$^{-1}$ m$^{-3}$). These relationships report $D_m$ in mm, $R$ in mm h$^{-1}$, and $Z$ in mm$^6$ m$^{-3}$.

In a scenario under which sampling and physical variability are not present, each measurement would obtain the same (intrinsic) values for $D_m$, $R$, and $Z$. As the simulations reveal, however, sample-to-sample variability even in this simulated statistically stationary rainfall can still be quite substantial, due mostly to the breadth of $p_s(k|N)$.

### 2.5. Simulation Details

In an attempt to directly compare the results here to those in *Smith et al.* [1993], simulations for each event were run to match this study with $\bar{N} = \{10, 20, 50, 100, 200, 500, 800\}$ drops. Additional simulations were also run with $\bar{N} = 1000$ drops.

It appears that *Smith et al.* [1993] based their results on approximately 2000 samples for each value of $\bar{N}$ used. The studies of *Jameson and Kostinski* [2002a] and *Smith and Kliche* [2005] appear to have used 10,000 samples at each value of $\bar{N}$, and the study of *Smith et al.* [2009] used 20,000 different samples for each value of $\bar{N}$.

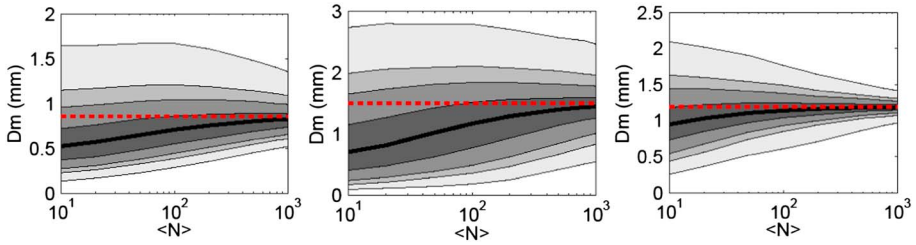


**Figure 5.** A detailed depiction of the convergence of $D_m$ to its true intrinsic value. In each panel, the intrinsic value of $D_m$ is marked with a dotted red line. The solid black line indicates the median of sampled values at each mean sample size. As the mean number of drops per sample ($\langle N \rangle$) increases, samples are more likely to take on values closer to the intrinsic $D_m$ value. The shading shows probability contours; half of all samples for each $\langle N \rangle$ lie within the darkest area. Contours indicate regions containing 50%, 80%, 90%, and 98% (darkest to lightest) of all samples. The panels show simulations based on (left) event 1, (middle) event 2, and (right) event 3. Simulations were run with $\langle N \rangle = 10, 20, 50, 100, 200, 500, 800$, and 1000 drops, with 110,000 samples for each $\langle N \rangle$.

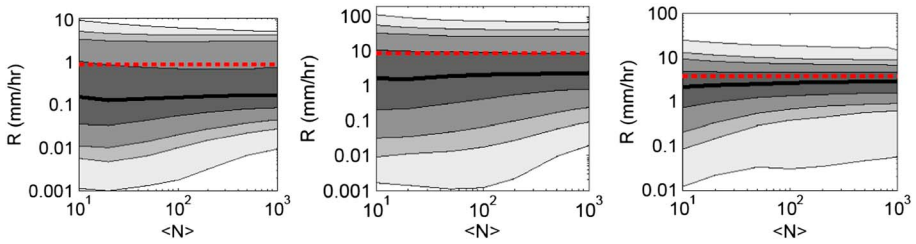


**Figure 6.** This figure is similar to Figure 5 but shows the convergence of $R$ to its intrinsic value in each event. Note that the $R$ axis is logarithmic and that in events 1 and 2, ~ 75% of all observed samples take on values smaller than the intrinsic value of $R$, even when $\langle N \rangle = 1000$ drops. Convergence of the median sampled value to the intrinsic value is clearly far slower for $R$ than for $D_m$.

Due to the breadth of the sample-size distribution driving the simulations here, sufficient statistical sampling was something of a concern. Thus, a total of 110,000 independent samples were used for each event and at each value of $\bar{N}$.

## 3. Data

Perhaps the most novel element of the approach utilized here is that the Monte Carlo simulation is partially driven with real data. The data used here were acquired by a two-dimensional video disdrometer manufactured by Joanneum and described in detail in *Kruger and Krajewski* [2002]. This disdrometer records precise arrival times and drop diameters for each drop in a sensing area of approximately 100 cm$^2$. The ability to resolve individual drop arrival times is crucial in the approach used here; the investigator must partition the data into time bins of specified duration in order to generate the sample-size distribution function $p_s(k|\bar{N})$ with the proper mean ($\bar{N}$). Other instruments that record summary data on a regular time interval (e.g., the frequently used Joss-Waldvogel or Parsivel disdrometers) would not be appropriate for this task.

The disdrometer used here is part of a disdrometer array located at Dixie Plantation near Charleston, South Carolina (32°44′26″N, 80°10′36″W). The other disdrometers in the array were used to confirm that two-dimensional video disdrometer was operating properly during each of the events studied here. A thorough description of the instrument array and site environment is presented in *Larsen et al.* [2014], *Jameson et al.* [2015a, 2015b], *Larsen and Teves* [2015], and *Larsen et al.* [2015], but a detailed site analysis is beyond the scope of this manuscript.

In this study, three separate storms (events) were used to drive the simulations. Basic summary information for each of these three events is presented in Table 1, along with the event-averaged intrinsic values of each of the three key variables. Although many rain events were available, these three were chosen because (i) each lasted several hours with at least several hundred thousand total detected drops, (ii) each had continuous accumulations over the duration of the event, and (iii) the events demonstrated a variety of different values for bulk properties $D_m$, $R$, and $Z$. Prior to the selection of these three events, event-averaged $p_s(k|\bar{N})$ and $p_d(D)$ were generated for each possible event recorded in 2014. The three events chosen here well represent the variety of different shapes that $p_s(k|\bar{N})$ and $p_d(D)$ take in the ensemble.

