Variability of the raindrop size distribution across scales in Mediterranean rainfall: characterisation and stochastic simulation

THÈSE N^{O} 7152 (2016)

PRÉSENTÉE LE 25 NOVEMBRE 2016

À LA FACULTÉ DE L'ENVIRONNEMENT NATUREL, ARCHITECTURAL ET CONSTRUIT
LABORATOIRE DE TÉLÉDÉTECTION ENVIRONNEMENTALE
PROGRAMME DOCTORAL EN GÉNIE CIVIL ET ENVIRONNEMENT

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Timothy Hugh RAUPACH

acceptée sur proposition du jury:

Prof. S. Takahama, président du jury
Prof. A. Berne, directeur de thèse
Prof. R. Uijlenhoet, rapporteur
Prof. F. Tapiador, rapporteur
Prof. M. Lehning, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE
Suisse
2016

This thesis is dedicated to my father
Michael Robin Raupach (1950–2015),
who has stars at elbow and foot.

Acknowledgements

Moving from Australia to Switzerland for a four-year PhD was a leap into the unknown. I consider myself very lucky, then, that this choice worked out so well. I will always remember my years of being a doctoral student in this wonderful and very mountainous country. But not only luck was on my side: there are many people I must thank for their advice, help, support, and friendship during this time. First, I thank my supervisor Alexis Berne, for his far-sighted vision, for the faith he put in me, and for his thorough and kind supervision. His positive attitude, scientific integrity, attention to detail, and push for highest-quality work are attributes that I hope to take with me into the future.

I thank my colleagues and friends in the LTE lab for the positive work atmosphere, lunches at Vinci, and beers at Satellite. I particularly want to thank Jacopo Grazioli for his patient and invaluable help, especially during field campaigns. My thanks go to all those from LTE who helped with data collection, whether it was fixing instruments, carrying boxes through the snow, spending hours dropping ball-bearings into the 2DVD, or working on the radar in sleet until our hands went numb. I thank the colleagues from other labs who collaborated on field campaigns: the LTHE group from Grenoble and the LaMP group from Clermont-Ferrand for their help with the HyMeX networks in France, and those from MeteoSwiss and SLF Davos who helped in Switzerland.

I'm grateful to my friends in Switzerland for all the fun times, both in town and in the mountains, and for their patience with my French. I thank my friend and colleague Luigi Renzullo, whose encouragement put me on the path to research. I thank my friends overseas for their friendship from afar, and their visits here. I want especially to thank Millie Rooney for her unwavering friendship and support. Big thanks go to the Ewing family of New Zealand! I send thanks and love to my astoundingly amazing family. My sister Anna and my brother Alex are wonderful and always have my infinite thanks. I thank my mum for her strength, love, wisdom, and for countless silly conversations. I'm so grateful to my incredible dad, to whom this thesis is dedicated, for his love, sage advice, optimism, and thoughtfulness. He is an inspiration always, and I remember him dearly. Finally, my heartfelt thanks go to my fantastic wife Tanya, to whom I am very grateful.

Lausanne, 2 September 2016 T. H. R.

This work was financed by the Swiss National Science Foundation, under grant 2000021_140669. All data acknowledgements are provided in Appendix A.

Résumé

La mesure de pluie est compliquée par la grande variabilité du phénomène, y compris à l'échelle des gouttes de pluie. Les mesures ponctuelles sont généralement précises, mais elles ont une représentativité spatiale très restreinte. À l'opposé, les radars météorologiques peuvent prendre des mesures indirectes sur des grands domaines, mais les propriétés microphysiques de la pluie doivent être connues ou déduites pour calculer les quantités de pluie. La distribution des gouttes de pluie (DSD) décrit statistiquement cette microstructure. Elle est souvent supposée être uniforme dans l'espace, mais elle a en fait une grande variabilité. Les nouvelles méthodes et les résultats associés qui composent cette thèse contribuent à la compréhension de la variabilité de la DSD sur les petites échelles, et son effet sur la mesure de la pluie.

Les méthodes présentées ont été développées à l'aide de données d'un réseau d'instruments sur un site de 13×7 km^2 en Ardèche, en France. Cette région est sujette aux fortes pluies méditerranéennes. Une technique pour améliorer la précision des mesures de la DSD par les disdromètres Parsivel est proposée. La méthode utilise un 2D-video-disdrometer comme instrument de référence. Une méthode géostatistique pour l'interpolation spatiale et la simulation stochastique de la DSD est donnée. Elle permet l'estimation ou la simulation de la DSD non-paramétrique en un point où il n'y a pas de mesure, en prenant compte des mesures dans un voisinage. Des essais ont montré que les estimations ont un biais minimal. Ces deux techniques ont été utilisées pour examiner la variabilité de la DSD sur le domaine étudié.

La variabilité de la DSD a été étudiée sur deux échelles typiques, qui correspondent à la taille d'un pixel du radar météorologique spatial Global Precipitation Mission (GPM) et à la taille d'un pixel d'un modèle météorologique opérationnel. Il est montré qu'une erreur se produit si une mesure ponctuelle est supposée représenter une surface. Cette erreur grandit avec l'aire et avec la taille des gouttes. Deux algorithmes pour calculer les quantités de pluie, qui correspondent aux deux échelles considerées, ont été testés. L'intensité de la pluie et la réflectivité radar étaient bien retrouvées, mais d'autres propriétés de la DSD étaient souvent peu représentatives du processus à une échelle plus petite que celle du pixel.

La normalisation de deux moments est un moyen de représenter la DSD dans une forme plus compacte. Pour l'utiliser, il faut supposer que la DSD normalisée est invariante. Cette supposition a été testée avec des données mesurées en France, en Suisse et aux États Unis. Il est montré que, pour l'utilisation pratique, la DSD normalisée peut être supposée invariante, pour les déplacements horizontaux et verticaux. En s'appuyant sur cette supposition, une

technique pour déduire la DSD d'après les variables polarimetriques est proposée. Elle a une performance similaire à, voir meilleur que, celle d'une méthode existante de référence. Une application de l'analyse multifractale à des mesures de neige des Alpes suisses à haute résolution, est présentée. Des comportements scalants ont été détectés dans des colonnes verticales reconstruites, pour les échelles entre 35 m et 2 m. Des faibles comportements scalants ont été trouvés dans des series de temps. Les resultats indiquent que pour les échelles plus petites que deux mètres ou une minute, la neige semble être distribuée d'une manière homogène.

Mots clefs : distribution des gouttes de pluie (DSD), microstructure de la pluie, variabilité des précipitations, simulation stochastique

Abstract

Measurement of rain is made difficult by the high variability of the precipitation process, down to raindrop scale. Point measurements are generally accurate, but their lack of spatial representativeness is a significant limitation. Weather radars indirectly measure rainfall over large regions, but the microphysical properties of the rain being measured must be known or inferred in order to compute rainfall quantities from radar data. The raindrop size distribution (DSD) statistically describes the microstructure of rain. While the DSD is often assumed to be uniform in space, it is in fact highly variable. The work in this thesis contributes to the understanding of the small-scale variability of the DSD and its effects on the measurement of rainfall.

The methods shown were developed using data from a network of disdrometers and radars over a 13×7 km^2 field site in Ardèche, France. This area experiences heavy Mediterranean rainfall. A technique to improve the accuracy of DSD measurements made by Parsivel disdrometers is proposed. The method uses a 2D-video-disdrometer as a reference instrument. A new geostatistical method for spatial interpolation and stochastic simulation of the experimental DSD is provided. It can estimate or simulate the non-parametric DSD at an unmeasured location, conditional on nearby measurements. Leave-one-out testing showed that estimates were produced with minimal bias. The correction and simulation techniques were used together to investigate the small-scale variability of the DSD in the study region.

DSD variability was studied in detail over two typical scales, corresponding to the footprint size of the Global Precipitation Mission (GPM) space-borne weather radar, and a typical size for an operational numerical weather model pixel. It is shown that the assumption that a point measurement of the DSD represents an areal estimation introduces error that increases with areal size and drop size. Satellite and weather model rainfall retrieval algorithms that correspond to these two typical domains were tested, and while it was found that rain intensity and radar reflectivity were well retrieved, other DSD properties were often not representative of the sub-grid process.

Double-moment normalisation provides a compact representation of the DSD, under the assumption that the normalised version DSD is invariant. This assumption was tested using instrument network data in France, Switzerland, and the United States. It is shown in this work that for practical purposes, the double-normalised DSD can be assumed invariant through horizontal and vertical displacement. Using this assumption, a new method for retrieval of the DSD from polarimetric radar data is proposed. The new DSD-retrieval technique performs

as well or better than an existing method. An application of multifractal analysis to high-resolution snowfall data from the Swiss Alps is presented. Scaling of snowfall was observed in reconstructed vertical columns, at scales from about 35 metres to two metres, with no scaling observed at smaller scales. Weak scaling was observed in time series. The results indicate that at small (sub-metre or sub-minute) scale, snowfall appears homogeneously distributed.

Key words: precipitation microstructure, precipitation variability, drop size distribution (DSD), stochastic simulation

Contents

List of figures

List of tables

List of symbols

a	range-defining parameter of dry drift model [m]	63
A	set of points in which a process is active	153
A_c	area of ellipse/circle that covers a snowflake [m^2]	158
A_e	area of pixels covered by a snowflake [m^2]	158
B_A	bulk variable defined on areal scale	91
B_P	bulk variable defined on point scale	91
c	generalised gamma model parameter [–]	116
c_0	dry drift model nugget [–]	63
c_d	dry drift model (partial) sill [–]	63
c_D	fractal co-dimension function	154
C	constant for estimation of K_{dp} [–]	134
C_1	mean intermittency [–]	154
C_n	DSD-normalisation constant for moment order n [–]	115
$C_{v,i}$	drops recorded for vth velocity class and ith diameter class [–]	22
C_F	fractal co-dimension [–]	153
$\mathbf{C}$	principal components matrix	55
$\mathbf{C}^*$	matrix of estimated components	58
d	a distance [m]	64
d_ϵ	spatial dimension of a field ϵ [–]	153
d_0	zero distance in Gaussian dry drift model [m]	64
$d(z)$	distance from location z to nearest dry point [m]	53
D	equivolume raindrop diameter [mm]	4
D_F	fractal dimension [–]	153
D_i	class-centre diameter of ith drop diameter class [mm]	22
D_k	class-minimum diameter of kth drop diameter class [mm]	52
D_m	mass-weighted mean drop diameter [mm]	5
$D_{\max}$	maximum considered drop diameter [mm]	6
$D_{\min}$	minimum considered drop diameter [mm]	6
D_0	median volume drop diameter [mm]	5
e	a moment order [–]	154
e_s	maximum moment order [–]	154
$\mathbb{E}$	expectation	54
E_A	absolute point-to-areal difference on drop concentration [mm^{-1} m^{-3}]	91

E_B	relative point-to-areal difference on bulk variable [%]	91
E_R	relative point-to-areal difference on drop concentration [%]	91
E_t	difference (observed − reference) for time t	26
f_{hh}	Forward scattering amplitude in horizontal polarisation [cm]	134
f_{vv}	Forward scattering amplitude in vertical polarisation [cm]	134
f_G	Gaussian dry drift model	65
f_k	dry drift function for the kth diameter class	54
f_S	spherical dry drift model	64
F_N	normalised DSD model shape factor [–]	83
$\mathscr{F}$	the Fourier transform	156
g	acceleration due to gravity [m s^{-2}]	158
$h(x)$	double-normalised DSD [–]	115
$\hat{h}(x)$	generalised gamma function as a fitted double-normalised DSD [–]	116
H	degree of non-conservation [–]	154
i	the complex unit	156
I	a mean-normalised field	156
I_b	N_w-normalised reflectivity [dB]	85
I_e	normalised specific attenuation [dB km^{-1} mm m^3]	85
$I(z)$	binary rainfall occurrence for location z [–]	53
k	specific attenuation [dB km^{-1}]	84
K	moment scaling function	154
K_{dp}	specific differential phase shift [°km^{-1}]	134
$\hat{K}_{\mathrm{dp}}$	estimated K_{dp} [°km^{-1}]	134
$\lvert K_\omega \rvert^2$	dielectric factor of water [–]	5
l	a distance lag	56
l_ϵ	observation scale of a field ϵ	153
L	number of spatial locations [–]	55
L_ϵ	outer scale of a field ϵ	153
m	the mass of a snowflake [g]	158
M	number of drops recorded within integration time [–]	22
M_n	DSD moment of order n [mm^n m^{-3}]	115
$\hat{M}_n$	estimated DSD moment of order n [mm^n m^{-3}]	137
$\mathbf{M}$	matrix of log-transformed and detrended drop concentration measurements	55
$\mathbf{M}^*$	estimated detrended drop concentrations matrix	58
$\tilde{\mathbf{M}}$	normalised version of $\mathbf{M}$	55
$\tilde{\mathbf{M}}^*$	estimated normalised detrended concentrations matrix	58
$\overline{\mathbf{M}_k}$	mean of kth column in $\mathbf{M}$ [–]	55
N_0	intercept parameter of gamma DSD model [mm$^{-1-\mu}$ m^{-3}]	83
N_A	areal estimate of drop concentration [mm^{-1} m^{-3}]	91
N_i	volumetric drop concentration for ith diameter class [mm^{-1} m^{-3}]	22
N_l	number of samples pairs of points for distance lag l [–]	56
N_κ	number of boxes to cover an entire field at resolution κ [–]	153

List of acronyms

2DVD	two-dimensional video disdrometer	6
COSMO	consortium for small-scale modelling	9
DSD	(rain)drop size distribution	1
DFIR	double fence intercomparison reference	157
DFR	dual-frequency ratio	85
DPR	dual-frequency precipitation radar	84
DTM	double trace moment	155
EPFL	École Polytechnique Fédéral de Lausanne	18
LTE	Laboratoire de Télédétection Environnementale	18
GPM	global precipitation measurement	10
GPS	global positioning system	19
HPE	heavy precipitation event	10
HyMeX	Hydrological Cycle in the Mediterranean Experiment	11
IQR	interquartile range	29
IFloodS	NASA Iowa Flood Studies	117
KED	kriging with external drift	50
MCS	mesoscale convective system	10
MRR	micro rain radar	7
NWP	numerical weather model	3
NASA	National Aeronautics and Space Administration	117
ORB	orographic rain band	10
Parsivel	particle size and velocity (disdrometer)	6
PCA	principal component analysis	51
PPI	plan position indicator	62
PVU	percentage of variance unexplained	126
QPE	quantitative precipitation estimation	2
radar	radio detection and ranging	2
RB	(median) relative bias	26
RMSE	root mean squared error	26
ROI	region of interest	89
RSE	residual standard error	126
SNR	signal to noise ratio	60
SCOP	self-consistent with optional parameterisation	133

SCOP-ME	SCOP attenuation correction and microphysics estimation	133
SOP	special observation period	18
TM	trace moment	155
TRMM	tropical rainfall measuring mission	2
US	United States	80
UM	universal multifractal	13
UTC	coordinated universal time	25
UTM	universal transverse Mercator	89
WGS84	world geodetic system 1984	20

1 Introduction

Precipitation is a profoundly important process. As a primary part of the Earth's hydrological cycle, it delivers fresh water to the land, dramatically shapes the landscape, and is crucial to sustaining life. Liquid precipitation is rain, and its extremes have significant impacts: droughts and floods threaten life and property (e.g. Pielke Jr and Downton, 2000) and have substantial societal and economic effects (e.g. Ciais et al., 2005). Climate change is likely to change the frequency, intensity, and duration of rainfall events (Easterling et al., 2000; Trenberth et al., 2003; Stocker et al., 2013). It is expected that most land areas will experience an increase in the frequency of heavy rainfall events by the end of this century (Stocker et al., 2013). Accurate measurement is the key to being able to understand, model, and predict precipitation, including rain. Precipitation is not an easy process to measure, however, because it is notoriously variable. This variability exists on large scales (e.g. Koster and Suarez, 1995), which is why we experience wet and dry summers, or flooding in one region while another is unscathed. But precipitation is also highly variable at the small scale (e.g. Fabry, 1996), down to the scale of individual falling water particles. These particles, called hydrometeors, exist in many different phases, shapes, and sizes. The study of precipitation microphysics, or the dynamics of precipitation at the particle scale, leads to improved understanding of related processes across scales. The aim of this thesis is to contribute to the understanding and characterisation of the small-scale variability of rainfall.

Rainfall, of course, is made up of falling drops of water. When we speak of small-scale variability of rainfall, what is meant is variability in the number of raindrops there are, and in their sizes. Because it is unfeasible to count and measure every single raindrop in a storm, statistics are used to summarise the information into a form that is more convenient to deal with. The raindrop size distribution (DSD) statistically describes the microstructure of liquid precipitation. It is defined as the number of falling raindrops of a certain size per unit volume of air. All rainfall variables of interest can be calculated as weighted statistical moments of the DSD (e.g. Ulbrich, 1983; Testud et al., 2001). These include the total drop concentration, the characteristic drop diameter, the liquid water content, and the rain intensity. Information about the DSD is needed to calculate the interactions of electromagnetic waves with inho-

mogeneous collections of hydrometeors in the atmosphere, making DSD properties required knowledge for weather radar applications (Marshall et al., 1947; Bringi and Chandrasekar, 2001). While the DSD describes what is happening at the raindrop scale, it is fundamental to processes that occur on much larger scales (e.g. Uijlenhoet and Sempere Torres, 2006). It is used in investigations into rainfall microphysical processes (e.g. Rosenfeld and Ulbrich, 2003); interactions of raindrops with surface soil (e.g. van Dijk et al., 2002), vegetation canopies (e.g. Calder, 1986), and built environments (e.g. Blocken and Carmeliet, 2004); the cleaning effect rain has in removing aerosols from the atmosphere (e.g. Andronache, 2004); weather prediction models (e.g. Baldauf et al., 2011); and the effects of rainfall on telecommunication links (e.g. Crane, 1971; Schleiss and Berne, 2010).

Since the discovery that the power of electromagnetic radiation reflected off rain relates to the rain's intensity (Marshall et al., 1947), weather radar (radio detection and ranging) has revolutionised the measurement of precipitation. Marshall et al. (1947) foresaw the possibility of quantitative precipitation estimation (QPE) in the 1940s. In the 1970s, it was recognised that polarimetric radar, in which the waves are vertically and horizontally polarised, could offer greater insight into precipitation microstructure (Seliga and Bringi, 1976). Nowadays, radar has become an essential tool for the study of precipitation (Bringi and Chandrasekar, 2001; Berne and Krajewski, 2013), and operational networks provide near-real-time precipitation observations in many countries (e.g. Berne and Krajewski, 2013). Radars do not measure the rain intensity directly, but rather they make indirect and integrated measurements of the electromagnetic properties of hydrometeors in a measurement volume (Berne and Krajewski, 2013). The basic problem in QPE is determining the relationship between the radar measurements and the rain rate in the measured volume.

Conventional radars measure radar reflectivity Z (measured in $\mathrm{mm}^6\ \mathrm{m}^{-3}$ but often expressed in dBZ), which is non-linearly related to the rain rate R [$\mathrm{mm\ h}^{-1}$] via the DSD (e.g. Marshall and Palmer, 1948; Uijlenhoet, 2001). The Z–R relationship is usually modelled as a power law that is scale dependent (Verrier et al., 2013; Sassi et al., 2014) and affected by DSD variability (Chapon et al., 2008; Jaffrain and Berne, 2012a). Polarimetric radars measure the reflectivity and phase change of horizontally and vertically polarised waves, and use this information to retrieve DSD properties (Krajewski and Smith, 2002). In both cases, determining R from radar measurements is made complex by sampling issues (e.g. Andrieu et al., 1997), comparisons of incompatible scales, and instrumental uncertainty (Krajewski and Smith, 2002). With the advent of satellite-based weather radars (e.g. Kawanishi et al., 2000; Hou et al., 2014) and passive sensors, precipitation can now be observed on a global scale (e.g. Huffman et al., 2007; Hou et al., 2008). Algorithms for satellite-based QPE generally use a model of the DSD with parameters that are either fixed or derived from observations (e.g. Iguchi et al., 2000; Seto et al., 2013; Liao et al., 2014). The assumption of the DSD from a model was classed as a primary uncertainty factor for the Tropical Rainfall Measuring Mission (TRMM) satellite weather radar (Iguchi et al., 2009). Variability in the DSD affects the accuracy of radar observations of rainfall, whether they are made from the ground or from space.

