Hydrol. Earth Syst. Sci. Discuss., 8, 5605–5629, 2011
www.hydrol-earth-syst-sci-discuss.net/8/5605/2011/
doi:10.5194/hessd-8-5605-2011

Hydrology and Earth System Sciences Discussions

Skewness as measure of the invariance of instantaneous renormalized drop diameter distributions – Part 1: Convective vs. stratiform precipitation

M. Ignaccolo and C. De Michele

DIIAR, Politecnico di Milano, Milano, Italy

Received: 26 May 2011 – Accepted: 27 May 2011 – Published: 14 June 2011

Correspondence to: M. Ignaccolo (m.ignaccolo@diiar.polimi.it)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

We investigate the variability of the instantaneous distribution shape of the renormalized drop diameter making use of the third order central moment: the *skewness*. Disdrometer data, collected at Darwin Australia, are considered either as whole or as divided in convective and stratiform precipitation intervals. We show that in all cases the distribution of the skewness is strongly peaked around 0.64. This allows to identify a most common distribution of renormalized drop diameters and two main variations, one with larger and one with smaller skewness. The distributions' shapes are independent from the stratiform vs. convective classification.

1 Introduction

The term drop size distribution (DSD) acknowledges the stochastic nature of the occurrence of drop diameters. In particular the DSD is the concentration $\mathcal{N}(D)$ of drops per unit volume and unit diameter, namely

$$\mathcal{N}(D) = N_V p(D) \tag{1}$$

where N_V is the number of drops per cubic meter and $p(D)$ is the density per diameter millimeter. Marshall and Palmer (1948) used an exponential functional form for the density $p(D)$. However it was soon clear that the exponential form was a result of the long (≥ 30 min) time intervals used for sampling drop diameters. The shape DSDs is highly variable even inside a single shower, a property which has led Joss and Gori (1978) to introduce the concept of instantaneous DSDs (distributions sampled over 1–2 min time intervals). Ulbrich (1983) introduced the gamma distribution as functional form for fitting the instantaneous DSDs. Although other functional forms have been proposed (Feingold and Levin, 1987), the gamma distribution is the one adopted in the overwhelming majority of cases.

HESSD

8, 5605–5629, 2011

Aside from the search of a proper functional form to best describe the DSD, investigations have been made to link the high variability of DSDs to few parameters following the introduction of "renormalized" DSDs by Sekhon and Srivastava (1971). The goal of this type of analysis is to describe DSDs in terms of rainfall bulk variables (e.g. rainfall rate, liquid water content, mean volume diameter) and/or identify a possible "universal", invariant in space and time, shape for the renormalized DSD. So far many different renormalization procedures have been proposed (Willis, 1984; Sempere Torres et al., 1994; Maki et al., 2001; Testud et al., 2001; Uijlenhoet et al., 2003; Campos et al., 2006; Hazenber et al., 2011). The results are not conclusive in the sense that an "universal" shape has not been identified, yet the existing evidences indicate that indeed the great variability observed in the distributions of drop sizes could be described with few parameters.

A novel renormalization procedure has been introduced recently by Ignaccolo et al. (2009). The renormalized spectra have been shown (Ignaccolo and De Michele, 2010) to posses the following properties. (1) Synoptic origin invariance for a fixed observation site (Darwin, AUS): convective and stratiform precipitation databases have same distribution. (2) Rainfall rate invariance for a fixed observation site (Darwin, AUS): databases built according to different rainfall rate classes (Tokay and Short, 1996) share a common distribution. (3) Cross invariance: the distributions in (1) and (2) are essentially *identical*.

Do these results indicate the possible existence of a "universal" drop diameters distribution? In this manuscript we investigate this matter. More in particular we want to discuss the possible invariance of instantaneous, 1 min sampling, renormalized drop spectra. In fact, the properties (1), (2), and (3), above mentioned, refer to "averaged" renormalized spectra (see Sect. 2 for details). Are these results just due to "averaging" or they reflect an intrinsic dynamical property of rainfall. If the instantaneous renormalized spectra are all equal (strong equality → existence of a universal distribution) the invariance of "averaged" renormalized spectra reported in (Ignaccolo et al., 2009; Ignaccolo and De Michele, 2010) is a trivial consequence of the existence of a "universal"

HESSD
8, 5605–5629, 2011

distribution. If the instantaneous renormalized spectra manifest a high degree of variability not reducible to a definite criterion (weak/no equality) the invariant properties reported in Ignaccolo et al. (2009); Ignaccolo and De Michele (2010) are merely an accident. A third possibility is the occurrence of a case in between these two extremes: moderate invariance.

To answer these questions, we investigate the skewness γ of instantaneous drop distributions in addition to the two renormalization parameters adopted in Ignaccolo et al. (2009), the mean drop diameter μ and the standard deviation of drop diameters, σ. In the case of strong equality the skewness of renormalized instantaneous spectra is a fixed value, while in the case of weak/no equality one expects a flat distribution of skewness value. The analysis of disdrometer data collected in Darwin, either as whole or as divided in convective and stratiform precipitation intervals, shows that renormalized instantaneous spectra possess a moderate/strong degree of equality. The distribution of skewness values is strongly peaked around the value $\gamma \simeq 0.64$, both for convective and stratiform precipitation. This result indicates the existence of a *most common* distribution. However deviations with both higher and lower skewness do exist also if they are not predominant. In this sense the renormalization parameters μ and σ (average and standard deviation of drop diameters) capture a large part of, the variability of drop size distributions. Mean, standard deviation, and skewness are not totally independent from each other although the relationships can not be expressed in analytical form but are statistical in nature. For a given value of the mean diameter, large (small) skewness values are associates to large (small) values of the standard deviation.