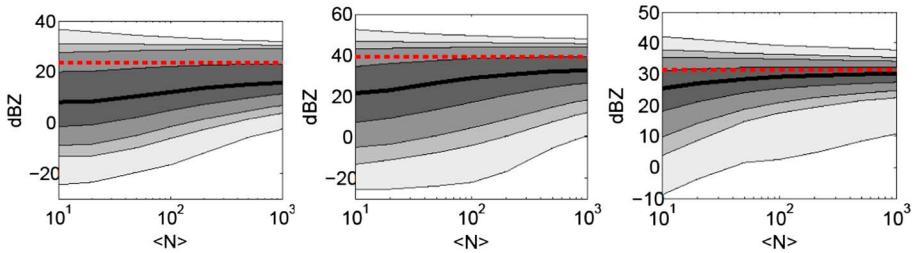


**Figure 7.** This figure is similar to Figures 5 and 6 but shows the convergence of $Z$ to its intrinsic value in each event. Consistent with *Smith et al.* [1993], we see that the median sampled value substantially underestimates the mean. The degree of this underestimation, however, seems more pronounced with this realistic sampling structure. In particular, the contours for events 2 and 3 imply that many more than 1000 drops must be expected in each sample to adequately estimate $Z$.

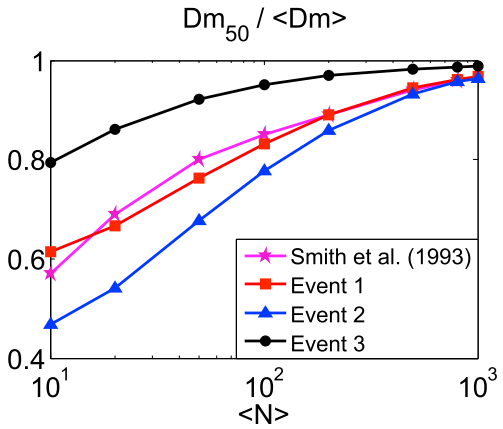


**Figure 8.** The median sampled value of $D_m$ divided by the intrinsic ensemble value of $D_m$ as a function of mean sample size for each of the three events, as well as the data from *Smith et al.* [1993, Figure 3].

Figure 4 shows $p_d(D)$ detected for each of the three events. (Recall that this is a plot of the "size-assignment" distribution instead of a volume-based drop size distribution. The differences are discussed at length in *Jameson and Kostinski* [2001, 2002a].) Calculation of ensemble values of $D_m$, $R$, and $Z$ was done by utilizing equations (3)–(6).

## 4. Simulation Results

### 4.1. Moment Estimation

Perhaps the central question of interest is "how long does a single sample have to last so that enough drops are likely detected to reliably estimate $D_m$, $R$, or $Z$?" In *Smith et al.* [1993], it was stated that "For most of the quantities, the skewness becomes small enough that the median sample values approach reasonably close to the population values for sample volumes yielding (expected) sample sizes greater than 100 or so drops." The results presented here clearly demonstrate that this conclusion is an unfortunate consequence of using a sample-size distribution much too narrow to replicate reality.

There are a number of different ways to visualize the simulated sample-to-sample variability in each of the associated variables. Figure 5 shows one of these methods. A subplot is presented for each of the three simulated events. The horizontal red dotted line in each panel shows the intrinsic value of $D_m$ for the event. The solid black line shows the median value of $D_m$ for the (nonempty) samples. As expected, the median approaches the intrinsic value with increasing sample size. The shaded contours in this figure show regions that include 50% (darkest), 80%, 90%, and 98% (lightest) of all nonempty observed samples. (In other words, 98% of all measured samples including at least one drop had $D_m$ values between about 0.5 and 1.3 mm for event 1 when $\bar{N} = 1000$; 2% of the measured values lied outside of these bounds).

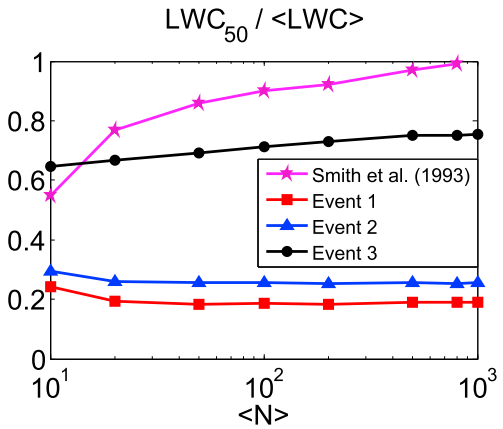


**Figure 9.** The median sampled value of LWC divided by the intrinsic ensemble value of LWC as a function of mean sample size for each of the three events, as well as the data from *Smith et al.* [1993, Figure 7]. Note how much more slowly the realistic sampling influences convergence to the intrinsic sampling when compared to $D_m$.

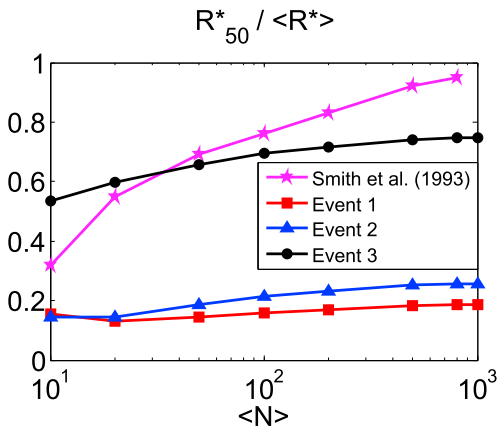


**Figure 10.** The median sampled value of $R^*$ (the fourth moment of the drop size distribution) divided by the intrinsic ensemble value of $R^*$ as a function of mean sample size for each of the three events, as well as the data from *Smith et al.* [1993, Figure 9].