Improving our understanding of the DSD and its small-scale variability will lead to better understanding of the physical processes at play in rainfall, reduced uncertainty in precipitation measurements, and more accurate numerical weather prediction (NWP). In this thesis we present new techniques for the measurement and stochastic simulation of the DSD, and we use them to quantify DSD variability and test DSD-retrieval techniques in Mediterranean rainfall. This introductory chapter sets the scene for the rest of the thesis, by briefly introducing the main topics and outlining the current state of the art. In Section 1.1 the DSD and its related bulk variables are introduced in more detail. Measurement and estimation of the DSD are discussed in Section 1.2. The effects of DSD variability and the change of support problem are discussed in Section 1.3. The bulk of the data used in this thesis were collected in Ardèche, France, a region that experiences heavy Mediterranean rainfall. The meteorological processes of this region are briefly introduced in Section 1.4. In Section 1.5, the outline of the rest of the thesis is shown.

1.1 The raindrop size distribution

Let us imagine a rainstorm, frozen in time. We take one cubic metre of space within the storm, and within this space we collect all the raindrops, count them and measure their sizes. On average, there would be about 10^3 falling raindrops in this cubic metre (Uijlenhoet and Sempere Torres, 2006). Most of the drops would be small, between about 0.1 and 1 mm in diameter, and close to spherical (Pruppacher and Klett, 2000). There would also be some larger drops, which would be affected by air resistance, and thus not spherical. The bottom of these drops flattens, giving them an oblate shape that can be predicted as a function of the drop's volume (e.g. Beard and Chuang, 1987; Andsager et al., 1999; Pruppacher and Klett, 2000; Thurai and Bringi, 2005; Thurai et al., 2007). Because not all raindrops are spherical, we speak of their size in terms of their equivolume diameter: the diameter of a sphere that contains the same amount of water as the drop. The great majority of raindrops have equivolume diameters between 0.1 and 6 mm (Uijlenhoet and Sempere Torres, 2006).

Now in our imaginary situation, let us allow the system to fall into motion. Each drop falls at a velocity that depends on its mass plus atmospheric conditions. The still-air terminal velocity of raindrops can be accurately predicted (e.g. Atlas et al., 1973; Beard, 1976; Brandes et al., 2002) and ranges from 0.1 to more than 9 m s^{-1} (Uijlenhoet and Sempere Torres, 2006; Roe, 2005). As the rain falls, the number and sizes of the drops in our cubic metre of air are constantly changing. Evaporation causes the loss of some (mostly small) drops (e.g. Rosenfeld and Ulbrich, 2003). Drops collide with each other, with smaller drops joining together and coalescing to form larger drops, and larger drops breaking up (e.g. Pruppacher and Klett, 2000). Raindrops can only reach a certain size – about 10 mm – before they break up into smaller drops solely due to aerodynamic forces (Pruppacher and Klett, 2000), and it is for this reason that there are always more small drops than larger ones.

The DSD is used to statistically summarise the information above. The volumetric DSD,

written $N(D)$ [mm^{-1} m^{-3}] , is the number of raindrops with equivolume diameter in the range $[D, D+\delta)$ mm, per unit volume of air (Marshall and Palmer, 1948). The DSD describes the microstructure of liquid precipitation. Integral parameters of rainfall, also known as bulk rainfall variables, can be derived as weighted moments of the DSD (e.g. Ulbrich, 1983; Testud et al., 2001). Any bulk variable P can be expressed as

$$P = a_P \int_0^\infty w_P D^p N(D) dD, \tag{1.1}$$

where p and a_p are constants and w_p is a weight that may depend on D (Ulbrich, 1985). In this section the most commonly used bulk variables are briefly defined in increasing moment order (for a detailed review, see e.g. Bringi and Chandrasekar, 2001).

The zeroth moment of the DSD is the total drop concentration N_t [m^{-3}], defined simply as

$$N_\mathrm{t} = \int_0^\infty N(D) dD. \tag{1.2}$$

The DSD can be expressed as the total drop concentration multiplied by a probability density function $f(D)$ [mm^{-1}], such that $N(D) = N_\mathrm{t} f(D)$. The liquid water content, W [g m^{-3}] is related to the third moment of the DSD:

$$W = \frac{\pi 10^{-3} \rho_\omega}{6} \int_0^\infty D^3 N(D) dD, \tag{1.3}$$

where ρ_ω [g cm^{-3}] is the density of water. The flux of rainwater at a surface is expressed by the rain rate R [mm h^{-1}], defined as

$$R = 6\pi 10^{-4} \int_0^\infty N(D) v(D) D^3 dD, \tag{1.4}$$

where $v(D)$ [m s^{-1}] is the still-air fall velocity for a drop with equivolume diameter D. For the work presented in this thesis, we used the model of Beard (1976) to calculate $v(D)$. R is the variable that is of most interest to hydrologists.

The median-volume drop diameter, D_0 [mm], is the diameter that divides the DSD into two portions of equal water volume. More commonly used as a characteristic drop diameter, however, is the mass-weighted mean drop diameter, D_m [mm]. It is defined as the fourth divided by the third moment of the DSD, such that

$$D_m = \frac{\int_0^\infty D^4 N(D) dD}{\int_0^\infty D^3 N(D) dD}. \tag{1.5}$$

Weather radars emit electromagnetic radiation and measure what is reflected back off hydrometeors in the atmosphere. Radar reflectivity, the quantity measured by conventional radars, can be derived from the DSD (Marshall and Palmer, 1948). When the particles scattering the radiation are much smaller than the radar wavelength, the scattering properties are governed by the Rayleigh regime, and radar reflectivity Z [dBZ] is equal to the sixth DSD moment (Marshall and Palmer, 1948; Bringi and Chandrasekar, 2001), such that

$$Z = 10\log_{10}\left[\int_{D_{\min}}^{D_{\max}} N(D) D^6 dD\right]. \tag{1.6}$$

It is often the case, however, that the particles are similar in size to the wavelength. In this case the reflectivity occurs in the Mie regime, and it can be calculated from the DSD using

$$Z = 10\log_{10}\left[\frac{10^6 \lambda^4}{\pi^5 |K_\omega|^2} \int_0^\infty \sigma_b(D) N(D) dD\right], \tag{1.7}$$

where λ [cm] is the radar wavelength, $|K_\omega|^2$ [-] is the dielectric factor of water, and $\sigma_b(D)$ [cm^2] is the back-scattering cross-section for a drop with equivolume diameter D (e.g. Bringi and Chandrasekar, 2001). The scattering properties of water droplets can be calculated using the T-matrix codes of Mishchenko and Travis (1998).

In the case of polarimetric radars, in which the electromagnetic waves are horizontally and vertically polarised, the horizontal reflectivity Z_H [dBZ] is calculated by replacing $\sigma_b(D)$ in Equation 1.7 with $\sigma_{bH}(D)$ [cm^2], the back-scattering cross section in horizontal polarisation. Radar reflectivity in vertical polarisation, Z_V [dBZ], is defined in the same way, using $\sigma_{bV}(D)$

[cm^2], the back-scattering cross-section in vertical polarisation. It is common practice to refer to radar reflectivity in dBZ as defined above. At times it is also used, however, in its linear units, in which case we have horizontal reflectivity Z_h [$mm^6\ m^{-3}$] and vertical reflectivity Z_v [$mm^6\ m^{-3}$], defined as $Z_h = 10^{Z_H/10}$ and $Z_v = 10^{Z_V/10}$ respectively. Differential reflectivity, Z_{DR} [dB], defined as $Z_H - Z_V$, is the ratio of horizontal to vertical reflectivity. Z_{DR} is a useful variable in rain, because with large drops being more oblate it is related to drop size (Seliga and Bringi, 1976).

In this section, all the integrals have been written assuming a continuous DSD function and drop sizes ranging from zero to infinity. This is idealised, because in reality, not only are drops finite in size, but they are usually measured in discrete classes of equivolume drop diameter. When working using measured data, therefore, the integrals in calculations of bulk variables convert to sums over the drop size classes from D_{min} [mm] to D_{max} [mm], the smallest and largest considered class-centre drop sizes. dD becomes the width of each class, and D is the centre diameter of each class. Studies on DSD truncation and the calculation of bulk variables have concluded that the effects of truncation are negligible as long as the included range of diameters is large enough around D_0 (Willis, 1984; Ulbrich, 1985; Vivekanandan et al., 2004).

It is common for the DSD to be summarised using a functional form defined by only a few parameters. The first proposed form was the exponential function of Marshall and Palmer (1948). The Gamma DSD (Ulbrich, 1983) is an extension of the exponential form that is more appropriate for instantaneous measurements of the DSD. Other compact forms of the DSD include the normalised DSD of Willis (1984), and normalisation approaches in which the DSD is expressed using one or more of its statistical moments and a normalised DSD function that describes the shape of the distribution (e.g. Sempere-Torres et al., 1994; Testud et al., 2001; Lee et al., 2004). In this thesis we provide a DSD interpolation method that requires no functional form (Chapter 3), show results of using it to test areal rainfall retrieval functions that do use a DSD model (Chapter 4), and study the spatial invariance of a normalised DSD function (Chapter 5).

1.2 Measurement and estimation of the DSD

Disdrometers are instruments that measure the DSD at ground level, with collection areas that are small enough that they are usually considered to be point measurements. There are various types of disdrometers. Impact-type disdrometers (e.g. the Distromet Joss-Waldvogel disdrometer, Joss and Waldvogel, 1967) measure the impact forces of raindrops that hit a sensor. Laser optical disdrometers use a sheet of light that is interrupted by falling drops. The width of the shadow and length of the interruption give the size and fall speed of the observed particle. The primary data set used in this thesis is from a network of OTT Particle Size and Velocity (Parsivel, Löffler-Mang and Joss, 2000) laser optical disdrometers. Another type of disdrometer is the two-dimensional video-disdrometer (2DVD) (Schönhuber et al., 2008) that uses two orthogonally-facing cameras to take images of individual falling particles and to

directly measure their velocities. While the 2DVD has limitations (Tokay et al., 2001, 2013), it provides high-resolution data on individual hydrometeors, and matches rain rates from collocated rain gauges better than other disdrometers, including Parsivels (Tokay et al., 2001; Thurai et al., 2011; Tokay et al., 2013). Disdrometer data require careful treatment, because they can be affected by splashing, wind turbulence, multiple particles, particles that are not water (insects, spiderwebs), and instrumentation uncertainties (e.g. Kruger and Krajewski, 2002; Thurai and Bringi, 2005; Jaffrain and Berne, 2011; Tokay et al., 2013).

Networks of disdrometers have been the preferred way to study the variability of the DSD across space and time (e.g. Miriovsky et al., 2004; Lee et al., 2009; Tapiador et al., 2010; Tokay and Bashor, 2010; Jaffrain et al., 2011; Jameson et al., 2015a). A difficulty is that disdrometers provide point measurements that may be too sparsely distributed or not numerous enough to fully capture the variability of the rainfall process. It has been estimated that at least six disdrometers would be required per square kilometre to properly sample the variability of the DSD (Tapiador et al., 2010). Interpolation using geostatistics (Matheron, 1971; Chilès and Delfiner, 1999) offers a possible solution to this problem, through the estimation of the DSD at unmeasured locations, conditioned on nearby measurements. To date, rainfall interpolation methods that produce gridded outputs have generally worked with individual bulk variables such as rain rate (e.g. Creutin and Obled, 1982; Chua and Bras, 1982; Goovaerts, 2000; Tobin et al., 2011; Masson and Frei, 2014; Haberlandt, 2007; Velasco-Forero et al., 2008; Tobin et al., 2011). However, the full variability of the DSD can not be captured using only integral variables, because it is possible for multiple DSDs to produce the same bulk variable values (Jameson et al., 2015b).

Stochastic simulation can be used within a geostatistical framework to produce many equally probable realisations of a process. This approach has been used to estimate DSD model parameters and investigate DSD variability (Jaffrain and Berne, 2012b; Schleiss et al., 2012). The assumption of either second-order or at least intrinsic stationarity is required for geostatistical approaches (Chilès and Delfiner, 1999). Rainfall, however, is a non-stationary process (Barancourt et al., 1992; Schleiss et al., 2014a). This non-stationarity is caused largely by the high intermittency (patchiness) of rainfall fields (Schleiss et al., 2014a). While intermittency can be modelled through the use of a rainfall occurrence map (Barancourt et al., 1992), this does not solve the problem of the non-stationarity. Schleiss et al. (2014a) offered a solution through taking into account the so-called “dry-drift”, or the tendency of rainfall to be heavier at a point that is further from a dry region (Barancourt et al., 1992; Braud et al., 1994; Emmanuel et al., 2012).

The techniques mentioned so far in this section have included direct sampling of the DSD with disdrometers, and interpolation and simulation of DSD parameters in order to investigate horizontal variability. The DSD varies in the vertical, as well. Vertically profiling radars that measure the Doppler spectrum are able to infer some properties of the DSD at height, using measured fall speeds towards the radar. This is the method used by micro rain radars (MRRs, Peters et al., 2002, 2005; Tridon et al., 2011). A significant limitation of these methods is

that vertical wind and turbulence are ignored (Peters et al., 2002). Schleiss and Smith (2015) proposed a geostatistical technique that uses disdrometer time series from the ground, and radar data, to estimate 3D–time variograms for DSD model parameters. Retrieval of the DSD from (non-Doppler) polarimetric radar data has been a long-standing goal, but has proven difficult. Some microphysical properties, such as the median drop diameter D_0, can be retrieved from polarimetric information (Seliga and Bringi, 1976). Since this was discovered, many DSD-retrieval methods have been developed to try to retrieve other properties of the DSD (e.g. Zhang et al., 2001; Gorgucci et al., 2002; Park et al., 2005b; Anagnostou et al., 2009, 2010; Thurai et al., 2012; Bringi et al., 2015; Kalogiros et al., 2013).

In this thesis we present new ways to correct possibly inaccurate measurements of the DSD (Chapter 2), and new ways to infer the DSD or its properties from nearby or remote measurements. A new geostatistical interpolation technique and stochastic simulation technique for the experimental (i.e. non-parametric) DSD is presented (Chapter 3), and we introduce a new DSD-retrieval technique that uses polarimetric radar data (Chapter 6). These techniques were developed using measurements of the DSD made with disdrometers and radars. The disdrometers used were Parsivel laser optical disdrometers and a 2DVD. The main data set, which is introduced in Chapter 2, was collected by a network of disdrometers collocated with rain gauges, plus a weather station, MRRs, and an X-band polarimetric weather radar (MXPol, see Schneebeli et al., 2013). Examples of these instruments deployed in the field are shown in Figure 1.1.

1.3 DSD variability and the change of support problem

The support of a measurement is the region in space over which is it taken. The change of support problem refers to the non-equivalence of measurements taken with different supports. DSD measurements are affected by the change of support problem because of the high variability of the DSD. DSD variability is known to be at least as great within rainfall events as between them (Tapiador et al., 2010; Jaffrain and Berne, 2012b), and it has been shown that the variability of the DSD is greater across larger domains than smaller ones (Jaffrain and Berne, 2012b; Jameson et al., 2015b), and to be greater for larger (and therefore rarer) drops than smaller ones (Jameson et al., 2015b). Variability is greater between measurements taken with short integration times, as compared to longer ones where the integration causes smoothing of the process (Jaffrain and Berne, 2012b; Tokay and Bashor, 2010). Given this high variability, an areal measurement of the DSD can therefore not be assumed to equal a point measurement. Yet, at times, only areal measurements or only point measurements are available. In this thesis we quantify the error introduced by assuming a point represents an area, and investigate whether areal rainfall retrieval algorithms properly represent the sub-grid DSD.

The sub-grid variability of the DSD affects the areal retrieval of rainfall. For example, in radar QPE, the properties of the link between radar reflectivity and rain intensity varies with the

Figure 1.1 – Examples of instruments used in this thesis, as deployed in the field. The top panel shows Pradel Grainage in Ardèche, France, with (L–R) the Vaisala weather station, 2DVD, first-generation Parsivel and Parsivel^2. Below left shows the Montbrun, also in Ardèche, with (L–R) the MRR, first-generation Parsivel, and MXPol. Below right shows Payerne Station SwissMetNet, in Payerne, Switzerland, with (L–R) the 2DVD and a first-generation Parsivel disdrometer.

DSD. As well as the scales of the measurements being different, any comparison between Z aloft and R on the ground is subject to additional uncertainty due to the vertical evolution of the rainfall (Zawadzki, 1975). The grid size of rainfall products varies. For example, a typical ground-based weather radar pixel size is 1×1 km^2 (Berne and Krajewski, 2013). NWP models typically use slightly larger pixels; for example the high-resolution Consortium for Small-scale Modelling (COSMO)[1] atmospheric model uses an operational pixel size of 2.8×2.8 km^2 (in Germany, Baldauf et al., 2011). Space-borne weather radars use a larger pixel size; for example

[1] See `http://www.cosmo-model.org`.

the Global Precipitation Measurement (GPM) satellite has a footprint of about 5×5 km^2 (Hou et al., 2014). The rainfall retrieval algorithms used by systems such as GPM and COSMO rely on first deriving DSD properties such as model parameters, then calculating bulk variables. The change of support problem means that DSD properties derived at the areal scale may not always be representative of the sub-grid rainfall process.

1.4 Mediterranean rainfall

The region of Ardèche, in France, was the study region for most of the work in this thesis. The western Mediterranean, in which Ardèche is located, is plagued by heavy precipitation events (HPEs) that threaten both life and property (Delrieu et al., 2005; Nuissier et al., 2008). The broad meteorological conditions that cause these events are well understood at a synoptic scale (Nuissier et al., 2011), so forecasting of likely HPEs across a large region is relatively accurate. However, forecasting and modelling the precise location of an HPE and the amount of rainfall likely to result remain difficult problems (Ricard et al., 2012; Delrieu et al., 2005). For example, in 2002 an HPE in the Gard region of France cost 24 lives and 1.2 billion euros in damage, but warnings produced before and during the event significantly underestimated the expected rainfall amount, and the prediction error in the location of the cell was about 100 km (Delrieu et al., 2005). Improved understanding of the small-scale variability of rainfall is important for the reduction of these sorts of problems.

The atmospheric conditions that cause heavy precipitation events in the Mediterranean are well known (e.g. Miniscloux et al., 2001; Lin et al., 2001). They can be divided into three meteorological ingredients. First, large masses of warm, moist air in the lower atmosphere are produced by the Mediterranean Sea, particularly during the end of summer and the start of autumn (Miniscloux et al., 2001). Second, synoptic-scale systems push this air onto the land. Typically, an upper-level cold trough forms between the Atlantic and the Mediterranean, with low pressure over the Gulf of Biscay and high pressure over central Europe: this trough-ridge pattern generates a southerly flow that destabilises the warm air and pushes it northwards (Miniscloux et al., 2001; Ricard et al., 2012). Third, the mountainous landscape of the Mediterranean coast triggers convective activity and heavy rainfall follows (Delrieu et al., 2005). The heaviest rainfall is on the south-eastern flanks of mountainous regions which lie perpendicular to the airflow from the sea (Nuissier et al., 2008). The synoptic pattern is often quasi-stationary and slowly shifts to the east while maintaining similar conditions, thus leading to high rainfall totals because events can persist for several hours to days (Ricard et al., 2012).

