P14R.5 A SIMULATION APPROACH TO SAMPLING EFFECTS IN RAINDROP SIZE DISTRIBUTION MEASUREMENTS IN NON-STATIONARY RAIN

Alexis Berne*and Remko Uijlenhoet

Environmental Sciences, Wageningen University, The Netherlands

1 INTRODUCTION

The raindrop size distribution (DSD) is fundamental for the interpretation of radar rainfall measurements, especially with respect to the parameterization of the power laws used to describe the relations between the bulk rain variables (e.g. radar reflecvity Z, rainfall intensity R).

The different types of sensors used to estimate the DSD (e.g. Joss-Waldvogel disdrometer, spectropluviometer and video-disdrometer) have a limited sampling volume or a limited sampling surface. The issue of uncertainty in DSD characterization due to sampling errors (independently of measurement errors) is crucial and has been recognized for a long time (e.g. Joss and Waldvogel, 1969). To quantify the sampling error, Gage et al. (2004, GCWT hereafter) have proposed to use time series from two collocated disdrometers.

The objectives of this paper are to investigate, using a stochastic DSD model, (i) the accuracy of the estimator proposed by GCWT and (ii) to quantify the influence of the disdrometer sampling error on the derived Z-R relations. First, the stochastic DSD model is presented. Then the disdrometer sampling error is quantified. Finally, the uncertainty on the parameters of the derived Z-R relations is analyzed.

2 DSD SIMULATOR

The DSD simulator used in the following has been proposed by Berne and Uijlenhoet (2005). It enables to generate DSD profiles in time corresponding to non-stationary rainfall. It is based on the exponential DSD, which two parameters N_t and λ are considered to be random variables

$$N(D|N_t, \lambda) = N_t \, \lambda \, e^{-\lambda D} \; , \qquad (1)$$

where $N(D|N_t, \lambda)dD$ denotes the drop concentration in the diameter interval $[D, D + dD]$ given N_t and λ. The latter are assumed to be jointly lognormally distributed and their logarithms are assumed to follow a bivariate first order vector auto-regressive process. The parameters have been derived from DSD measurements for a 4-hour rain event, collected during the HIRE'98 experiment in Marseille, France (see Table 1). To be consistent with the data presented by GCWT, we generate a profile corresponding to 8 hours of disdrometer measurements, with a time resolution of 60 s.

* *Corresponding author address:* Alexis Berne, Environmental Sciences, Nieuwe Kanaal 11, 6709 PA Wageningen, The Netherlands; e-mail: Alexis.Berne@wur.nl

Table 1: Mean, standard deviation and characteristic time of $N' = \ln N_t$ and $\lambda' = \ln \lambda$ deduced from HIRE'98 data at a 60 s time step.

	Mean	Std	θ (s)
N'	6.52	0.57	390
λ'	0.87	0.27	390

To simulate the sampling process of a JW disdrometer, Poissonian fluctuations are added to every time interval of the profile. The simulator is used here to generate 100 sampled profiles corresponding to the same reference profile. This is equivalent to having 100 collocated disdrometers sampling the same rain event.

The simulator provides profiles of N_t and λ, with a given resolution, from which the bulk rain variables can be derived. We focus in this paper on the radar reflectivity Z for a 10 cm wavelength weather radar (i.e. S-band, so attenuation effects due to precipitation are negligible). The main advantage of a simulation approach is the possibility to generate a large number of sampled time series, and therefore the possibility to use a Monte Carlo approach to derive robust statistics on the sampling effects.

3 QUANTIFICATION OF THE SAMPLING ERROR ON THE RADAR REFLECTIVITY

The quantification of the sampling error for a non-stationary rain event is complex due to the mixing of sampling fluctuations and natural variability. The variance of Z over the population of N_t and λ can be written as

$$\mathrm{Var}\,[Z] = \mathrm{Var}\,[\mathrm{E}\,[Z|N_t, \lambda]] + \mathrm{E}\,[\mathrm{Var}\,[Z|N_t, \lambda]] \; , \quad (2)$$

where E denotes the expectation and Var the variance. Equation (2) shows that the total variability, Var[Z], is the sum of a first term, Var[E[$Z|N_t, \lambda$]], which represents the natural variability, and a second term, E[Var[$Z|N_t, \lambda$]], which represents the variability due to the sampling effect. Of course, in Eq. (2) Z can be replaced by any bulk rain variable (e.g. the rain rate R).