Figure 6 is similar to Figure 5, except that Figure 6 shows the convergence of $R$ to its intrinsic value. Note that the $R$ axis is logarithmic due to the wide variety of different values each sample may take. Even for $\bar{N}$=1000 drops, sample-to-sample variability in $R$ of 3 orders of magnitude is still possible. Note also that in events 1 and 2, approximately 75% of all samples exhibit $R$ values smaller than the intrinsic value, more or less independent of $\bar{N}$. This is consistent with the qualitative result from *Smith et al.* [1993] that asserted that due to the skewness of the drop size distribution, the median underestimates the true value. The relative insensitivity to $\bar{N}$, however, is surprising. The combined effects of the drop size distribution and a broad (and skewed) sample-size distribution magnify the effect seen in *Smith et al.* [1993] substantially and result in much slower convergence to the median.

Figure 7 shows the approach of $Z$ to its intrinsic value with increasing $\bar{N}$. Note that convergence is expected to be slower for $Z$ since it is related to a higher moment of the drop size distribution. Each event shows a large number of individual observations that severely underestimate the true intrinsic value of $Z$.

Figures 8–11 were generated to more directly compare the results of these simulations with those in *Smith et al.* [1993]. These figures show the median sampled values of $D_m$, LWC, $R^*$, and $Z$ (respectively) normalized to the intrinsic ensemble values. These plots show not only the results from the three events simulated here but also include the results from *Smith et al.* [1993] (taken from the plots in their study). Note that convergence to the intrinsic values of LWC, $R^*$, and $Z$ is much slower here than in the *Smith et al.*'s [1993] study. There is also *substantial* event-to-event variability as evidenced, for example, by noting the different rates of convergence to $Z$ with increasing sample size for events 2 and 3.

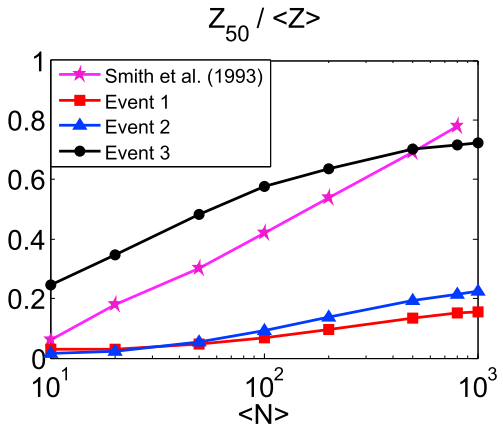


**Figure 11.** The median sampled value of $Z$ divided by the intrinsic ensemble value of $Z$ as a function of mean sample size for each of the three events, as well as the data from *Smith et al.* [1993, Figure 11].

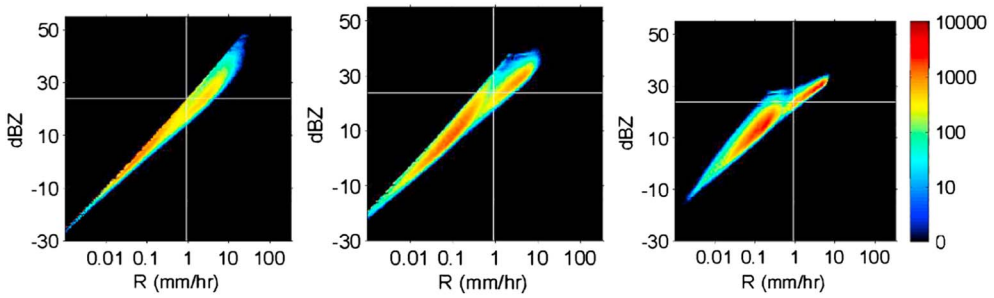


**Figure 12.** Plots of $Z$ and $R$ pairs for each simulated sample drawing on the data from event 1. Each panel shows one million $Z$-$R$ pairs. (left) $Z$-$R$ pairs with an average of 10 drops/sample, (middle) pairs with 100 drops/sample, and (right) pairs with 1000 drops/sample. Coloring is used to indicate the number of observations with the associated $Z$-$R$ pairs. In each panel, the intrinsic (underlying and constant) values of $Z$ and $R$ are indicated by the intersection of the white lines.

The qualitative observation of *Smith et al.* [1993] that higher moments converge more slowly than lower moments does seem to hold. However, convergence for *all* moments depends drastically on the details of the sampling statistics of the event and generally is far slower than convergence based on a Poisson assumption for sample size. Based on the results shown here, the suggestion that $\langle N \rangle$ of a few hundred drops may be sufficient to ensure accurate moment estimation is definitely brought into serious doubt.

Attempts were made to try and drive the simulations with even more drops per sample to determine if there is a mean sample size sufficiently large to ensure reliable estimation of each moment. Unfortunately, even sample sizes as small as $\bar{N} = 10000$ result in an increasing lack of sample-to-sample independence. Thus, the implied requirements of tens of thousands to millions of drops as asserted in *Jameson and Kostinski* [2002a] could not be explicitly tested via this method.

### 4.2. Apparent *Z-R* Relationships

The use of disdrometer measurements to derive $Z$-$R$ relationships of the form $Z = aR^b$ is very common (an exhaustive list is impossible to assemble, but a partial list of studies that use disdrometer measurements to determine $a$ and $b$ values in that relationship include *Donnadieu* [1980], *Tokay and Short* [1996], *Campos and Zawadzki* [2000], *Steiner and Smith* [2000], *Jameson and Kostinski* [2001], *Salles and Creutin* [2003], *Nzeukou et al.* [2004], *Steiner and Smith* [2004], *Lee and Zawadzki* [2005a], and *Lee et al.* [2009]. Both *Smith et al.* [1993] and *Jameson and Kostinski* [2002a] pointed out that "realistic-looking" $Z$-$R$ scatterplots can be developed for these simulations as well. It is important to reiterate that in these simulations these scatterplots are *purely statistical artifacts*! All measurements of $Z$ and $R$ would (in the absence of sampling variability) be equal to the intrinsic values of $Z$ and $R$ and a $Z$-$R$ "scatterplot" should just be a single point.