Mediterranean HPEs can be classified as deep-convective or shallow-convective, or a mixture of these two types. At the deep-convective end of the classification are quasi-stationary mesoscale convective systems (MCSs), which form a characteristic V shape with convective cells continuously generated at the point of the V. A quasi-stationary MCS can generate up to 200-600 mm of rainfall in under 24 hours (Delrieu et al., 2005). At the shallow-convective end are orographic events, that in stable conditions form orographic rain bands (ORBs). These

events are associated with stationary flow, and form stable and active bands of precipitation which are parallel to the direction of the wind and can last for several hours (Miniscloux et al., 2001; Godart et al., 2011). Typically, deep-convective events are temporally shorter, show more intermittency ("patchiness"), and have higher rain rates than shallow-convective events (Godart et al., 2011). Orographic rainfall is a complex process (Houze, 2012) and precipitation rates in mountainous regions remain poorly known as a result (Roe, 2005).

The Hydrological Cycle in the Mediterranean Experiment[2] (HyMeX, Drobinski et al., 2014; Ducrocq et al., 2014) is a long-term, multi-disciplinary, international project that brings together teams of scientists with the common goal of providing better understanding of the complex water cycle of the Mediterranean. It has a particular emphasis on extreme and high-impact weather events such as HPEs. The main study region for this thesis was in Cévennes, France, in which an instrument network was deployed as part of HyMeX. The Cévennes region lies to the southeast of the Massif Central, and forms a large south-easterly facing slope to the sea, dissected by deep and narrow northwest to southeast valleys (Miniscloux et al., 2001; Godart et al., 2011). It is one of the five rainiest areas in the region (Nuissier et al., 2008). Rain occurs particularly during the autumn - there is a well defined precipitation maximum in October (Frei and Schär, 1998). The region is subject to HPEs (Ricard et al., 2012) and ORBs (Godart et al., 2011), with an average 7.6 days per year with rainfall over 150 mm (between 1967 and 2006, Ricard et al., 2012). Data from two autumn field campaigns in this region were used in this work.

1.5 Thesis outline

The work in this thesis follows a logical arc, from collection of accurate DSD data from a network of disdrometers, to technique development for the study of the variability of the DSD using spatial interpolation and stochastic simulation, to analysis of DSD variability in the horizontal and vertical. Here, a brief overview of the following chapters is provided. Each chapter has been adapted from a published or submitted research article.

Chapter 2 relates to collection of the high-quality and accurate point measurements of the DSD. Most of the DSD data used in this thesis were collected by a network of Parsivel disdrometers, which is introduced in this chapter. Like any instrument, disdrometers have collection errors that must be taken into account when their data are used. In particular, these disdrometers classify drops into "bins" of drop size, and may misestimate the number of drops in each class. In Chapter 2 we present a method for the correction of disdrometer data recorded using Parsivel disdrometers, that uses a 2DVD as a reference. In the first part of the method, drop velocities are corrected with reference to a theoretical model of raindrop terminal velocity. Second, raw disdrometer measurements are filtered to remove particles that are unlikely to be real drops. Third, concentrations per Parsivel diameter class are corrected such that on average they match those recorded by the 2DVD. The correction method improves the match

[2] See `http://www.hymex.org`

of DSD moments between the Parsivel and the 2DVD, and test results are shown in which, in the majority of cases, the Parsivel-derived rain rate was closer to that of a collocated rain gauge after the correction was applied. The corrected Parsivel data were then used for technique development and analyses of DSD variability.

In Chapter 3, we present a new method for spatial interpolation and stochastic simulation of experimental DSD spectra. Disdrometers such as Parsivels provide non-parametric experimental DSD spectra in which concentrations are provided per drop diameter class. In a network of disdrometers, the DSD is measured at discrete point locations with varying inter-measurement distance. Our interpolation and simulation approach uses geostatistics to estimate or simulate the DSD, in the same non-parametric classes, at unmeasured points in space. Careful processing is required in order to comply with the assumptions of the geostatistical technique. Non-stationarity of the DSD field is taken into account using the dry-drift (Schleiss et al., 2014a), which we show exists on DSD concentrations. Principal component analysis is used to find uncorrelated components of the DSD, which are then interpolated or simulated for the requested points. A back-transformation process produces the estimates of the DSD. Results of leave-one-out testing show low bias on the estimated DSDs.

In Chapter 4, the data correction and DSD stochastic simulation techniques are brought together to investigate the small-scale variability of the DSD. In this chapter we report on a study in which simulated high-resolution grids of DSDs, conditional on measurements, were used to test the effects of DSD variability on areal rainfall retrieval. DSD variability is studied over two scales that were chosen for their similarity to areal scales used in rainfall products. The first scale, 5×5 km^2, is the ground footprint size of the Global Precipitation Measurement (GPM) satellite weather radar. The second scale, 2.8×2.8 km^2, is the size of a pixel in the operational Consortium for Small-scale Modelling (COSMO) numerical weather model. The error introduced by assuming that a point measurement represents an areal measurement is quantified by areal size, and the retrieval algorithms for both GPM and COSMO are evaluated with respect to whether their pixel-scale results are representative of the sub-pixel process.

In Chapter 5 we move to investigation of the double-moment normalisation of the DSD. Normalisation allows for the compact representation of the DSD and is also useful for the investigation of DSD variability, but it relies on the idea that the normalised DSD is the same everywhere. We show that the double-moment normalisation method of Lee et al. (2004) is effective at collapsing the DSD into a mean shape. We show the results of tests on the invariance of the double-normalised DSD in three climatic regions, through both horizontal and vertical displacement of the normalised DSD. This is the first test of the invariance of the normalised DSD in the vertical and the first over a large horizontal range of up to 100 km in one region. The normalised DSD trained in one region is shown to be applicable to another region more than 7000 km away.

In Chapter 6, the assumption that the normalised DSD is invariant is used to develop a new technique for the retrieval of the DSD from polarimetric radar data. The assumption

of an invariant normalised DSD means that the goal of the new technique is essentially to recover two moments of the DSD from radar information, because with these values the whole DSD can be reconstructed at any point using double-moment normalisation. We show how moments three and six of the DSD can be recovered from polarimetric variables, and show results of testing the retrieval technique against a state-of-the-art method. The technique performs as well or better than the existing technique. Further, we introduce a new method for the treatment of noisy radar variables that substantially improves the performance of both techniques when they are applied to real radar data.

In Chapter 7, we turn from the study of small-scale variability of rain to that of snow, and present an investigation into the multifractal properties of snowfall at high temporal and spatial resolutions. Universal multifractal (UM) analysis is a method for describing the spatio-temporal variability of a process over many scales. We applied it to high-resolution snowfall data collected in the Swiss Alps with a 2DVD and cut into time series, reconstructed vertical columns, and small-scale snowflake accumulation maps. This is the first application of full UM to vertical columns of snowfall structure, and to snowfall-only accumulation maps and time series at such high resolution.

In the final chapter, Chapter 8, conclusions are drawn and perspectives for future research are explored. The work in the thesis is summarised and put into perspective.

2 Correction of Parsivel drop size distribution measurements

This chapter is adapted from:

- T. H. Raupach and A. Berne. Correction of raindrop size distributions measured by Parsivel disdrometers, using a two-dimensional video disdrometer as a reference. *Atmospheric Measurement Techniques*, 8(1):343–365, 2015. doi: 10.5194/amt-8-343-2015. URL http://www.atmos-meas-tech.net/8/343/2015/. Distributed under Creative Commons Attribution 3.0 License.
- T. H. Raupach and A. Berne. Corrigendum to "Correction of raindrop size distributions measured by Parsivel disdrometers, using a two-dimensional video disdrometer as a reference" published in *Atmospheric Measurement Techniques*, 8, 343-365, 2015. Atmos. Meas. Tech., 2015. doi: 10.5194/ amt-8-343-2015-corrigendum. URL http://www.atmos-meas-tech.net/8/343/2015/. Distributed under Creative Commons Attribution 3.0 License.

This work was completed by T. Raupach under the supervision of A. Berne. Research, analyses, and writing are by T. Raupach. For data acknowledgements, see Appendix A.

2.1 Introduction

In order to study rainfall microstructure effectively, we require accurate measurements of the DSD. Disdrometers are instruments that measure the DSD at a point location. There are various types, but in this chapter we are concerned with the OTT Hydromet particle size and velocity (Parsivel) disdrometer, and the two-dimensional video disdrometer (2DVD) from Joanneum Research. The original Parsivel was by PM Tech Inc. OTT Hydromet purchased the rights to the instrument and redesigned it in 2005; the result was the first-generation Parsivel. The second-generation Parsivel^2 was introduced in 2011, and provided improvements over the first-generation model (Tokay et al., 2014). The Parsivel is a laser optical disdrometer that uses a sheet of light through which drops fall. The diameter and velocity of a drop is then determined by sensing the shadow it casts and for how long it casts it (Löffler-Mang and Joss, 2000). Parsivels bin drops into classes of velocity and diameter and record the number of drops measured per class over an integration time. Parsivel disdrometers have been shown to be susceptible to errors in the recorded drop concentrations, particularly for small and large drops (Krajewski et al., 2006; Tokay et al., 2013). The Parsivel measurement technique assumes properties of the precipitation that are far more appropriate for rain than for solid precipitation; for example, that particles will be spheroidal, have a horizontal orientation of their major axis, and that only one particle will be in the beam at once (Yuter et al., 2006; Battaglia et al., 2010). The Parsivel is, however, a low cost, durable, and reliable instrument that makes it particularly well-suited for deployment in networks to study the small-scale variability of the DSD (e.g. Tapiador et al., 2010; Jaffrain et al., 2011).

The 2DVD [1] uses two perpendicular high-speed line-scan cameras, each with an opposing light source, to measure particles from orthogonal angles and thus record their shape (e.g. Thurai and Bringi, 2005; Thurai et al., 2007) as well as their size and velocity (Kruger and Krajewski, 2002; Schönhuber et al., 2008). Information on each individual particle that falls through the measurement area of the 2DVD is recorded. A particle's fall speed is determined by the difference in time between its detection in the two camera planes, which are offset vertically by 6.2–7 mm. Thus, the 2DVD uses no literature-derived estimates for raindrop shape or velocity; these parameters are measured directly (Schönhuber et al., 2008). Some drawbacks of the 2DVD have been noted. In particular, drops with diameters smaller than 0.2 mm have been found to be unreliably measured (Tokay et al., 2001); Tokay et al. (2013) recommend taking 0.3 mm as a minimum measured diameter in 2DVD data due to underestimation of drop counts below this diameter. In earlier designs of the instrument, the reliability of measurements decreased with increasing wind speed (Nešpor et al., 2000). This has subsequently been addressed through design improvements (Schönhuber et al., 2007).

Several comparisons between 2DVD and Parsivel disdrometers have been reported on in the literature. In experimental trials the 2DVD has been found to produce better matches to rain

[1] The 2DVD was called the two-dimensional video distrometer by Schönhuber et al. (2008), to emphasise that the instrument collects information on the distribution of particles. To avoid confusion we use the standard spelling of disdrometer.

gauges than Joss and Waldvogel (Tokay et al., 2001) and Parsivel (Thurai et al., 2011; Tokay et al., 2013) disdrometers. Krajewski et al. (2006) showed that PM Tech Parsivel disdrometers measured higher numbers of small drops (0.2 to 0.4 mm) than the 2DVD and generally reported higher rain rates. In a study in Alabama, USA, using first-generation Parsivels, Tokay et al. (2013) found that Parsivel disdrometers were less sensitive to small drops than the 2DVD, and that they overestimated the numbers of drops over 2.44 mm in diameter, while underestimating the numbers of drops under 0.76 mm in diameter. Furthermore, they found that Parsivels measured fall velocities lower than the expected terminal fall speeds for drops larger than 2.44 mm in diameter. Tokay et al. (2013) concluded that inhomogeneous laser beams in first-generation Parsivel disdrometers were the cause of the misestimation of drop counts. Thurai et al. (2011) found that first-generation Parsivels recorded higher mass-weighted mean diameter and rain rate than 2DVD, mostly when the rain rate exceeded 20 mm h^{-1}.

Disdrometers can record erroneous measurements due to wind turbulence, splashing, mismatching between cameras (in the case of the 2DVD), multiple drops appearing at the same time, margin-fallers, or external interference from, for example, insects or spiderwebs. Minimal data treatment for disdrometer measurements usually involves removing outlier points by reference to expected terminal fall velocity (e.g. Tokay et al., 2001; Kruger and Krajewski, 2002; Thurai and Bringi, 2005). For example, Tokay et al. (2013) removed drops exceeding ±50 % of the expected terminal fall speed, while Jaffrain and Berne (2011) used a threshold of ±60 % of the expected fall speed. This existing approach removes particles that are obviously erroneous, but it has some shortcomings. By only removing measurements, it does not allow for the fact that the disdrometer may underestimate the number of drops falling. Most importantly, the treatment is based solely on bulk variables such as rain rate, and does not test whether the resulting DSDs after the correction are physically viable.

In this chapter, we present a correction method for DSD measurements provided by Parsivel disdrometers, using a 2DVD as a reference instrument. The correction is designed to ensure that the DSDs recorded by Parsivel disdrometers are accurate, in terms of both the raw DSD and its moments. The correction method adjusts two properties of the recorded DSDs. First, drop velocities per diameter class are shifted such that the mean velocity per diameter class aligns with the theoretical terminal drop velocity for raindrops of that diameter; these raw measurements can then be screened for implausible measurements. Second, per-diameter-class volumetric drop concentrations are scaled such that they match, in a statistical way, the concentrations measured by a collocated 2DVD.

The rest of this chapter is organised as follows: the data used are described in Section 2.2, where we introduce the Parsivel network that provided data for most of the work in this thesis. Measurement of the DSD and the instruments we are concerned with in this work are discussed in Section 2.3. The correction is introduced in Section 2.4. The results of the correction applied to the data are shown in Section 2.5 for first-generation Parsivels, and in Section 2.6 for Parsivel^2. The application of the technique to another climatology is addressed in Section 2.7. Concluding remarks are made in Section 2.8.

2.2 Data

The Parsivel DSD correction was developed and tested on first-generation Parsivel data collected during two consecutive autumns in Ardèche, France, as part of HyMeX (Drobinski et al., 2014; Ducrocq et al., 2014). The method was then also tested on second-generation Parsivel (Parsivel^2 hereafter) data collected in the same region in autumn 2013. To test the method in a different climatology and region, we used data from Payerne, Switzerland, using a first-generation Parsivel and the same 2DVD used in the HyMeX campaign. In this section these data sets are briefly described.

2.2.1 HyMeX SOPs 2012 and 2013

Two autumn campaigns in the same region in Ardèche, France, provided the primary data used in this thesis. The campaigns were special observation periods (SOPs) run between September and November in both 2012 (SOP2012) and 2013 (SOP2013). The field site was a roughly 13 × 7 km^2 area in the Cévennes region. Cévennes has a Köppen–Geiger Cfa climate type, which indicates that it has a temperate climate with no dry season and a hot summer (Peel et al., 2007). The town of Montélimar, about 18 km from the SOP2012 and SOP2013 field area, records an average annual rainfall of 905 mm, with 77 rainy days per year on average (MeteoFrance, 2014). Cévennes experiences Mediterranean rainfall and has a well-defined precipitation maximum in October (Frei and Schär, 1998). It is subject to heavy precipitation events that can produce large rainfall totals (greater than 150 mm) in a day (Ricard et al., 2012).

The instrument network can be split into two sub-networks that were separately managed. The first, deployed by EPFL LTE, consisted of seven (nine) first-generation Parsivel disdrometers in 2012 (2013), a transportable, polarimetric X-Band weather radar (MXPol, for instrument details see Schneebeli et al., 2013), a Vaisalla weather station (2013 only), and the 2DVD. The second (HPicoNet, see Appendix A) was composed of five (four in 2012) Parsivel^2 disdrometers, two first-generation Parsivel disdrometers, and a network of rain gauges. Collocated rain gauge measurements were available for all disdrometers across both networks, with the exception of Montbrun in 2013. In addition, in 2013, two micro rain radars (MRRs, Peters et al., 2002, 2005; Tridon et al., 2011) collected vertical profiles of precipitation. A map of the disdrometer locations is shown in Figure 2.1, and station information is shown in Table 2.1.

In this chapter, we used data from the LTE network and rain gauge network in 2012 and 2013, and the HPicoNet disdrometers in 2013 only. Due to data availability, only the HPicoNet stations shown in Table 2.15 were used in this chapter. Further, because there was no collocated rain gauge at Montbrun, the disdrometer data from Montbrun was not used in this chapter. The 2DVD was collocated with a Parsivel and a tipping-bucket rain gauge in 2013. For our purposes here, the main difference between the setup of the two years is that in SOP2013, there was a Parsivel, Parsivel^2, and rain gauge collocated with the 2DVD at the Pradel Grainage site. In SOP2012, the closest Parsivel and rain gauge to the 2DVD were at the site of Pradel 1 and Pradel 2, about 480 m away. For some analyses we combined data from the LTE network

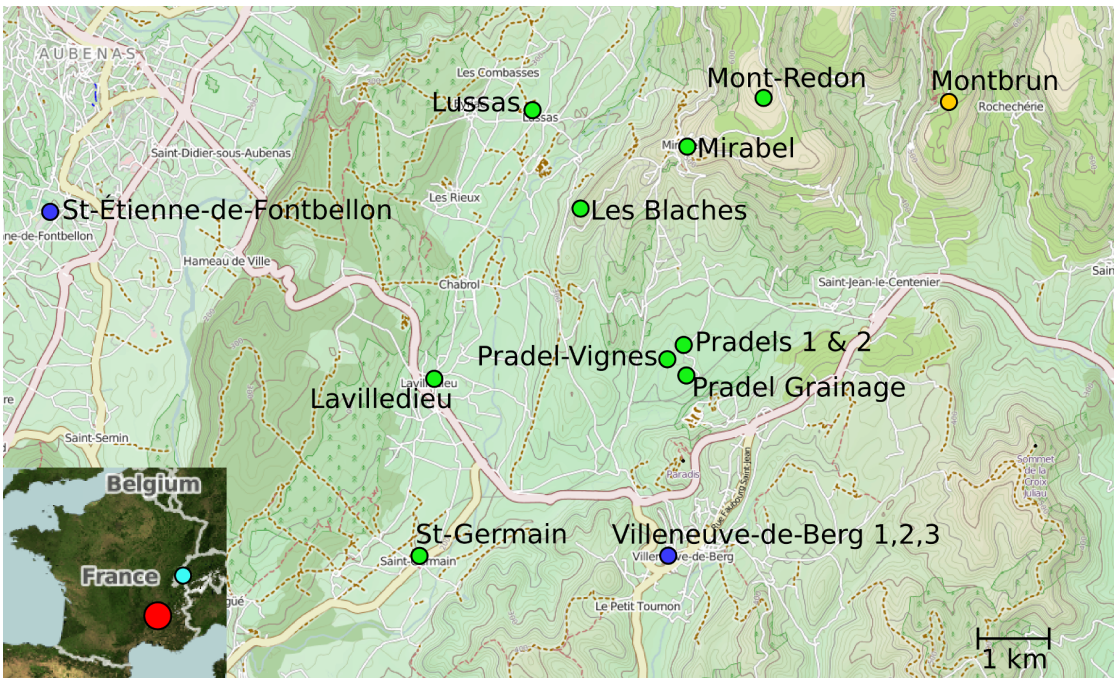


Figure 2.1 – A map of the HyMeX station network, showing Parsivel stations (green) and Parsivel^2 stations (dark blue). The MXPol weather radar was located, with a first-generation Parsivel, at Montbrun (yellow). The MRRs were at Montbrun and Pradel Grainage. Montbrun and Pradel Grainage disdrometers were deployed only in 2013. Pradel Grainage was the location for the 2DVD, and for both first and second-generation instruments in 2013. The inset map shows the location of the HyMeX network (red, not to scale) and Payerne (light blue, not to scale). Maps ©Thunderforest (CC BY-SA, http://www.thunderforest.org/), map data ©OpenStreetMap (ODbL, http://www.openstreetmap.org/copyright).

in SOP2012 and SOP2013 into a single data set, which we refer to as the “combined SOPs” data set.