HESSD
8, 5605–5629, 2011

2 Methodology and data processing

2.1 Renormalization

Given a renormalization time interval I, of length typically in the range 1–5 min, the renormalization procedures usually adopted in Literature (Willis, 1984; Sempere Torres et al., 1994; Maki et al., 2001; Testud et al., 2001; Uijlenhoet et al., 2003; Campos et al., 2006; Hazenber et al., 2011) operate a rescaling of both the diameter D and the instantaneous drop size distribution $\mathcal{N}_\mathrm{I}(D)$, the DSD relative to the drops inside the renormalization time interval:

$$\begin{cases} (D,\mathcal{N}_\mathrm{I}(D)) \rightarrow (D^* = D/X_\mathrm{I},\mathcal{N}_\mathrm{I}^*(D*) = \mathcal{N}_\mathrm{I}(D)/Y_\mathrm{I}) \\ \mathcal{N}_\mathrm{I}(D) = \dfrac{N_\mathrm{I}}{AT}\dfrac{p_{\mathrm{G,I}}(D)}{v(D)} \end{cases} \tag{2}$$

The variables X_I and Y_I are the instantaneous bulk variables (e.g. volume mean diameter and liquid water content, respectively) used to obtain the renormalized diameter D^* and renormalized instantaneous spectra $\mathcal{N}_\mathrm{I}^*(D^*)$. The existence of an invariant spectrum would result in a "single" shape for the graph $(D^*,\mathcal{N}_\mathrm{I}^*(D))$. The bottom equality of Eq. (2) is the relationship between the instantaneous DSD and the instantaneous probability density of diameter observed at the ground $p_{\mathrm{G,I}}(D)$. The parameter A is the base area, in m^2, of the unit volume where $\mathcal{N}_\mathrm{I}(D)$ is the concentration (for disdrometer data A is the catchment area of the instrument), T is the length of the renormalization time interval in seconds, while N_I is the time interval drop count. Finally, $v(D)$ is the falling speed, in m s^{-1}, of a drop of diameter D.

The renormalization procedure proposed by Ignaccolo et al. (2009) operates as follows. For each renormalization time interval I:

$$\begin{cases} D \rightarrow D_\mathrm{R} = \dfrac{D-\mu_\mathrm{I}}{\sigma_\mathrm{I}} \\ p_{\mathrm{G,I}}(D) \rightarrow p_{\mathrm{G,I}}(D_\mathrm{R}) = \sigma_\mathrm{I} p_{\mathrm{G,I}}(\sigma_\mathrm{I} D_\mathrm{R} + \mu_\mathrm{I}) \end{cases} \tag{3}$$

HESSD

8, 5605–5629, 2011

Skewness of instantaneous renormalized spectra

M. Ignaccolo and C. De Michele

where μ_I and σ_I are the mean, and standard deviation of the drop diameter observed at the ground. This renormalization procedure operates a change of variable from the diameter D to the renormalized diameter D_R having zero mean and unit variance. The instantaneous renormalized spectrum is in this case the instantaneous probability density $p_{\mathrm{G,I}}(D_\mathrm{R})$ of the renormalized drop diameter. This density is obtained from that of drop diameter observed at the ground $p_{\mathrm{G,I}}(D)$ using the second identity of Eq. (3). The rationale behind this choice of Eq. (3) is that the time series $\{D_k\}$, $k = 1,2,3,\ldots$, of drop diameters is derived from a stochastic process which would be stationary if it were not for a variable mean and a variable standard deviation (e.g., the works of Kostinski and Jameson, 1997 and Smith, 1993 support this ansatz). If so the renormalization procedure removes the non-stationarity, so that sequences of renormalized diameters relative to two different datasets should have the same probability density $p_\mathrm{G}(D_\mathrm{R})$: check of self consistency. This is precisely what is observed in Ignaccolo et al. (2009); Ignaccolo and De Michele (2010).

Disdrometers categorize drops in diameter classes leading to a quantization error in the sequence of drop diameters. A drop in the j-th class has to be considered as the occurrence of a drop with a random diameter value D uniformly distributed in the range $[D_j - \Delta_j/2,\ D_j + \Delta_j/2]$, where D_j and Δ_j are the central value and the width of the class. The corresponding renormalized diameter is then an uniformly distributed random number in the interval $[(D_j - \Delta_j/2 - \mu_\mathrm{I})/\sigma_\mathrm{I},\ (D_j + \Delta_j/2 + \mu_\mathrm{I})/\sigma_\mathrm{I}]$. Therefore given a data set of disdrometer counts there is not an unique sequence of renormalized diameters corresponding to it and as a consequence there is not an unique probability density $p_\mathrm{G}(D_\mathrm{R})$. However one can repeat the renormalization procedure M times and define the probability density $p_\mathrm{G}(D_\mathrm{R})$ of a particular dataset as the average density of the M realizations. Ignaccolo and De Michele (2010) show that already for $M = 100$ one obtains a fairly stable average value and the fluctuations around the average are negligible down to frequency of $\sim 100/N$, with N being the total number of drops in the dataset. Alternatively, the probability density of renormalized drop diameter obtained with a single run of the renormalization procedure sensibly deviates from the average

HESSD
8, 5605–5629, 2011

$p_G(D_R)$ obtained from many realizations only for frequency values $\lesssim 100/N$. Hereby we will denote by $p_G(D_R)$ the average density for an infinite number of realizations of the renormalization procedure since an analytical formula can be derived for this quantity. Let us consider a single realization of the renormalization procedure. For each renormalization time interval I

$$N_I p_{G,I}(D_R) dD_R + \epsilon_I(D_R) dD_R \tag{4}$$

is the number of drops in the infinitesimal interval dD_R centered around the value D_R. The first term of Eq. (4) is the expected value from the particular drop counts observed in the interval, N_I is the total drop count in the interval. The second term of Eq. (4) expresses the statistical fluctuations due to the quantization error of the disdrometers: a drop in the j-th diameter class is assigned a random value of the renormalized diameter in the range $[(D_j - \Delta_j/2 - \mu_I)/\sigma_I, (D_j + \Delta_j/2 + \mu_I)/\sigma_I]$. Note that $\int \epsilon_I(D_R) dD_R = 0$ so that the integration in D_R of Eq. (4) simply returns the number N_I. The probability density relative to a single realization of the renormalization procedure is obtained performing the sum over all renormalization time intervals. Thus the difference between two realizations of the renormalization procedure is the second term in Eq. (4). Averaging over an infinite number of realizations eliminates the statistical fluctuations so that

$$p_G(D_R) = \frac{1}{N} \sum_I N_I \, p_{G,I}(D_R) \tag{5}$$

where N is the total number of drops in the data set.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |
HESSD
8, 5605–5629, 2011
Skewness of instantaneous renormalized spectra
M. Ignaccolo and C. De Michele