In the simulations conducted here, sampling variability again develops a quite realistic-looking $Z$-$R$ relationship (see Figures 12–14). To explore this further, one million samples were generated for each of the three events with $\bar{N} = 10$, 100, and 1000. Because so many samples were utilized, a density map is used instead of a scatterplot. The scatter in this density map is very representative of a "typical" disdrometer-derived

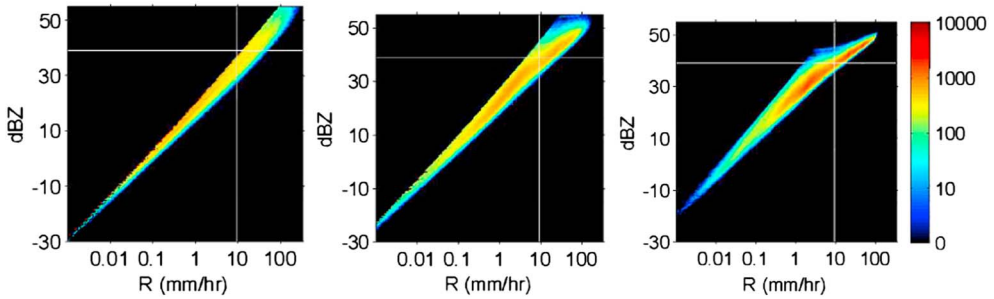


**Figure 13.** Similar to Figure 12 but these plots are for simulations using the data from event 2. Again, each panel shows one million $Z$-$R$ pairs. (left) $Z$-$R$ pairs with an average of 10 drops/sample, (middle) pairs with 100 drops/sample, and (right) pairs with 1000 drops/sample. Coloring is used to indicate the number of observations with the associated $Z$-$R$ pairs. In each panel, the intrinsic (underlying and constant) values of $Z$ and $R$ are indicated by the intersection of the white lines.

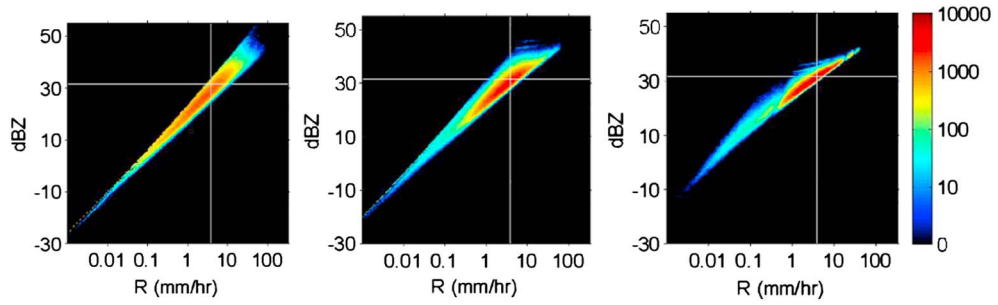


**Figure 14.** Similar to Figures 12 and 13 but these plots are for simulations using the data from event 3. Each panel shows one million *Z*-*R* pairs. (left) *Z*-*R* pairs with an average of 10 drops/sample, (middle) pairs with 100 drops/sample, and (right) pairs with 1000 drops/sample. Coloring is used to indicate the number of observations with the associated *Z*-*R* pairs. In each panel, the intrinsic (underlying and constant) values of *Z* and *R* are indicated by the intersection of the white lines.

*Z*-*R* relationship. Perhaps surprising is that the scatter does not seem to diminish very much even after $\bar{N}$ is increased by 2 orders of magnitude; the skewness of the sampling distribution may suggest that even $10^4$ drops (which, for this 2DVD data, corresponds to individual data points of nominally 1–10 min) still leaves substantial *Z*-*R* scatter.

Figure 15 shows how the fitted *a* and *b* parameters change depending on $\bar{N}$ for each of the three simulated events. As has been noted several places [e.g., *Campos and Zawadzki*, 2000; *Jameson and Kostinski*, 2002a; *Steiner and Smith*, 2000, 2004], the numerical values of *a* and *b* depend on (i) if *Z* or *R* is treated as the independent variable and (ii) if the fitting is done through a linear fit on the log-transformed data or via a least squares power law fit on the raw data. Since conducting a linear fit on the log-transformed data seems to be more common, this was the convention used here. Both variables were used as the independent variable, but the general nature of the results remained unchanged. The fitting parameters when rain rate was used as the independent variable are plotted.

Figure 15 unequivocally shows that higher $\bar{N}$ values yield larger *a* values and *b* values closer to unity. This is consistent with what was seen by *Jameson and Kostinski* [2002a] but inconsistent with *Smith et al.* [1993], where greater departures of *b* from unity were observed with increasing $\bar{N}$. However, when Poisson sampling is used (as in the *Smith et al.* [1993] study), larger $\bar{N}$ samples yield much smaller sample-to-sample variability in all moments (including *Z* and *R*) resulting in what were called "implausible" *Z*-*R*-like relationships in the *Smith et al.* [1993] study. Here,"realistic-looking" *Z*-*R* relationships are still retained for $\bar{N} = 1000$.

The fact that the *Z*-*R* relationships developed from this study mimic fairly closely those found from previous disdrometer studies perhaps should not be quite so surprising. Studies that use sequential disdrometer measurements to infer temporally linked *Z*-*R* pairs from a single storm that evolves over time are implicitly

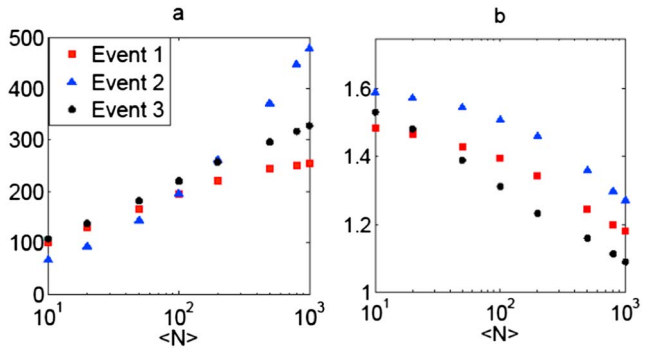


**Figure 15.** A plot showing the inferred parameters *a* and *b* in the relationship $Z = aR^b$ for the three events. Each displayed point is the result of a fit through 10,000 simulated *Z* and *R* pairs with the mean number of drops as shown. Results shown here are consistent with those seen in *Jameson and Kostinski* [2002b] (where uniform sample-size distribution functions were used and $b \to 1$ as $\bar{N} \to \infty$) but inconsistent with those seen in *Smith et al.* [1993] (where Poisson sample-size distribution functions were used and *b* increases with increasing *N*).