Due to a clock error with the 2DVD, a variable clock drift was present in the 2DVD data. During the campaign, Parsivel clocks were synchronised using inbuilt global positioning system (GPS) receivers and were thus more reliable than the 2DVD clock. Adjustments were made to the 2DVD data for SOP2013 in order to synchronise the clocks of the instruments, for events where it was possible to do so. This synchronisation was done manually, by comparing time series of the rain rate from the 2DVD and a collocated Parsivel. The 2DVD time series was shifted forward in time to match the Parsivel time series as closely as possible, at 30 s temporal resolution. The adjustment was then applied to the series of individual 2DVD drops. Table 2.2 shows the adjustments made per event in SOP2013, which are between 30 and 60 s. For SOP2012, no adjustments were made because the 2DVD was not collocated with any instruments.

Inst	Name	Lat (°N)	Long (°E)	Alt	H12	H13	A12	A13	T
P1	Lavilledieu	44.5772	4.4532	227	108	70	233	193	30
P1	Les Blaches	44.6008	4.4810	429	103	67	233	148	30
P1	Lussas	44.6123	4.4706	289	84	64	202	149	30
P1	Mirabel	44.6069	4.4987	496	108	72	217	193	30
P1	Pradel 1	44.5829	4.4987	278	105	72	233	144	30
P1	Pradel 2	44.5829	4.4987	278	108	64	258	115	30
P1	Pradel Grainage	44.5790	4.5011	271		68		146	30
P1	St-Germain	44.5551	4.4497	204	110	55	263	76	30
P1	Montbrun	44.6141	4.5460	602		44		85	30
2DVD	2DVD	44.5790	4.5011	271	102	66	229	197	
P1	Mont-Redon	44.6141	4.5148	636	60	94	97	190	60
P1	Pradel-Vignes	44.5801	4.4950	256	1	37	0	106	60
P2	Pradel Grainage	44.5790	4.5011	271		79		179	60
P2	Villeneuve de Berg	44.5548	4.4953	301	65	71	115	179	60
P2	Villeneuve de Berg 2	44.5547	4.4954	301	68	73	141	188	60
P2	Villeneuve de Berg 3	44.5548	4.4955	301	68		134		60
P2	St Etienne de Fontbellon	44.6000	4.3826	302	73		121		60

Table 2.1 – Disdrometer station information for the HyMeX campaigns, showing the instrument (P1 – first-generation Parsivel, P2 – Parsivel^2, 2DVD – two-dimensional video disdrometer), the WSG84 (World Geodetic System 1984) coordinates of each station, its altitude [m] above sea level, the number of hours it recorded liquid precipitation with all quality control flags positive during 2012 (H12) and 2013 (H13), the total amount [mm] it recorded for those times in 2012 (A12) and in 2013 (A13), and its raw measurement integration time (T) [s]. Note that hours and amounts shown in this table were calculated using 5-minutes resolution and $R > 0.1$ mm h^{-1}, after the correction had been applied. The station at St-Germain had technical problems in 2013. A horizontal line separates the two networks that were separately maintained and had different integration times.

2.2.2 Payerne

To test the method on data collected in a different region and a different climatology, we used data collected in Payerne, Switzerland. Payerne has a Köppen–Geiger Cfb climate type, which indicates that it has a temperate climate, without a dry season, and with a warm summer. It has an average annual rainfall of 891 mm, with an average of 114 rainy days per year (MeteoSwiss, 2013), meaning that the precipitation in Payerne is more evenly spread across the year than in Ardèche. In contrast to the drier summers of Ardèche, Payerne's rainfall is higher during the summer months (MeteoSwiss, 2013). A first-generation Parsivel disdrometer was collocated with a 2DVD at Station SwissMetNet in Payerne between April and June 2014 (see Table 5.2 for station coordinates). In this time the Parsivel recorded 263 mm of liquid precipitation ($R > 0.01$ mm h^{-1}) in a total of 208 hours of rain, while the 2DVD, which was deployed later, recorded 84 mm in a total of 130 hours of rain. In this chapter, we describe the development of the Parsivel correction technique by focusing on data from the HyMeX SOP2012 and SOP2013 campaigns. We then discuss the application of the technique to data from Payerne, to examine its applicability to a different climatological region.

Event	From (UTC)	To (UTC)	A [s]
24,25	20-10-2013, 00:00:00	24-10-2013, 00:00:00	60
26	27-10-2013, 00:00:00	28-10-2013, 00:00:00	30
27	02-11-2013, 00:00:00	03-11-2013, 00:00:00	60
28	04-11-2013, 00:00:00	05-11-2013, 00:00:00	60
29	05-11-2013, 00:00:00	06-11-2013, 00:00:00	30
30	18-11-2013, 00:00:00	19-11-2013, 00:00:00	30

Table 2.2 – Clock adjustments (A) for 2DVD events in HyMeX SOP2013.

2.3 Processing of disdrometer measurements

Disdrometer measurements must be processed to convert raw measurements into more useful forms. In this section we describe the processing of data from the Parsivel disdrometer and 2DVD.

2.3.1 Parsivel

Parsivel disdrometers bin measured particles into particle counts per velocity and diameter class. There are 32 velocity classes and 32 diameter classes, with varying widths. Parsivels also determine the rainfall intensity (or rain rate), and two status flags: one provides an indication of the type of precipitation being observed (liquid or solid, for example), and another provides information on the quality of the measurement. For example, if the glass in front of the Parsivel's laser beam is dirty and reliable measurements are no longer possible, that will be indicated by a quality flag with value of 2. Value 0 indicates normal operation, while value 1 indicates dirty glass but that measurements are still possible. Value 3 indicates that the laser is damaged. We make use of these flags to restrict our analysis to high-quality measurements.

The effective sampling area of the Parsivel disdrometer is about 54 cm^2, but is different for different diameter classes, due to the fact that the whole drop diameter must be included in the sampling area for the drop to be counted. So-called "margin-fallers" are automatically removed, which reduces the effective sampling area. For the ith class, the sampling area is (Löffler-Mang and Joss, 2000; Battaglia et al., 2010)

$$S_i^{\mathrm{Pars}} = 10^{-6} \times P_l \left(P_w - \frac{D_i}{2} \right), \tag{2.1}$$

where S_i^{Pars} [m^2] is the effective sampling area, D_i [mm] is the class-centre equivolume drop diameter for the ith diameter class, P_l [mm] is the length of the Parsivel beam (180 mm), and P_w [mm] is the width of the beam (30 mm).

Let $C_{v,i}$ [–] be the raw number of particles recorded by the Parsivel for the vth velocity class and the ith equivolume drop diameter class. Let Δt [s] be the measurement integration time,

V_v [m s^{-1}] the class-centre velocity of the vth velocity class, and δ_i [mm] the width of the ith diameter class. Then we can convert the raw number of particles into a per-diameter-class volumetric drop concentration N_i^{Pars} [m^{-3} mm^{-1}] using

$$N_i^{\text{Pars}} = \frac{1}{S_i^{\text{Pars}} \delta_i \Delta t} \sum_{v=1}^{32} \frac{C_{v,i}}{V_v}. \tag{2.2}$$

It is worth noting that the Parsivel instrument itself calculates and provides an estimate of the rain intensity. In this chapter we always refer to the estimate of rain rate provided by the Parsivel as the "Parsivel-derived intensity", to avoid confusion with the DSD-derived rain rate R, which is defined by Equation 1.4. The values of these two variables are usually very similar, but they are not exactly the same; differences are possibly due to peculiarities of the implemented Parsivel processing algorithm that is not public.

2.3.2 Two-dimensional video disdrometer

The 2DVD records details of individual drops, including the diameter and velocity of each and the effective sampling area of the instrument at the moment the drop was recorded. For our purposes it is practical to bin the drops into diameter classes. Let M be the number of drops that were recorded within one integration time of length Δt, and let S_j^{2DVD} [m^2] and V_j [m s^{-1}] be respectively the effective sampling area and fall velocity for the jth recorded particle. Then the ith equivolume diameter class, where the class width is δ_i [mm], will have a drop concentration N_i^{2DVD} [m^{-3} mm^{-1}] of

$$N_i^{\text{2DVD}} = \frac{1}{\delta_i \Delta t} \sum_{j=1}^{M} \frac{1}{S_j^{\text{2DVD}} V_j}. \tag{2.3}$$

While most 2DVD-derived bulk rainfall variables are calculated using this N_i^{2DVD}, the rainfall rate R [m h^{-1}] for a given time step can be calculated directly from the individual drop measurements without binning the drops into classes. The rain rate is given by

$$R^{\text{2DVD}} = \frac{6\pi \times 10^{-4}}{\Delta t} \sum_{j=1}^{M} \frac{D_j^3}{S_j^{\text{2DVD}}}, \tag{2.4}$$

where D_j [mm] is the equivolume diameter of the jth recorded drop. The difference between the drop-wise rain rate and the rain rate calculated from a binned DSD is very small; in the 2DVD data used in this chapter, the mean relative difference between DSD-derived rain rate

and rain rate calculated drop-wise was less than 0.5%.

While the classes for the Parsivel disdrometer are predefined, we can choose any class definition for the 2DVD data. For comparisons of drop concentrations and DSD moments with the Parsivel records, we used Parsivel diameter classes for the 2DVD. For computation of the rain rate R from 2DVD data we used diameter classes with a constant width of 0.2 mm, corresponding to the resolution of the 2DVD.

2.3.3 Criteria for suspicious particles

Before converting our raw drop counts into per-diameter-class volumetric drop counts, we perform some data processing, the aim of which is to filter out particles recorded by the Parsivels and the 2DVD that are very unlikely to be raindrops. These measurements are assumed to be caused by external interferences such as insects, or droplets of water caught in spiderwebs inside the measurement area. We use simple thresholds to exclude classes of velocity and diameter which are unfeasible. To decide on the values for the thresholds, the 2DVD was used as the reference because it is not as easily affected by these external factors as Parsivel disdrometers.

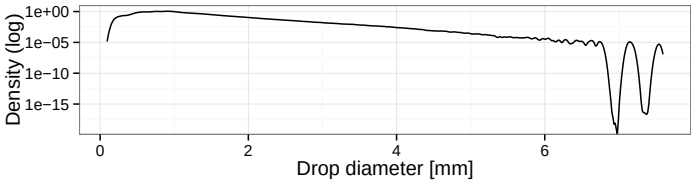


Figure 2.2 – Distribution of drop diameters recorded by the 2DVD in SOP2012 and SOP2013 events, with the y axis on a log scale.

Drops can only reach a certain size (about 10 mm) before they break up into smaller drops due to aerodynamic forces (Pruppacher and Klett, 2000). Figure 2.2 shows the distribution of drop sizes recorded in rain events from HyMeX (2012 and 2013) by the 2DVD. Table 2.4 shows the number of drops per diameter class for larger drops. Based on this information, and by looking at the velocity/diameter combinations that the 2DVD hardly ever recorded, we chose a filter that removes a drop with diameter D [mm] and velocity V [m s^{-1}] if any of the following conditions are true:

$$D > 7.5, \tag{2.5}$$

$$V > v(D) + 4, \tag{2.6}$$

$$V < v(D) - 3, \tag{2.7}$$

where $v(D)$ is the terminal velocity for a drop of equivolume diameter D as defined by Beard

(1976). Figure 2.3 shows the occurrence of velocity/diameter combinations recorded by the 2DVD during the combined SOPs. Figure 2.5 shows similar plots for sums of drop counts per Parsivel diameter and velocity class, for both the 2DVD and Parsivel. In these figures, the grey area is the region in which drops will be removed. Over the combined SOPs data set, the filter removed 0.2% of the drops recorded by the 2DVD. This filtering of suspicious records was applied to both Parsivel and 2DVD data before resampling to any different time resolutions. To resample Parsivel records, the mean DSD was found over each new time period and bulk rainfall variables were then calculated from each mean DSD. The Parsivel precipitation type flag was resampled to give an indication of the proportion of the time period for which solid precipitation was recorded. Note that here, solid precipitation refers to any precipitation that does not fit into the Parsivel instrument's criteria for liquid precipitation, which is based on the velocity and size of the particle (see Löffler-Mang and Joss, 2000). The worst quality flag was kept for each resampled time step, to give an indication of whether any low-quality flags were raised during the resampled integration time. Drop size classes that contained drops after the filter was applied are shown in Table 2.3. First-generation Parsivels do not record drops in the first two diameter classes (up to 0.2495 mm); we also considered these classes unreliable in Parsivel^2 records and did not use them.

$D_{\text{class min}}$ [mm]	$D_{\text{class max}}$ [mm]	D [mm]	δ [mm]
0.2495	0.3745	0.31	0.1250
0.3745	0.4995	0.44	0.1250
0.4995	0.6245	0.56	0.1250
0.6245	0.7495	0.69	0.1250
0.7495	0.8745	0.81	0.1250
0.8745	0.9995	0.94	0.1250
0.9995	1.1245	1.06	0.1250
1.1245	1.2500	1.19	0.1255
1.2500	1.5000	1.38	0.2500
1.5000	1.7500	1.62	0.2500
1.7500	2.0000	1.88	0.2500
2.0000	2.2500	2.12	0.2500
2.2500	2.5000	2.38	0.2500
2.5000	3.0000	2.75	0.5000
3.0000	3.5000	3.25	0.5000
3.5000	4.0000	3.75	0.5000
4.0000	4.5000	4.25	0.5000
4.5000	5.0000	4.75	0.5000
5.0000	6.0000	5.50	1.0000
6.0000	7.0000	6.50	1.0000

Table 2.3 – Parsivel equivolume drop diameter classes for which drops were recorded. D is the centre of each class to two decimal places, while δ is the width of each class. Class definitions are from the OTT Parsivel operating instructions. This table was originally included in Raupach and Berne (2016) (Chapter 4 of this thesis).

2.3.4 2DVD as reference instrument

Given that the 2DVD has previously been shown to produce better matches to independent rain rate measurements than Parsivel (e.g. Tokay et al., 2001; Krajewski et al., 2006), and that

Diameter class [mm]	Number of drops	% total drops
(5, 5.5]	273	0.00531
(5.5, 6]	97	0.00189
(6, 6.5]	36	0.00070
(6.5, 7]	10	0.00019
(7, 7.5]	3	0.00006
(7.5, 8]	1	0.00002

Table 2.4 – Numbers of large drops recorded by the 2DVD during the combined SOP event times.

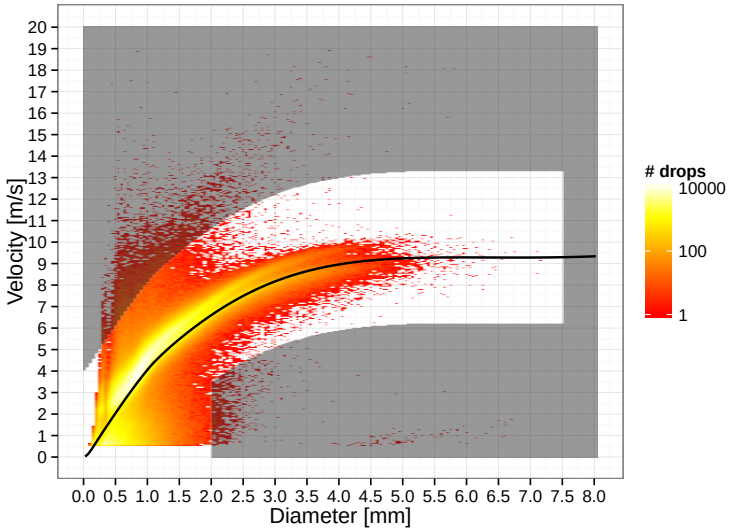


Figure 2.3 – Occurrence of velocity/diameter combinations recorded by 2DVD during the HyMeX campaigns in the autumns of 2012 and 2013. The physical-drop filter is overlaid in grey. The black line indicates the expected terminal drop velocity of Beard (1976).

it provides higher-resolution DSD measurements than Parsivel, both temporally and in the drop sizes it can discern, we used the 2DVD as the reference instrument for this work. To test the reliability of the 2DVD we compared the 2DVD measurements to collocated rain gauges for the HyMeX SOP2013 campaign. Two separate instruments were collocated with the 2DVD during SOP2013: a Vaisala weather station equipped with a rain cap, and a tipping-bucket rain gauge. We compared the rain rate derived from the 2DVD drop data (Equation 2.4) to rain gauge records. To remove solid particles we considered only time steps for which the collocated Parsivel recorded at least 90% liquid precipitation, and for which the 2DVD and rain gauge both recorded a rain rate greater than or equal to 0.1 mm h^{-1}. One outlier time step, for which the 2DVD was only partially working (4 October 2013, 18:00 UTC – Coordinated Universal Time), was removed. The comparisons are shown in Figure 2.6.

It is worth noting here the performance statistics we use. In all scatter plots in this chapter, the one-to-one line is shown in red dashes, while the blue line indicates the line of best fit found using linear least squares regression, with standard error shaded in grey. The reference instrument is always on the x axis. The regression slope (reg. slope) is the slope of the

regression line. For a given time t, let the reference value be R_t and the observed value be O_t. Let the total number of time steps be T. The mean ratio is defined as the reference mean divided by the observed mean. Let E_t be the difference for the tth time step, defined as $E_t = O_t - R_t$. RMSE is the root mean squared error,

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{t=1}^{T} E_t^2}{T}}. \tag{2.8}$$

r^2 is the squared Pearson correlation coefficient between reference and observed data sets. Bias is the mean of the differences, $\langle E \rangle$. Relative bias (RB) is the median of the relative errors, a percentage defined as

$$\mathrm{RB} = \mathrm{median}\left\{(O_t - R_t)/R_t \times 100\right\}. \tag{2.9}$$

We are only concerned with liquid precipitation in this chapter, so we subset time steps to those in which the Parsivel recorded no solid precipitation (for five-minute resolution) or at most 10% solid precipitation (for one-hour resolution), and for which the Parsivels recorded no non-zero quality status flags. Furthermore, we only compared time steps for which both instruments being compared recorded non-zero rain amounts. We take 0.01 mm h^{-1} as the minimum rain rate the Parsivel can record in one 30 s integration time. Thus, we use a non-zero rain rate threshold of 0.001 mm h^{-1} at five-minute resolution and of 8.3×10^{-5} mm h^{-1} at one-hour resolution. Because each tip of the tipping-bucket rain gauges indicates 0.1 mm of accumulated precipitation, the minimum rain rate that a rain gauge can measure in five minutes is 1.2 mm h^{-1}, and in one hour the minimum is 0.1 mm h^{-1}. When comparing to rain gauges, the non-zero rain rate threshold therefore becomes 1.2 mm h^{-1} for five-minute resolution and 0.1 mm h^{-1} for one-hour resolution. Because our correction affects the DSD-derived rain rates from the Parsivels, we use the Parsivel-derived intensity when applying the non-zero threshold to Parsivel data. We refer to time steps that satisfy these criteria as those with non-zero liquid DSDs.

The 2DVD showed excellent agreement with the tipping-bucket rain gauge and Vaisala weather station, with high correlation coefficients (r^2 at least 0.98) and low bias amounts for both comparisons (absolute bias less than or equal to 0.2 mm h^{-1}). In both cases the 2DVD tended to slightly underestimate the rain amount given by the other gauge. We conclude, however, that the 2DVD provided reliable measurements of the rain rate. Note that the relative bias between 2DVD and gauge was −14%, and between 2DVD and Vaisala it was 9%. The difference in these relative biases can be explained largely by differences in measurements of very small rain rates. This is equivalent to the relative bias we observed between two collocated Parsivels

(Pradel 1 and Pradel 2) using the same constraints to choose comparison time steps, also at one-hour resolution, after filtering for unfeasible records (but before any other correction was applied). Using Pradel 1 as reference, the relative bias was −6% (16%) in 2013 (2012). Using Pradel 2 as reference, the relative bias was 7% (−14%) in 2013 (2012). This means that when we compare Parsivel rain rates to rain gauges or to the 2DVD, we cannot distinguish the level of agreement when the relative bias is less than about 10%, due to instrumental uncertainty.