2.2 Calculation of renormalization parameters

Given a renormalization time interval, the n-th instantaneous moment $M_{\mathrm{I},n}$ of the probability density $p_{\mathrm{G,I}}(D)$ of drop diameters at the ground is

$$M_{\mathrm{I},n} \underset{\mathrm{def}}{=} \int_0^{+\infty} D^n p_{\mathrm{G,I}}(D) dD. \tag{6}$$

Since disdrometers classify drop diameters per classes, quantization error, the probability density is a step function so that

$$M_{\mathrm{I},n} = \sum_j p^j_{\mathrm{G,I}} \int_{D_{j,\mathrm{L}}}^{D_{j,\mathrm{R}}} D^n dD = \sum_j \frac{n_{\mathrm{I},j}}{N_\mathrm{I}\Delta_j} \frac{(D_{j,\mathrm{R}})^{n+1} - (D_{j,\mathrm{L}})^{n+1}}{n+1} \tag{7}$$

where the sum is taken over the diameter classes, index j. The symbol $p^j_{\mathrm{G,I}}$ indicates the value of $p_{\mathrm{G,I}}(D)$ inside the j-th diameter class, while $n_{\mathrm{I},j}$, Δ_j, $D_{j,\mathrm{L}}$, and $D_{j,\mathrm{R}}$ are respectively the drop count, width, left limit, and right limit of the j-th diameter class. Finally N_I is the drop count inside the renormalization time interval considered. The renormalization procedure adopted here requires the calculation of two parameters: the mean μ_I and the standard deviation σ_I of drop diameters. In addition to the renormalization parameters, the skewness γ_I is evaluated and used as parameter to quantify the "equality" of instantaneous renormalized spectra. All these parameters can be obtained from the calculation of the n-th instantaneous moment $M_{\mathrm{I},n}$, Eq. (7), as follows

HESSD
8, 5605–5629, 2011

M. Ignaccolo and C. De Michele

$$
\begin{cases}
\mu_{\mathrm{I}} = M_{\mathrm{I},1} \\
\sigma_{\mathrm{I}} = \sqrt{M_{\mathrm{I},2} - (M_{\mathrm{I},1})^2} \\
\gamma_{\mathrm{I}} = \dfrac{\left[M_{\mathrm{I},3} + 2(M_{\mathrm{I},1})^3 - 3M_{\mathrm{I},1}M_{\mathrm{I},2}\right]}{\left[M_{\mathrm{I},2} - (M_{\mathrm{I},1})^2\right]^{3/2}}
\end{cases}
\tag{8}
$$

2.3 Skewness as measure of invariance

To "measure" the invariance of two different instantaneous renormalized spectra $p_{\mathrm{G,I}}(D_{\mathrm{R}})$, one could consider a statistical test of equality between sample distributions, e.g. *Kolmogorov-Smirnov* (KS). This approach is not feasible in our case. (1) If M is the total number of renormalization intervals, one has to test $M(M-1)/2$ couples of renormalization time intervals, in our case $M = 6863 \Rightarrow 23\,546\,953$ couples. (2) Statistical tests of equality have no transitive property. If the couples of renormalization time intervals (I_1, I_2) and (I_2, I_3) pass the test, nothing can be implied about the couple (I_1, I_3). Thus one cannot simply use the number of couples which have the same spectra as an indication of the existence of an invariant distribution. (3) Inadequacy of tests of equality between sample distributions. Let us suppose that there is an universal distribution so that for each renormalization time interval $p_{\mathrm{G,I}}(D_{\mathrm{R}}) = f(D_{\mathrm{R}})$, then the instantaneous spectra $p_{\mathrm{G,I}}(D)$ are obtained from $f(D_{\mathrm{R}})$ inverting the relations in Eq. (3) with a given value $\mu_{\mathrm{I}}^{\mathrm{th}}$ and $\sigma_{\mathrm{I}}^{\mathrm{th}}$. The superscript "th" indicates theoretical values. We cannot directly observe $p_{\mathrm{G,I}}(D_{\mathrm{R}})$ but only $p_{\mathrm{G,I}}(D)$, and use Eq. (8) to have an estimate of $\mu_{\mathrm{I}}^{\mathrm{th}}$ and $\sigma_{\mathrm{I}}^{\mathrm{th}}$. Thus the possible statistical differences between the renormalized spectra of two renormalization time intervals are due to (A) sampling fluctuations, drop counts, $< \infty$; (B) errors in estimating the theoretical values $\mu_{\mathrm{I}}^{\mathrm{th}}$ and $\sigma_{\mathrm{I}}^{\mathrm{th}}$, which introduce consequently errors in $p_{\mathrm{G,I}}(D_{\mathrm{R}})$. However, tests of equality between sample distributions

HESSD
8, 5605–5629, 2011

M. Ignaccolo and C. De Michele

take in account only (A) and not (B) as source of statistical differences. Therefore these tests are susceptible to failure even in the case when an universal distribution exists.

Due to these limitations, we follow an alternative approach and consider the skewness γ_I as the parameter to characterize the equality of instantaneous renormalized drop spectra. The rationale for this choice are (1) the densities $p_{G,I}(D_R)$ have all zero mean and unit variance in virtue of the renormalization procedure. Thus, the skewness is the next standardized moment which can be used to describe the distribution. (2) The renormalization procedure, Eq. (3), preserves the skewness: the instantaneous density $p_{G,I}(D)$ and $p_{G,I}(D_R)$ have the same skewness. In fact

$$\int_0^{+\infty}\left(\frac{D-\mu_I}{\sigma_I}\right)^k p_{G,I}(D)dD = \int_{-\frac{\mu_I}{\sigma_I}}^{+\infty} D_R^k \sigma_I p_{G,I}(\sigma_I D_R + \mu_I)dD_R = \int_{-\frac{\mu_I}{\sigma_I}}^{+\infty} D_R^k p_{G,I}(D_R)dD_R \tag{9}$$

where k is any real number. Thus in general the renormalization procedure of Eq. (3) preserves not only the skewness ($k=3$), but all the standardized moments (k integer).