engaging in the same types of event-averaging that are explicitly employed here. The only difference in the approach created here compared to disdrometer-driven *Z-R* studies is (i) the explicit limitation to a single event in this study and (ii) the explicit separations of the distributions governing drop size and the distributions governing drop numbers in each sample. Either way, both traditional disdrometer-driven *Z-R* studies and the study presented here suffer from the same limitations in trying to infer properties of a statistical relationship ultimately desired for use in a context (e.g., radar measurements) where the associated sample sizes (and, consequently, the sampling fluctuations) are *drastically* different than the measured system. This study once again highlights the well-known issues associated with the degree to which disdrometer-derived data can be appropriately used to infer radar-scale information.

## 5. Discussion

### 5.1. Validity of These Simulations

It should be emphasized that these simulations, though driven by real data, still are just Monte Carlo simulations. In particular, the homogenization process that results in the creation of the sample size distribution function is unphysical. Figures 1–3 clearly demonstrate that the sample size distributions used to drive the simulations presented in this manuscript are much broader than the Poisson distributions utilized in *Smith et al.* [1993], *Smith and Kliche* [2005], and *Smith et al.* [2009] and much more skewed than the uniform distributions used in *Jameson and Kostinski* [2002a].

In reality, though it is expected that the sampling distribution likely should be broader than Poisson and skewed, it is invalid to assume that the actual detection of different numbers of drops in different samples is purely an issue of sampling variability. On storm scales, rain is typically statistically nonstationary, with deterministic trends to the drop arrival rate over time that exist independent of any sampling considerations. The degree to which underlying drop arrival rates change on minute to minute or second to second scales is still very much in question (see, e.g., *Jameson et al.* [2015a] for just one example of many papers investigating this question), but truly homogeneous rain does not seem to exist over time scales of hours or longer except under extremely rare conditions [*Larsen et al.*, 2005].

As such, the quantitative results of these simulations should not be accepted as "the new standard." Even if the results presented here are limited to being applied to "homogeneous rain," the underlying sampling statistics that were used to drive the simulation were not homogeneous from the outset but rather homogenized in a rather artificial way. This unphysical "homogenization" manifests itself in a number of ways, including (i) the larger than physically expected value of $p_s(0|\bar{N})$ for all events, (ii) the artificial removal of any temporal structure on time scales longer than $\bar{N}\tau$, and (iii) the explicit decoupling of $p_s(k|\bar{N})$ from $p(D)$ in the simulation structure. The degree to which each of these may influence the quantitative results presented here is, at present, uncertain.

Because of the above caveats, this study merely attempts to establish a new upper bound for the effect of sampling variability. If the conclusion of $\bar{N} \sim 100$ drops for convergence developed in the *Smith et al.* [1993] study may have been a bit optimistic, any conclusions drawn based on the simulations here may be a bit pessimistic—with the "truth" associated with sampling variability likely lying somewhere between.

### 5.2. Sample-Based Estimation of Moments

Simulation-based studies of sampling variability, like the one conducted here, attempt to determine necessary conditions to ensure that a real sample adequately represents the true state of the system. Ideally, an individual sample would be both an unbiased and an accurate estimator of the desired quantity. In an effort to determine whether a sampled system obeys these criteria, one can explore the median sample value (which identifies whether the estimator is biased) as well as the breadth of the observed samples (which helps to determine the accuracy with which a single sample could estimate a quantity). The results in Figures 5–11 clearly and strongly suggest that even statistically stationary rain requires more than $\bar{N} = 1000$ drops for reliable parameter estimation.

Although several of the qualitative results of both *Smith et al.* [1993] and *Jameson and Kostinski* [2002a] hold, use of the quantitative estimates of the *Smith et al.* [1993] study (as has been done in, for example, *Smith and Kliche* [2005], *Zhang et al.* [2006], *Brandes et al.* [2007], and *Brandes et al.* [2008] may be rather misleading and overly optimistic.

At this point, it appears that the only quantity that can be estimated in realistic sampling to reliable levels of estimation with fewer than $10^4$ drops may be $D_m$. Future work will hopefully reveal if there is a point at which LWC, $R$, $R^*$, and $Z$ can be reliably estimated in homogeneous rain.

This, of course, still leaves the practical practitioner in a bind, as most instruments are incapable of measuring a large enough sample area to obtain sufficient numbers of drops to reliably estimate the desired rain properties. Adding more instruments to a site to increase the sample area is often cost prohibitive, even if the instrument-to-instrument issues revealed in other studies [e.g., *Donnadieu*, 1980; *Campos and Zawadzki*, 2000; *Habib et al.*, 2001; *Ciach*, 2003; *Salles and Creutin*, 2003; *Miriovsky et al.*, 2004; *Tokay et al.*, 2005; *Krajewski et al.*, 2006; *Jaffrain and Berne*, 2011; *Thurai et al.*, 2011; *Tokay et al.*, 2013] could be overcome.

There do exist some possible ways out of the conundrum. Use of dual-polarimetric radar to determine rain properties may enable the investigator to infer bulk statistics without being subject to such limited sample areas. Additionally, the Precipitation Occurrence Sensor System (POSS) described and utilized in *Sheppard* [1990], *Sheppard and Joe* [1994], *Campos and Zawadzki* [2000], and *Lee and Zawadzki* [2005a] may have sufficiently large sample area to avoid several of the sampling issues explored in this manuscript. Alternatively, there may be some ways to process or analyze the data in such a way as to minimize (or at least better bound) the effect that sampling issues can have on inferring particular rain properties like rain rate or reflectivity factor. (One method that has already been utilized to attempt to do this can be found in the Sequential Intensity Filtering Technique described and used in *Lee and Zawadzki* [2005b, 2005a].)

