Journal of Hydrology (2006) **328**, 65–82

available at www.sciencedirect.com

SCIENCE DIRECT®

journal homepage: www.elsevier.com/locate/jhydrol

Analytical solutions to sampling effects in drop size distribution measurements during stationary rainfall: Estimation of bulk rainfall variables

Remko Uijlenhoet [a,*], Josep M. Porrà [b], Daniel Sempere Torres [c], Jean-Dominique Creutin [d]

[a] *Chair of Hydrology and Quantitative Water Management, Environmental Sciences Group, Wageningen University, Nieuwe Kanaal 11, 6709 PA Wageningen, The Netherlands*
[b] *VENCA, Barcelona, Spain*
[c] *Grup de Recerca Aplicada en Hidrometeorologia, Universitat Politècnica de Catalunya, Barcelona, Spain*
[d] *Laboratoire d'étude des Transferts en Hydrologie et Environnement (UMR 5564, CNRS-INPG-IRD-UJF), Grenoble, France*

Received 24 October 2004; received in revised form 10 December 2004; accepted 30 November 2005

KEYWORDS
Raindrop size distribution;
Disdrometer;
Sampling error;
Marked point process;
Edgeworth expansion

Summary A stochastic model of the microstructure of rainfall is used to derive explicit expressions for the magnitude of the sampling fluctuations in rainfall properties estimated from raindrop size measurements in stationary rainfall. The model is a marked point process, in which the points represent the drop centers, assumed to be uniformly distributed in space. This assumption, which is supported both by theoretical and by empirical evidence, implies that during periods of stationary rainfall the number of drops in a sample volume follows a Poisson distribution. The marks represent the drop sizes, assumed to be distributed independent of their positions according to some general drop size distribution.

Within this framework, it is shown analytically how the sampling distribution of the estimator of any bulk rainfall variable (such as liquid water content, rain rate, or radar reflectivity) in stationary rainfall converges from a strongly skewed distribution to a (symmetrical) Gaussian distribution with increasing sample size. The relevant parameter controlling this evolution is the average number of drops in the sample n_s. For a given sample size, the skewness of the sampling distribution is found to be more pronounced for higher order moments of the drop size distribution. For instance, the sampling distribution of the normalized mean diameter becomes nearly Gaussian for $n_s > 10$, while the sampling distribution of the normalized rain rate remains skewed for $n_s \sim 500$.

* Corresponding author. Tel.: +31 317 485760; fax: +31 317 484885.
E-mail address: remko.uijlenhoet@wur.nl (R. Uijlenhoet).

0022-1694/$ - see front matter
doi:10.1016/j.jhydrol.2005.11.043

Additionally, it is shown analytically that, as a result of the mentioned skewness, the median Q_{50} as an estimator of a bulk rainfall variable always underestimates its population value Q_p in stationary rainfall. The ratio of the former to the latter is found to be $Q_{50}/Q_p = 1 - b/n_s + O(n_s^{-2})$, where b is a constant depending on the drop size distribution. For bulk rainfall variables this constant is positive and therefore the median always underestimates the population value. This provides a theoretical confirmation and explanation of previously published simulation results.Finally, relationships between the expected number of raindrops in the sample n_s and the rain rate are established for different parametric forms of the raindrop size distribution. These relationships are first compared to experimental results and then used to provide examples of sampling distributions of bulk rainfall variables (in this case rain rate) for different values of the average rain rate and different integration times of the disdrometric device involved (in this case a Joss–Waldvogel disdrometer). The practical relevance of these results is (1) that they provide *exact* solutions to the sampling problem during (relatively rare) periods of stationary rainfall (e.g., drizzle), and (2) that they provide a *lower bound* to the magnitude of the sampling problem in the general situation where sampling fluctuations and natural variability co-exist.

Introduction

There are many different types of instruments available to estimate raindrop size distributions. Besides the intrinsic interest these measurements have from a meteorological point of view, they also serve to establish relationships between different bulk rainfall variables such as the liquid water content W, the rain rate R, and the radar reflectivity factor Z, hereafter simply referred to as ''reflectivity'' (see Jameson and Kostinski, 2001b, for a recent discussion of some fundamental issues involved in Z–R relations). Relationships determined in this way are a crucial step in algorithms to translate weather radar measurements aloft into rain rate estimates at the ground (e.g., Battan, 1973). During a rainfall event, bulk variables change as a consequence of natural event evolution (often referred to as ''natural variability''), but also because disdrometric instruments have limited sampling volumes (referred to as ''sampling fluctuations''). In other words, apart from specific instrumental uncertainties (such as deadtime effects), observed fluctuations must be due 'both to statistical sampling errors and to real fine-scale physical variations which are not readily separable from the statistical ones' (Gertzman and Atlas, 1977). Although it is in principle not possible to distinguish between both types of variation (as argued by Jameson and Kostinski, 2001c), a better knowledge of the magnitude of the variability associated with sampling problems alone, that is of the fluctuations in bulk rainfall variables in the absence of natural variability, would certainly provide a step in the right direction. The objective of this paper is to provide a mathematical framework to quantify such fluctuations analytically. This will provide a lower bound to the magnitude of sampling problems in the general situation where sampling fluctuations and natural variability co-exist.

The aim behind the various studies that have been devoted to sampling fluctuations in disdrometric measurements is precisely the one just mentioned, namely to try to separate between natural variability and sampling fluctuations (Gertzman and Atlas, 1977). In this study, we analyze sampling fluctuations by modelling rainfall as a so-called marked point process, which is possibly the simplest model that is able to account for the discrete raindrop structure of rainfall (Smith, 1993). The mathematical description of the model is reviewed in detail in the following section. The basic model hypotheses are spatial homogeneity and the absence of drop interaction. This corresponds to the assumption of stationary (steady) rainfall, where all natural variability is absent and only sampling fluctuations are present (see Jameson and Kostinski, 2002 for a recent discussion of the conditions for ''steady rain''). Besides having an intrinsic interest in that the steady rain hypothesis provides a lower bound to sampling problems in general, as explained above, it will be shown in the following section that periods of prolonged quasi-steady rain do indeed occur in nature. This shows that our derivations are not merely an academic exercise. Our statistical framework not only allows a unification of previous theoretical studies on sampling fluctuations in rainfall observations (e.g., Cornford, 1967, 1968; Joss and Waldvogel, 1969; Gertzman and Atlas, 1977), but also provides analytical results that were known previously only from Monte-Carlo simulations on the basis of the marked point process model (Smith et al., 1993). Among the most important results we obtain are (1) the exact behavior, in stationary rainfall, of the sampling distribution of the estimator of any bulk rainfall property (such as liquid water content, rain rate, or radar reflectivity) as the sample size increases, and (2) the dependence of the median estimate of such a bulk rainfall variable from a disdrometric measurement on the corresponding sample size, again in stationary rainfall.

Our results apply to any disdrometric instrument. Nevertheless, a distinction between two types of devices must be considered because the interpretation of the sampling volume is different in each case. The first type of device is that with a constant sampling volume V_s independent of drop diameter. This type of apparatus can be looked at as making

an instantaneous image (a ''snapshot'') of a particular rainfall volume. Consequently, it is directly suited to estimate concentrations (such as the liquid water content W and the reflectivity Z). The second type of device, on the other hand, detects drops arriving a particular sampling surface. This type is therefore directly suited to estimate fluxes (such as the rain rate R). In this case, the effective sampling volume depends on the size of the raindrops because for a given sampling time (that is the time that the surface is exposed to the rainfall) drops with greater velocities explore a greater volume than those moving slower (Gertzman and Atlas, 1977). For both types of instruments, the relevant quantity that characterizes the sampling fluctuations is the average number of drops in the sampling volume (1st type) or the average number of drops arriving at the sampling surface during the sampling interval (2nd type). It is important to mention here that the drop concentration measured by both instruments is not the same, although they are related, as will be shown in ''Statistical model''.

One property that has been assumed frequently is the Gaussianity (normality) of the probability density of the rainfall variable estimators obtained from disdrometric data (e.g., Joss and Waldvogel, 1969; Gertzman and Atlas, 1977). In the (symmetric) Gaussian case, the coefficient of variation (the ratio between the standard deviation and the mean value) completely determines the relative error due to sampling uncertainties. However, it is well known that for small sampling volumes, even during stationary rainfall, the probability density function shows an amount of asymmetry (e.g., Smith et al., 1993). In this paper, we show analytically how the distribution of the estimator of any bulk rainfall variable obtained from a disdrometric measurement converges to a Gaussian distribution when the average number of drops sampled approaches infinity. Although one might have anticipated this result on the basis of the central limit theorem, it has to the best of our knowledge never been demonstrated before in this context. This may be due to the relatively advanced level of statistics involved in the problem. As opposed to standard sampling situations, where the number of samples is known (e.g., Kendall and Stuart, 1979), in the case of raindrop sampling we deal with a situation where the number of samples itself is a random variable. This gives rise to so-called *random sums of random variables* (Cramér, 1946). The expansion employed is particularly useful to obtain the dependence of the median of the estimator as a function of the average number of drops sampled. The paper by Smith et al. (1993) has stimulated us to calculate this dependence. The simulation results of that paper show that the skewness of the sampling distributions for small samples produces an underestimation of the bulk rainfall properties. Our calculations provide a theoretical confirmation of these results.

Statistical model

Poisson hypothesis

In most precipitation studies, raindrop populations have been characterized by the drop size distribution $N(D)$ only. This quantity gives the average number of drops per unit volume of air and per unit diameter [L^{-4}] (see Porrà et al., 1998; Jameson and Kostinski, 2001c for fundamental discussions on the interpretation of the concept of a raindrop size distribution). It is an average because both the number of drops in a volume and their diameters fluctuate. Hence, any description of rainfall at the level of raindrops has to include two types of fluctuations: the statistics of the distribution of drops in space and the distribution of their sizes.

We assume that drops are distributed uniformly (in a statistical sense, that is 'as uniformly as randomness allows') in a given volume or, equivalently, that the number of drops is distributed according to a Poisson distribution. This is the simplest model that is possible. The Poisson homogeneity hypothesis forms the basis of many previous studies dealing with sampling fluctuations in rainfall observations (e.g., Cornford, 1967, 1968; Joss and Waldvogel, 1969; de Bruin, 1977; Gertzman and Atlas, 1977; Stow and Jones, 1981; Wirth et al., 1983; Wong and Chidambaram, 1985; Chandrasekar and Bringi, 1987; Hosking and Stow, 1987; Chandrasekar and Gori, 1991; Smith et al., 1993; Bardsley, 1995). There exist traces of empirical evidence for the Poisson homogeneity hypothesis during stationary rainfall. For instance, Kostinski and Jameson (1997) find indications for Poisson behavior during 'a time of unusually constant flux'. Uijlenhoet et al. (1999) find that 'for moderate rain rates the arrival rate fluctuations of the raindrops which contribute most to rain rate and radar reflectivity factor behave according to Poisson statistics [in stationary rainfall]'.