2.4 Data processing

We use Joss Waldvogel disdrometer data at 1 min time resolution recorded in Darwin, Australia (12.45° S, 130.83° E, 2 m a.m.s.l.) for 97 consecutive days, from 4 November 2005 to 10 Februar 2006. This is the same database as that of Ignaccolo and De Michele (2010). Drop diameters are classified in 20 different classes covering the range 0.3–5.6 mm. Moreover counts are corrected against the instrument dead time, Sauvageot and Lacaux (1995). The total number of minutes in this dataset is 139 680 of which only 26 595 ($\simeq$ 19 %) display at least one drop count. The total drop count is 2 943 435. Reflectivity maps are available for the time intervals 9 November to 6 December 2005, and 6 Januar to 10 Februar 2006, allowing for stratiform versus convective classification through the identification of the bright band. A total of 19 stratiform and 33 convective time intervals were identified with this method (details are provided

HESSD
8, 5605–5629, 2011

in the online material of Ignaccolo and De Michele, 2010). Parsing together all these intervals we obtain the stratiform subsets (4669 min of which 4264 with at least a drop count for a total of 407 277 drops) and the convective subsets (2931 min of which 2267 with at least a drop count for a total of 1 077 488 drop counts).

We select the length of the renormalization time intervals I to be 1 min, the same as the time resolution of our dataset. One "artifact" afflicting the renormalization procedure is the statistical errors in calculating the mean and standard deviation of the drop diameters for each renormalization time interval. To mitigate this issue, we exclude minutes with counts ≤ 60, or with a number of occupied diameter classes < 3 from the renormalization procedure. The rationale behind these threshold values is: 1) They identify a dynamical property of the rainfall phenomenon as the quiescent time intervals, intervals of sparse precipitation, with negligible contribution to the overall precipitated volume (we refer the reader to Ignaccolo et al. (2009) for a detailed discussion). 2) They allow a "reasonable" (law of large numbers) estimate of the mean and standard deviation of drop diameters. Once these minutes are removed, we are left with 6863 non-quiescent minutes and 2 758 320 drops for Darwin database (1844 non-quiescent minutes and 355 545 drops for the stratiform subset, and 1536 non-quiescent minutes and 1 066 299 drops for the convective subset).

Another possible source of artifacts is the outlier drop counts. With a choice of 1 min for the length of the renormalization time interval and 20 diameter classes the drop counts are mostly distributed in such a way to cover continuously all the diameter classes in between an initial and a final class: e.g. classes 4 to 16 have non zero counts and classes 1 to 3 and 17 to 20 are all empty. Occasionally there are class gaps, that is two non adjacent classes with non zero count are separated by one or more classes with null counts. These gaps are due to sampling fluctuations occurring for the classes with small probability of occurrence (if N_I is the number of drops inside the renormalization time interval, as rule of thumb probabilities of the order of $1/N_I - 10/N_I$, $N_I \gg 10$, are going to be affected by sampling fluctuations). Thus one can observe, e.g., a zero count for the classes 1 to 3, non zero counts from class 4 to 7, a count

HESSD
8, 5605–5629, 2011

of 1 on class 8 followed by zeros from class 9 to 20. To verify that the class gaps are due to sampling fluctuations and not to some real dynamical property of rainfall we proceed as follows. For each 1 min time interval with class gaps we seek the class with the maximum count, and then we move both to the left and right of this class until we either reach a zero count or the minimum (1) or maximum (20) diameter class of the instrument. With this procedure we identify the no gap region containing the max count. We then calculate the fraction of the total number of drops inside the renormalization time interval belonging to the no gap region containing the max count. If the class gap is due to sampling fluctuations we expect this fraction to be large (e.g. 90 % or larger). For the Darwin database 1091 of the 6863 ($\simeq$ 16 %) renormalization time intervals with class gaps only for 86 intervals the no gap region contains less than 90 % of the total drop count.

Even if the drop counts following a gap constitute a small portion of total drop count inside a renormalization time interval, their effect amount to 1) larger values for the parameters μ_I, σ_I, and γ_I 2) fatter left and right tail for the instantaneous probability densities function at ground $p_{\mathrm{G,I}}(D)$ and $p_{\mathrm{G,I}}(D_\mathrm{R})$ and for the probability $p_\mathrm{G}(D_\mathrm{R})$ relative to an entire data set. Thereby, we set to zero the outlier drop counts before applying the renormalization procedure which reduces the total number of drops of the 6863 non-quiescent renormalization time intervals considered from 2 758 320 to 2 753 796, ~ -0.16 % (from 355 545 to 354 743, ~ -0.22 %, for the stratiform subset, and from 1 066 299 to 1 064 561, ~ -0.16 %, for the convective subset).

3 Results

Figure 1 shows the probability $\Pr(\gamma_\mathrm{I})$ of having a renormalized drop spectrum $p_{\mathrm{G,I}}(D_\mathrm{R})$ with skewness γ_I for the entire Darwin database and for the stratiform and convective subsets. We see how the distribution for the entire dataset is peaked around the skewness value 0.64. With respect the entire database the distribution of the stratiform (convective) subset is more peaked (flat) around a slightly larger (smaller) skewness

HESSD
8, 5605–5629, 2011

value. This figure indicates the existence of a substantial degree of invariance which we quantify using the concept of skewness class. A renormalization time interval I belongs to the skewness class r if the relative difference of the corresponding skewness γ_I with respect the most probable value 0.64 is within the percentage range $[(r-1/2)\times 100, (r+1/2)\times 100]$: e.g. skewness class zero (s0) implies γ_I in the range [0.32, 0.96], skewness class plus-one (s + 1) implies γ_I in the range [0.96, 1.60], while skewness class minus-two (s − 2) implies γ_I in the range [−0.96, −0.32]. Skewness classes are indicated in Fig. 1 with shadowed regions separated by vertical lines.