### 5.3. Nature of Convergence

Although definitive conclusions are difficult to draw with only three events, it is interesting that the strongest and weakest events (in terms of intrinsic $D_m$, $R$, and $Z$) converge more slowly with increasing $\bar{N}$ than the intermediate storm. This suggests that no simple relationship that estimates the rate of convergence to adequate sampling based on something as simple as $\langle R \rangle$ is likely; a complicated interplay between an event-averaged variable like $\langle R \rangle$ and measured values of $p_s(k|\bar{N})$ and $p_d(D)$ may be present.

### 5.4. *Z-R* Relationships as Statistical Artifacts

The *Z-R* "curve" for each simulated event should, in the presence of perfect sampling, be a single point. As noted earlier, however, realistic-looking *Z-R* scatterplots and very reasonable looking power law fits result after simulating these measurements. With sampling variability potentially accounting for 5 or more orders of magnitude of scatter along both the $R$ and $Z$ axes even for $\bar{N} = 10^3$ drops, the true utility of such *Z-R* scatterplots in real applications in the presence of other sources of uncertainty (instrumental uncertainty and physical changes) is naturally brought into question. Efforts have been made by some investigators [e.g., *Hagen and Yuter*, 2003; *Yuter and Houze*, 2003; *Houze et al.*, 2004] to use longer time period samples to develop estimates for $Z$ and $R$, but the results presented here suggest that even this may not remove sampling variability effects.

Of course, countless manuscripts have been attributing differing $a$ and $b$ parameters to physical differences in rainfall type (and hence the size distribution) for decades. Perhaps the results here may not be inconsistent with the results of these studies—the ultimate conclusions, though similar, come about via a statistical rather than meteorological reasoning. For example, strong summer convective storms often will have a larger density of drops and thus typically will have larger numbers of measured drops per sample if samples are made on a "fixed-time" basis. These stronger (higher $\bar{N}$ events) naturally have a smaller sampling effect and—via the mechanism presented in Figure 15—end up with lower $b$ values. Thus, it is unclear as to whether the lowering of the "b" value associated with heavy rainfall is primarily due to a physical change in the relative abundance of drops of different sizes, or if it is because the statistics of sampling variability are changed. It is probably most likely that both factors contribute in some (unknown) proportion.

It has been argued by some investigators that if rain is truly statistically homogeneous, then the underlying *Z-R* relationship should be linear. Modeling and experimental studies [see, e.g., *List*, 1988; *Jameson and Kostinski*, 2002a; *Prat and Barros*, 2009] have revealed this linear behavior for high rain rates. This is consistent with the results in Figure 15 and can be explained physically as well. As $\bar{N}$ increases, the bias in all moments decreases and an individual sample becomes a more reliable estimator of both $R$ and $Z$. If $R$ and $Z$ are adequately sampled in homogeneous rain, the resulting *Z-R* relationships physically *must* be linear [*Jameson and Kostinski*, 2001]. Although this is not necessarily conclusive, the results from these simulations suggest that perhaps nonlinear *Z-R* relationships derived from disdrometer data say more about sampling variability than any underlying meteorology.

## 6. Conclusions

Some of the qualitative results from these simulations are consistent with the results seen in *Smith et al.* [1993] and *Jameson and Kostinski* [2002a]. In particular are the following:

1. A single observation (sample) will (on average) do a better job of estimating all quantities if $\bar{N}$ is larger. (This is consistent with both previous studies.)
2. Due to sampling variability alone, realistic-looking *Z*-*R* scatterplots can be generated and, for some values of $\bar{N}$, both the *a* and *b* parameters take on realistic values. (This is consistent with both previous studies.)
3. The median of sampled values underestimates the true intrinsic value of all quantities, with a more severe underestimation occurring when $\bar{N}$ is small. (This is consistent with the findings in *Smith et al.* [1993].)
4. The exponent of the fit to the simulated *Z*-*R* scatterplot converges to $b = 1$ when $\bar{N}$ increases. (This is consistent with the findings in *Jameson and Kostinski* [2002a].)

The work done here helps to crystallize these conclusions, which may be applicable under a wide variety of conditions.

In general, however, the quantitative results presented in *Smith et al.* [1993] are significantly altered when put into the context of non-Poisson sampling. Though there are situations where homogeneous rainfall may exist over substantial time intervals with Poisson-like statistics [see, e.g., *Larsen et al.*, 2005], every indication is that such rainfall is quite rare. A more common (though perhaps somewhat artificially extreme) scenario is explored with the simulations conducted here, where the distribution of the number of drops in each sample is quite broad. In these scenarios, convergence to the underlying intrinsic values is slower than previously estimated, and sample-to-sample variability is broader than previously estimated.

Given that disdrometers have physical limitations on their sample volumes based on their physical design, and given that real rain may have substantial underlying variability on scales shorter than the time required to acquire $10^4$ drops per sample, the results here suggest that it is at least possible that using a single disdrometer observation to estimate "true" rainfall properties may be a fool's errand; it may be the case that to obtain enough drops to reliably estimate any of the rainfall properties, measurements of such long duration would be required that the assumption of statistical homogeneity no longer applies.

The ultimate impact on the community for rain measurement research is in understanding and quantifying uncertainty. The *Smith et al.* [1993] study is often cited to note that "sampling variability exists," and it is also sometimes used to bound the uncertainty introduced by this variability. Those bounds are too low, and rain data should be reevaluated in light of this realization. Current efforts with small-scale multidisdrometer arrays [e.g., *Tapiador et al.*, 2010; *Jaffrain et al.*, 2011; *Larsen and Teves*, 2015], instruments with larger sample areas like the Precipitation Occurrence Sensor System (POSS), and different processing techniques like the sequential intensity filtering technique (SIFT) will hopefully reveal definitive estimates of the uncertainty associated with bulk parameter estimation from individual disdrometer measurements.