2.4 Correction of Parsivel DSDs

The correction of Parsivel DSDs is made in two steps. The two steps were chosen so that both the velocity and diameter measurements made by Parsivel disdrometers are addressed. First, the raw Parsivel data is corrected so that per-diameter-class mean velocities match the expected terminal velocity for each class. At this point the raw data can be screened for unfeasible measurements as described in Section 2.3.3. Second, a per-diameter-class adjustment factor is applied to Parsivel classes, in order to make the drop size distribution match, in a statistical way, that recorded by a 2DVD. This adjustment of drop concentrations in effect changes the drop equivolume diameters measured by the Parsivel. In this section we address each correction in turn.

2.4.1 Correction of per-diameter-class drop velocities

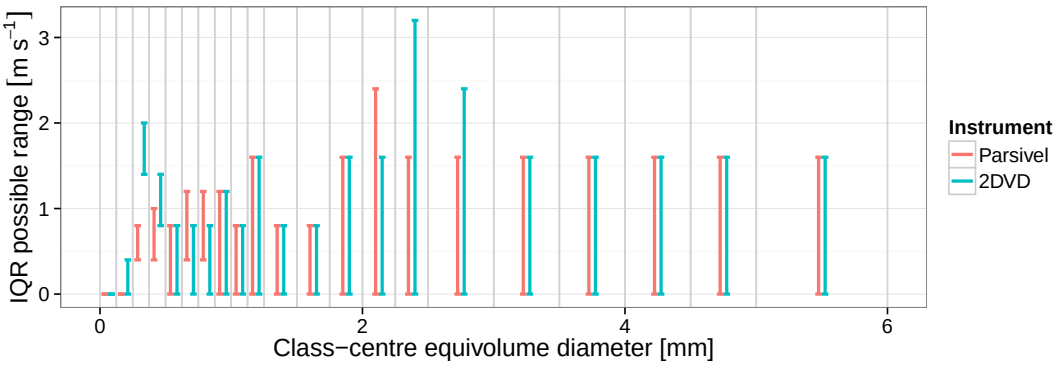


Figure 2.4 – Velocity IQR possible ranges, by Parsivel diameter and velocity class, for mean drop counts for 2DVD and the collocated Parsivel in HyMeX 2013. Only time steps for which both instruments recorded a value and the Parsivel-recorded liquid rain were included. The grey vertical bars indicate the Parsivel diameter class boundaries. Above the 21st diameter class (drops larger than 6 mm), there were not enough drops to meaningfully calculate a velocity range; for this reason the plot is truncated to 6 mm.

Figures 2.3 and 2.5 show the density of particles recorded at each diameter/velocity combination, by the 2DVD and Parsivel disdrometers. Both the 2DVD and Parsivel record drops at a range of velocities for a given equivolume diameter or diameter class. In these plots, the black line is the expected terminal velocity per drop diameter, calculated using the method of Beard (1976). The 2DVD recorded the highest concentrations of drops on and very near the expected

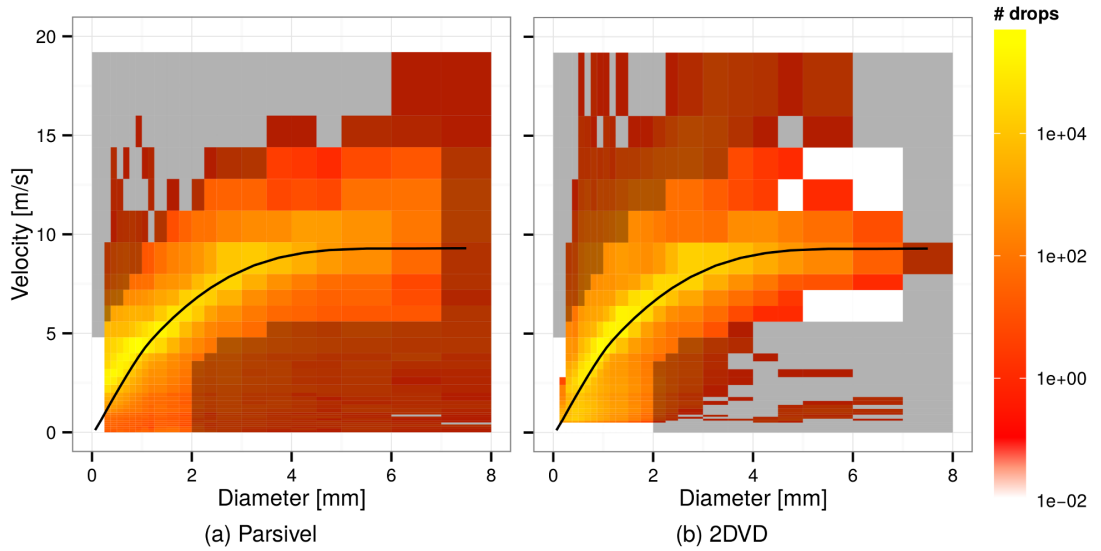
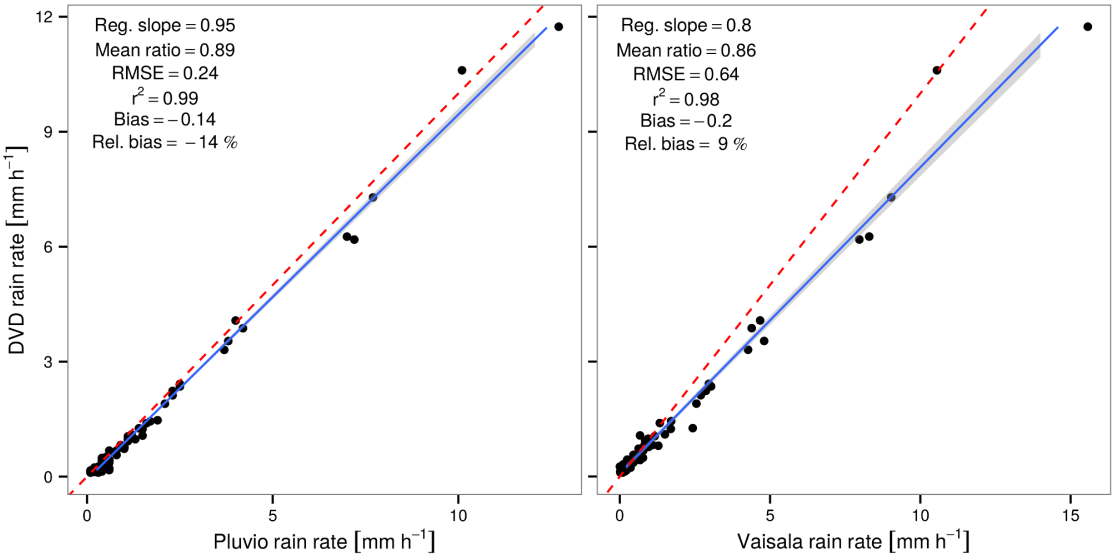


Figure 2.5 – Sum of raw drop occurrences per Parsivel class, for the 2012 and 2013 campaigns. Parsivel counts are summed at stations Pradel 1 (for 2012) and Pradel Grainage (for 2013). The filtered areas are overlaid in grey. The black line is the expected terminal drop velocity calculated by Beard (1976).

terminal velocities. Indeed in SOP2013, for time steps for which the nearest Parsivel recorded liquid precipitation, the bias between expected terminal velocity and velocity recorded by the 2DVD was 0.05 m s^{-1} and the relative bias was 2% (over the combined SOPs the bias was 0.2 m s^{-1} and relative bias was 6.5%). We hence consider the terminal fall velocity from Beard (1976) as the reference value for fall velocity. The Parsivel tends to overestimate the velocities of small drops.

To correct the velocities in the Parsivel data, we take the set of recorded velocities for each drop diameter class, and shift the values such that the mean velocity is equal to the expected terminal velocity as calculated by the algorithm of Beard (1976). Because the velocity classes do not have constant width, the classes are first sub-sampled into classes of width 0.1 m s^{-1}, then shifted and regrouped into the original class sizes. Except when some drops were counted in very low-velocity classes and are shifted out of the valid velocity range, the number of drops per diameter class remains the same before and after the velocity shift. An example plot of drop counts per velocity and diameter class before and after the velocity shift is shown in Figure 2.7. The velocity shift is equivalent to shifting each column up or down such that the mean velocity for each column (which is usually close to the brightest point) aligns with the line that indicates the expected terminal velocity. As an example, for the average drop counts per velocity and diameter class for SOP2013, using the Parsivel at Pradel Grainage, the mean shift required per diameter class from 0 to 5 mm was -0.34 m s^{-1}. Once the velocities are corrected in the raw Parsivel data, any suspicious particles are removed using the criteria shown in Section 2.3.3, and the volumetric drop concentrations per diameter class are found using Equation (2.2). This correction and filtering was applied before resampling to any lower time resolutions.



(a) Pluvio to 2DVD, 2013. (b) Vaisala to 2DVD, 2013.

Figure 2.6 – Scatter plots showing the comparison between the 2DVD and (a) a collocated tipping-bucket rain gauge (Pluvio), and (b) a collocated Vaisala weather station. Time steps compared are at one-hour resolution from HyMeX SOP2013, and include only those times for which the collocated Parsivel recorded a Parsivel-derived rain rate ≥ 0.1 mm h^{-1}, $\leq 10\%$ of the time step was marked as solid precipitation, and for which both the 2DVD and gauge recorded a rain rate ≥ 0.1 mm h^{-1}.

Figure 2.4 shows a comparison between the interquartile ranges (IQRs) of the recorded velocities, by diameter class. We calculated the mean drop counts per Parsivel velocity and diameter class for SOP2013, for 2DVD and Parsivel. We then applied the velocity correction to the Parsivel drop counts, and calculated the velocity class quantiles, weighted by the drop counts per diameter class, for each set. Given that the velocity classes can be quite large, it is possible that the quantiles both fall within one velocity class; in this case, the IQR could be between 0 and the width of the velocity class. We calculated the minimum and maximum possible IQR range for each diameter class. The plot shows that when binned into the Parsivel velocity classes, the spread of velocities was of the same order of magnitude between the two instruments. We conclude that to correct the velocities measured by Parsivel, it is sufficient to shift the mean velocity closer to the expected terminal velocity as described above, then remove suspicious particles (see Equations 2.5, 2.6, and 2.7).

2.4.2 Correction of diameter-class concentrations

We now turn to correcting the drop concentrations per diameter class with reference to the 2DVD. Let $P(i)$ be the ratio of 2DVD drop concentration to Parsivel drop concentration, defined such that for the ith equivolume diameter class with centre-diameter D_i [mm], at any given

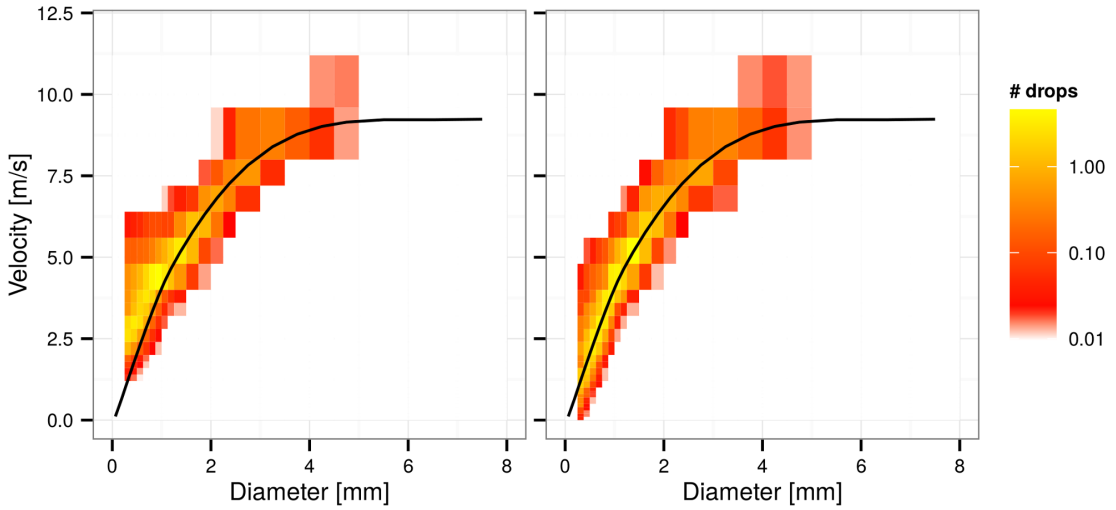


Figure 2.7 – An example of the velocity correction. Average drop counts for liquid rain from the Parsivel at Pradel Grainage for SOP2013, shown (a) before the velocity correction and (b) afterwards.

time step,

$$P(i) = \frac{N_i^{\text{2DVD}}}{N_i^{\text{Pars}}}. \tag{2.10}$$

$P(i)$ is thus the correction factor for that time step: when the Parsivel drop concentration for class i is multiplied by $P(i)$ it will match the 2DVD drop concentration for class i. To "train" the correction for a given data set, we find median values of $P(i)$ per class of Parsivel-derived rain intensity. Parsivel-derived intensity is used as it is a measurement of the rain intensity that is always available with Parsivel disdrometers, and is independent of our DSD correction. It is hence easily accessible to all potential users. The result is a collection of correction factors for each Parsivel-derived intensity class. When Parsivel records are multiplied by these correction factors, the per-diameter-class drop counts are scaled to match the corresponding 2DVD drop counts.

To explain the correction in more detail, we take as an example the HyMeX 2012 and 2013 SOPs and show each step of the correction calibration. We used data from SOP2013 to train the correction, because there was a Parsivel collocated with the 2DVD at Pradel Grainage in that campaign. We used a time resolution of one hour, in order to increase the chance of a time step sampling large drops, and in order to smooth outliers. Assuming the obtained correction is not dependent on the temporal resolution, it will be applied at resolutions higher than one hour in order to have reliable Parsivel DSD measurements for studies of small-scale DSD variability. A strict set of criteria was used to choose which time steps the comparison should

Class (i)	D_i [mm]	[0, 0.5)	[0.5, 1)	[1, 2)	[2, 200)
3	0.31	0.04	0.06	0.08	0.11
4	0.44	0.11	0.14	0.21	0.26
5	0.56	0.36	0.41	0.57	0.60
6	0.69	0.44	0.50	0.66	0.78
7	0.81	0.64	0.71	0.86	1.04
8	0.94	0.68	0.68	0.89	1.00
9	1.06	0.78	0.77	0.95	1.16
10	1.19	0.82	0.77	0.95	1.16
11	1.38	0.78	0.74	0.91	1.11
12	1.62	0.68	0.65	0.80	0.94
13	1.88	0.67	0.53	0.70	0.87
14	2.12	0.60	0.50	0.65	0.80
15	2.38	0.44	0.50	0.56	0.74
16	2.75	0.41	0.40	0.43	0.69
17	3.25	0.39	0.42	0.35	0.65
18	3.75	0.42		0.43	0.49
19	4.25				0.39
20	4.75				0.17
21	5.50				0.34

Table 2.5 – Calibrated first-generation Parsivel correction factors for Parsivel-derived intensity classes for the SOP2013 campaign. Each row contains the class number, the centre equivolume diameter for the class (D_i), and the calibrated factors $P(i)$ for each class of Parsivel-derived intensity. Intensity class boundaries are provided in millimetres per hour [mm h^{-1}].

be performed on. We used time steps for which the 2DVD and the collocated Parsivel recorded a non-zero liquid DSD. For all of SOP2013 there were 234 such time steps, corresponding to 234 h of rainfall over which we trained the correction factors. For each valid time step, we compared the mean DSD recorded by the 2DVD and collocated Parsivel.

Values of $P(i)$ were calculated for each time step in the training set, by comparing the Parsivel DSD to the 2DVD DSD. The result is a distribution of P values for each drop diameter class. To investigate the effect of rain intensity on the values of $P(i)$, we divided the time steps into classes of intensity, using the Parsivel-derived intensity modelled by the sensor. The median $P(i)$ values of each intensity and diameter class are shown in Figure 2.8. There is clearly a dependency between the values in the $P(i)$ curve and the rain intensity.

The most notable feature of Figure 2.8 is that the numbers of small drops (under about 0.7 mm) were overestimated by the Parsivel. For these classes, the values of $P(i)$ are low, indicating that the Parsivel drop counts need to be scaled down to match the corresponding 2DVD drop counts. For low rain rates, below 1 mm h^{-1}, the Parsivel overestimated drop counts in all classes up to 4 mm. Note that large drops are very rare in these rain rate classes and, as we will see, the values of $P(i)$ are more reliable for smaller drop sizes. We identified groups of behaviour of $P(i)$ by ranges of Parsivel-derived intensity, and thus divided the intensities into four classes for ranges $[0, 0.5)$, $[0.5, 1)$, $[1, 2)$ and $[2, 200)$ mm h^{-1}. Using these ranges as the class definitions for Parsivel-derived intensity, we obtained distributions of $P(i)$ per drop diameter and Parsivel-derived intensity class that are shown in Figure 2.9. The distributions are over

Moment	Bias bef.	Bias aft.	RB bef. [%]	RB aft. [%]	RMSE bef.	RMSE aft.	r^2 bef.	r^2 aft.
0	130.57	10.65	161.72	15.70	221.35	43.02	0.57	0.91
1	62.37	9.35	103.95	18.08	102.07	36.31	0.78	0.93
2	40.65	9.47	81.94	21.19	64.28	40.68	0.93	0.95
3	46.55	12.33	77.70	25.93	93.29	62.36	0.96	0.95
4	96.61	23.15	82.40	32.51	284.39	131.71	0.95	0.95
5	286.19	66.03	97.28	40.65	1053.62	384.56	0.93	0.93
6	1022.93	261.95	117.29	49.89	4266.92	1532.92	0.89	0.88
7	4137.77	1268.72	144.14	60.91	18732.09	7692.04	0.84	0.80

Table 2.6 – Time series statistics per moment, comparing Parsivel data (at Pradel Grainage) before (bef.) and after (aft.) the correction is applied, to the 2DVD (the reference), at five-minute resolution, for event times. Units of bias and RMSE are m^{-3} mm^p, where p is the moment order.

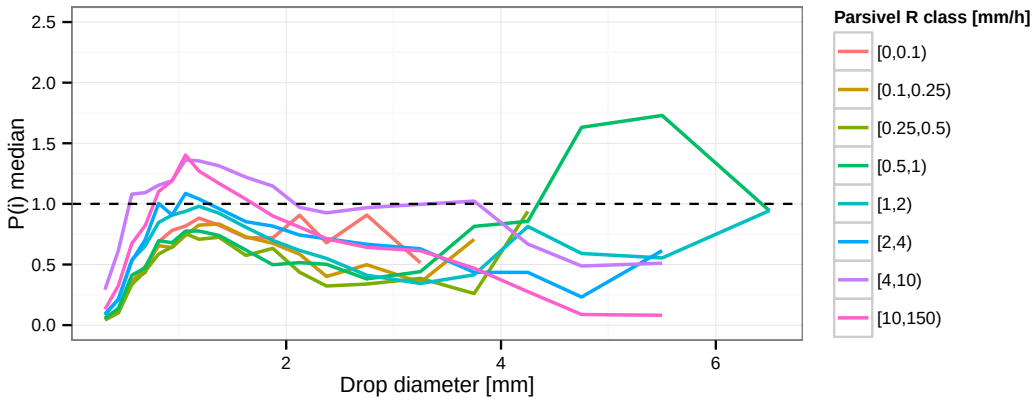


Figure 2.8 – Median $P(i)$ values classed by Parsivel-derived intensity.

all time steps and they get larger as the drop diameter increases, which shows that there was much more uncertainty in the correction factors for large diameters than for small diameters.

Across these rain rate classes there was a tendency for the Parsivel to overestimate the numbers of drops smaller than 0.81 mm in diameter and greater than 1.62 mm in diameter, with the best performance occurring in the 1–2 mm drop diameter range. For rain rates above 2 mm h^{-1}, the Parsivel more closely matched the 2DVD and indeed underestimates the numbers of drops between about 0.8 and 1.38 mm in diameter.