For two collocated disdrometers close enough to sample the same DSD population, the two sampled values Z_1 and Z_2 are independent and identically distributed for given values of N_t and λ, which implies

$$\begin{cases} \mathrm{E}\left[\Delta Z|N_t,\lambda\right] &= 0 \\ \mathrm{Var}\left[\Delta Z|N_t,\lambda\right] &= 2\,\mathrm{Var}\left[Z|N_t,\lambda\right] \end{cases}, \quad (3)$$

where $\Delta Z = Z_1 - Z_2$. Writing Eq. (2) with ΔZ instead of Z and using Eq. (3) yields

$$\mathrm{Var}\left[\Delta Z\right] = \mathrm{E}\left[\mathrm{Var}\left[\Delta Z|N_t,\lambda\right]\right] = 2\,\mathrm{E}\left[\mathrm{Var}\left[Z|N_t,\lambda\right]\right] . \quad (4)$$

Equation (4) shows that taking the difference of the two sampled Z values, as proposed by GCWT, removes the natural variability and allows to quantify the mean sampling variability alone. These expressions are valid no matter if Z is expresssed in linear (mm^6 m^{-3}) or in logarithmic (dBZ) units. In the sequel, Z will be expressed in dBZ.

3.1 Estimation of Var[ΔZ]

Note that Eq. (4) has been derived for the expectation calculated over the population of N_t and λ. In practice, we only have access to a subset of the population of N_t and λ, through the measured Z time series. Therefore, the validity of Eq. (4) over a profile (i.e. one realization of the bivariate (N_t, λ)-process) must be investigated, as well as the accuracy of the estimation of Var[ΔZ] from the measurements of two collocated disdrometers.

To test the validity of Eq. (4) along a profile, we analyze the distribution of (Var[ΔZ]-2E[Var[$Z|N_t, \lambda$]])/Var[ΔZ] for 100 simulated reference profiles (representing 100 rain events). Recall that conditional upon each reference profile, we simulate 100 sampled profiles (representing 100 disdrometers). Hence, for a given reference profile, Var[ΔZ] can be estimated as the average of the variance of ΔZ calculated over the 4950 (100×99/2) possible pairs. E[Var[$Z|N_t, \lambda$]] can be estimated directly as we dispose of 100 sampled Z values for every time interval of the reference profile. In this manner, we obtain one value of (Var[ΔZ]-2E[Var[$Z|N_t, \lambda$]])/Var[ΔZ] per reference profile. The mean over the 100 reference profiles is found to be about 10^{-5} and 80% of the values are within the interval $\pm 3 \times 10^{-4}$. Therefore Eq. (4) can be considered valid along a profile (i.e. per realization) for the given ratio between the length of the profile and the characteristic time scale. The mean sampling error is accurately quantified by Var[ΔZ]/2 (if derived from a large number of sampled profiles).

To assess the accuracy of the estimation of Var[ΔZ] with only two sampled profiles (corresponding to two disdrometers), the probability distribution of Var[ΔZ] is studied. First, the distribution of Var[ΔZ] calculated using the 4950 pairs of sampled profiles corresponding to one given reference profile is plotted in the top panel of Fig. 1. The difference between an estimate from one single pair of disdrometers and the mean value of Var[ΔZ] appears to be limited (80% of the values are within an interval of ± 10% around the mean).

To quantify the variability of this distribution for different reference profiles, the distribution of the coefficient of variation of the distribution of Var[ΔZ] for the 100 reference profiles is plotted in the bottom panel of Fig. 1. The mean is about 0.079 and the 10% and 90% quantiles are about 0.074 and 0.083, respectively. This indicates that the distribution plotted in the top panel of Fig. 1 is representative for different reference profiles for the employed parameters of the stochastic DSD model. In summary, Var[ΔZ]/2, calculated from two collocated disdrometers and hence closely related to the estimator proposed by GCWT, provides a relatively accurate estimate of the mean sampling error affecting radar reflectivity time series derived from JW disdrometers.

3.2 Influence of the length of the profile

The results presented in the previous section have been derived from simulated rain profiles of 8 hours, which corresponds to about 74 times the characteristic time scale of the studied rainfall (see Table 1). This section is devoted to the analysis of the influence of the length of the profile on the accuracy of the sampling error estimation. Figure 2 presents the evolution of the mean of the coefficient of variation of the distributions of Var[ΔZ] as a function of the ratio of the length of the profile and the characteristic time scale. It is clear from Fig. 2 that the length of the measurement series must be significantly larger than the characteristic time scale (i.e. the decorrelation time) in order to obtain an accurate estimation of the mean sampling error. For instance, to achieve an accuracy corresponding to a coefficient of vari-