Table 1 reports for each database (all, stratiform, and convective) the percentages I % of the database number of renormalization time intervals belonging to a given skewness class, and the percentage d % of the database total number of drops belonging to renormalization time intervals in a given skewness class. We see how the percentage I % for the skewness class s0 is ∼ 64 % for the stratiform database but ∼ 51 % for the convective database. The other two skewness classes which are appreciably populated are s + 1 and s − 1, with s + 1 more predominant in the stratiform case and s − 1 more predominant in the convective case. However if we consider the percentage of d % of total databases drops, we see a more balanced repartition among the skewness classes: d % is in the range 61–62 % for skewness class s0, 14–18 % for s − 1, and 16–19 % for s + 1. The origin of this balance is depicted in Fig. 2 which illustrates the occupancy $O(N_I, \gamma_I)$, the number of couples (N_I, γ_I) inside a given box in the $N_I\gamma_I$–plane. For the convective database higher value of the occupancy $O(N_I, \gamma_I)$ are observed for skewness class s + 1, however the drop count N_I is not as high as for the skewness classes s0 and s − 1. Moreover, we notice how large ($N_I > 1000$) drop counts occur almost exclusively for skewness classes s0, s + 1, and s − 1: renormalized drop spectra with "extreme" skewness values are rare and with relatively small drop counts.

Next, we divide the entire, stratiform and convective Darwin databases in subsets according to the skewness class of each renormalization time interval. We calculate the probability density $p_G(D_R)$ of the renormalized drop diameter D_R for each subsets using Eq. (5). Figure 3 shows the results for each skewness class in comparison with the density of the entire Darwin database. The classes s + 4 and s – 3 are not shown because of poor statistics. The class s + 4 (s – 3) has 8 (13), 1 (0), and 3 (5) renormalization time intervals in the entire, stratiform, and convective databases, respectively. For each skewness class, the probability density $p_G(D_R)$ does not depend from the particular dataset considered (entire, stratiform, or convective). The continuous line in all panels of Fig. 3 is the probability density $p_G(D_R)$ of the whole Darwin dataset (including all skewness classes). Comparing this density with the densities per skewness class, we see how the class s0 essentially defines the central part of the whole Darwin data set density, while the skewness classes of degree $r \gtrless 0$ largely influence of the tails of the distribution. The percentage d % of the database total number of drops belonging to renormalization time intervals in a given skewness class determines the shape of the probability density $p_G(D_R)$ of the entire dataset considered. From table I we see that the variability of the percentage d % with respect the databases considered is relatively small so that we expect an a substantial invariance in distribution which is precisely what is observed in Ignaccolo and De Michele (2010).

The results depicted in Fig. 1 indicate that the probability density $p_G(D_R)$ relative to the skewness class s0 could be considered as a "standard" distribution $S(D_R)$: the most probable distribution. Two main deviations from the most common distribution are observed, one with smaller ($p_G(D_R)$ relative to the skewness class s – 1) and one with larger skewness ($p_G(D_R)$ relative to the skewness class s + 1).

3.1 Sample variability of the skewness

Even if there was an unique renormalized spectra one can expect the observed values for the skewness inside a renormalization time interval to show some variability due to sampling fluctuations. It is thus possible that part of the deviation from the most probable value $\gamma_I = 0.64$ shown in Fig. 1 might be due to sampling errors. To test this possibility, we use the standard distribution $S(D_R)$ to create an artificial sequence of drop counts which are supposed to have all the same skewness. For each renormalization time intervals we extract N_I renormalized diameters and use the mean μ_I

and the standard deviation σ_I of the renormalization time interval to obtain the corresponding simulated (sim) instantaneous renormalized spectrum $p^{\mathrm{sim}}_{\mathrm{G,I}}(D_{\mathrm{R}})$. Starting from $p^{\mathrm{sim}}_{\mathrm{G,I}}(D_{\mathrm{R}})$, and inverting the relationships of Eq. (3) drop counts are obtained having care of rejecting, if any occur, drops with a simulated diameter $D < 0.3$ mm since this is the minimum drop diameter detectable by JW disdrometer.

We apply this procedure to the entire Darwin data set to obtain an artificial sequence. We then calculate the skewness values of the artificial sequence and compare its distribution with that observed in reality. Figure 4 shows the results. We see how the distribution of skewness values of the artificial sequence is much more peaked than the real one. For the artificial sequence, almost 92 % of the renormalization time intervals are inside the skewness class s0, with ~ 5 % and ~ 3 % on the $s-1$ and $s+1$ skewness classes. Thus the observed spreading of skewness values depicted in Fig. 1 can be ascribed to sample fluctuation only partially (compare 5 % and 3 % with the values of %I in Table 1).

3.2 Relationship among mean, standard deviation, and skewness

Our statistical description associates to each renormalization time intervals four parameters: the drop count N_I, the mean μ_I, the standard deviation σ_I and the skewness γ_I of drop diameters. Only the mean and the standard deviation are used in the renormalization procedure of Eq. (3). More in general these four parameters are related to each other albeit in a statistical way. E.g. Ignaccolo and De Michele (2010) show that there is an approximate linear relation between the averages values of the parameters μ_I and σ_I associated with a given rainfall rate class: $\langle \mu_I \rangle \propto \langle \sigma_I \rangle$ where the symbol $\langle \ldots \rangle$ indicates the rainfall rate class average.

Hereby, we show that given a fixed range of values for the mean diameter μ_I, the more negative (positive) is the skewness class to which the renormalization time interval belongs the smaller (larger) is the standard deviation σ_I of the drop diameters. We divide the entire Darwin dataset in subsets according to the skewness class. Then we fix a range of values for the mean diameter μ_I and for each subset we evaluate the median of the observed values of σ_I when μ_I is in the given range. The results are reported in Fig. 5. The vertical lines depicts the range of μ_I values considered. The median of the σ_I values is calculated only when at least 10 time intervals are in the range considered. We see how the median value of σ_I increases as the skewness class goes from −2 to +2.