To train the correction factors, we randomly selected sets of 80% of the valid training time steps. To determine the impact of sampling effect, we reran the calibration 100 times with different randomly chosen calibration time steps, taking the median of the per-class $P(i)$ distribution each time, and recording the range of resulting values. These ranges are shown per Parsivel-derived intensity class in Figure 2.10. We see that the sampling effect for small drops was very small, but that it was larger for larger drop size classes. To ensure a more robust correction, we want to only apply the correction to drop diameter classes for which the training sampling effect (the spread) on $P(i)$ is small. However, in order for the correction to

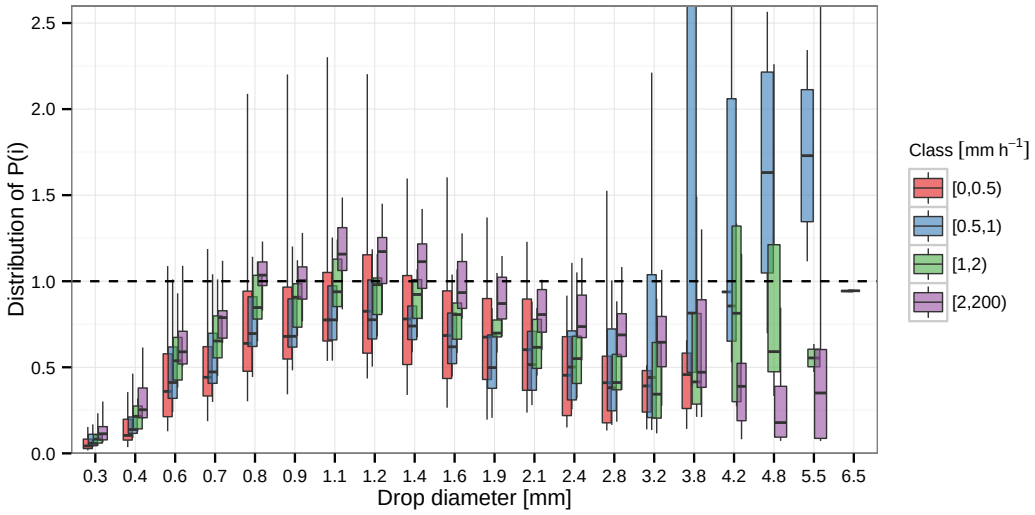


Figure 2.9 – Distributions of $P(i)$ values classed by Parsivel-derived intensity. The correction factors used are the medians of the distributions. The boxes show the interquartile ranges while lines show the 10–90% quantiles of each distribution. The y axis is cut at 2.6.

affect all moments of the DSD it is important that it is applied to larger drops as well as smaller ones. We decided to apply a threshold on the spread of the sampling effect. The correction was kept for increasing drop diameters until the sampling effect first surpassed this threshold.

There are hence two threshold values that must be chosen to train correction factors. The first is the minimum-allowed volumetric drop concentration for which 2DVD and Parsivel classes will be compared; let this threshold be Q [mm^{-1} m^{-3}]. The second threshold is the maximum-allowed spread in values of $P(i)$ over 100 training iterations of the filter; let this threshold be A. Q was set to 1×10^{-5} and acted simply to stop diminishingly small drop concentrations from adversely affecting the correction calibration. A was set to 0.7. The spread was also required to be larger than 1×10^{-6}, to ensure enough samples were available to give a representative calibration for each diameter class. A sensitivity analysis showed that the values of Q and A did not affect greatly the outcome of the calibration, so long as Q was low enough and A was large enough to allow for sampling and therefore correction of larger drops sizes.

To derive the final correction factors we iterated over 100 sets of training time steps, selecting randomly 80% of the available times for each iteration. The per-diameter and per-intensity class correction factor is the mean value of $P(i)$ medians for each class over all iterations. The calibrated correction factors for SOP2013 are shown in Table 2.5. The largest drop diameter class affected by the correction was the 21st Parsivel class, with a centre size of 5.5 mm. Drops up to the 17th class (class-centre diameter 3.25 mm) are corrected no matter the Parsivel-derived intensity. To apply the correction, each time step was taken separately and, depending on the Parsivel-derived intensity of the time step, the appropriate scaling factors were applied to each equivolume drop diameter amount.

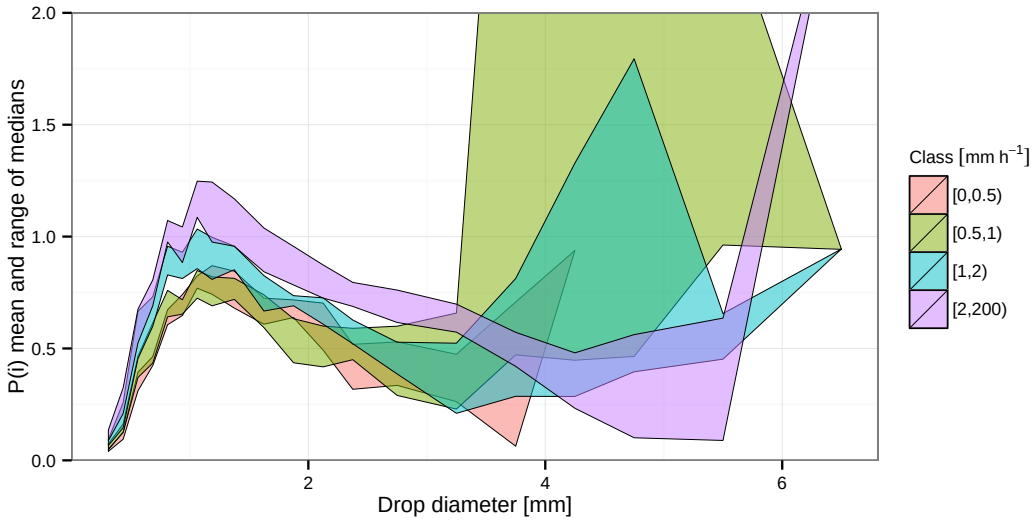


Figure 2.10 – Sampling effect per diameter class, for different classes of Parsivel-derived intensity. The coloured regions represent the minimum and maximum median $P(i)$ per equivolume drop diameter class observed over 100 iterations. The y axis is cut at 2.0.

The correction ensures that the corrected DSD more closely matches the DSD recorded by the 2DVD. For example, for the HyMeX SOP2013 data, Figure 2.11 shows the distributions of $P(i)$ after the correction, for one example validation set of 20% of the one hour time steps in SOP2013. After the correction the DSD much more closely matched that of the 2DVD, especially for small drop diameter classes. For larger drops of greater than about 3 mm the match was not as close, but note this is 20% of the data and sampling effect changes large drop comparisons much more than small ones. The fact that the large drops differed from the $1:1$ line reflects the difficulty in training a correction for classes in which there are not many drops to use as a training set, and demonstrates why we chose to train on one hour time steps and to use the mean $P(i)$ values over many iterations.

2.5 Drop concentration correction results

In this section we explore the effect of the correction on the moments of the DSD, including the derived rain rate. Our goal in this work is to have reliable DSD measurements from networks of Parsivel disdrometers, in order to be able to study the small-scale variability of the DSD in space and time. We are therefore interested in higher time resolutions than the one hour resolution we used to train the correction factors. Recall that the choice of one hour resolution for the training set was made to increase the numbers of sampled large drops, but that we aim to have a correction that is independent of the time resolution. We thus applied the trained correction to five-minute time resolution data to evaluate its effects, for all first-generation Parsivels in the SOP2013 campaign. We also applied the correction to data from SOP2012, as an independent validation data set, and to the combined SOPs. Recall that because we are

only interested in liquid precipitation, we subset the available time steps for each Parsivel station to those that contained no Parsivel warning flags regarding data quality, and no solid precipitation markers, and we only compared time steps for which both instruments being compared measured non-zero rain rates.

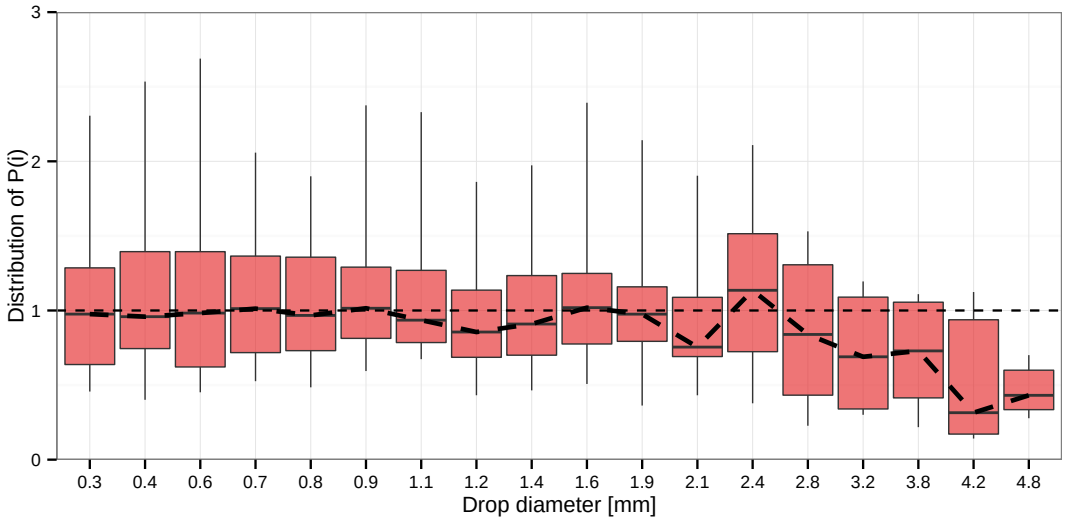


Figure 2.11 – The distributions of $P(i)$ values for the corrected DSD, on an example set of validation time steps from SOP2013, and for all classes of Parsivel-derived intensity, for the Parsivel collocated with the 2DVD. The y axis is cut at 3.

2.5.1 DSD moments

To demonstrate the effect of the Parsivel DSD correction on the moments of the DSD, we compare the first seven moments of the DSD recorded by the 2DVD, to the same moments derived from Parsivel DSDs before and after the correction is applied. For these comparisons we used HyMeX SOP2013 event time steps at five-minute resolution, and the Parsivel collocated with the 2DVD at Pradel Grainage. Comparisons of moments of orders 0, 1, 4, and 6 are displayed in Figure 2.12, Q–Q plots for these moments are shown in Figure 2.13, and time series statistics are shown in Table 2.6. We see from the densities and Q–Q plots that the correction shifted the distributions of all the moments towards those of the 2DVD. The statistics show an improvement in the relative bias of all moments, by a maximum of 146% for moment zero and a minimum of 49% for moment four. RMSE was improved for all moments. r^2 was improved for moments of orders 3, 4, 6, and 7, and remained very similar for the other moments. These results demonstrate that the correction improves Parsivel DSDs at high temporal resolution even when it is trained from one-hour DSD spectra.

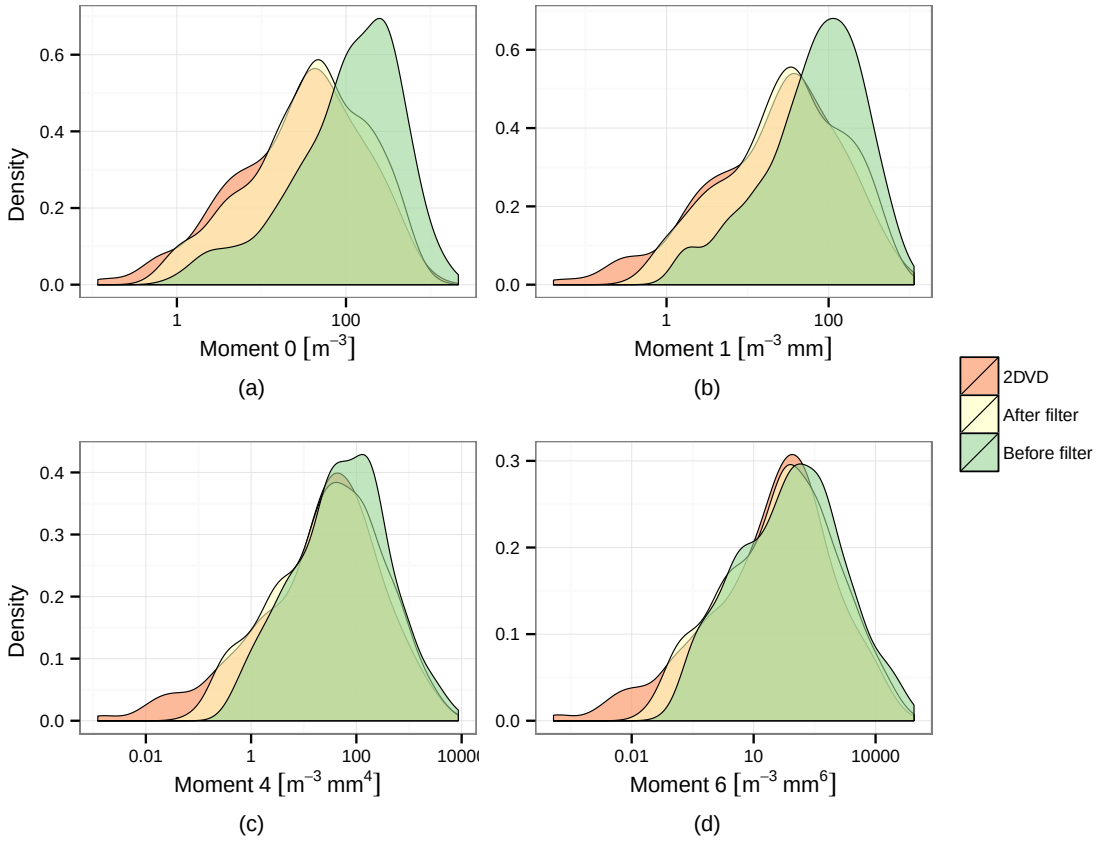


Figure 2.12 – The effect of the correction on DSD moments (a) 0 and (b) 1, (c) 4, and (d) 6, showing the densities of the Parsivel-derived DSD moments before and after the correction was applied, and the 2DVD moments, for HyMeX SOP2013 event time steps at Pradel Grainage.

2.5.2 Effect on rain rates

Having confirmed that the correction shifts the densities of the DSD moments towards those of the 2DVD, we used independent instruments – collocated tipping-bucket rain gauges – to test the effect of the correction on the rain rates produced by Parsivel DSDs. Two of the rain gauges provided measurements that we considered to be suspicious. The station at Mirabel-Pradel-Ferme-2, which is physically closest to our Parsivels Pradel 1 and Pradel 2, produced a marked overestimate of the rain amounts compared to those Parsivels, the 2DVD, and the rain gauge at Mirabel-Pradel-Ferme-1. For this reason we used Mirabel-Pradel-Ferme-1 as the reference gauge at this location. Mirabel-Pradel-Ferme-1 was located approximately 12 m away from Mirabel-Pradel-Ferme-2. Similarly, the rain gauge at Lavilledieu-Ecole-2 was physically closest to our Parsivel at Lavilledieu but, for a period of 1.5 h on 18 September 2012, this rain gauge produced rain rates that were markedly smaller than the rain rates produced by our Parsivel and the nearby rain gauge Lavilledieu-Ecole-1. This gauge, which was approximately 12 m away, provided measurements that more closely matched the Parsivel during this time. We thus used Lavilledieu-Ecole-1 as the reference rain gauge for this station.

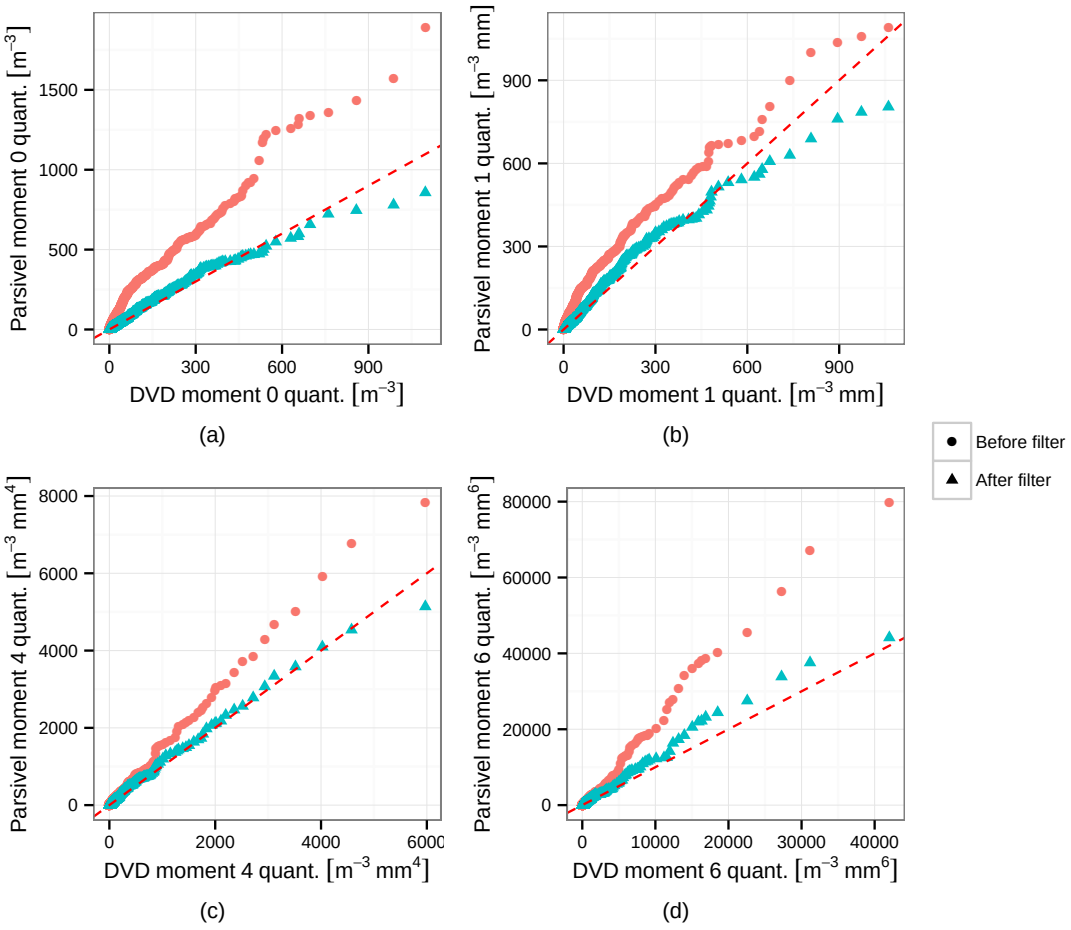


Figure 2.13 – Quantile-to-quantile plots showing the effect of the correction on Parsivel DSD moments (a) 0, (b) 1, (c) 4, and (d) 6, by comparing to the 2DVD moments, for HyMeX SOP2013 event time steps at Pradel Grainage.

We compiled performance statistics for each of the first-generation Parsivel stations, before and after the correction was applied, for a five-minute time resolution. As an example, Figure 2.14 shows a scatter plot of rain rates compared to a collocated rain gauge for Pradel 1, the Parsivel that was closest to the 2DVD and deployed in both 2013 and 2012, for five-minute time resolution across both campaigns. The statistics for this station show that the correction produced a clear improvement in the rain rate; the relative bias was reduced by 13%, the mean ratio and regression slope were both closer to 1 and the RMSE was reduced.

Given that the correction was trained only on SOP2013 data, it makes sense to look at the results from SOP2012 and SOP2013 separately as well as together. For SOP2012 only, the performance effects per statistic are shown in Table 2.7. For SOP2013 only, the performance effects are shown in Table 2.8. The differences shown are between the performance statistics after the corrections (velocity and concentration) had both been made, minus the statistics when no correction had been made. The before and after sets were both screened to remove implausible measurements. For SOP2012, the correction improved the RMSE, bias, and relative

Parsivel	Pluvio	ΔRMSE	Δ|bias|	Δ|RB|	Δr^2	N
Mirabel	Mirabel-Mairie	0.712	0.574	1.384	0.003	271
Lussas	Lussas-Salle-Polyvalente	0.365	0.472	10.447	−0.002	290
St-Germain	Saint-Germain-Ecole	−0.114	−0.239	−6.898	−0.010	655
Lavilledieu	Lavilledieu-Ecole-2	0.320	0.382	5.576	−0.005	638
Les Blaches	Mirabel-Les-Blaches	−0.213	−0.474	−11.221	−0.006	299
Pradel 1	Mirabel-Pradel-Ferme-1	−0.162	−0.215	−2.319	−0.010	301
Pradel 2	Mirabel-Pradel-Ferme-1	−0.790	−0.556	−12.484	−0.007	326

Table 2.7 – Performance effects of the proposed correction on Parsivel data, and stations on which comparisons were performed, for SOP2012 only at five-minute time resolution. N is the number of time steps on which comparison was possible (high quality, liquid precipitation only). ΔRMSE and Δ|bias| are in mm h^{-1}, while Δ|RB| is in percentage points.