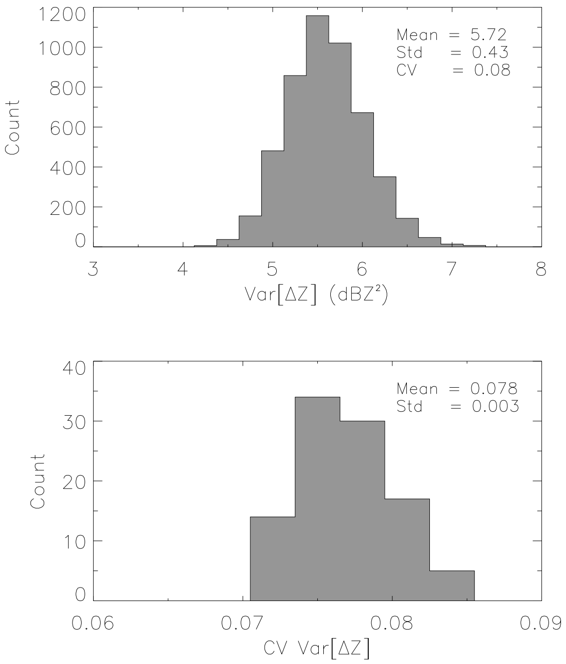

Figure 1: Top panel: distribution of Var[ΔZ] values from 4950 simulated pairs of disdrometers for a given reference profile. Bottom panel: distribution of the coefficient of variation of the distributions of Var[ΔZ] from 100 reference profiles.

ation of about 0.15 (0.10), one needs a time series which is about 20 (50) times longer than the characteristic time scale. In practice, these estimates provide a lower bound, as mixing of different types of precipitation is likely to occur more often when the time series become longer.

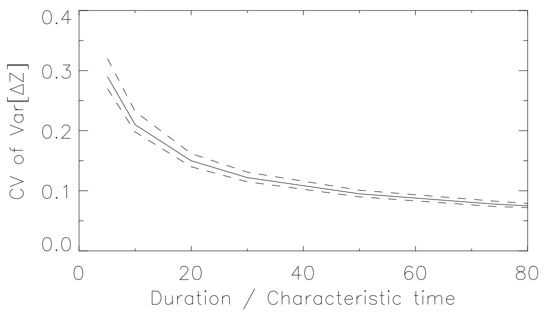

Figure 2: Coefficient of variation of the distibutions of Var[ΔZ] as a function of the ratio of the length of the time series and the characteristic time scale. The solid line indicates the mean; the dashed lines indicate the 10% and 90% quantiles.

4 UNCERTAINTY DUE TO SAMPLING EFFECTS IN THE Z-R RELATION

DSD measurements are often used to derive the parameter values of a power law of the form $Z = a\,R^b$ to relate the radar reflectvity Z to the rainfall intensity R.

4.1 Variability of the parameters a and b

For a given reference profile, a reference power law Z-R relation can be fitted (with a non-linear regression technique). A power law can also be fitted on each sampled profile and the relative deviation from the reference relation can be calculated. In this manner, we obtain the distribution of the "relative error" on the prefactor a and the exponent b for each reference profile. For instance, Figure 3 presents the distribution of the relative error on a and b for a given reference profile.

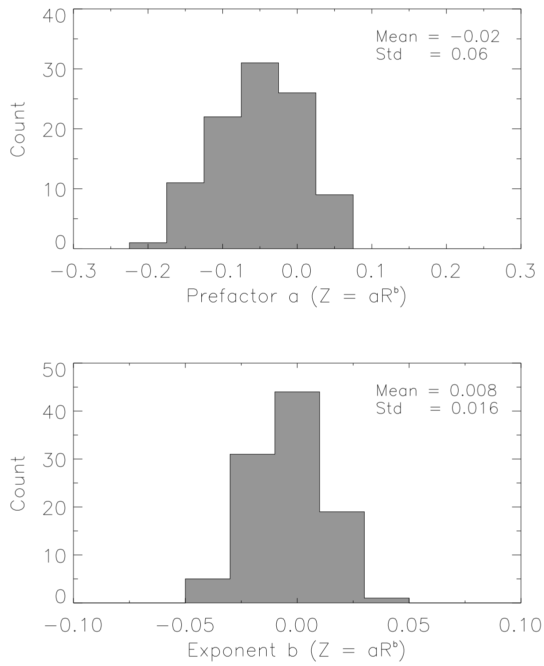

Figure 3: Distribution of the relative error on the prefactor a (top panel) and on the exponent b (bottom panel), from 100 sampled profiles corresponding to the same reference profile.

The uncertainty on the Z-R relation due to sampling errors can be quantified as the standard deviation of this distribution for the parameters a and b. To have more robust statistics, we analyze the distribution of this standard deviation for 100 reference profiles. Figure 4 shows that in case of an exponential DSD, the disdrometer sampling error generates a limited uncertainty in the parameters of the Z-R relation: about 10% for the prefactor and about 3% for the exponent. It must be noted that the prefactor is more affected than the exponent. These values correspond to mean estimates, and the individual deviations can be larger.