As additional empirical evidence for the Poisson hypothesis in stationary rainfall, Figs. 1 and 2 present the results of a re-analysis of the dataset of Uijlenhoet et al. (1999). This dataset, which has been kindly provided to us by Mr. R.J. Moore (Centre for Ecology and Hydrology, UK), has been collected as part of the NERC (National Environment Research Council) Special Topic HYREX, a hydrological radar experiment organized in the United Kingdom, at the Bridge Farm Orchard site on 14 February 1995 (Moore et al., 2000). The dataset concerns a 35 min period of roughly uncorrelated fluctuations around a constant mean rain rate of about 3.5 mm h^{-1}, containing 210 consecutive 10 s raindrop size distributions (comprising a total of 6281 raindrops). The raindrop counts, collected using a 50 cm^2 Paired-Pulse Optical Disdrometer (Illingworth and Stevens, 1987), are distributed among 16 diameter intervals of 0.21 mm width. The average wind speed during the event amounted approximately 3 m s^{-1}. The empirical frequency function calculated from the 210 observations has been compared for each raindrop diameter interval with the theoretical frequency function expected for a Poisson distribution with the same mean. Fig. 1 shows the results for the first 6 intervals, corresponding to diameters from 0.51 to 1.77 mm (Uijlenhoet, 1999). This analysis demonstrates that the Poisson hypothesis is closely obeyed for all diameter intervals except the first three, which show significant deviations from the Poisson distribution based on the values of the dispersion index (see Eq. (3)) and the χ^2 goodness-of-fit statistic. The fit with the Poisson frequency function becomes nearly perfect for the last diameter intervals (which are not shown here as they provide little extra information). Fig. 2 shows the corresponding empirical autocorrelation functions, which indicate that the raindrop counts in each of the diameter intervals exhibit nearly no temporal correlation, as

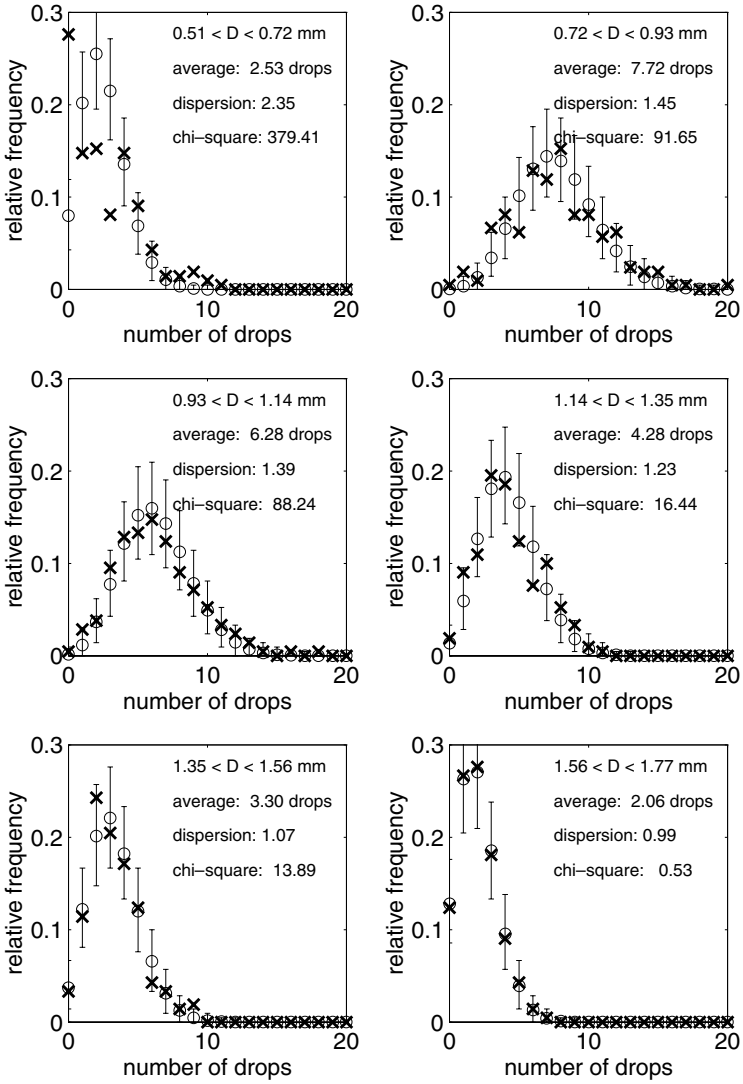

Figure 1 Empirical (crosses) and theoretical Poisson (circles, Eq. (2)) frequency functions of raindrop counts for diameters between 0.51 and 1.77 mm (24 degrees of freedom). Error bars indicate 95% confidence limits. Also indicated are the average number of raindrops per 10 s interval, the Poisson dispersion index (see Eq. (3)) and the χ^2 goodness-of-fit statistic (adapted from Uijlenhoet et al., 1999).

would be the expected for a homogeneous Poisson process (only some first-order, that is 10 s, autocorrelations show significant deviations from zero). As such, the analyzed dataset can be considered to represent more than half an hour of quasi-Poissonian rainfall.

Although this example shows that rainfall observations may occasionally behave according to Poisson statistics during rare periods of exceptional stationarity, it now becomes more and more clear that rainfall exhibits pronounced spatial and temporal drop clustering. Since the homogeneous Poisson process is not able to cope with these types of clustering, a more versatile description of raindrop statistics is needed. There exist basically two approaches to tackling this problem. The first consists of generalizing the restrictive homogeneous Poisson process (which has a constant mean) to a Poisson process with a randomly varying mean, that is a so-called doubly stochastic Poisson process or Cox process (e.g., Cox and Isham, 1980 for a summary of the properties of this type of stochastic point process). This approach was pioneered by Sasyo (1965) and has more recently been used by Smith (1993). It has been put in an entirely new perspective in a recent series of articles (Kostinski and Jameson, 1997, 1999; Jameson and Kostinski, 1998, 1999, 2000; Jameson et al., 1999). An alternative approach is to abandon the Poisson process framework altogether and replace it with a (multi-)fractal approach. Examples of the latter are a box counting analysis of the spatial distribution of raindrops (Lovejoy and Schertzer, 1990) and fractal analyses of the temporal distribution of raindrop arrivals (Zawadzki, 1995; Lavergnat and Golé, 1998). It should be noted that the statistical significance of the empirical support in favor of the (multi-)fractal hypothesis provided by Lovejoy and Schertzer (1990) has recently been put into question by several authors (Jameson and Kostinski, 1998; Gabella et al., 2001; Jameson and Kostinski, 2001a; Gabella and Perona, 2001). In fact, spatial distributions of raindrops measured using stereo-photography indicate that the (multi-)fractal character of rain, if

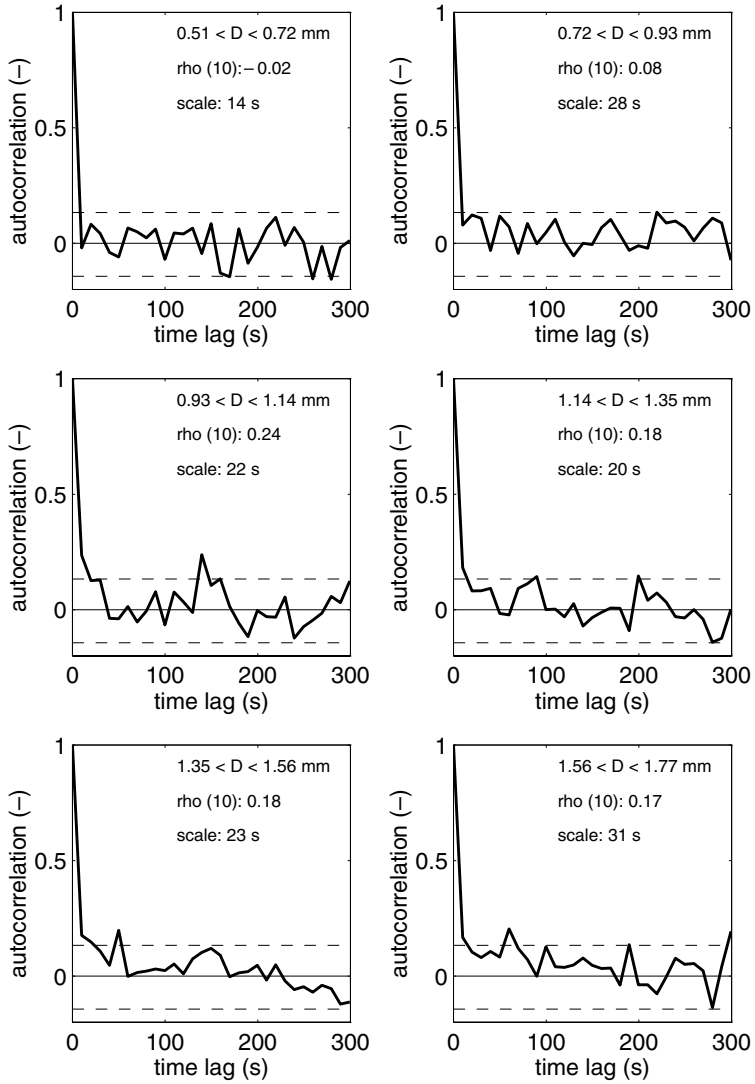

Figure 2 Empirical autocorrelation functions of 10 s raindrop counts for diameters between 0.51 and 1.77 mm (solid lines) together with approximate 95% confidence limits (dashed lines). Also indicated are the first order autocorrelation coefficient and the scale of fluctuation (which can be interpreted as a decorrelation time), calculated as the integral under the autocorrelation function (Vanmarcke, 1983).

present at all, is more clearly detectable in the spatial distribution of the liquid (rain)water content than in the spatial distribution of the drops per se — which has been found to be relatively homogeneous (Desaulniers-Soucy et al., 2001; Lovejoy et al., 2003). The main difference between the two approaches is that doubly stochastic Poisson process models tend to produce clustering of raindrops on certain distinct, predefined spatial and/or temporal scales, whereas (multi-)fractal processes are associated with clustering of raindrops on *all* scales. Although the evidence in support of the former has recently been more convincing than that in support of the latter, it seems that the discussion between the two has not yet come to a definite conclusion. Fabry (1996) argues that gravitational mixing (associated with the differential fall speeds of the raindrops) tends to lead to homogeneous rainfall at the smallest spatial scales, whereas (fractal) scaling regimes exist at larger scales.

However, Jameson and Kostinski (1998) have recently made an important point by arguing that 'evidence of non-clustering, Poissonian structure conflicts with any ubiquitous fractal description of rain'. The logical consequence of this is that both their (Kostinski and Jameson, 1997) and our (Uijlenhoet, 1999; Uijlenhoet et al., 1999) previously mentioned traces of empirical evidence in favor of a Poissonian structure during conditions of stationary rainfall provide arguments to reject the (multi-)fractal hypothesis. This shows that there is no conflict whatsoever between our approach and that put forward in the previously mentioned series of articles by Jameson and Kostinski, as one might be tempted to think at first sight. Our (admittedly more restrictive) homogeneous Poisson process framework merely provides a special (limiting) case of their (more general) doubly stochastic Poisson process framework. The practical relevance of our results is therefore (1) that they provide *exact* solutions to the sampling

problem during (relatively rare) periods of stationary rainfall (e.g., drizzle), and (2) that they provide a *lower bound* to the magnitude of the sampling problem in the general situation where sampling fluctuations and natural variability co-exist. In the near future, we intend to generalize the mathematical framework presented in this paper to be able to cope with non-homogeneous Poisson processes.

Mathematical formulation

Within the homogeneous Poisson model, the density of drops per unit volume of air N_T, $[L^{-3}]$, is the only parameter required to characterize the spatial distribution of drops. The expected number of drops in a sample volume V_s, $[L^3]$, becomes

$$n_s = V_s N_T, \tag{1}$$

and the probability that k drops are contained in that volume is given by the Poisson distribution $P_{n_s}(k)$, that is (Mood et al., 1974),

$$\text{Prob}\{k \text{ drops in } V_s\} = P_{n_s}(k) = \frac{n_s^k e^{-n_s}}{k!}. \tag{2}$$

An important property of the homogeneous Poisson process is that for non-overlapping sample volumes of the same size V_s, these distributions are independent and identical (Cox and Isham, 1980). Another property of the Poisson model is that the mean of number of drops is equal to its variance, that is,

$$E[k] = \text{Var}[k] = n_s. \tag{3}$$

This property of the Poisson distribution can be used to define a so-called *dispersion index* as the ratio of the variance of the number of drops to the corresponding mean (e.g., Cox and Isham, 1980). For a homogeneous Poisson process this index obviously equals one. Significant deviations from one observed in real data can then be interpreted as indications for deviations from homogeneous Poisson behavior (see Fig. 1).

The distribution of drop diameters will be characterized, in general, by a probability density function $p(D)$, $[L^{-1}]$, such that $p(D)\,dD$ yields the probability that the diameter of a drop falls in the interval $(D, D + dD]$. For instance, Marshall and Palmer (1948) suggested that drop diameters follow an exponential distribution. The drop size distribution $N(D)$ can be written in terms of N_T and the distribution of sizes $p(D)$ as

$$N(D) = N_T p(D), \tag{4}$$

because the drop density N_T is the integral of the drop size distribution (DSD) over all drop diameters. This obvious notion was first employed by Chandrasekar and Bringi (1987) for the special case of a gamma form for $N(D)$ and by Sempere Torres et al. (1998, see the appendix in that paper) for the general case of any parametric form for $N(D)$. Porrà et al. (1998) discuss the fundamental hypotheses on which Eq. (4) is based. More recently, this product representation of $N(D)$ has been employed by Kostinski and Jameson (1999) and by Uijlenhoet (1999).