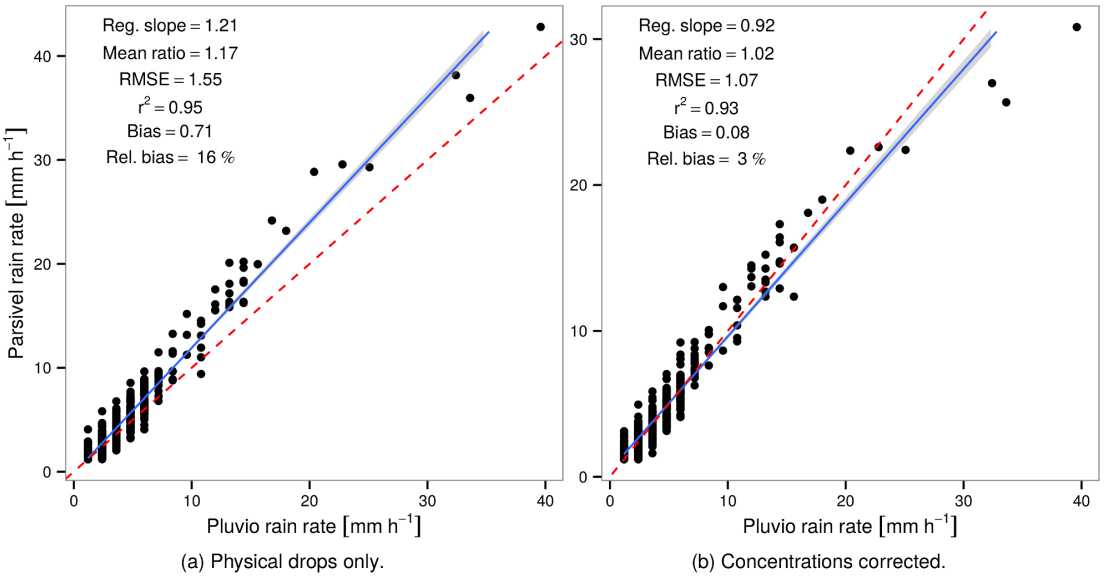


Figure 2.14 – Scatter plots showing the effect of the proposed correction, for the combined SOPs, with liquid precipitation only and rain rates over 1.2 mm h^{-1}, for Pradel 1, the closest station to the 2DVD that was present in both 2012 and 2013.

bias at four of the seven stations. At Mirabel the relative bias performance was hardly affected by the DSD correction. At the remaining two stations (Lussas and Lavilledieu) the relative bias was degraded, leaving the final relative bias at these stations as −11 and −9% respectively; both these relative biases are close to the instrumental variability we observed in Section 2.3.4. Recall that the 2DVD slightly underestimated the rain rate with respect to collocated gauges. For SOP2013, the RMSE and bias were improved at six of eight stations, relative bias was improved at seven of eight stations, and r^2 was hardly changed. The remaining station (Mirabel) showed a degradation of relative bias to an after-correction relative bias of −21%. At Mirabel, the Parsivel was placed on the edge of a retaining wall, which may have introduced turbulence and affected the Parsivel measurements.

For the combined SOPs data set, the Parsivel performance statistics before any correction are

Parsivel	Pluvio	ΔRMSE	Δ|bias|	Δ|RB|	Δr^2	N
Mirabel	Mirabel-Mairie	2.075	1.497	9.078	0.000	133
Lussas	Lussas-Salle-Polyvalente	−0.881	−0.724	−12.553	−0.006	374
St-Germain	Saint-Germain-Ecole	0.550	0.105	−3.676	−0.017	204
Lavilledieu	Lavilledieu-Ecole-2	−0.599	−0.629	−13.442	−0.006	387
Pradel Grainage	Pradel Grainage	−0.948	−0.783	−15.543	−0.009	376
Les Blaches	Mirabel-Les-Blaches	−0.319	−0.216	−3.133	−0.006	194
Pradel 1	Mirabel-Pradel-Ferme-1	−0.895	−0.725	−16.393	−0.004	218
Pradel 2	Mirabel-Pradel-Ferme-1	−0.956	−0.724	−17.183	−0.004	166

Table 2.8 – Performance effects of the proposed correction on Parsivel data, and stations on which comparisons were performed, for SOP2013 only at five-minute time resolution. N is the number of time steps on which comparison was possible (high quality, liquid precipitation only). ΔRMSE and Δ|bias| are in mm h^{-1}, while Δ|RB| is in percentage points.

Parsivel station	Rain gauge	ΔRMSE	Δ|bias|	Δ|RB|	Δr^2	N
Mirabel	Mirabel-Mairie	1.271	0.884	5.540	0.003	404
Lussas	Lussas-Salle-Polyvalente	−0.411	−0.389	−11.578	−0.005	664
St-Germain	Saint-Germain-Ecole	0.034	−0.159	−6.758	−0.024	859
Lavilledieu	Lavilledieu-Ecole-2	−0.135	0.208	2.834	−0.001	1025
Pradel Grainage	Pradel Grainage	−0.948	−0.783	−15.543	−0.018	376
Les Blaches	Mirabel-Les-Blaches	−0.257	−0.370	−9.545	−0.011	493
Pradel 1	Mirabel-Pradel-Ferme-1	−0.474	−0.621	−12.590	−0.018	519
Pradel 2	Mirabel-Pradel-Ferme-1	−0.845	−0.614	−13.897	−0.012	492

Table 2.9 – Performance effects of the proposed correction on Parsivel data, and stations on which comparisons were performed, for combined SOPs at five-minute time resolution. N is the number of time steps on which comparison was possible. ΔRMSE and Δ|bias| are in mm h^{-1}, while Δ|RB| is in percentage points.

shown in Table 2.10, after both velocity and concentration corrections in Table 2.11, and the changes made to the performance by the DSD correction are shown in Table 2.9. Again, all data sets were screened for implausible particles. From these data we can see that RMSE, bias, and relative bias were improved at six of the eight stations. At two stations (Lavilledieu and Mirabel) there was a degradation of performance in terms of rain rate, by a maximum of about 5% in terms of relative bias. The relative bias at Lavilledieu after the correction was applied was −6%, which is within the instrumental error limits. Mirabel may have suffered from turbulence effects; its relative bias across the combined SOPs was already −15% before any correction was performed. Despite degradations in R bias that were limited to two disdrometers, this analysis of the influence of the correction on the combined SOPs data set confirms its overall benefit to the DSD recorded by Parsivel disdrometers, even at high temporal resolution.

2.5.3 Results at lower temporal resolution

To further test the effects of the correction on Parsivel DSD-derived rain rates compared to collocated rain gauges, and to test the applicability of the filter to different time resolutions, we performed the same analysis as in the previous section but for one-hour temporal resolution on the combined SOPs data set. The differences made by the correction to the DSD moments at

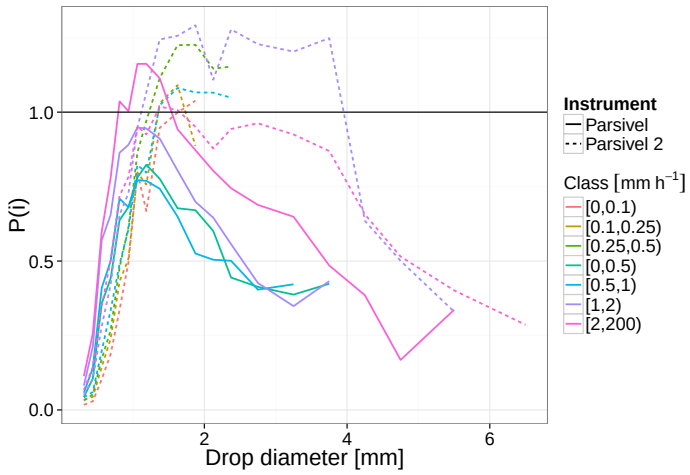


Figure 2.15 – Comparison between correction factors for different generations of Parsivel disdrometers in SOP2013.

Station	RMSE	Bias	r^2	RB [%]	Fit slope	Mean ratio
Mirabel	1.51	−0.64	0.98	−15.32	1.04	0.90
Lussas	1.55	0.51	0.94	11.68	1.18	1.12
St-Germain	1.08	0.35	0.96	7.56	1.12	1.08
Lavilledieu	1.25	0.21	0.96	3.59	1.14	1.05
Pradel Grainage	2.04	1.12	0.97	25.88	1.29	1.28
Les Blaches	1.35	0.50	0.95	10.25	1.16	1.13
Pradel 1	1.55	0.71	0.95	15.88	1.21	1.17
Pradel 2	1.69	0.88	0.97	20.26	1.29	1.22

Table 2.10 – Performance statistics for rain rate per Parsivel station for the combined SOPs at five-minute resolution, before the DSD correction is applied. RMSE and bias are in mm h^{-1}.

Station	RMSE	Bias	r^2	RB [%]	Fit slope	Mean ratio
Mirabel	2.78	−1.52	0.98	−20.86	0.75	0.78
Lussas	1.14	−0.12	0.94	0.10	0.90	0.97
St-Germain	1.11	−0.19	0.94	−0.80	0.85	0.96
Lavilledieu	1.11	−0.42	0.96	−6.42	0.88	0.92
Pradel Grainage	1.09	0.34	0.96	10.33	0.94	1.08
Les Blaches	1.09	−0.13	0.94	−0.70	0.86	0.97
Pradel 1	1.07	0.08	0.93	3.29	0.92	1.02
Pradel 2	0.85	0.26	0.96	6.36	0.97	1.06

Table 2.11 – Performance statistics for rain rate per Parsivel station for the combined SOPs at five-minute resolution, after the DSD correction is applied. RMSE and bias are in mm h^{-1}.

one-hour time resolution are shown in Table 2.12. At one-hour time resolution, the correction improved the bias, relative bias, and RMSE on all moment orders. r^2 between moment orders before and after the correction was improved for moments of order 1–3, and maintained at the same level for the other moments. The differences made to the rain rate to gauge comparisons at one-hour resolution are shown in Table 2.13. RMSE was improved at five of the eight stations, and bias at three of the eight. Relative bias was degraded in all but two

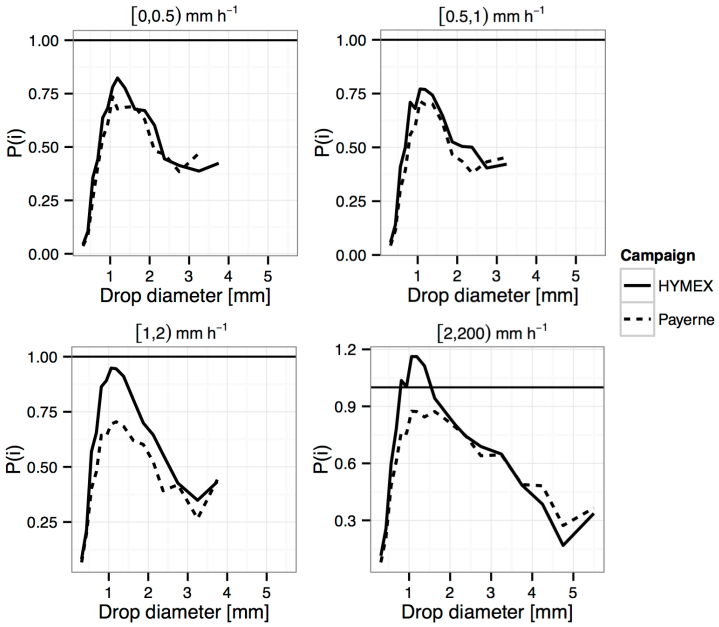


Figure 2.16 – Comparison between correction factors for different campaigns, Payerne and SOP2013. Both sets were trained on data at one-hour time resolution.

cases. This degradation is attributable to very small rain rates; indeed, when we selected time steps for which the rain rate was greater than or equal to 1.2 mm h^{-1}, the relative bias was improved at five of the eight stations, and only the station at Mirabel had an after-correction relative bias that was greater than the instrumental variability. For rain rates between 0.1 and 1.2 mm h^{-1}, the after-correction bias was negative at all stations and the per-station mean bias was −0.12 mm h^{-1}. This bias is similar to the bias of the 2DVD compared to the gauge for the same rain rates (−0.08 mm h^{-1}). We conclude that our correction procedures result in Parsivel measurements that better match those of the 2DVD, which itself underestimated rain rate for low rain rates when compared to a collocated gauge. We recommend that care is taken with the application of this correction to rain rates below 1.2 mm h^{-1}.

2.6 Application to Parsivel^2

We applied our method to second-generation Parsivels (Parsivel^2) that were also deployed in the HyMeX 2013 campaign. To train the correction for Parsivel^2 we followed the same method of comparing Parsivel records for the station at Pradel Grainage to the collocated 2DVD to train the correction factors per Parsivel-derived rain intensity class. The only difference was that, due to changes between the first- and second-generation Parsivels, the curves of $P(i)$ per Parsivel-derived intensity class showed different and more complex behaviour to those of the first-generation Parsivel. The classes we used were $[0,0.1)$, $[0.1,0.25)$, $[0.25,0.5)$, $[0.5,1)$, $[1,2)$ and $[2,200)$ mm h^{-1}.

Moment	Bias bef.	Bias aft.	RB bef. [%]	RB aft. [%]	RMSE bef.	RMSE aft.	r^2 bef.	r^2 aft.
0	82.85	−0.37	124.17	−1.19	147.85	43.96	0.45	0.77
1	38.51	0.52	76.19	0.06	73.19	37.89	0.67	0.82
2	24.19	−0.22	54.63	1.16	57.30	46.37	0.82	0.83
3	27.53	−3.08	51.47	6.94	89.83	79.95	0.84	0.82
4	59.27	−11.86	62.99	10.56	229.91	189.98	0.80	0.78
5	183.42	−37.16	79.02	20.14	751.41	579.54	0.72	0.71
6	684.13	−95.17	100.05	25.21	2836.62	2098.63	0.61	0.62
7	2892.55	−98.86	121.78	42.33	11915.72	8589.15	0.50	0.52

Table 2.12 – Time series statistics per moment, comparing Parsivel data (at Pradel Grainage) before (bef.) and after (aft.) the correction is applied, to 2DVD data, at one-hour resolution, for event times in SOP2013. The 2DVD is taken as the reference. Units of bias and RMSE are m^{-3} mm^p where p is the moment order.

Parsivel station	Rain gauge	ΔRMSE	Δ|bias|	Δ|RB|	Δr^2	N
Mirabel	Mirabel-Mairie	0.519	0.351	16.345	0.004	122
Lussas	Lussas-Salle-Polyvalente	−0.086	0.141	14.954	0.006	183
St-Germain	Saint-Germain-Ecole	0.019	0.064	15.162	−0.011	223
Lavilledieu	Lavilledieu-Ecole-2	0.149	0.291	23.399	−0.005	277
Pradel Grainage	Pradel Grainage	−0.110	−0.094	−13.390	−0.019	131
Les Blaches	Mirabel-Les-Blaches	−0.096	0.043	13.500	−0.003	117
Pradel 1	Mirabel-Pradel-Ferme-1	−0.138	−0.028	12.987	−0.011	130
Pradel 2	Mirabel-Pradel-Ferme-1	−0.223	−0.152	−4.520	−0.013	118

Table 2.13 – Performance effects of the proposed correction on Parsivel data, and stations on which comparisons were performed, at one-hour time resolution, for the combined SOPs. N is the number of time steps on which comparison was possible. ΔRMSE and Δ|bias| are in mm h^{-1}, while Δ|RB| is in percentage points.

Apart from the different Parsivel-derived rain intensity class definitions, the training process was identical to that shown in Section 2.4. The resulting correction factors are shown for the HyMeX Parsivel^2 data set in Table 2.14. A comparison of the corrections for first- and second-generation Parsivel is shown in Figure 2.15 and shows significant differences. Differences are expected, given that at a minimum the laser is different between the two instruments. Both filters were similar for drops up to about 1 mm in diameter, in that they both show the Parsivel overestimating drops in comparison to the 2DVD. Parsivel^2 is shown to underestimate the numbers of drops between 1.38 and 3.25 mm diameter. Drops larger than 3.5 mm were overestimated by both generations of Parsivel, but less so by Parsivel^2. After training the correction factors we applied them to Parsivel^2 data for all available stations. Due to small differences in clock times between the rain gauges and Parsivel^2 stations we used one-hour time resolution. We first compared the moments to the 2DVD moments for event time steps only; these results are shown in Table 2.16. The bias, relative bias, RMSE, and r^2 were improved for all moment orders.

We compared the rain rates after the correction of Parsivel^2 to those recorded by collocated rain gauges, for all available time steps. Due to timing errors with the Parsivel^2 network, we applied the correction to one-hour time steps. The results are shown in Table 2.15. Absolute and relative bias were improved at one Parsivel^2 station, but degraded at the two others. Again, there appears to be an effect of low rain rate on these performance statistics. When we counted

Class (i)	D_i [mm]	[0, 0.1)	[0.1, 0.25)	[0.25, 0.5)	[0.5, 1)	[1, 2)	[2, 200)
3	0.31	0.02	0.03	0.03	0.04	0.05	0.07
4	0.44	0.03	0.04	0.05	0.06	0.10	0.14
5	0.56	0.10	0.14	0.17	0.20	0.28	0.35
6	0.69	0.19	0.24	0.27	0.33	0.42	0.50
7	0.81	0.34	0.43	0.48	0.49	0.65	0.72
8	0.94	0.50	0.51	0.61	0.61	0.73	0.78
9	1.06	0.80	0.80	0.86	0.82	0.95	0.95
10	1.19	0.67	0.76	0.97	0.80	1.07	0.93
11	1.38	0.95	1.03	1.11	1.02	1.24	1.02
12	1.62	1.00	1.09	1.23	1.08	1.26	1.01
13	1.88	1.04	0.89	1.23	1.07	1.29	0.95
14	2.12			1.15	1.07	1.11	0.88
15	2.38			1.15	1.05	1.28	0.94
16	2.75					1.23	0.96
17	3.25					1.20	0.93
18	3.75					1.25	0.87
19	4.25					0.63	0.66
20	4.75					0.50	0.52
21	5.50					0.33	0.40
22	6.50						0.29

Table 2.14 – Calibrated Parsivel^2 correction factors for Parsivel-derived intensity classes for the HyMeX 2013 campaigns. Each row contains the class number, the centre equivolume diameter for the class (D_i), and the calibrated factors P(i) for each class of Parsivel-derived intensity. Intensity class boundaries are provided in millimetres per hour.

only time steps with rain rates at or above 1.2 mm h^{-1}, the worst degradation in relative bias dropped from 17 to 15%. There were many outliers in these data sets, and work is ongoing to further refine the correction on these Parsivel^2 data. Despite this degradation, the correction improved the DSD moments compared to the 2DVD. We hypothesise that training the Parsivel^2 correction factors using more data and therefore a lower time resolution, plus fixing potential clock issues in this data set, would improve the performance of the correction on Parsivel^2 data.

2.7 Application to another climatology

Finally, we applied our technique to data collected in a different region and climatology (see Section 2.2.2). In the Payerne 2014 campaign, a 2DVD and first-generation Parsivel were collocated. When compared to a collocated rain gauge at one-hour resolution, the 2DVD recorded rain rates with a performance very similar to that shown in the SOP2013 campaign (see Section 2.3.4), with a mean ratio of 0.84, an RMSE of 0.24 mm h^{-1}, an r^2 of 0.98, bias of -0.15 mm h^{-1}, and a relative bias of -19%.