4 Conclusions

"Averaged" renormalized spectra relative to stratiform and convective precipitation at Darwin (AU) possess a common shape (Ignaccolo and De Michele, 2010). Is this due to a more general invariance: the invariance of single instantaneous renormalized spectra? In this manuscript, we investigated this possibility providing a positive answer. Since instantaneous renormalized spectra have all zero mean and unitary variance, the skewness was taken as a parameter to test invariance. The probability $\Pr(\gamma_\mathrm{I})$ of skewness values is peaked around the values 0.64. The dispersion of skewness values around the most probable one is quantified in terms of skewness classes. The skewness class s0 centered around the most probable value contains $\sim 57\,\%$ of the renormalization time intervals corresponding to $\sim 60\,\%$ of the total drops in the database. The remaining renormalization time intervals essentially belong to the skewness classes s + 1 and s − 1 with only $\sim 2\,\%$ of intervals in the skewness classes associated with the tails of the probability $\Pr(\gamma_\mathrm{I})$. The results of this classification allow us (using Eq. 5) to define the most common renormalized spectra $S(D_\mathrm{R})$ as the density of renormalized drop diameters D_R relative to the subsets of renormalization time intervals inside the skewness class s0. We use the most common distribution to produce an artificial database and prove that part of the observed dispersion of the skewness around the most probable value is due to sampling limitation since instantaneous spectra are derived from 1 min drop counts (Fig. 4). About 8 % of the renormalization time intervals of the artificial database belong to the skewness classes

HESSD
8, 5605–5629, 2011

Skewness of instantaneous renormalized spectra

M. Ignaccolo and C. De Michele

s + 1 and s − 1 compared to ∼ 40 % for the real data. In this sense one can estimate the sampling effects to be responsible for approximatively one fifth (8/40) of the observed dispersion of skewness values.

Another issue, we address is how the skewness of renormalized instantaneous spectra is dependent from the two renormalization parameters: mean and standard deviation of drop diameters. For a given range of mean diameter values, we calculate the median of the observed values of the standard deviation for each skewness class. We find that the median increases as the skewness class goes from −2 to +2. That is given a value of the mean diameter, larger (smaller) values of the skewness are associated with large (small) values of the standard deviation.

The probability densities $p_G(D_R)$ for the databases of a given skewness class are independent from the stratiform versus convective classification: stratiform and convective instantaneous spectra with "same" skewness value have the "same" renormalized spectra. More in general the adoption of the skewness as metric to measure the equality of renormalized spectra can be useful in comparing instantaneous renormalized spectra at different locations with different meteorological regimes. E.g. in the companion paper we will discuss the properties of instantaneous renormalized spectra in the case of orographic precipitation.

Acknowledgements. We are thankful to C. R. Williams and the National Oceanic and Atmospheric Administration for the public availability of the data set. M. I. thankfully acknowledges the support of Accademia Nazionale dei Lincei (Italy) through "B. Segre" scholarship. C. D. M. thankfully acknowledges Comune di Milano (Italy) through BIODESCESA project.

References

Campos, E. F., Zawadzki, I., Petitdidier, M., and Fernandez, W.: Measurement of raindrop size distributions in tropical rain at Costa Rica, J. Hydrol., 328, 98–109, 2006. 5607, 5609

Feingold, G. and Levin, Z.: Application of the lognormal rain-drop distribution to differential reflectivity radar measurements (ZDR), J. Atmos. Ocean. Tech., 4, 377–382, 1987. 5606

Hazenberg, P., Nan, Y., Boudevillain, B., Delrieu, G., and Uijlenhoet, R.: Scaling of raindrop size distributions and classification of radar reflectivity – rain rate relations in intense Mediterranean precipitation, J. Hydrol., 402, 179–192, 2011. 5607, 5609

Ignaccolo, M. and De Michele, C.: Statistical collapse of stratiform and convective drop diameter distributions at the ground, Geophys. Res. Lett., 37, L24402, doi:10.1029/2010GL045454, 2010. 5607, 5608, 5610, 5614, 5615, 5618, 5619, 5620

Ignaccolo, M., De Michele, C., and Bianco, S.: The drop-like nature of rain and its invariant statistical properties, J. Hydrometeorol., 10(1), 79–95, 2009. 5607, 5608, 5609, 5610, 5615

Joss, J. and Gori, E. G.: Shapes of raindrop size distributions, J. Appl. Meteorol., 17, 1054–1061, 1978. 5606

Konstinski, A. B. and Jameson, A. R.: Fluctuation properties of precipitation, Part I: On deviations of single-size drop counts from the poisson distribution, J. Atmos. Sci., 54, 2174–2186, 1997. 5610

Maki, M., Keenan, T. D., Sasaki, Y., and Nakamura, K.: Characteristics of the raindrop size distribution in tropical continental squall lines observed in Darwin, Australia, J. Appl. Meteorol., 40, 1393–1412, 2001. 5607, 5609

Marshall, J. S. and Palmer, W.: The distribution of raindrops with size, J. Meteorol., 5, 165–166, 1948. 5606

Sauvageot, H. and Lacaux, J. P.: The shape of averaged drop size distributions, J. Atmos. Sci., 52, 1070–1083, 1995. 5614

Sekhon, R. S. and Srivastava, R. C.: Doppler radar observations of drop-size distributions in a thunderstorm, J. Atmos. Sci., 28, 983–994, 1971. 5607

Sempere Torres, D., Porrà, J. M., and Creutin, J. D.: A general formulation for raindrop size distribution, J. Appl. Meteorol., 33, 1494–1502, 1994. 5607, 5609

Smith, J. A.: Marked point process models of raindrop-size distributions, J. Appl. Meteorol., 32, 284–296, 1993. 5610

Testud, J., Oury, S., Black, R. A., Amayenc, P., and Dou, X.: The concept of normalized distribution to describe raindrop spectra: a tool for cloud physics and cloud remote sensing, J. Appl. Meteorol., 40, 1118–1140, 2001. 5607, 5609

Tokay, A. and Short, D. A.: Evidence from tropical raindrop spectra of the origin of rain from stratiform versus convective clouds, J. Appl. Meteorol., 35(3), 355–371, 1996. 5607

Uijlenhoet, R., Smith, J. A., and Steiner, M.: The microphysical structure of extreme precipitation as inferred from ground-based raindrop spectra, J. Atmos. Sci., 60, 1220–1238, 2003. 5607, 5609

Ulbrich, C. W.: Natural variations in the analytical form of the raindrop size distribution, J. Clim. Appl. Meteorol., 22, 1764–1775, 1983. 5606

Willis, P. T.: Functional fits to some observed dropsize distributions and parameterization of rain, J. Atmos. Sci., 41, 1648–1661, 1984. 5607, 5609

HESSD

8, 5605–5629, 2011

Skewness of instantaneous renormalized spectra

M. Ignaccolo and C. De Michele

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◀| |▶

◀ ▶

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

HESSD

8, 5605–5629, 2011

Skewness of instantaneous renormalized spectra

M. Ignaccolo and C. De Michele

Table 1. Classification of the databases (respectively all, stratiform, and convective) in classes of skewness using the number of time intervals (*I* %), and number of drops (*d* %).