The model just introduced is a marked point process in which the point process represents the positions of drops in the sample volume and the mark associated with a drop represents its diameter (Smith, 1993). This sample-volume process can be transformed into an arrival process by considering the arrival of drops at a given surface during a given period of time. The connection is established by assuming (1) that drops fall vertically downward at a terminal velocity that depends exclusively on their diameter and (2) that drops do not interact with each other. The absence of interaction together with the Poisson distribution of drops in space implies that the inter arrival times of drops are exponentially distributed (Smith, 1993). Thus, the probability density function $\phi(t)$, $[T^{-1}]$, of the interval between the arrivals at the sampling surface of two consecutive drops reads

$$\phi(t) = \lambda S e^{-\lambda S t}, \tag{5}$$

where λ, $[L^{-2}T^{-1}]$, is the arrival rate and S, $[L^2]$, is the sampling surface. The parameter that characterizes the arrival process is then the mean arrival rate λ, which is the expected number of drops arriving at the sampling surface per unit area and per unit time. The arrival rate can be computed in terms of the drop size distribution $N(D)$ and the drop terminal fall velocity $v(D)$ as

$$\lambda = \int v(D) N(D)\, dD = N_T \langle v \rangle, \tag{6}$$

where the integral extends over all drop diameters and $\langle v \rangle$ is the average terminal velocity of drops in the sample volume

$$\langle v \rangle = \int v(D) p(D)\, dD. \tag{7}$$

The stochastic process of the arrival of drops at a given surface during a given period of time obtained in this way is again a marked point process, in which the arrival times now constitute the point process (in time) and the diameter of each drop is its corresponding mark. The exponential distribution (Eq. (5)) completely determines the statistics of the arrival point process. In fact, this process is also of the Poisson type because the number of drops arriving at the sampling surface S during a given period of time t is distributed according to a Poisson distribution as well. The expected number of drops n_s that will reach the surface during the period t is given by

$$n_s = \lambda S t \tag{8}$$

and the probability that k drops reach the surface over that period of time is also given by Eq. (2), with n_s calculated by Eq. (8).

The probability density of the sizes of drops arriving at the sampling surface differs from $p(D)$ and will be denoted by $p_A(D)$. The relation between them reads

$$p_A(D) = \frac{v(D)}{\langle v \rangle} p(D). \tag{9}$$

It is now possible to define a raindrop size distribution for flux processes, which will be denoted by $N_A(D)$ to distinguish it from $N(D)$, such that $N_A(D)\,dD$ represents the expected number of raindrops with diameters between D and $D + dD$ arriving at a surface per unit area and per unit time (Uijlen-

hoet and Stricker, 1999; Uijlenhoet, 2001). In terms of $p_A(D)$ and λ, the mean arrival rate, the raindrop size distribution $N_A(D)$ is

$$N_A(D) = \lambda p_A(D), \tag{10}$$

which is the flux process equivalent of Eq. (4). Although there is no reason to privilege one of the two DSD definitions over the other, the reality is that $N(D)$ is the most widely used expression. It seems logical to use $N(D)$ when analyzing processes related to a sample volume (such as measuring radar reflectivity) and $N_A(D)$ when analyzing processes involving fluxes (such as rain rate measurement).

All instruments used to measure raindrop diameters can be attached naturally to one of the two forms of the raindrop size distribution. In general, disdrometric devices belong to two types, as was discussed in the ''Introduction'': (1) volume integrating instruments and (2) time integrating devices. The first type of instrument yields the number of drops and their diameters in a particular sample volume at a given moment. The raindrop camera (Mueller, 1966; Smith, 1993), the optical array probe (Knollenberg, 1970), hydrometeor detection and ranging using stereophotography (Desaulniers-Soucy et al., 2001; Lovejoy et al., 2003), and the VPR (vertically pointing Doppler radar) (Hauser and Amayenc, 1981; Kollias et al., 2002) all work in this way. These instruments are suited to provide direct estimates of $N(D)$. The other type of instrument measures the diameters of raindrops that reach a sampling surface during a period of time. The flour pans and sieves method (Laws and Parsons, 1943), the dyed filter (or blotting) paper method (Marshall and Palmer, 1948), the disdrometer RD-69 (Joss and Waldvogel, 1967), the optical spectro-pluviometer (OSP) (Salles et al., 1998), and the 2-D video disdrometer (2DVD) (Kruger and Krajewski, 2002; Bringi et al., 2003), belong to this kind. The drop size distribution $N_A(D)$ can be estimated more easily from this type of instrument than $N(D)$.

General formalism

In homogeneous rainfall, measured rainfall properties become random variables due to sampling fluctuations. The aim of disdrometric measurements is to infer these properties using the appropriate statistical estimator. In general, the estimator of any bulk rainfall variable Q can be expressed as

$$Q = \sum_{i=1}^{n} h_Q(D_i), \tag{11}$$

where D_i is the diameter of the ith drop measured in a sample of n drops and h_Q is a function describing the contribution of one drop with a given diameter to Q (e.g., Eq. (12)). The variable Q is random because n and each D_i are random variables.

We assume that the random variables D_i are independent and identically distributed (iid), and independent of n. The probability density function of D_i depends on the type of device used to measure the diameter. For volume integrating instruments (type 1), the diameters D_i are distributed according to the probability density function of raindrop diameter $p(D)$, defined in Eq. (4), whereas for time integrating devices (type 2) this density is given by $p_A(D)$, Eq. (9). We have justified in the previous section that the number of drops sampled n is a Poisson random variable. The mean of n, n_s, depends on the sampling volume V_s, Eq. (1), for instruments of type 1, or on the sampling interval t and surface S, Eq. (8), for the second type of device.

For example, the liquid water content W of a sample taken by a drop-camera is measured simply by summing up the masses of all drops in the sample. The number of drops counted gives n and the function h is the expression of drop mass as a function of diameter divided by the sampling volume V_s. In this example, $Q = W$ and the function h_Q reads

$$h_W(D) = \frac{\pi \rho_w D^3}{6 V_s}, \tag{12}$$

where ρ_w [ML^{-3}] is the liquid water density. For given units of the quantities in Eq. (12) a numerical factor may have to be included in the definition of $h_W(D)$. Given the definition of $h_W(D)$, Eq. (11) reduces to the usual W-estimator

$$W_s = \frac{\pi \rho_w}{6 V_s} \sum_{i=1}^{n} D_i^3. \tag{13}$$

The subscript s means that this quantity provides a sample estimate of W. The estimators of most common bulk rainfall properties obtained from disdrometric measurements (rain rate, reflectivity and kinetic energy flux density, among others) can be expressed in the form of Eq. (11) by using a convenient function h_Q (Uijlenhoet and Stricker, 1999).

The aim of this paper is to study the properties of the random variables Q that follow from the associated probability density function $f(q)$. The method to study the function f is based on the hypothesis that the diameters of the sampled drops are completely independent of each other and of n. This assumption appears in addition to the hypothesis of the homogeneous drop distribution in space used to obtain the Poisson law for the number of drops sampled. According to these hypotheses, Q is the sum of a random number n of iid random variables $h = h(D)$. The probability density of D will be denoted by $f_d(D)$; thus $f_d(D)\,dD$ simply gives the probability that a raindrop has a diameter that falls in the interval $(D, D + dD)$. As we have explained in ''Statistical model'', $f_d(D)$ coincides with $p(D)$ for type 1 instruments and with $p_A(D)$ for devices of type 2. In terms of $f_d(D)$, the probability density of h can be written as

$$p(h) = \left|\frac{dD}{dh}\right| f_d(D), \tag{14}$$

where the diameter D on the right-hand side has to be taken as a function of h. We have assumed that h is a monotonous function of the diameter D, such as in Eq. (12), so that it can be inverted to provide $D(h)$, otherwise Eq. (14) becomes more complicated.

As the distribution of raindrop sizes is continuous, the random variables Q are continuous random variables as well. However, the probability density function of Q, $f(q)$, has an atom at $q = 0$, corresponding to the case where no drop is detected ($n = 0$). In other words, there is a finite probability that the random variable Q takes exactly the value 0. Therefore, the probability density function (pdf) of Q will have the form

$$f(q) = p_0\delta(q) + f_c(q). \quad (15)$$

The first term on the right-hand side, with the Diraq delta function, accounts for the probability that $Q = 0$ and p_0 gives the corresponding probability that no drop is detected. The probability p_0 decreases exponentially as e^{-n_s} because of the Poisson distribution assumption — see Eq. (2). For $n_s \geqslant 10$, p_0 can be neglected as it is less 5×10^{-3}%. The second term on the right-hand side gives the continuous part of $f(q)$, which as we have just shown becomes the unique contribution to $f(q)$ for $n_s \geqslant 10$.

Sampling distributions

Appendix A is devoted to the derivation of the characteristic function (cf) of the general estimator of any bulk rainfall variable Q (Eq. (11)), that is the Fourier transform of its pdf $f(q)$ (Eq. (15)). The cf is useful for calculating the moments and cumulants of the estimator. In Appendix B, the so-called Edgeworth expansion of $f(q)$ is derived (Eq. (B.10)). As an example of the utility of the Edgeworth expansion of the probability density of an estimator Q, we will analyze the sampling distribution of the estimator of D_1, the first moment of the drop size distribution $N(D)$, and of R^*, the fourth moment of the drop size distribution, a quantity closely related to the rain rate (Joss and Gori, 1978). We assume that the drop size distribution $N(D)$ corresponds to a Marshall—Palmer exponential distribution (Marshall and Palmer, 1948) without diameter truncation,

$$N(D) = N_0 e^{-\Lambda D}, \quad D \in (0, \infty). \quad (16)$$

From $N(D)$, the probability density function of the drop diameter, $f_d(D)$, and the drop concentration N_T are readily obtained (Uijlenhoet and Stricker, 1999), namely

$$f_d(D) = \Lambda e^{-\Lambda D}, \quad (17)$$

and

$$N_T = \frac{N_0}{\Lambda}. \quad (18)$$

The pdf $f_d(D)$ coincides with $p(D)$ because we are dealing with a sample-volume process. The drop concentration times the sampling volume V_s yields the expected number of drops in the sample, $n_s = V_s N_T$, which will be the parameter characterizing the sample size.

The population values of the quantities measured are computed as moments of the drop size distribution $N(D)$, Eq. (16). Let us first analyze the first moment D_1. The population value of this property, denoted by the subscript p, is

$$D_{1,p} = \int_0^\infty D N(D)\, dD = \frac{N_T}{\Lambda} = \frac{n_s}{\Lambda V_s}. \quad (19)$$

The estimator

$$D_{1,s} = \frac{\sum_{i=1}^n D_i}{V_s} \quad (20)$$

provides a sample estimate of D_1 when a population is sampled by observing the drops in some sampling volume V_s. The ratio of sample to population value has the following simpler expression:

$$\frac{D_{1,s}}{D_{1,p}} = \frac{\sum_{i=1}^n \Lambda D_i}{n_s}. \quad (21)$$

This ratio will be denoted by η.

We now obtain the sampling distribution of η. Note that η is an estimator of the class defined in Eq. (11) with the following expression for the function h

$$h_\eta(D) = \frac{\Lambda D}{n_s}. \quad (22)$$

The probability density function of h_η is obtained from the pdf of D, $f_d(D)$, and reads (Eq. (14))

$$p(h) = n_s e^{-n_s h}. \quad (23)$$

The sampling distribution of η can be calculated exactly because the convolution of n exponential densities is a gamma density of order n. This is an exceptional case which we use here to test how well the Edgeworth expansion performs. The exact expression for $p(\eta)$ is

$$p(\eta) = e^{-n_s}\delta(\eta) + \frac{n_s}{\sqrt{\eta}} e^{-n_s(1+\eta)} I_1(2n_s\sqrt{\eta}), \quad (24)$$

where I_1 is the modified Bessel function of order 1 (Abramowitz and Stegun, 1972). The term with the Diraq delta function takes into account the probability that there is no drop in the sample, and therefore $\eta = 0$ (Eq. (15)). As discussed in ''General formalism'', this term is negligible for n_s greater than 10. When n_s goes to infinity, $p(\eta)$ converges to a Gaussian with mean equal to 1 and standard deviation σ equal to $\sqrt{2/n_s}$

$$p(\eta) \overset{n_s \to \infty}{\to} \sqrt{\frac{n_s}{4\pi}} \exp\left[-\frac{n_s}{4}(\eta - 1)^2\right]. \quad (25)$$

The mean and the variance were computed from Eqs. (A.11) and (A.12) after calculating the moments of function h,

$$\langle h^k \rangle = \int_0^\infty h^k p(h)\, dh = \frac{k!}{n_s^k}. \quad (26)$$

The limiting sampling distribution of η becomes a delta function at $\eta = 1$ because the standard deviation converges to zero.