## Appendix A: Sampling Strategies Used in the Literature

This works expand upon and extend similar studies conducted by *Smith et al.* [1993], *Jameson and Kostinski* [2002a], *Smith and Kliche* [2005], and *Smith et al.* [2009]. These four studies—in combination with the one

**Table A1.** Sampling Strategies Utilized in This Study and Similar Other Computational Sampling Studies

| Study | Number of Samples at Each $< N >$ | Distribution for $< N >$ | Form of DSD |
|---|---|---|---|
| *Smith et al.* [1993] | 2000[a] | Poisson | Exponential |
| *Jameson and Kostinski* [2002a] | 10,000[b] | Uniform | Exponential |
| *Smith and Kliche* [2005] | 10,000[c] | Poisson | Exponential |
| *Smith et al.* [2009] | 20,000[d] | Poisson | Gamma |
| Larsen and O'Dell (Three separate events) | 110,000 (each event) | Homogenized Data (empirical) | Homogenized Data (empirical) |

[a]Estimated from looking at the histograms in *Smith et al.* [1993].
[b]From caption of Figure 5 in *Jameson and Kostinski* [2002a].
[c]From *y* axis of Figures 9–11 in *Smith and Kliche* [2005].
[d]From *y* axis of Figures 4–6 in *Smith et al.* [2009].

presented in this manuscript—all utilized rather similar sampling strategies in computational simulations. Table A1 is supplied to make very clear the similarities and differences between the methods used in these studies.

**Acknowledgments**
The data used in this study can be supplied via request to the corresponding author of this manuscript. This research was supported by the U.S. National Science Foundation under grants AGS-1230240 and AGS-1532977. The authors would like to thank Michael Chute, Joerael Harris, and Robert Lemasters for their efforts in constructing and maintaining the instrument array and Joshua Teves for assistance with data processing. The authors would also like to thank Arthur Jameson, Alexander Kostinski, and the anonymous reviewers for their helpful insights.

## References

Brandes, E. A., K. Ikeda, G. Zhang, M. Schönhuber, and R. M. Rasmussen (2007), A statistical and physical description of hydrometeor distributions in Colorado snowstorms using a video disdrometer, *J. Appl. Meteorol. Climatol.*, *46*, 634–650, doi:10.1175/JAM2489.

Brandes, E. A., K. Ikeda, G. Thompson, and M. Schoönhuber (2008), Aggregate terminal velocity/temperature relations, *J. Appl. Meteorol. Climatol.*, *47*, 2729–2736, doi:10.1175/2008JAMC1869.

Campos, E., and I. Zawadzki (2000), Instrumental uncertainties in $Z$-$R$ relations, *J. Appl. Meteorol.*, *39*, 1088–1102.

Ciach, G. J. (2003), Local random errors in tipping-bucket rain gauge measurements, *J. Atmos. Oceanic Technol.*, *20*, 752–759.

Donnadieu, G. (1980), Comparison of results obtained with the VIDIAZ spectropluviometer and the Joss-Waldvogel rainfall disdrometer in a "rain of a thundery type", *J. Appl. Meterol.*, *19*, 593–597.

Foote, G. B., and P. S. du Toit (1969), Terminal velocity of raindrops aloft, *J. Appl. Meteorol.*, *8*, 249–253.

Gires, A., I. Tchiguirinskaia, D. Schertzer, and A. Berne (2015), 2DVD data revisited: Multifractal insights into cuts of the spatiotemporal rainfall process, *J. Hydrometeorol.*, *16*, 548–562, doi:10.1175/JHM-D-14-0127.

Habib, E., W. F. Krajewski, and A. Kruger (2001), Sampling errors of tipping-bucket rain gauge measurements, *J. Hydrol. Eng.*, *6*, 159–166.

Hagen, M., and S. E. Yuter (2003), Relations between radar reflectivity, liquid-water content, and rainfall rate during the MAP SOP, *Q. J. R. Meteorolog. Soc.*, *129*, 477–493, doi:10.1256/qj.02.23.

Houze, R. A., S. Brodzik, C. Schumacher, S. E. Yuter, and C. R. Williams (2004), Uncertainties in oceanic radar rain maps at Kwajelein and implications for satellite validation, *J. Appl. Meteorol.*, *43*, 1114–1132.

Jaffrain, J., and A. Berne (2011), Experimental quantification of the sampling uncertainty associated with measurements from PARSIVEL disdrometers, *J. Hydrometeorol.*, *12*, 352–370, doi:10.1175/2010JHM1244.

Jaffrain, J., A. Studzinski, and A. Berne (2011), A network of disdrometers to quantify the small-scale variability of the raindrop size distribution, *Water Resour. Res.*, *47*, W00H06, doi:10.1029/2010WR009872.

Jameson, A., and M. Larsen (2016), The variability of rainfall rate as a function of area, *J. Geophys. Res. Atmos.*, *121*(2), 746–758, doi:10.1002/2015JD024126.

Jameson, A. R., and A. B. Kostinski (2000), Fluctuation properties of precipitation. Part VI: Observiations of hyperfine clustering and drop size distribution structures in three-dimensional rain, *J. Atmos. Sci.*, *57*, 373–388.

Jameson, A. R., and A. B. Kostinski (2001), Reconsideration of the physical and empirical origins of $Z$-$R$ relations in radar meteorology, *Q. J. R. Meteorolog. Soc.*, *127*, 517–538.

Jameson, A. R., and A. B. Kostinski (2002a), Spurious power-law relations among rainfall and radar parameters, *Q. J. R. Meteorolog. Soc.*, *128*, 2045–2058.

Jameson, A. R., and A. B. Kostinski (2002b), When is rain steady?, *J. Appl. Meteorol.*, *41*, 83–90.

Jameson, A. R., A. B. Kostinski, and A. Kruger (1999), Fluctuation properties of precipitation. Part IV: Finescale clustering of drops in variable rain, *J. Atmos. Sci.*, *56*, 82–91.

Jameson, A. R., M. L. Larsen, and A. B. Kostinski (2015a), Disdrometer network observations of finescale spatial-temporal clustering in rain, *J. Atmos. Sci.*, *72*, 1648–1666.

Jameson, A. R., M. L. Larsen, and A. B. Kostinski (2015b), On the variability of drop size distributions over areas, *J. Atmos. Sci.*, *72*, 1386–1397.