Database		All		Stratiform		Convective	
Class of γ		*I* %	*d* %	*I* %	*d* %	*I* %	*d* %
s − 3	[−0.96, −1.60]	0.19	0.11	0	0	0.32	0.15
s − 2	[−0.96, −0.32]	1.74	1.14	0.21	0.08	2.40	1.99
s − 1	[−0.32, 0.32]	20.99	18.61	14.48	14.12	25.91	18.33
s0	[0.32, 0.96]	57.57	61.44	64.31	62.06	51.23	61.13
s + 1	[0.96, 1.60]	16.37	15.87	18.6	19.11	16.53	17.33
s + 2	[1.60, 2.24]	2.66	2.48	1.95	3.67	2.93	1.99
s + 3	[2.24, 2.88]	0.32	0.25	0.37	0.81	0.45	0.17
s + 4	[2.88, 3.52]	0.11	0.07	0.05	0.12	0.19	0.06

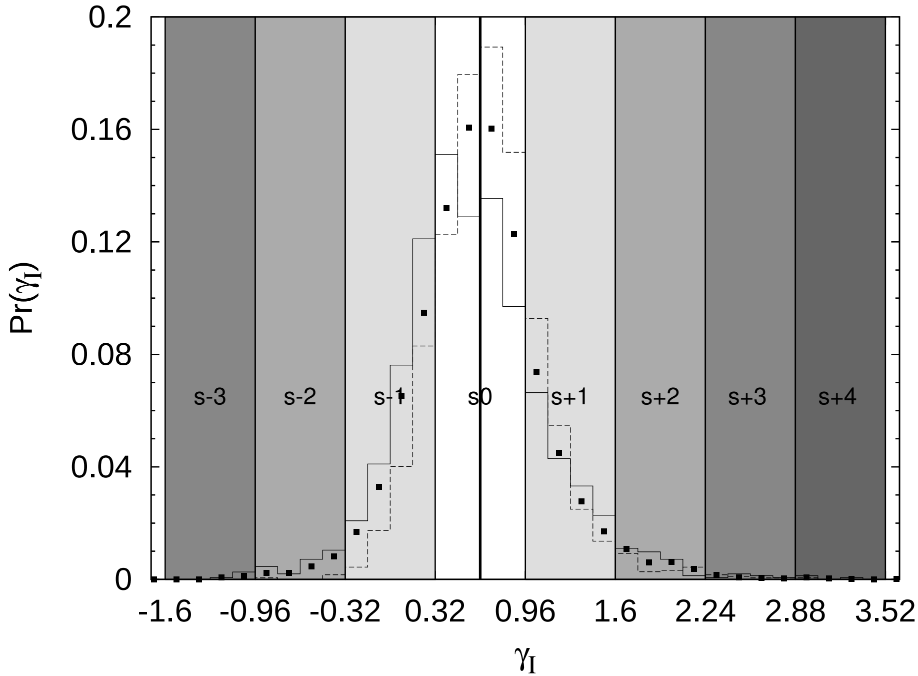

Fig. 1. The probability Pr(γ_I) for the entire (full squares), stratiform (dashed line), and convective (full line) datasets at Darwin. The shaded regions indicate different skewness classes.

HESSD
8, 5605–5629, 2011

Skewness of instantaneous renormalized spectra

M. Ignaccolo and C. De Michele

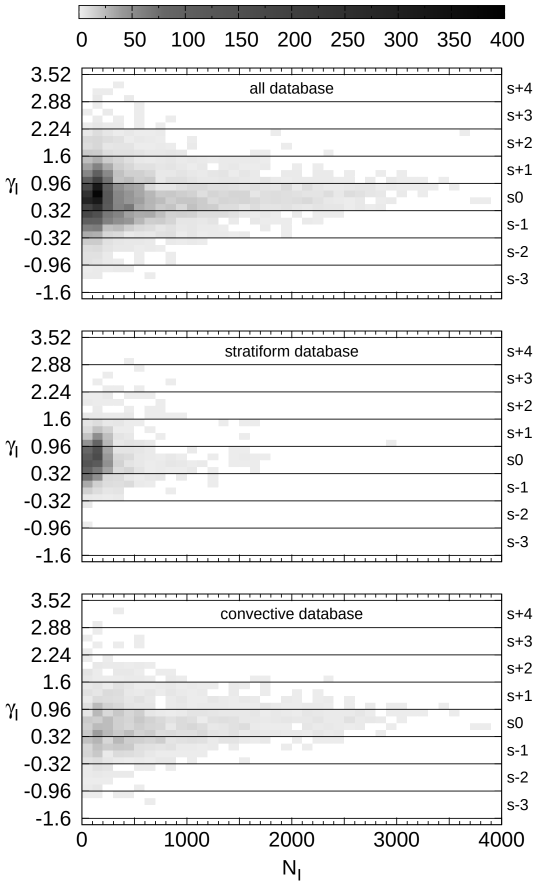

Fig. 2. The occupancy $O(N_I, \gamma_I)$ for the entire, stratiform, and convective datasets at Darwin. Horizontal lines denote the skewness classes.