Fig. 3 displays the exact pdf of η (solid line, Eq. (24)), the Edgeworth expansion of it, up to order $n_s^{-3/2}$ (dotted line, Eq. (B.10)), and the Gaussian approximation (long dashed line, Eq. (25)) for $n_s = 10$. Note that the skewness of the exact distribution is marked. The Edgeworth expansion (dotted line) reproduces it quite well, whereas the Gaussian approximation clearly cannot account for this asymmetry. Note, however, the anomalous behavior of the Edgeworth expansion in the tails, where it becomes even negative for $\eta < 0$ (Cramér, 1946; Kendall and Stuart, 1977). This fact does not weaken the utility of the series as an approximation to the actual distribution in the central region, as becomes evident from the plot. In addition, the quantile values can also be estimated much better from this series than from the Gaussian approximation provided the quantile is not too close to 0 (say 10%) or to 1 (say 90%).

The same figure also shows the exact sampling pdf of η for $n_s = 40$ (short dashed line). For this sample size, the sampling distribution already becomes nearly Gaussian. The estimator of η , Eq. (21), is unbiased because its expected value is 1 regardless of the sample size. The median of the estimate, however, depends on n_s. The discussion of the properties of the median of an estimator Q is postponed to the following section.

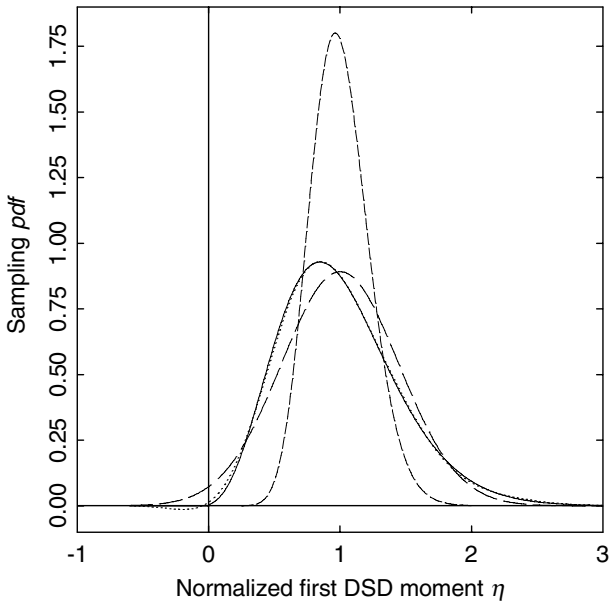

Figure 3 The exact sampling probability density function of η (−), the ratio of the sample to the population value of the first moment of the raindrop size distribution $N(D)$ (solid line, Eq. (24)), the Edgeworth series expansion of it up to order $n_s^{-3/2}$ (dotted line, Eq. (B.10)), and the Gaussian approximation (long dashed line, Eq. (25)) when n_s (−), the expected number of raindrops in the sample, equals 10. In addition, the exact sampling probability density function of η for $n_s = 40$ (short dashed line).

We have explained above that the Edgeworth expansion is asymptotic and therefore there is no warranty that it will give a useful approximation for all finite n_s. Indeed, problems (regions with negative values and non-unimodal behavior of the probability density function) appear when the skewness γ of the sampling distribution, as defined by Eq. (A.14), is too high ($\gamma > 1$) (Cramér, 1946; Kendall and Stuart, 1977). It follows from Eq. (A.14) that the estimators of the moments of the drop size distribution show this handicap more strongly as the order of the moment increases. As an example, we analyze the case of the 4th order moment, D_4, which will be denoted R^* because of its similarity to the rain rate (Smith et al., 1993).

The ''normalized rain rate'' (R_s^*/R_p^*) can be written as an estimator Q with a function h given by

$$h(D) = \frac{\Lambda^4 D^4}{24 n_s}. \tag{27}$$

Using Eq. (A.14), the coefficient of skewness of the estimator of the normalized rain rate, when the drop size distribution is exponential without truncation, reads

$$\gamma = \frac{12!}{8!^{3/2}\sqrt{n_s}} = \frac{59.16}{\sqrt{n_s}}. \tag{28}$$

This value implies that the Edgeworth expansion of the probability density of the estimator of the 4th order normalized moment starts to provide practical results for sample sizes as large as $n_s = 59.16^2 \sim 3500$ (for which $\gamma \approx 1$). The reason for such a big number lies in the tail of the exponential drop size distribution, which extends to infinity. To obtain more realistic results, it is necessary to impose a maximum diameter and truncate the exponential $f_d(D)$, Eq. (17). We choose

$$D_{max} = \frac{10}{\Lambda}, \tag{29}$$

following the study on sampling variability of Smith et al. (1993). Hence, D_{max} is chosen such that ΛD_{max} is a constant (independent of rain rate, for example). In this way, the pdf of the drop diameters, $f_d(D)$, is hardly modified and reads

$$f_d(D) = \begin{cases} \frac{\Lambda e^{-\Lambda D}}{1-e^{-10}}, & 0 < D < 10/\Lambda, \\ 0, & \text{otherwise.} \end{cases} \tag{30}$$

Note that the denominator $(1 - e^{-10})$ is the normalization factor due to the truncation effect. The moments of the diameter according to this distribution are much smaller than those obtained when calculated using the untruncated exponential function, particularly those of higher orders (such as R^* and Z).

The diameter truncation, Eq. (29), changes the population value of R^* to

$$R_p^* = \int_0^{10/\Lambda} D^4 N_T \frac{\Lambda e^{-\Lambda D}}{1-e^{-10}} dD = \frac{\Gamma(5,10) n_s}{\Gamma(1,10) V_s \Lambda^4} = \frac{23.29 n_s}{V_s \Lambda^4}, \tag{31}$$

where $\Gamma(n,x)$ represents the incomplete gamma function (Gradshteyn and Ryzhik, 1980)

$$\Gamma(n,x) = \int_0^x t^{n-1} e^{-t} dt. \tag{32}$$

It is worth mentioning that if D_{max} would be infinitely large (the untruncated case), the coefficient would be 24 (4!) instead of 23.29, that is a difference of only 3%. The normalization factor $1 - e^{-10}$ has been written as $\Gamma(1,10)$ in the expression of R_p^*. Thus, the normalized rain rate becomes

$$\frac{R_s^*}{R_p^*} = \frac{\sum_{i=1}^n \Lambda^4 D_i^4}{23.29 n_s}. \tag{33}$$

The coefficient of skewness of this estimator is

$$\gamma = \frac{\Gamma(13,10)\sqrt{\Gamma(1,10)}}{\Gamma(9,10)^{3/2}\sqrt{n_s}} = \frac{22.63}{\sqrt{n_s}}, \tag{34}$$

which is less than half the value obtained when no truncation was considered, Eq. (28). From the new coefficient of skewness, it follows that the Edgeworth expansion will yield reasonable results for samples sizes around $n_s = 22.63^2 \sim 500$ or greater (for which $\gamma < 1$). Fig. 4 shows the Edgeworth series of the pdf of the normalized rain rate for $n_s = 500$ as a dotted line. When compared with the Gaussian approximation (dashed line), it becomes clear that the skewness of the estimator cannot be neglected, not even for a mean sample size as large as 500.

It has been shown by simulation (Smith et al., 1993) that the sampling distribution of the normalized rain rate closely resembles a lognormal distribution for small sample sizes. Although we know now that these distributions are *not* exactly lognormal, we have plotted in Fig. 4 the lognormal pdf

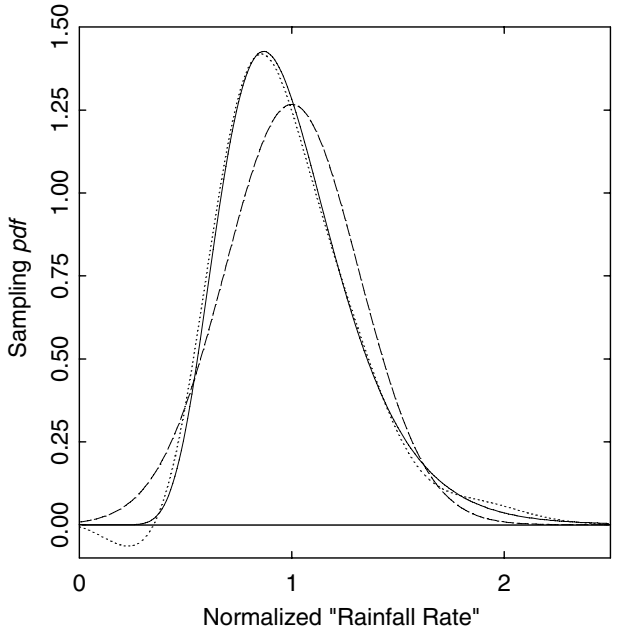

Figure 4 The Edgeworth series expansion of the sampling probability density function of the normalized rain rate for an expected sample size $n_s = 500$ (dotted line), the Gaussian approximation (dashed line) and the lognormal approximation (solid line, Eq. (35)).

$$f_{ap}(q) = \frac{1}{\sqrt{2\pi}\sigma q} \exp\left[-\frac{1}{2}\left(\frac{\ln q - \mu}{\sigma}\right)^2\right] \tag{35}$$

with $\mu = -0.04725$ and $\sigma = 0.3074$, which imply the same mean (1.00) and variance (0.09911) as the normalized rain rate for $n_s = 500$. Apart from the tail for values less than 0.4, the lognormal approximation (solid line) and the Edgeworth expansion fit quite well, particularly in the region around the mode.

The median of Q

We have shown in Appendix A how to calculate the moments and the cumulants of the estimator Q. In some cases, the median Q_{50} is preferred to the mean even though it is a biased estimator of the population values. Unfortunately, the computation for the median is not as simple as that for the mean. Nevertheless, using the Edgeworth expansion, it is possible to deduce the behavior of Q_{50} as a function of n_s (see Appendix C). An immediate consequence of the obtained expression for the median (Eq. (C.10)) is that Q_{50} always underestimates the population value because $\langle h^3 \rangle$ is positive for all rainfall variable estimators. The normalized ratio between Q_{50} and $\langle Q \rangle$ reads

$$\frac{Q_{50}}{\langle Q \rangle} = 1 - \frac{\langle h^3 \rangle}{3! \langle h^2 \rangle \langle h \rangle n_s} + O(n_s^{-2}), \tag{36}$$

where we have taken into account that $\langle Q \rangle = n_s \langle h \rangle$.

Smith et al. (1993) have simulated a raindrop population to estimate the ratio between the median and the mean for several bulk rainfall properties. Our theoretical result can be compared with their simulations when the sample size n_s is large enough to justify the use of the Edgeworth expansion. Indeed, we will demonstrate this convergence for two rainfall variables: the liquid water content (Eq. (13)) and the rain rate R^* defined in the previous section.

The normalized values of the liquid water content and of R^* are defined, according to Smith et al. (1993), as follows:

$$\frac{W_s}{W_p} = \frac{\sum_{i=1}^{n} \Lambda^3 D_i^3}{6 n_s}, \tag{37}$$

$$\frac{R_s^*}{R_p^*} = \frac{\sum_{i=1}^{n} \Lambda^4 D_i^4}{24 n_s}. \tag{38}$$

In these expressions, the subscripts s and p denote the sample and the population values, respectively. Drop diameters are distributed according to the truncated exponential given by Eq. (30). Using this distribution, the population value of the rain rate R^* was already obtained (Eq. (31)). The population value of the liquid water content is obtained by the same method and reads

$$W_p = 5.938 \frac{\pi \rho_w n_s}{6 V_s \Lambda^3}. \tag{39}$$

Note that the ratios given by Eqs. (37) and (38) are not exactly the ratio of sample to population values but the ratio calculated without taking into account the truncation of the drop size distribution. This feature makes the expected values of the normalized ratios, Eqs. (37) and (38), slightly lower than 1. Precisely, these values are:

$$\left\langle \frac{W_s}{W_p} \right\rangle = \frac{\Gamma(4,10)}{6\Gamma(1,10)} = 0.9897, \tag{40}$$

$$\left\langle \frac{R_s^*}{R_p^*} \right\rangle = \frac{\Gamma(5,10)}{24\Gamma(1,10)} = 0.9708, \tag{41}$$

where $\Gamma(n,x)$ is the incomplete gamma function defined previously (Eq. (32)).