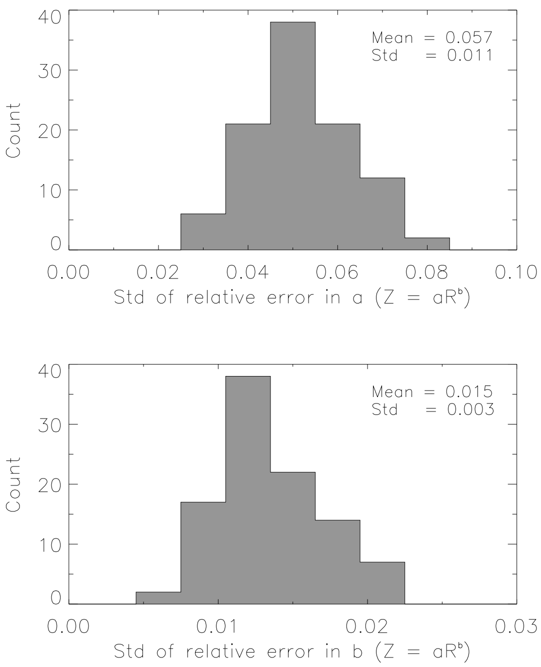

Figure 4: Distribution of σ_a-values (top panel) and σ_b-values (bottom panel) from 100 reference profiles.

4.2 Influence of the length of the profile

The length of the profile has an influence on the values of the fitted power laws. Figure 5 presents the evolution of the mean of the standard deviation of the relative error on the parameters a and b as a function of the ratio between the length of the profile and the characteristic time scale. When the length of the profile increases, the uncertainty on both parameters decreases, as well as the dispersion of the distribution (indicated by the 10% and 90% quantiles). For instance, the mean standard deviation of the relative error on a is about 10% for a length of 20 times the characteristic time scale.

5 CONCLUSIONS

A simulation framework, based on a stochastic model capable of simulating the sampling process of a JW disdrometer in non-stationary rain, has been used to show that the estimator proposed by

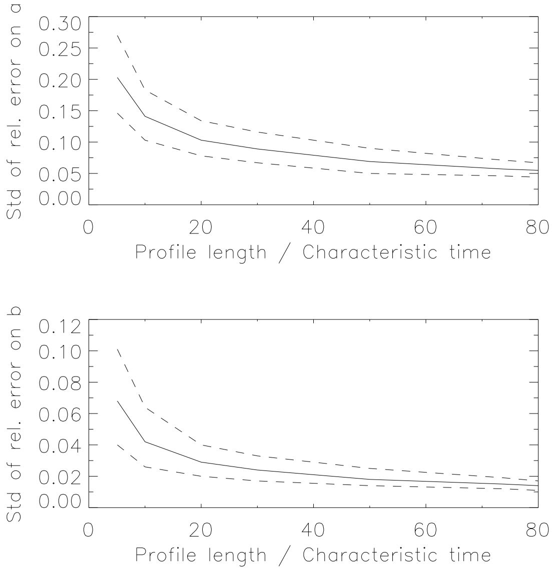

Figure 5: Mean standard deviation of the relative error on the prefactor (top panel) and the exponent (bottom panel) of the Z-R relation as a function of the ratio of the length of the time series and the characteristic time scale. The solid line indicates the mean; the dashed lines indicate the 10% and 90% quantiles.

GCWT is reliable, but that the ratio between the length of the profile and the characteristic time scale has a significant influence on its accuracy. This approach also allows to quantify the uncertainty due to the sampling error in the parameterization of the Z-R relation, which appears to be limited for the exponential DSD and the given parameters of the stochastic DSD model. The prefactor is more affected than the exponent of the power law $Z = a\,R^b$. Further investigation is needed to assess the influence of alternative DSD models (gamma, lognormal).

Acknowledgments. This research is supported by the EU projects FLOODSITE and VOLTAIRE. The second author is also supported by the Netherlands Organization for Scientific Research (NWO).

REFERENCES

Berne, A. and R. Uijlenhoet, 2005: A stochastic model of range profiles of raindrop size distributions: application to radar attenuation correction. *Geophys. Res. Lett.*, **32**, doi:10.1029/2004GL021899.

Gage, K., W. Clark, C. Williams, and A. Tokay, 2004: Determining reflectivity measurement error from serial measurements using paired disdrometers and profilers. *Geophys. Res. Lett.*, **31**, doi:10.1029/2004GL020591.

Joss, J. and A. Waldvogel, 1969: Raindrop size distribution and sampling size errors. *J. Atmos. Sci.*, **26**, 566–569.