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

Kostinski, A. B., and A. R. Jameson (1999), Fluctuation properties of precipitation. Part III: On the ubiquity and emergence of the exponential drop size spectra, *J. Atmos. Sci.*, *56*, 111–121.

Krajewski, W. F., A. Kruger, C. Caracciolo, P. Gole, L. Barthes, J.-D. Creutin, J.-Y. Delahaye, E. I. Nikolopoulos, F. Ogden, and J.-P. Vinson (2006), DEVEX-disdrometer evaluation experiment: Basic results and implications for hydrologic studies, *Adv. Water Res.*, *29*, 311–325.

Kruger, A., and W. F. Krajewski (2002), Two-dimensional video disdrometer: A description, *J. Atmos. Oceanic Technol.*, *19*, 602–617.

Larsen, M. L., and J. B. Teves (2015), Identifying individual rain events with a dense disdrometer network, *Adv. Meteorol.*, *2015*, 582782, doi:10.1155/2015/582782.

Larsen, M. L., A. B. Kostinski, and A. Tokay (2005), Observation and analysis of uncorrelated rain, *J. Atmos. Sci.*, *62*, 4071–4083.

Larsen, M. L., A. B. Kostinski, and A. R. Jameson (2014), Further evidence for superterminal raindrops, *Geophys. Res. Lett.*, *41*(19), 6914–6918, doi:10.1002/2014GL061397.

Larsen, M. L., T. B. Hayward, and J. B. Teves (2015), Scaling properties of raindrop size distributions as measured by a dense array of optical disdrometers, *J. Hydrol.*, *521*, 424–432, doi:10.1016/j.jhydrol.2014.12.016.

Lee, C. K., G.-W. Lee, I. Zawadzki, and K.-E. Kim (2009), A preliminary analysis of spatial variability of raindrop size distributions during stratiform rain events, *J. Appl. Meteorol. Climatol.*, *48*, 270–283, doi:10.1175/2008JAMC1877.

Lee, G.-W., and I. Zawadzki (2005a), Variability of drop size distributions: Time-scale dependence of the variability and its effects on rain estimation, *J. Appl. Meteorl.*, *44*, 241–255.

Lee, G.-W., and I. Zawadzki (2005b), Variability of drop size distributions: Noise and noise filtering in disdrometric data, *J. Appl. Meteorol.*, *44*, 634–652.

List, R. (1988), A linear radar reflectivity-rainrate relationship for steady tropical rain, *J. Atmos. Sci.*, *45*, 3564–3572.

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

Miriovsky, B. J., et al. (2004), An experimental study of small-scale variability of radar reflectivity using disdrometer observations, *J. Appl. Meteorol.*, *43*, 106–118.

Nzeukou, A., H. Sauvageot, A. D. Ochou, and C. M. F. Kebe (2004), Raindrop size distribution and radar parameters at Cape Verde, *J. Appl. Meteorol.*, *43*, 90–105.

Prat, O. P., and A. P. Barros (2009), Exploring the transient behavior of $Z$-$R$ relationships: Implications for radar rainfall estimation, *J. Appl. Meteorol. Climatol.*, *48*, 2127–2143, doi:10.1175/2009JAMS2165.

Salles, C., and J.-D. Creutin (2003), Instrumental uncertainties in $Z$-$R$ relationships and raindrop fall velocities, *J. Appl. Meteorol.*, *42*, 279–290.

Sheppard, B. (1990), Measurement of raindrop size distribution using a small Doppler radar, *J. Atmos. Oceanic Technol.*, *7*, 255–268.

Sheppard, B., and P. Joe (1994), Comparison of raindrop size distribution measurements by a Joss-Waldvogel disdrometer, a PMS 2DG spectrometer, and a POSS Doppler radar, *J. Atmos. Oceanic Technol.*, *11*, 874–887.

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. Meteorol.*, *44*, 1195–1205.

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

Smith, P. L., Z. Liu, and J. Joss (1993), A study of sampling-variability effects in raindrop size observations, *J. Appl. Meteorol.*, *32*, 1259–1269.

Steiner, M., and J. A. Smith (2000), Reflectivity, rain rate, and kinetic energy flux relationships based on raindrop spectra, *J. Appl. Meteorol.*, *39*, 1923–1940.

Steiner, M., and J. A. Smith (2004), Scale dependence of radar-rainfall rates—An assessment based on raindrop spectra, *J. Hydrometeorol.*, *5*, 1171–1180.

Tapiador, F. J., R. Checa, and M. de Castro (2010), An experiment to measure the spatial variability of rain drop size distribution using sixteen laser disdrometers, *Geophys. Res. Lett.*, *37*, L16803, doi:10.1029/2010GL044120.

Thurai, M., W. A. Petersen, A. Tokay, C. Shultz, and P. Gatlin (2011), Drop size distribution comparisons between Parsivel and 2-D video disdrometers, *Adv. Geosci.*, *30*, 3–9, doi:10.5194/adgeo-30-3-2011.

Tokay, A., and D. A. Short (1996), Evidence from tropical raindrop spectra of the origin of rain from stratiform versus convective clouds, *J. Appl. Meteorol.*, *35*, 355–371.

Tokay, A., P. G. Bashor, and K. R. Wolff (2005), Error characteristics of rainfall measurements by collocated Joss-Waldvogel disdrometers, *J. Atmos. Oceanic Technol.*, *22*, 513–527.

Tokay, A., W. A. Petersen, P. Gatlin, and M. Wingo (2013), Comparison of raindrop size distribution measurements by collocated disdromters, *J. Atmos. Oceanic Technol.*, *30*, 1672–1690, doi:10.1175/JTECH-D-12-00163.

Yuter, S. E., and R. A. Houze (2003), Microphysical modes of precipitation growth determined by S-band vertically pointing radar in orographic precipitation during MAP, *Q. J. R. Meteorolog. Soc.*, *129*, 455–476.

Zhang, G., J. Sun, and E. A. Brandes (2006), Improving parameterization of rain microphysics with disdrometer and radar observations, *J. Atmos. Sci.*, *63*, 1273–1290, doi:10.1175/JAS3680.1.