In Fig. 5, we compare the simulation values obtained in (Smith et al., 1993) for $n_s = 10$, 20, 50, 100, 200, 500, and 800 with the theoretical asymptotic expression Eq. (36). For the liquid water content, the theoretical result (solid line)

$$\text{median}(W_s/W_p) = 0.9897 - \frac{8.724}{n_s} + O(n_s^{-2}) \tag{42}$$

fits the simulation data (squares) very well within the statistical fluctuations when $n_s > 100$. The results are plotted in dB, defined as

$$\text{dB value} = 10 \log[\text{median}(W_s/W_p)]. \tag{43}$$

In the same plot, the simulation values for the normalized rain rate are represented as circles together with the theoretical calculation

$$\text{median}(R_s^*/R_p^*) = 0.9708 - \frac{25.78}{n_s} + O(n_s^{-2}) \tag{44}$$

(short dashed line). The minimum value of n_s that makes this expression agree with the simulation data lies around 200. This value is higher than that for the liquid water content because the probability density function of the estima-

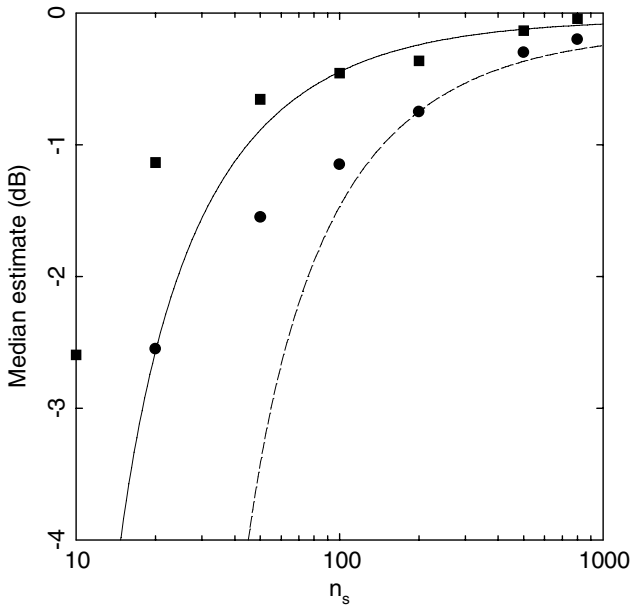

Figure 5 Comparison of the simulation values of the ratio between the median and the mean obtained in Smith et al. (1993) for n_s = 10, 20, 50, 100, 200, 500, and 800 with the theoretical asymptotic expression (Eq. (36)) for the liquid water content (simulation data: squares; theoretical result, Eq. (42): solid line) and the rain rate (simulation data: circles; theoretical result, Eq. (44): short dashed line). The results are plotted in dB, as defined by Eq. (43).

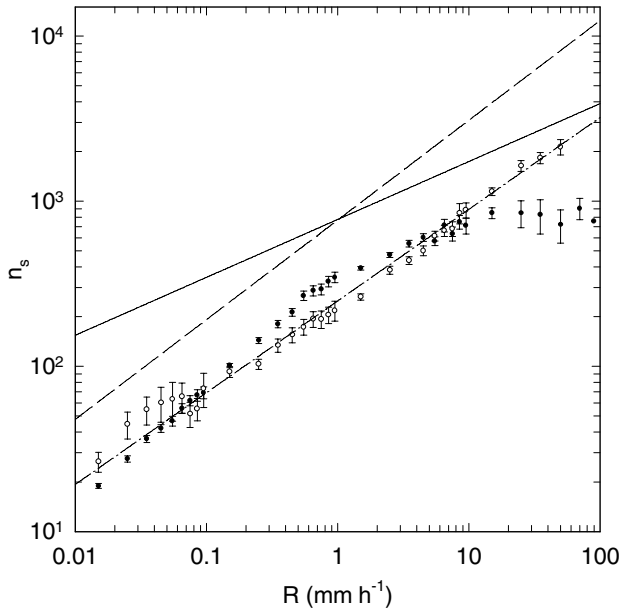

Figure 6 The theoretically expected number of drops n_s collected by a RD-69 disdrometer (for which the product of the sampling surface S times the integrating period t equals St = 3000 cm^2 s) as a function of the rain rate for three families of raindrop size distributions: Marshall—Palmer (solid line, Eq. (D.9)); Sekhon—Srivastava (dashed line, Eq. (D.10)); Willis—Tattelman (dash-dotted line, Eq. (D.12)). Circles with error bars are experimental data from Switzerland, obtained using a RD-69 disdrometer (closed circles) and France, obtained using an optical spectro-pluviometer (OSP, open circles).

tor of R^* is more skewed than the one for W and therefore the Edgeworth expansion works satisfactorily only for higher values of n_s.

Rain rate dependence of sampling distributions

Up to here, we have discussed sampling variability as a function of the expected number of drops sampled. However, for practical purposes one may be interested in the effect of sampling as a function of rain rate. The study of this dependence requires the analysis of the relation between the drop size distribution and rain rate. It is clear that the number of drops sampled by a disdrometric instrument also depends on the sampling properties of the device. The smaller these properties are (the sampling volume V_s or the surface S and the integration time t), the bigger the sampling effects will be. As an application of our general results, we will analyze the sampling effects on the Joss and Waldvogel disdrometer (RD-69) because it is one of the most widely used disdrometric instruments.

In Fig. 6, we have plotted the expression of the expected number of drops n_s collected by a RD-69 disdrometer as a function of the rain rate for the three families of raindrop size distributions discussed in Appendix D: Marshall—Palmer (solid line, Eq. (D.9)); Sekhon—Srivastava (dashed line, Eq. (D.10)); Willis—Tattelman (dash-dotted line, Eq. (D.12)). The product of the sampling surface S times the integrating period t plays the role of the sampling volume in a flux measurement device. From the properties of the RD-69 disdrometer, it follows that St = 3000 cm^2 s. Fig. 6 shows that there is significant variability in n_s depending on the DSD expression used to calculate the arrival rate. However, the actual number of drops detected will usually be less than that calculated theoretically because the latter overestimates the number of small drops (particularly in the exponential case). Furthermore, we have assumed that the disdrometer works as an ideal instrument that is able to detect all drops captured. In fact, deadtime problems make the RD-69 disdrometer to be far from perfect (e.g., Uijlenhoet et al., 2002). In addition, the disdrometer is less sensitive to drops of small diameters. As a consequence of these effects, the theoretical calculations above should only be considered as a first approximation to the actual number of drops sampled by a disdrometer.

In addition to the theoretical expressions for n_s, we have included in Fig. 6 two experimental datasets. In one, data were gathered from May to November 1988 at the Hoenggerberg in Zurich, Switzerland, by a RD-69 disdrometer. These have been kindly provided to us by Dr. M. Steiner of Princeton University (Steiner, 1991). This dataset includes about 30,000 1-min raindrop spectra which have been classified in 35 bins according to the precipitation intensity. For each rain rate class, the average number of drops detected by the disdrometer is plotted (closed circles). The deadtime correction was already applied to the original dataset. The error bar at each point corresponds to 2 standard deviations from the mean, which is roughly equal to $2\sqrt{n_s}$ divided by

the square root of the number of minutes in the class. The second dataset was collected by an optical spectro-pluviometer (OSP) from September 1992 to November 1993 in Marseille, France, and includes 6429 1-min spectra (Salles, 1995). The sampling surface of this device is 100 cm^2. Data have been organized in the same rain rate classes as in the first dataset (open circles). The error bars for this second dataset are larger because the number of minutes registered is less than for the first dataset. Note that the two relations based on the exponential DSD parameterization (solid line and dashed line) significantly overestimate the actual number of drops detected using both devices. This can be partly explained by the fact that both the RD-69 and the OSP are known to have problems detecting small drops. The ''saturation'' of the RD-69 (closed circles) for rain rates in excess of 20 mm h^{-1} is likely due to deadtime effects, which the applied correction procedure is not able to account for.

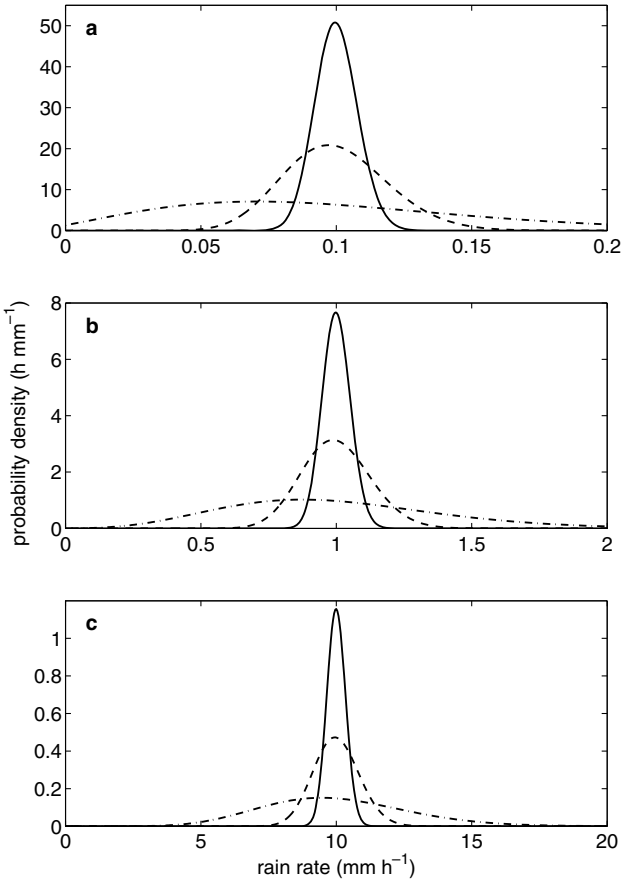

Figure 7 Exact form of the continuous part of the sampling distribution of rain rate in case the arrival process of raindrops at a surface obeys a Poisson process and the volumes of the arriving raindrops are exponentially distributed (functionally equivalent to Eq. (24)). The rain rate dependence of the arrival rate (Eq. (D.9)) and the mean volume of the raindrops are those which follow from a combination of the Marshall and Palmer (1948) distribution and the Atlas and Ulbrich (1977) raindrop terminal fall speed parameterization (Uijlenhoet and Stricker, 1999). (a) Mean rain rate is 0.1 mm h^{-1}; (b) mean rain rate is 1 mm h^{-1}; (c) mean rain rate is 10 mm h^{-1}. Dash-dotted line: integration time is 1 s; dashed line: integration time is 10 s; solid line: integration time is 60 s. In all cases the sampling surface is 50 cm^2, corresponding to a Joss and Waldvogel (1967) disdrometer.

Finally, in Fig. 7 the results of this section (where relations between n_s and R have been established) are combined with those of ''Sampling distributions'' (where the sampling distributions of bulk rainfall variables are derived as a function of n_s). This figure provides examples of (the continuous part of) the sampling distributions of the rain rate R for different mean (stationary) rainfall rates and different integration times for a 50 cm^2, Joss–Waldvogel RD-69 disdrometer. The previously noted evolution from skewed sampling distributions for small mean sample sizes to symmetrical (Gaussian) distributions for larger mean sample sizes (in stationary rainfall) is clearly observed. These results demonstrate that in many cases a Gaussian sampling distribution will probably be a reasonable assumption in stationary rainfall. The practical relevance of our theoretical analysis is (1) that it provides *exact* solutions to the sampling problem during (relatively rare) periods of stationary rainfall (e.g., drizzle), and (2) that it provides a *lower bound* to the magnitude of the sampling problem in the general situation where sampling fluctuations and natural variability exist side by side.

Conclusions

Rainfall properties estimated from raindrop size measurements show great variability. Generally speaking, three factors can explain this variability: (1) climatological factors (as different kinds of rainfall have different properties), (2) physical factors (as meteorological conditions change during rainfall events), and (3) instrumental factors (those associated with the device used to measure drop diameters). The latter includes any instrument malfunctioning, device sensitivity, and sample size effect. In this paper, we have focused on the analysis of the magnitude of the variability caused by sample size: the so-called sampling fluctuations. We have restricted our analysis to the special (limiting) case of stationary (steady) rainfall.

The first step in the analysis of sampling fluctuations involves the choice of a statistical rainfall model that includes both the statistics of drop positions in space and the distribution of drop sizes. We have chosen a marked point process model in which drop centers are assumed to be uniformly distributed in space as if they were points, which has been shown to be a plausible assumption during periods of stationary rainfall. This model implies that the number of drops in a sample volume is distributed according to a Poisson distribution. As far as the drop size distribution is concerned, our calculations apply for a general distribution, although Marshall and Palmer's exponential distribution is used to obtain numerical results.

Within this framework, it is shown how the sampling distribution of the estimator of any bulk rainfall variable converges to a Gaussian distribution in stationary rainfall. The relevant parameter controlling this evolution turns out to be the average number of drops in the sample n_s. The skewness of the distribution is more pronounced for higher order

moments of the drop size distribution. For instance, the sampling distribution of the normalized mean diameter becomes nearly Gaussian for n_s greater than 10 while the sampling distribution of the normalized rain rate remains skewed for n_s as large as 500.