HESSD
8, 5605–5629, 2011

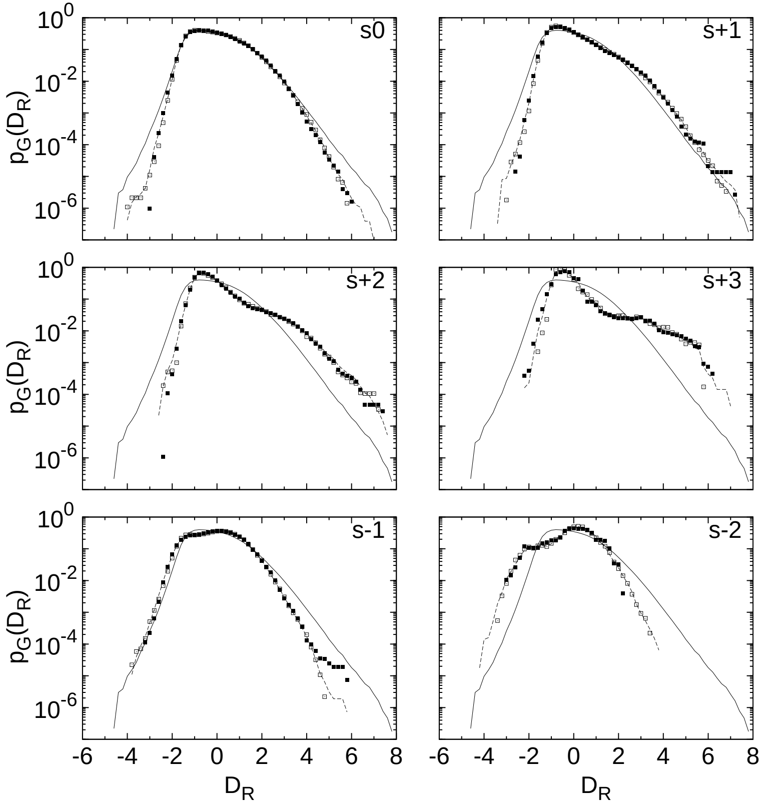

Fig. 3. The value of the probability density $P_G(D_R)$ for the skewness classes subsets obtained from the entire database (dashed line), the stratiform database (full squares), and the convective database (open squares). The label on the top right corner of each panel indicates the skewness class: s0, s + 1, s + 2, s + 3, s − 1, s − 2. The full line denotes the probability density $P_G(D_R)$ for the entire Darwin database.

HESSD
8, 5605–5629, 2011

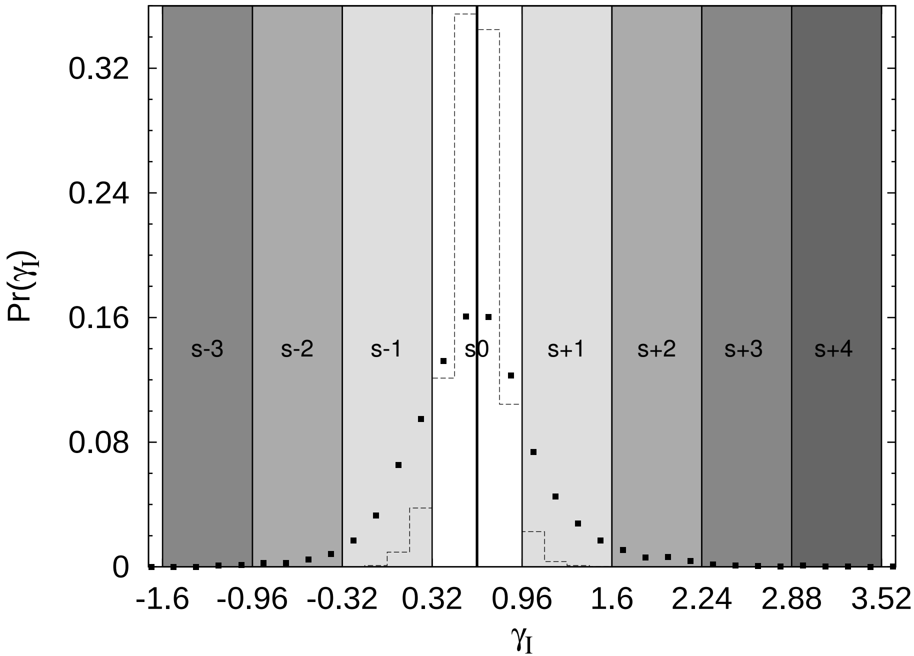

Fig. 4. The probability $\Pr(\gamma_\mathrm{I})$ for the entire dataset at Darwin (full squares) and for an artificial realization (dashed line) created using the standard probability density $S(D_\mathrm{R})$. The shaded regions indicate different skewness classes.

HESSD
8, 5605–5629, 2011

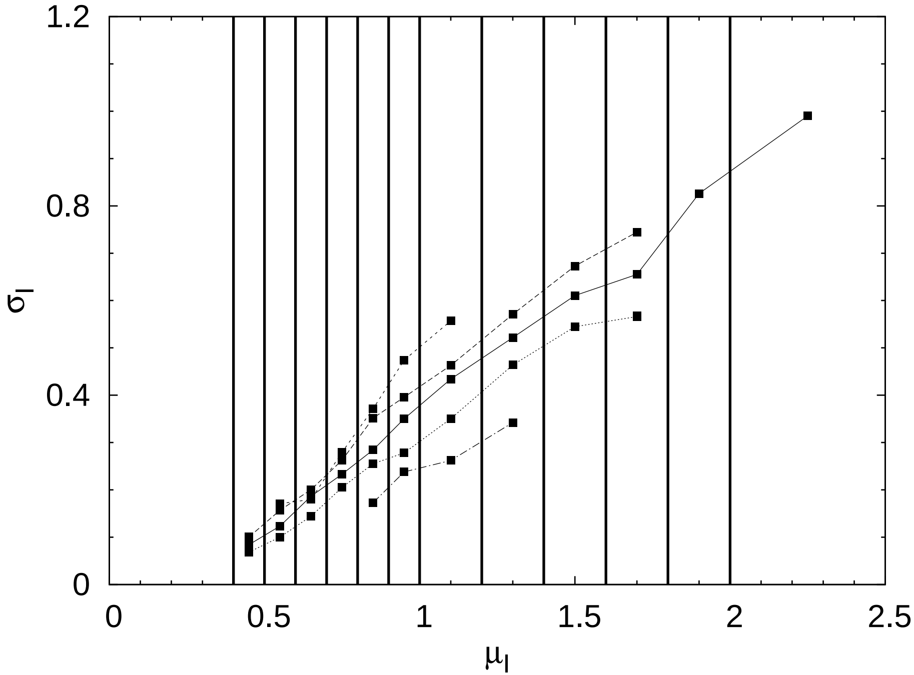

Fig. 5. The median value of the parameter σ_l for a given range of μ_l. The results relative to five different skewness classes are shown: s0 (solid line), s + 1 (long-dashed line), s + 2 (short-dashed line), s−1 (dotted line), and s − 2 (dotted-dashed line). Vertical lines indicate the different range of values of μ_l used to calculate the median.

HESSD
8, 5605–5629, 2011

Skewness of instantaneous renormalized spectra

M. Ignaccolo and C. De Michele