A second result of our analysis is the conclusion that the median as an estimator of a bulk rainfall variable always underestimates the population value of that bulk variable. According to our analysis, the ratio between the median of an estimator of a bulk rainfall variable and its population value reads $Q_{50}/Q_p = 1 - b/n_s + O(n_s^{-2})$, following Eq. (36), where b is a constant which depends on the drop size distribution. For positively defined rainfall variables, this constant is positive and therefore the median Q_{50} always underestimates the population value Q_p. This result provides a theoretical confirmation and explanation of results in Smith et al. (1993), who had already obtained this result via Monte-Carlo simulation.

The practical relevance of the results of this paper lies in the possibility of estimating the effect of the sample size associated with a given disdrometric instrument on a particular rainfall estimator. For instance, it has been shown that for the disdrometer (RD69) and for the optical spectro-pluviometer (OSP) the average number of drops sampled for rain rates below 0.1 mm h^{-1} is less than 100 and that therefore the corresponding reflectivity estimates will be strongly influenced by sampling fluctuations. The approximation to the sampling distribution given in this paper, Eq. (B.10), provides a method to better establish the accuracy of an estimate of a bulk rainfall variable obtained from disdrometric measurements. Moreover, the developed method can be generalized to provide an estimate of the correlation between estimators of different rainfall variables induced by sampling fluctuations only. In the near future we intend to generalize the mathematical framework presented in this paper to be able to cope with non-homogeneous Poisson processes (and the resulting non-stationary rainfall processes). This will allow us to tackle the sampling problem in the general situation where sampling fluctuations and natural variability coexist.

Acknowledgments

Thanks are due to R.J. Moore (Centre for Ecology and Hydrology, UK) for providing the dataset from the UK and to Dr. M. Steiner (Princeton University, USA) for providing the dataset from Zurich. The work reported in this paper has been supported financially by EU projects VOLTAIRE (EVK2-2001-00273) and FLOODsite (GOCE-CT-2004-505420). R.U. acknowledges the EU for a Marie Curie Postdoctoral Fellowship (ENV4-CT96-5030) and the Netherlands Organization for Scientific Research (NWO) for a grant (016.021.003) in the framework of the Innovational Research Incentives Scheme (*Vernieuwingsimpuls*). This paper was finalized while R.U. was a Visiting Professor at the Grup de Recerca Aplicada en Hidrometeorologia (GRAHI), Universitat Politècnica de Catalunya, Barcelona, Spain, in the framework of a *Distinció de Recerca de la Generalitat de Catalunya*.

Appendix A. Characteristic function

The characteristic function (cf) of Q (Eq. (11)),

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} e^{i\omega q} f(q)\, dq, \tag{A.1}$$

can generally be computed easier than the probability density function $f(q)$. Note that by definition $\hat{f}(\omega)$ is the average (expectation) of $\exp(i\omega q)$ with respect to $f(q)$ and that consequently it can be written as $\hat{f}(\omega) = \langle \exp(i\omega q) \rangle$. To obtain the characteristic function $\hat{f}(\omega)$, we first expand $f(q)$ as

$$f(q) = \sum_{k=0}^{\infty} p(n=k) f(q|k), \tag{A.2}$$

using the total probability theorem, or in other words, using the property that events with different n are mutually exclusive. The function $f(q|k)$ is the probability density of Q conditional on the event $n = k$, that is, given that exactly k drops have been detected. Hence, $f(q|k)$ is the pdf of the sum of k independent random variables, that is the convolution of their individual densities. Since the cf of the convolution of a number of densities equals the product of the cfs of the individual densities (a special case of a general property of the Fourier transform), the cf of $f(q|k)$ can be expressed in terms of the characteristic functions of the random variables h as (Kendall and Stuart, 1977)

$$\hat{f}(\omega|k) = \int_{-\infty}^{\infty} e^{i\omega q} f(q|k)\, dq = [\hat{h}(\omega)]^k, \tag{A.3}$$

where

$$\hat{h}(\omega) = \int_{-\infty}^{\infty} e^{i\omega h} p(h)\, dh = \int_{-\infty}^{\infty} e^{i\omega h(D)} f_d(D)\, dD \tag{A.4}$$

is the cf of h calculated using its probability density, Eq. (14). If Eq. (A.2) is multiplied by $\exp(i\omega q)$ and integrated over q from $-\infty$ to ∞, the following relation between $\hat{f}(\omega)$ and $\hat{f}(\omega|k)$ results:

$$\hat{f}(\omega) = \sum_{k=0}^{\infty} p(n=k) \hat{f}(\omega|k). \tag{A.5}$$

Introducing in this expression the assumption that the number of drops sampled is distributed according to a Poisson law (Eq. (2)) we obtain the cf of Q

$$\hat{f}(\omega) = \sum_{k=0}^{\infty} \frac{n_s^k}{k!} e^{-n_s} [\hat{h}(\omega)]^k = \exp\left\{-n_s[1 - \hat{h}(\omega)]\right\}. \tag{A.6}$$

From this expression, it is possible to obtain all moments of Q as functions of the cf $\hat{h}(\omega)$, that is, the characteristic function of the probability density of h, Eq. (14). Indeed, the rth order moment of the estimator Q,

$$\langle Q^r \rangle = \int_{-\infty}^{\infty} q^r f(q)\, dq, \tag{A.7}$$

can be computed directly from the cf $\hat{f}(\omega)$ as

$$\langle Q^r \rangle = (-i)^r \frac{d^r}{d\omega^r} \hat{f}(\omega)\bigg|_{\omega=0}, \tag{A.8}$$

as can be seen from Eq. (A.1). In other words, $\hat{f}(\omega)$ is the moment generating function of Q (Mood et al., 1974). An

alternative, in this case more convenient, set of descriptors of $f(q)$ are the so-called *cumulants* of the estimator Q, $\kappa_r(Q)$. In general, cumulants of order r are related to moments of order r and lower (Kendall and Stuart, 1977). Here, they can be readily computed in terms of the moments of h. By definition, the natural logarithm of the cf $\hat{h}(\omega)$ yields the cumulant generating function of Q (Mood et al., 1974). The cumulants are then obtained from Eq. (A.6) according to

$$\kappa_r(Q) = (-\mathrm{i})^r \frac{\mathrm{d}^r}{\mathrm{d}\omega^r} \ln \hat{f}(\omega)\bigg|_{\omega=0} = n_s \langle h^r \rangle, \tag{A.9}$$

where $\langle h^r \rangle$ is the rth order moment of h,

$$\langle h^r \rangle = \int_{-\infty}^{\infty} h^r p(h)\,\mathrm{d}h = \int_{-\infty}^{\infty} h(D)^r f_d(D)\,\mathrm{d}D. \tag{A.10}$$

Hence, for the Poisson marked point process considered here, the effect of sample size on the cumulants of the sampling distribution of a bulk rainfall variable Q is entirely contained in the average number of drops in the sample n_s, whereas the effect of the drop size distribution is entirely contained in the moments of h.

The mean value of Q, μ_Q, and its variance, σ_Q^2, are particular cases of this expression. The mean value of Q is simply the average value of the random variable h, $\langle h \rangle$, times the average number of drops n_s,

$$\mu_Q = \int_{-\infty}^{\infty} q f(q)\,\mathrm{d}q = n_s \langle h \rangle. \tag{A.11}$$

This result shows that Q is a non-biased estimator of the properties of the parent population $f_d(D)$, which means that the average of the estimator of a property of the parent population (Eq. (11)) is equal to the population value. For instance, given that the population value of the liquid water content is W_0, the average of the estimator (Eq. (13)) is precisely $W_0 = \pi \rho_w N_T \langle D^3 \rangle / 6$ (from Eqs. (1) and (12)). The variance equals the second order cumulant and reads

$$\sigma_Q^2 = \langle Q^2 \rangle - \langle Q \rangle^2 = n_s \langle h^2 \rangle. \tag{A.12}$$

An important consequence that follows from the expressions of the mean and the variance given above is that the coefficient of variation of the estimator Q,

$$\mathrm{CV} = \frac{\sigma_Q}{\mu_Q} = \frac{\sqrt{\langle h^2 \rangle}}{\langle h \rangle} \frac{1}{\sqrt{n_s}}, \tag{A.13}$$

converges to zero as $1/\sqrt{n_s}$ when $n_s \to \infty$. Therefore, the magnitude of the sampling fluctuations decreases as $1/\sqrt{n_s}$ when the mean sampling size goes to infinity. Joss and Waldvogel (1969) proved this property for the estimator of the mean number of drops, the rain rate and the reflectivity and it was generalized to other moments of the drop size distribution by Gertzman and Atlas (1977). In these references, it was assumed that the DSD had an exponential shape whereas we did not make any assumption of this kind.

The coefficient of skewness γ of the estimator Q can also be computed from Eq. (A.9) and reads

$$\gamma = \frac{\kappa_3(Q)}{\sigma_Q^3} = \frac{\langle h^3 \rangle}{\langle h^2 \rangle^{3/2}} \frac{1}{\sqrt{n_s}}. \tag{A.14}$$

This coefficient gives an idea of the asymmetry of the probability density function of the estimator Q. Hence, $f(q)$ becomes symmetrical as $n_s \to \infty$. In Appendix B, we exploit the behavior of the cumulants of Q to draw some more detailed conclusions about the shape of the pdf of the random variable Q.

Appendix B. Edgeworth expansion

The probability density of the estimator Q could in principle be obtained by inverting the characteristic function $\hat{f}(\omega)$ (Eq. (A.6)), although this task cannot be carried out explicitly in general. Nevertheless, when the sample size increases, it is expected that the estimator Q becomes a Gaussian random variable. This assumption, implicit in most analyses of sampling fluctuations, is based on the central limit theorem (Cramér, 1946).

The analysis of the cumulants of the standardized variable

$$z = \frac{Q - \mu_Q}{\sigma_Q} \tag{B.1}$$

suggests indeed that the probability density function of z, as the mean sample size n_s goes to infinity, converges to the standard normal pdf

$$G(z) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-z^2/2}. \tag{B.2}$$

It follows from Eq. (B.1) and the definition of the corresponding cf (equivalent to Eq. (A.1)) that the cumulant generating function of z is related to that of Q according to $-\mathrm{i}\omega\mu_Q/\sigma_Q + \ln \hat{f}(\omega/\sigma_Q)$. The cumulants of the random variable z now follow directly from subsequent derivatives of this expression with respect to ω (Eq. (A.9)) and read $\kappa_1 = 0$, $\kappa_2 = 1$, and

$$\kappa_r = \frac{\kappa_r(Q)}{\sigma_Q^r} = \frac{\langle h^r \rangle}{\langle h^2 \rangle^{r/2}} \frac{1}{n_s^{r/2-1}} = a_r \frac{1}{n_s^{r/2-1}} \quad (r \geqslant 3), \tag{B.3}$$

where the coefficients

$$a_r = \frac{\langle h^r \rangle}{\langle h^2 \rangle^{r/2}} \tag{B.4}$$

have been defined for later use (note that $\kappa_3 = \gamma$, Eq. (A.14)). This result shows that, when $n_s \to \infty$, all cumulants of z of order greater than 2 converge to zero. Thus, the cumulant generating function of z becomes, in this limit, $-\omega^2/2$, which implies that the characteristic function of z becomes that of a Gaussian random variable with zero mean and unit standard deviation

$$\hat{G}(\omega) = \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}\omega z} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-z^2/2}\,\mathrm{d}z = \mathrm{e}^{-\omega^2/2}. \tag{B.5}$$

Consequently, the probability density $p(z)$ of the standardized random variable z is Gaussian and given by Eq. (B.2). As Q is a linear combination of z, $Q = \sigma_Q z + \mu_Q$, the random variable Q also becomes Gaussian in the limit $n_s \to \infty$.

Nevertheless, for small samples, the distribution of Q shows the effect of the skewness, given by Eq. (A.14). As a consequence, the probability density function of Q will not be Gaussian for finite samples. It is then natural to ask which are the corrections to the Gaussian shape required

when n_s is finite. The answer to this question is the so-called Edgeworth expansion (Cramér, 1946; Kendall and Stuart, 1977), which yields an asymptotic development of the pdf $p(z)$ in powers of the mean sample size n_s.

The derivation of the Edgeworth expansion for the estimator Q when the number n of random variables h_Q to be added is not random, Eq. (11), can be found in many references (Cramér, 1946; Abramowitz and Stegun, 1972; Kendall and Stuart, 1977). When the number n becomes a Poisson random variable, the demonstration follows a similar reasoning. We reproduce here the main steps for the sake of completeness.

The characteristic function of $p(z)$, in terms of its cumulants κ_r, Eq. (B.3), is

$$\hat{p}(\omega) = \int_{-\infty}^{\infty} e^{i\omega z} p(z)\, dz = \exp\left[-\frac{\omega^2}{2} + \sum_{r=3}^{\infty} \kappa_r \frac{(i\omega)^r}{r!}\right]. \tag{B.6}$$

In order to invert this cf, $e^{\omega^2/2}\hat{p}(\omega)$ is expanded in powers of ω and the series' terms are collected in powers of $1/\sqrt{n_s}$. This method leads to the following expansion up to order $n_s^{-3/2}$:

$$\hat{p}(\omega) = e^{-\omega^2/2}\left\{1 + \frac{\kappa_3}{3!}(i\omega)^3 + \left[\frac{\kappa_4}{4!}(i\omega)^4 + \frac{\kappa_3^2}{2!3!^2}(i\omega)^6\right] + O(n_s^{-3/2})\right\}. \tag{B.7}$$

According to the cumulant values (Eq. (B.3)) the term containing κ_3 in this development is of order $n_s^{-1/2}$ whereas the terms inside the square brackets are of order n_s^{-1}. This expansion can be readily inverted term by term because the Fourier transform of the rth derivative of the standardized Gaussian is

$$\int_{-\infty}^{\infty} e^{i\omega z} G^{(r)}(z)\, dz = (-i\omega)^r e^{-\omega^2/2}, \tag{B.8}$$

and the rth derivative of $G(z)$ is related to the Hermite polynomial $He_r(z)$ according to

$$G^{(r)}(z) = \frac{1}{\sqrt{2\pi}} \frac{d^r}{dz^r} e^{-z^2/2} = (-1)^r He_r(z) G(z) \tag{B.9}$$

(Abramowitz and Stegun, 1972). Consequently, as n_s goes to infinity, $p(z)$ can be expanded as

$$p(z) \sim G(z)\left\{1 + \frac{a_3}{3!}He_3(z)\frac{1}{\sqrt{n_s}} + \left[\frac{a_4}{4!}He_4(z) + \frac{10a_3^2}{6!}He_6(z)\right]\frac{1}{n_s} + O(n_s^{-3/2})\right\}, \tag{B.10}$$

where the coefficients a_r are defined by Eq. (B.4). Several comments are necessary at this point. First, when $n_s \to \infty$, the function $p(z)$ converges to $G(z)$, a Gaussian pdf with zero mean and unit variance, in agreement with the central limit theorem discussed above. For finite n_s, the following (higher order) terms give corrections to the normal limiting behavior. Note, however, that the expansion for $p(z)$ is asymptotic in the sense that the ratio between the series and $p(z)$ converges to 1 as $n_s \to \infty$. For finite n_s the series may even diverge (e.g., Blinnikov and Moessner, 1998). However, the Edgeworth expansion, when truncated at a given order, yields a representation of $p(z)$ that improves as n_s grows. Secondly, the expansion of $f(q)$ can be obtained from Eq. (B.10) taking into account the relation

$$f(q) = \frac{1}{\sigma_Q} p(z)\Big|_{z=(q-\mu_Q)/\sigma_Q}, \tag{B.11}$$

which is a consequence of the relation between z and Q (Eq. (B.1)). Thirdly, the Hermite polynomials of even order are symmetric about $z = 0$ whereas the rest, with odd indexes, are asymmetric (Abramowitz and Stegun, 1972). This feature implies that the first correction to the normal behavior, of order $1/\sqrt{n_s}$, introduces an asymmetry into the density of the estimator Q. Finally, note that the original drop size distribution only appears through the function $f_d(D)$ used to calculate the moments $\langle h^k \rangle$ in the coefficients a_r (Eq. (B.4)). Since we model rainfall as a marked point process, it is interesting to mention that the properties of the point process (Poisson) are entirely contained in n_s, whereas the properties of the marks are completely included in the coefficients a_r.

Appendix C. Expansion of the median

The median is defined as

$$\int_{-\infty}^{Q_{50}} f(q)\, dq = \int_{Q_{50}}^{\infty} f(q)\, dq = \frac{1}{2}. \tag{C.1}$$

After the change of variables $z = (q - \mu_Q)/\sigma_Q$, we obtain the following equation for $z_0 = (Q_{50} - \mu_Q)/\sigma_Q$,

$$\int_{-\infty}^{z_0} p(z)\, dz = \int_{z_0}^{\infty} p(z)\, dz. \tag{C.2}$$

Introduction of the Edgeworth expansion, Eq. (B.10), into this equation yields the following equation for z_0 (to order $n_s^{-3/2}$):

$$e^{z_0^2/2}\,\mathrm{erf}(z_0/\sqrt{2}) = \frac{2a_3}{3!\sqrt{2\pi}}He_2(z_0)\frac{1}{\sqrt{n_s}} + \left[\frac{2a_4}{4!\sqrt{2\pi}}He_3(z_0) + \frac{20a_3^2}{6!\sqrt{2\pi}}He_5(z_0)\right]\frac{1}{n_s} + O(n_s^{-3/2}), \tag{C.3}$$

where $\mathrm{erf}(z)$ is the error function (Abramowitz and Stegun, 1972)

$$\mathrm{erf}(z) = \frac{2}{\sqrt{\pi}}\int_0^z e^{-t^2}\, dt. \tag{C.4}$$

The solution to Eq. (C.3) can be obtained perturbationally by expanding z_0 in powers of $1/\sqrt{n_s}$,

$$z_0 = \frac{\beta_1}{\sqrt{n_s}} + \frac{\beta_2}{n_s} + O(n_s^{-3/2}), \tag{C.5}$$

for $n_s \to \infty$. Note that z_0 converges to 0 when $n_s \to \infty$ because $f(q)$ becomes Gaussian in this limit and therefore $\langle Q \rangle = Q_{50}$, which implies $z_0 = 0$. After substituting the small z expansion of $\mathrm{erf}(z)$ (Abramowitz and Stegun, 1972)

$$\mathrm{erf}(z) = \frac{2}{\sqrt{\pi}}e^{-z^2}\left[z + \frac{2}{3}z^3 + O(z^5)\right], \tag{C.6}$$

the left-hand side of Eq. (C.3) can be expanded in powers of $1/\sqrt{n_s}$ according to

$$e^{z_0^2/2}\mathrm{erf}(z_0/\sqrt{2}) = \frac{2}{\sqrt{2\pi}}\left[\frac{\beta_1}{\sqrt{n_s}} + \frac{\beta_2}{n_s} + O(n_s^{-3/2})\right]. \quad (C.7)$$

To the order considered, $n_s^{-3/2}$, the first term on the right-hand side of Eq. (C.3) gives the contribution

$$\frac{2a_3}{3!\sqrt{2\pi}}\mathrm{He}_2(z_0)\frac{1}{\sqrt{n_s}} = -\frac{2}{\sqrt{2\pi}}\frac{a_3}{3!\sqrt{n_s}} + O(n_s^{-3/2}) \quad (C.8)$$

and the rest of the terms does not contribute to this order. Equating the expansions given by Eqs. (C.7) and (C.8), we obtain $\beta_1 = -a_3/3!$ and $\beta_2 = 0$. Therefore, the solution to the equation for z_0 is

$$z_0 = -\frac{a_3}{3!\sqrt{n_s}} + O(n_s^{-3/2}), \quad (C.9)$$

which recalling the meaning of z_0 and the fact that $\sigma_Q = \sqrt{n_s\langle h^2\rangle}$ (Eq. (A.12)), yields

$$Q_{50} = \langle Q\rangle - \frac{1}{3!}\frac{\langle h^3\rangle}{\langle h^2\rangle} + O(n_s^{-1}). \quad (C.10)$$

The expression for a_r (Eq. (B.4)) for $r = 3$ has been used to obtain this result. Another approach to obtain the same expression would have been the so-called Cornish–Fisher expansion (Abramowitz and Stegun, 1972; Kendall and Stuart, 1977). However, we have used the perturbational method to show an application of the Edgeworth expansion.

Appendix D. The expected number of drops sampled

Many of the studies devoted to the drop size distribution (DSD) have analyzed the parameterization of this function using global rainfall variables as parameters. Since Marshall and Palmer (1948), it has been customary to use just one variable as a parameter to describe the variation of the raindrop size distribution (provided the type of rain does not change). The rain rate is the variable most often used as a reference, although any other would apply. In this way, we could write the DSD as $N(D,R)$ to explicitly point out that it depends both on the diameter and on the reference variable, which we take as the rain rate R. It has been shown that most DSD expressions proposed in the literature can be written as the scaling law (Sempere Torres et al., 1994, 1998)

$$N(D,R) = R^{\alpha}g(D/R^{\beta}), \quad (D.1)$$

where α and β are two exponents and $g(x)$ is the function that characterizes the DSD shape (see Uijlenhoet et al., 2003a,b for recent applications of this formalism). Both the exponents and the function g are independent of R. This formulation implies that the drop concentration, N_T, reads

$$N_T = R^{\alpha+\beta}\int_0^{\infty} g(x)\,dx = C_V R^{\alpha+\beta}, \quad (D.2)$$

and that the probability density function of the drop diameters, $p(D) = N(D,R)/N_T$, scales as

$$p(D) = R^{-\beta}\tilde{g}(D/R^{\beta}), \quad (D.3)$$

where $\tilde{g}(x) = g(x)/C_V$ and C_V was defined in Eq. (D.2) (Uijlenhoet et al., 2003a).

The expression for the drop concentration, Eq. (D.2), allows us to write the expected number of drops measured by a disdrometric instrument of volume integrating type as a function of the rain rate as

$$n_s = V_s N_T = C_V V_s R^{\alpha+\beta}, \quad (D.4)$$

where V_s is the sampling volume. Therefore, the number of drops sampled grows as a power of the rain rate. For a time integrating instrument, the same behavior is obtained provided the drop terminal velocity depends on a power of the diameter,

$$v(D) = aD^b. \quad (D.5)$$

In the power-law model of Atlas and Ulbrich (1977), based on the data of Gunn and Kinzer (1949), $a = 3.78\ \mathrm{m\ s^{-1}\ mm^{-b}}$, $b = 0.67$, D is expressed in mm, and v in $\mathrm{m\ s^{-1}}$. The arrival rate computed using this model and Eq. (6) reads

$$\lambda = R^{\alpha+\beta(1+b)}\int_0^{\infty} ax^b g(x)\,dx = C_A R^{\alpha+\beta(1+b)}, \quad (D.6)$$

when $v(D)$ is the power law Eq. (D.5) and the DSD is given by Eq. (D.1). This expression can be simplified further by taking into account the fact that the exponents α and β satisfy the equation

$$\alpha + \beta(4+b) = 1, \quad (D.7)$$

as required by the self-consistency of the expression for $N(D,R)$ (Sempere Torres et al., 1994, 1998). Thus, the exponent of R in the expression for λ becomes $1-3\beta$ and the expected number of drops sampled by a time integrating instrument is

$$n_s = \lambda St = C_A St R^{1-3\beta}. \quad (D.8)$$

The expression for n_s as a function of R allows us to estimate the range of rain rates for which the sampling fluctuations are of concern. We will concentrate on the RD-69 disdrometer (Joss and Waldvogel, 1967), which has a sampling surface of $50\ \mathrm{cm^2}$ and a sampling time of 60 s. In addition to these properties, the raindrop size distribution $N(D)$, or even better $N_A(D)$, must be provided to calculate the arrival rate λ. We consider three different cases:

1. The Marshall and Palmer distribution (Eq. (16)), with $N_0 = 0.08\ \mathrm{cm^{-4}}$, $\Lambda = 41R^{-0.21}\ \mathrm{cm^{-1}}$, and D in cm, thought to be representative for stratiform conditions. Note that $\alpha = 0$, $\beta = 0.21$ and $b = 0.67$ do not satisfy Eq. (D.7) and that as a result the Marshall–Palmer distribution is not self-consistent (Uijlenhoet and Stricker, 1999; Uijlenhoet, 2001). From now on the units of the rain rate R are $\mathrm{mm\ h^{-1}}$. When the arrival rate is computed with this DSD, the following result is reached:

$$\lambda_1 = 0.259R^{0.35}\ \mathrm{drops\ cm^{-2}\ s^{-1}}. \quad (D.9)$$

   This expression was obtained without considering any maximum diameter, because it can be shown that the effect of the diameter truncation on the calculation of λ is less than 1% for $R < 250\ \mathrm{mm\ h^{-1}}$.
2. The exponential DSD proposed by Sekhon and Srivastava (1971) for thunderstorm (convective) conditions. In this distribution the N_0-factor (Eq. (16)) depends on R accord-

ing to $N_0 = 0.07R^{0.37}$ cm^{-4}. The parameter Λ in this case reads $\Lambda = 38R^{-0.14}$ cm^{-1}. The value of the arrival rate obtained with this drop size distribution is

$$\lambda_2 = 0.257R^{0.60} \text{ drops cm}^{-2}\text{ s}^{-1}. \quad (D.10)$$

3. The 'intense rainfall' DSD of Willis and Tattelman (1989). This distribution has the shape of a gamma function and reads

$$N(D,R) = 50.6R^{0.3}\left(\frac{D}{R^{0.153}}\right)^{2.16}\exp\left(-56.8\frac{D}{R^{0.153}}\right) \quad (D.11)$$

with D in cm and $N(D,R)$ in cm^{-4}. The arrival rate in this case reads

$$\lambda_3 = 0.0831R^{0.556} \text{ drops cm}^{-2}\text{ s}^{-1}. \quad (D.12)$$

Fig. 6 presents a comparison of these theoretical expressions with two experimental datasets.

References

Abramowitz, M., Stegun, I.A., 1972. Handbook of Mathematical Functions, eighth ed. Dover, New York, 1046pp.

Atlas, D., Ulbrich, C.W., 1977. Path- and area integrated rainfall measurement by microwave attenuation in the 1–3 cm band. J. Appl. Meteorol. 16, 1322–1331.

Bardsley, W.E., 1995. Comment on ''A stochastic model relating rainfall intensity to raindrop processes'' by Smith, J.A., De Veaux, R.D.. Water Resour. Res. 31, 1607–1609.

Battan, L.J., 1973. Radar Observation of the Atmosphere. The University of Chicago Press, Chicago, IL, 324pp.

Blinnikov, S., Moessner, R., 1998. Expansions for nearly Gaussian distributions. Astron. Astrophys. Suppl. Ser. 130, 193–205.

Bringi, V.N., Chandrasekar, V., Hubbert, J., Gorgucci, E., Randeu, W.L., Schoenhuber, M., 2003. Raindrop size distribution in different climatic regimes from disdrometer and dual-polarized radar analysis. J. Atmos. Sci. 60, 354–365.

Chandrasekar, V., Bringi, V.N., 1987. Simulation of radar reflectivity and surface measurements of rainfall. J. Atmos. Ocean. Technol. 4, 464–478.

Chandrasekar, V., Gori, E.G., 1991. Multiple disdrometer observations of rainfall. J. Appl. Meteorol. 30, 1514–1520.

Cornford, S.G., 1967. Sampling errors in measurements of raindrop and cloud droplet concentrations. Meteorol. Mag. 96, 271–282.

Cornford, S.G., 1968. Sampling errors in measurements of particle size distributions. Meteorol. Mag. 97, 12–16.

Cox, D.R., Isham, V., 1980. Point Processes. Chapman & Hall, London, 188pp.

Cramér, H., 1946. Mathematical Methods of Statistics. Princeton University Press, Princeton, pp. 208–232.

de Bruin, H.A.R., 1977. The accuracy of measuring areal precipitation with a rain gauge network. In: Proceedings and Informations, vol. 23, Precipitation and Measurements of Precipitation, Committee for Hydrological Research TNO, The Hague, pp. 17–46.

Desaulniers-Soucy, N., Lovejoy, S., Schertzer, D., 2001. The HYDROP experiment: an empirical method for the determination of the continuum limit in rain. Atmos. Res. 59–60, 163–197.

Fabry, F., 1996. On the determination of scale ranges for precipitation fields. J. Geophys. Res. D 101, 12819–12826.

Gabella, M., Pavone, S., Perona, G., 2001. Errors in the estimate of the fractal correlation dimension of raindrop spatial distribution. J. Appl. Meteorol. 40, 664–668.

Gabella, M., Perona, G., 2001. Comments on ''Errors in the estimate of the fractal correlation dimension of raindrop spatial distribution'' – Reply. J. Appl. Meteorol. 40, 2099.

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

Gradshteyn, I.S., Ryzhik, I.M., 1980. Tables of Integrals, Series, and Products. Academic Press, New York, 1160pp.

Gunn, R., Kinzer, G.D., 1949. The terminal velocity of fall for water droplets in stagnant air. J. Meteorol. 6, 243–248.

Hauser, D., Amayenc, P., 1981. A new method for deducing hydrometeor-size distributions and vertical air motions from Doppler radar measurements at vertical incidence. J. Appl. Meteorol. 20, 547–555.

Hosking, J.G., Stow, C.D., 1987. The arrival rate of raindrops at the ground. J. Clim. Appl. Meteorol. 26, 433–442.

Illingworth, A.J., Stevens, C.J., 1987. An optical disdrometer for the measurement of raindrop size spectra in windy conditions. J. Atmos. Ocean. Technol. 4, 411–421.

Jameson, A.R., Kostinski, A.B., 1998. Fluctuation properties of precipitation. Part II: Reconsideration of the meaning and measurement of raindrop size distributions. J. Atmos. Sci. 55, 283–294.

Jameson, A.R., Kostinski, A.B., 1999. Fluctuation properties of precipitation. Part V: Distribution of rain rates – theory and observations in clustered rain. J. Atmos. Sci. 56, 3920–3932.

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

Jameson, A.R., Kostinski, A.B., 2001a. Comments on ''Errors in the estimate of the fractal correlation dimension of raindrop spatial distribution''. J. Appl. Meteorol. 40, 2098.

Jameson, A.R., Kostinski, A.B., 2001b. Reconsideration of the physical and empirical origins of Z–R relations in radar meteorology. Quart. J. Roy. Meteorol. Soc. 127, 517–538.

Jameson, A.R., Kostinski, A.B., 2001c. What is a raindrop size distribution? Bull. Am. Meteorol. Soc. 82, 1169–1177.

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

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

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

Joss, J., Waldvogel, A., 1967. Ein spektrograph für niederschlagstropfen mit automatischer auswertung. Pageoph. 68, 240–246.

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

Kendall, M.G., Stuart, A., 1977. The advanced theory of statistics, fourth ed.Distribution Theory, vol. I Charles Griffin, London, 472pp.

Kendall, M.G., Stuart, A., 1979. The advanced theory of statistics, fourth ed.Inference and Relationship, vol. II Charles Griffin, London, 748pp.

Knollenberg, R.G., 1970. The optical array: an alternative to scattering or extinction for airborne particle size determination. J. Appl. Meteorol. 9, 86–103.

Kollias, P., Albrecht, B.A., Marks, F., 2002. Why Mie? Accurate observations of vertical air velocities and raindrops using a cloud radar. Bull. Am. Meteorol. Soc. 83, 1471–1483.

Kostinski, A.B., Jameson, A.R., 1997. Fluctuation properties of precipitation. Part I: On deviations of single-size drop counts from the Poisson distribution. J. Atmos. Sci. 54, 2174–2186.

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

Kruger, A., Krajewski, W.F., 2002. Two-dimensional video disdrometer: a description. J. Atmos. Ocean. Technol. 19, 602–617.

Lavergnat, J., Golé, P., 1998. A stochastic raindrop time distribution model. J. Appl. Meteorol. 37, 805–818.

Laws, J.O., Parsons, D.A., 1943. The relation of raindrop-size to intensity. Trans. Am. Geophys. Union 24, 452–460.

Lovejoy, S., Lilley, M., Desaulniers-Soucy, N., Schertzer, D., 2003. Large particle number limit in rain. Phys. Rev. E 68, 025301-1-4.

Lovejoy, S., Schertzer, D., 1990. Fractals, raindrops and resolution dependence of rain measurements. J. Appl. Meteorol. 29, 1167–1170.

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

Mood, A.M., Graybill, F.A., Boes, D.C., 1974. Introduction to the Theory of Statistics, third ed. McGraw-Hill, Singapore, 564pp.

Moore, R.J., Jones, D.A., Cox, D.R., Isham, V., 2000. Design of the HYREX raingauge network. Hydrol. Earth Syst. Sci. 4, 523–530.

Mueller, E.A., 1966. Radar Cross Sections from Drop Size Spectra. Doctoral Thesis, University of Illinois, Urbana-Champaign, 89pp.

Porrà, J.M., Sempere Torres, D., Creutin, J.-D., 1998. Modeling of drop size distribution and its applications to rainfall measurements from radar. In: Gupta, V.K., Barndorff-Nielsen, O.E., Perez-Abreu, V., Waymire, E. (Eds.), Stochastic Methods in Hydrology: Rain, Landforms and Floods. World Scientific, Singapore, pp. 73–84.

Salles, C., 1995. Analyse microphysique de la pluie au sol: Mesures par spectropluviomètre optique et méthodes statistiques d'analyse spectrale et de simulation numérique. Doctoral thesis, Université Joseph Fourier, Grenoble, 212pp.

Salles, C., Creutin, J.-D., Sempere Torres, D., 1998. The optical spectropluviometer revisited. J. Atmos. Ocean. Technol. 15, 1215–1222.

Sasyo, Y., 1965. On the probabilistic analysis of precipitation particles. In: Proceedings of the International Conference on Cloud Physics, International Association of Meteorological and Atmospheric Physics, pp. 254–259.

Sekhon, R.S., Srivastava, R.C., 1971. Doppler radar observations of drop-size distributions in a thunderstorm. J. Atmos. Sci. 28, 983–994.

Sempere Torres, D., Porrà, J.M., Creutin, J.-D., 1994. A general formulation for raindrop size distribution. J. Appl. Meteorol. 33, 1494–1502.

Sempere Torres, D., Porrà, J.M., Creutin, J.-D., 1998. Experimental evidence of a general description for raindrop size distribution properties. J. Geophys. Res. D 103, 1785–1797.

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

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

Steiner, M., 1991. Die kombination von Doppler- und polarisationsradarmessungen im niederschlag. Neue perspektiven für die radarmeteorologie. Doctoral Thesis, Swiss Federal Institute of Technology, Zürich, 116pp.

Stow, C.D., Jones, K., 1981. A self-evaluating disdrometer for the measurement of raindrop size and charge at the ground. J. Appl. Meteorol. 20, 1160–1176.

Uijlenhoet, R., 1999. Parameterization of rainfall microstructure for radar meteorology and hydrology. Doctoral Thesis, Wageningen University, Wageningen, 279pp.

Uijlenhoet, R., 2001. Raindrop size distributions and radar reflectivity-rain rate relationships for radar hydrology. Hydrol. Earth Syst. Sci. 5, 615–627.

Uijlenhoet, R., Smith, J.A., Steiner, M., 2003a. The microphysical structure of extreme precipitation as inferred from ground-based raindrop spectra. J. Atmos. Sci. 60, 1220–1238.

Uijlenhoet, R., Steiner, M., Smith, J.A., 2002. Influence of disdrometer deadtime correction on self-consistent analytical parameterizations for raindrop size distributions. Proc. ERAD 1, 104–112.

Uijlenhoet, R., Steiner, M., Smith, J.A., 2003b. Variability of raindrop size distributions in a squall line and implications for radar rainfall estimation. J. Hydrometeorol. 4, 43–61.

Uijlenhoet, R., Stricker, J.N.M., 1999. A consistent rainfall parameterization based on the exponential raindrop size distribution. J. Hydrol. 218, 101–127.

Uijlenhoet, R., Stricker, J.N.M., Torfs, P.J.J.F., Creutin, J.-D., 1999. Towards a stochastic model of rainfall for radar hydrology: testing the poisson homogeneity hypothesis. Phys. Chem. Earth B 24, 747–755.

Vanmarcke, E., 1983. Random Fields: Analysis and Synthesis. The MIT Press, Cambridge, MA, 382pp.

Willis, P.T., Tattelman, P., 1989. Drop-size distributions associated with intense rainfall. J. Appl. Meteorol. 28, 3–15.

Wirth, E., Zoltán, C., Székely, C., 1983. On a sampling error in hailpad measurements. J. Clim. Appl. Meteorol. 22, 2100–2102.

Wong, R.K.W., Chidambaram, N., 1985. Gamma size distribution and stochastic sampling errors. J. Clim. Appl. Meteorol. 24, 568–579.

Zawadzki, I., 1995. Is rain fractal? In: Kundzewicz, Z.W. (Ed.), New Uncertainty Concepts in Hydrology and Water Resources. Cambridge University Press, Cambridge, pp. 104–108.