We used the same technique as described in Section 2.4, using data with one-hour temporal resolution for training, and 10-minute resolution (the resolution of the reference rain gauge) for testing. We used the same Parsivel-derived intensity classes as for SOP2013. The resulting correction factors are shown in Table 2.17. The correction is compared to the correction found

Parsivel	Pluvio	ΔRMSE	Δ|bias|	Δ|RB|	Δr^2	N [ht]
Villeneuve de Berg	Villeneuve de Berg-1	−0.06	−0.18	−12.76	−0.01	129
Pradel Grainage	Pradel Grainage	0.30	0.33	8.21	−0.01	154
Villeneuve de Berg 2	Villeneuve de Berg-2	0.17	0.33	17.25	−0.00	132
Mont-Redon	Mirabel-Mont-Redon	-0.00	-0.28	-39.68	0.00	128
Pradel-Vignes	Mirabel-Pradel-Vignes	0.11	0.41	18.73	0.02	58

Table 2.15 – Performance effects of the proposed correction on Parsivel^2 network data, and stations on which comparisons were performed. N is the number of time steps on which comparison was possible. ΔRMSE and Δ|bias| are in mm h^{-1}, while Δ|RB| is a percentage. Mont-Redon and Pradel-Vignes were first-generation Parsivels to which the first-generation correction was applied, and are included for interest. They are listed in this table with the Parsivel^2 network because they were maintained as a group as part of the HPicoNet network.

Moment	Bias bef.	Bias aft.	RB bef. [%]	RB aft. [%]	RMSE bef.	RMSE aft.	r^2 bef.	r^2 aft.
0	187.29	−3.11	304.64	−0.35	279.82	44.28	0.40	0.79
1	86.54	−1.20	205.58	0.70	127.77	38.11	0.56	0.82
2	43.95	−1.48	107.63	0.81	73.55	45.19	0.77	0.85
3	25.43	−3.76	62.01	−0.02	75.55	70.85	0.87	0.88
4	23.61	−11.15	41.71	−0.07	146.80	144.76	0.89	0.90
5	67.11	−36.19	24.58	−1.17	445.86	380.29	0.89	0.91
6	341.81	−129.11	14.87	−3.99	1889.86	1250.60	0.88	0.92
7	1892.72	−502.11	8.02	−6.96	9514.94	4912.23	0.87	0.92

Table 2.16 – Time series statistics per moment, comparing Parsivel^2 data (at Pradel Grainage) before (bef.) and after (aft.) the correction is applied, to the 2DVD, at one-hour resolution, for event times. The 2DVD is taken as the reference. Units of bias and RMSE are m^{-3} mm^p, where p is the moment order.

in SOP2013, per Parsivel-derived rain rate class, in Figure 2.16. These plots show that the trained correction factors were very similar across climatologies, for classes of low rain rates (0–1 mm h^{-1}). For larger rain rates, there were some differences between the correction factors, but the general shape remained the same. It should be noted that in the Payerne 2014 data set, the precipitation sampled had much lower intensity than that found in SOP2013, and we hypothesise that the differences are due to these sampling effects. We tested the results using both one hour and 10-minute resolution data sets.

For consistency, we kept the threshold for the maximum-allowed spread in $P(i)$ at 0.7. In the Payerne 2014 data set, there appears to be greater sensitivity to this threshold, indicating that the sample size available for training was smaller than in SOP2013. For this reason we also tested the results when the correction trained using SOP2013 was applied to the Payerne 2014 data set. The effect on Parsivel rain rate performances are summarised in Table 2.18. The comparisons of moments with those of the 2DVD are included in the appendix, in Tables 2.19, 2.20, and 2.21.

The results show that when the Payerne data set was used to train the correction factors, there was a slight improvement in the Parsivel's rain rate estimation at 10-minute resolution. At one-hour resolution, the absolute bias was maintained but the relative bias was degraded. Correlations were maintained by the correction. When the HyMeX-trained correction was

Class (i)	D_i [mm]	[0,0.5)	[0.5,1)	[1,2)	[2,200)
3	0.31	0.04	0.04	0.07	0.08
4	0.44	0.08	0.12	0.20	0.21
5	0.56	0.26	0.31	0.40	0.48
6	0.69	0.41	0.39	0.48	0.61
7	0.81	0.54	0.56	0.64	0.77
8	0.94	0.61	0.60	0.64	0.76
9	1.06	0.73	0.71	0.69	0.87
10	1.19	0.68	0.70	0.70	0.87
11	1.38	0.69	0.70	0.68	0.84
12	1.62	0.69	0.61	0.62	0.87
13	1.88	0.63	0.47	0.60	0.83
14	2.12	0.48	0.43	0.52	0.79
15	2.38	0.47	0.38	0.39	0.75
16	2.75	0.38	0.43	0.42	0.64
17	3.25	0.47	0.45	0.27	0.64
18	3.75			0.44	0.49
19	4.25				0.48
20	4.75				0.27
21	5.50				0.37

Table 2.17 – Calibrated Parsivel correction factors for Parsivel-derived intensity classes for the Payerne campaign. Each row contains the class number, the centre equivolume diameter for the class (D_i), and the calibrated factors $P(i)$ for each class of Parsivel-derived intensity. Intensity class boundaries are provided in millimetres per hour.

Set	Res.	ΔRMSE	Δ|bias|	Δ|RB|	Δr^2	N
Payerne	10 min	−0.126	−0.061	−5.465	−0.000	435
Payerne	1 hour	−0.030	−0.015	1.927	−0.002	164
HyMeX SOP2013	10 min	−0.142	−0.152	−10.712	−0.003	435

Table 2.18 – Summary of performance effects of the Parsivel correction, for Payerne. Set indicates which data set was used to train the correction factors using one resolution, Res. is the temporal resolution to which the corrections were applied, and N is the number of time steps to which the correction was applied. ΔRMSE and Δ|bias| are in mm h^{-1}, while Δ|RB| (relative bias) is in percentage points.

applied to the Payerne data set, the performance was improved again. This indicates again that the sample size of the Payerne data set may have been smaller than required for a representative set of correction factors to be trained. Whether the Payerne-trained or HyMeX-trained correction factors were used, there was an improvement in the match between Parsivel and the 2DVD at Payerne for all moments. This suggests that the correction is robust and can be applied as such in different climatic regions.

2.8 Conclusions

We have developed a method to correct raindrop size distributions recorded by Parsivel disdrometers, using a two-dimensional video disdrometer as a reference instrument. The correction is made in two steps. First, raw Parsivel drop counts binned by velocity and diameter are shifted so that per-diameter-class mean velocities align with expected terminal velocities. The raw data can then be screened for particles that are unlikely to be raindrops,

Moment	Bias bef.	Bias aft.	RB bef. [%]	RB aft. [%]	RMSE bef.	RMSE aft.	r^2 bef.	r^2 aft.
0	149.20	2.21	253.21	4.19	219.53	24.34	0.67	0.94
1	74.29	2.29	173.22	5.50	107.87	17.41	0.79	0.96
2	45.37	2.20	125.64	7.78	66.06	16.11	0.90	0.97
3	36.28	2.25	103.45	10.36	56.31	20.04	0.96	0.97
4	41.13	2.74	93.07	14.23	80.52	35.23	0.98	0.97
5	70.68	5.59	95.20	19.66	203.91	91.82	0.97	0.95
6	181.65	23.49	103.58	26.42	729.63	345.07	0.95	0.90
7	632.93	137.31	120.77	35.22	3161.40	1705.47	0.91	0.77

Table 2.19 – Time series statistics per moment, comparing Parsivel data for Payerne 2014 before (bef.) and after (aft.) the correction is applied, to the 2DVD, at 10-minute resolution. The 2DVD is taken as the reference. Units of bias and RMSE are m^{-3} mm^p, where p is the moment order.

Moment	Bias bef.	Bias aft.	RB bef. [%]	RB aft. [%]	RMSE bef.	RMSE aft.	r^2 bef.	r^2 aft.
0	90.65	0.76	293.95	0.10	144.30	11.71	0.77	0.97
1	45.00	0.90	194.23	3.19	71.92	8.52	0.85	0.97
2	27.23	0.84	138.27	4.62	44.23	7.61	0.91	0.98
3	21.20	0.79	95.80	5.06	35.99	8.67	0.95	0.98
4	22.62	0.93	79.72	6.86	43.83	13.12	0.98	0.98
5	35.32	2.50	82.12	9.79	91.51	27.24	0.99	0.98
6	82.91	13.68	90.01	15.39	291.96	91.10	0.99	0.96
7	276.01	85.50	102.78	23.30	1173.24	478.91	0.97	0.89

Table 2.20 – Time series statistics per moment, comparing Parsivel data for Payerne 2014 before (bef.) and after (aft.) the correction is applied, to the 2DVD, at one-hour resolution. The 2DVD is taken as the reference. Units of bias and RMSE are m^{-3} mm^p, where p is the moment order.

and per-diameter-class volumetric drop concentrations can then be calculated. Second, these volumetric drop concentrations are adjusted by factors trained by reference to the 2DVD. The adjustment causes the drop concentrations to match those of the 2DVD in a statistical way.

The correction was applied to Parsivel and Parsivel^2 data from two autumn field campaigns in Ardèche, France. The results showed an improvement in the accuracy of moments of the DSD, when compared to the 2DVD as the reference instrument. Comparison of rain rate with collocated rain gauges showed changes that are acceptable, given the overall improvement in the accuracy of the DSD afforded by the correction. It must be noted that because the 2DVD is used as the reference instrument, the adjusted Parsivel drop size distributions will be, at best, as accurate as the measurements obtained by the 2DVD. If a better reference becomes available, exactly the same approach could be applied to correct the Parsivel (or indeed any other disdrometer) and to improve the agreement with the reference.

The correction was shown to be timescale-independent through application to both five-minute and one-hour Parsivel records. While in this case the correction was trained on data sets containing mainly light to intermediate rain rates (mostly below 20 mm h^{-1}), the method is flexible because it is conditioned on the Parsivel-derived rain intensity, and could be trained for higher rain rate classes as required. The method does not involve changing the hardware or software of the instrument, and it can be applied retrospectively to existing

Moment	Bias bef.	Bias aft.	RB bef. [%]	RB aft. [%]	RMSE bef.	RMSE aft.	r^2 bef.	r^2 aft.
0	149.20	21.67	253.21	30.02	219.53	41.87	0.67	0.94
1	74.29	17.60	173.22	31.27	107.87	34.62	0.79	0.95
2	45.37	15.68	125.64	32.75	66.06	33.03	0.90	0.96
3	36.28	15.41	103.45	33.43	56.31	36.59	0.96	0.95
4	41.13	16.82	93.07	35.20	80.52	50.80	0.98	0.95
5	70.68	21.75	95.20	38.35	203.91	107.89	0.97	0.93
6	181.65	41.73	103.58	44.27	729.63	374.43	0.95	0.88
7	632.93	149.21	120.77	52.35	3161.40	1782.65	0.91	0.75

Table 2.21 – Time series statistics per moment, comparing Parsivel data for Payerne 2014 before (bef.) and after (aft.) the correction is applied, to the 2DVD, at 10-minute resolution. In this case the SOP2013 correction is applied to the Payerne 2014 data set. The 2DVD is taken as the reference. Units of bias and RMSE are m^{-3} mm^p, where p is the moment order.

data sets. The correction offers the ability to improve the accuracy of the DSDs recorded by Parsivel disdrometers, which are instruments that are especially suitable for deployment in networks. High-quality DSD measurements from networks of Parsivel disdrometers can be used in valuable work on topics such as the small-scale variability of the drop size distribution. The correction has been demonstrated to work across two different climatologies in Europe.

3 Spatial interpolation of experimental raindrop size distribution spectra

This chapter is adapted from:

- T. H. Raupach and Alexis Berne. Spatial interpolation of experimental raindrop size distribution spectra. *Quarterly Journal of the Royal Meteorological Society*, 2016. doi: 10.1002/qj.2801.

This work was completed by T. Raupach under the supervision of A. Berne. Research, analyses, and writing are by T. Raupach. J. Grazioli provided information on the MXPol radar, including calculations of radar reflectivity per signal-to-noise threshold. For data acknowledgements, see Appendix A.

3.1 Introduction

The DSD is sometimes assumed to be uniform in space, for example across the measurement volume of a radar system or the pixel of a numerical weather model. In reality the DSD is highly variable, even at small scales (Jameson and Kostinski, 2001; Uijlenhoet et al., 2003b; Jaffrain and Berne, 2012b). To investigate the small-scale variability of the DSD and thus the error the assumption of a uniform DSD introduces, networks of disdrometers are often used to measure the DSD at point locations across the domain of interest (e.g. Tapiador et al., 2010; Jaffrain et al., 2011). These measurements sample the precipitation process at discrete points that may be sparsely and unevenly distributed. In this chapter we present a geostatistical interpolation and stochastic simulation method that can estimate or simulate the experimental DSD unmeasured locations, conditioned on nearby measurements.

Geostatistics provides spatial interpolation methods that can be used to estimate the value of a continuous process at unmeasured locations, given measurements at nearby points (Matheron, 1971). Geostatistics requires that the process being studied is second-order or at least intrinsically stationary (Chilès and Delfiner, 1999). Geostatistics methods have long been used to produce gridded precipitation fields, most often from networks of rain gauges (e.g. Creutin and Obled, 1982; Chua and Bras, 1982). The use of kriging with external drift (KED) is especially common. In KED, an external "drift" variable is used to find the expected values of the process, and kriging is used to estimate the residuals around these expected values (Chilès and Delfiner, 1999). For precipitation, altitude above sea level has been of particular importance as an external drift variable (e.g. Goovaerts, 2000; Tobin et al., 2011; Masson and Frei, 2014). Other studies have used radar data (Haberlandt, 2007; Velasco-Forero et al., 2008) and numerical weather model outputs (Tobin et al., 2011) to provide external drift information. These previous uses of kriging have focused on single variables such as the rainfall intensity. Here we present a new technique that, instead of focusing on bulk variables, can interpolate measured DSD spectra in a statistically robust way.

A problem with the use of geostatistics on precipitation fields is that rainfall is a non-stationary process (Barancourt et al., 1992; Schleiss et al., 2014a). Schleiss et al. (2014a) showed that at small spatial and temporal scales (their study used a 100×80 km^2 region and five-minute temporal resolution) the non-stationarity is caused largely by rainfall intermittency. Intermittency is the patchiness of the rain process, which leads to the existence of distinct wet and dry areas that have to be modelled to accurately represent the precipitation field (Kundu and Siddani, 2011; Schleiss et al., 2011). To deal with intermittency, Barancourt et al. (1992) suggested the use of a binary occurrence (rain/dry) process which can itself be sampled and interpolated to produce an expected occurrence map. Interpolation outputs for the rain field are then found and set to zero in the expected dry regions. This technique has been frequently employed (e.g. Syed et al., 2003; De Oliveira, 2004; Berrocal et al., 2008). Using this method, both the occurrence process and rainfall process must still be assumed to be either second-order or intrinsically stationary.

In previous studies, precipitation amounts or intensities have been interpolated using geostatistics. It is also possible to apply the same geostatistical techniques to other bulk rainfall variables individually. However, all integral variables are linked via the DSD and are related to each other in complex ways (Sempere-Torres et al., 1994). Analysis of multiple bulk variables in isolation from each other fails to take into account these relationships. One way to solve this problem would be to perform multivariate geostatistics, and to interpolate several bulk variables at once, specifically taking into account the relationships between the chosen variables. We go further, and present a more general approach in which we interpolate non-parametric measured DSD spectra. Our approach avoids the problem of having to choose which bulk variables to interpolate, avoids the choice of DSD model and any associated uncertainty, keeps all available information, and preserves the relationships between the bulk variables.

It has been observed in multiple studies that rain intensity tends to be lower towards the edges of rainy areas (Barancourt et al., 1992; Braud et al., 1994; Emmanuel et al., 2012). Schleiss et al. (2014a) called this phenomenon the "dry drift". They posited that by modelling this effect and subtracting it from the rain field, one is left with a detrended field that contains only the random fluctuations of the precipitation around the dry drift. This detrended field has an expectation of zero everywhere and is assumed to be second-order stationary. Their study showed that intermittency is a significant source of non-stationarity in precipitation. The dry drift concept has been extended to the DSD. Schleiss et al. (2014a) noted that the total drop concentration and the probability distribution of drop sizes were affected by different dry drifts, implying that bulk DSD variables all have different dry drifts (Schleiss et al., 2014a). Schleiss et al. (2012) investigated the drift on parameters of a DSD represented by a gamma distribution. In this paper we investigate the dry drift on measured drop concentrations per equivolume diameter class, from measured DSD spectra, thus avoiding the parameter-fitting step and associated uncertainty.

Our new method for spatial interpolation of DSD spectra uses geostatistical techniques to estimate the DSD at unobserved locations. The method is novel because it works on the (non-parametric) measured DSD spectra, and thus estimates the DSD without any assumption of its functional form. This way we maintain all possible information, and avoid any potential errors and assumptions involved in modelling or normalisation of the DSD. Further, it is useful that the interpolation results are estimations of measured DSDs, in that they can be used to investigate DSD model fit, spatial effects on DSD functional forms, or to study the spatial variability of particular drop size classes. In the new technique, the dry drifts of DSD drop concentrations, not model parameters, are taken into account to deal with non-stationarity and intermittency. Principal components analysis (PCA) is used to describe the detrended DSD in terms of uncorrelated components, for which the use of univariate geostatistics is appropriate. PCA also allows for noise removal if required. The use of the dry drift and PCA is required to transform the data into a form that can be assumed to honour the requirements of univariate geostatistics. The method ensures that all measured microstructure information is kept and relationships between the bulk variables are preserved.

As a demonstration of the utility of this technique, we show results from its application to HyMeX disdrometer network data. The accuracy of the technique was analysed through leave-one-out testing. The rest of this chapter is organised as follows: in Section 3.2 our new interpolation technique is presented, and its assumptions are carefully addressed. In Section 3.3 we show how the technique was tested through application to HyMeX campaign data. Results of the testing are shown in Section 3.4. In Section 3.5 we show how the technique can easily be modified to perform stochastic simulation of the non-parametric DSD. Conclusions are drawn in Section 3.6.

3.2 Spatial interpolation of experimental DSD spectra

In this section the new method for spatial interpolation of the DSD is explained. We describe it in broad terms before delving into the details. As input data, we have measured DSD spectra, with concentrations binned by drop equivolume diameter, recorded by a number of disdrometers arranged in a known spatial configuration. The output of the process will be a prediction of the DSD at any given point(s). The steps of the method are as follows:

1. Data for each equivolume diameter class are treated to remove the non-stationarity introduced by the effect of rainfall intermittency; the so-called "dry drift" introduced by Schleiss et al. (2014a). The resulting fields are assumed to be second-order stationary.

2. Principal component analysis is used to transform the stationary drop concentration fields into uncorrelated principal components, on which univariate geostatistics can be applied.

3. Variograms are found for each principal component to characterise its spatial structure.

4. Ordinary kriging, informed by the variograms, is performed in order to predict the values and estimation variance of each component at the prediction point(s).

5. The predicted components are back-transformed to recover the DSD spectrum at the prediction point(s).

We will now address each of these points in order, carefully examining the assumptions of each processing step. At this stage, the interpolation process will be described mathematically. Application to real data is shown in Section 3.3. Let us start by writing the volumetric drop concentration for a diameter class D_k at a spatial location z as $N(D_k, z)$. $N(D_k, z)$ [$\mathrm{mm}^{-1}\ \mathrm{m}^{-3}$] is then the concentration, at z, of drops with equivolume diameters in the half-closed interval $[D_k, D_k + \delta_k)$ [mm]. There is no requirement for the class widths δ_k to be constant.

3.2.1 Subtraction of the dry drift

The concept of the "dry drift", introduced by Schleiss et al. (2014a), is used to deal with intermittency and non-stationarity of the rain field. We apply the dry-drift concept presented in Schleiss et al. (2014a) to measured drop concentrations. Our aim is to characterise the relationship between the value of per-diameter-bin drop concentration at a point, and that point's distance to the closest dry region. Once this relationship is characterised, the dry drift can be subtracted, thus removing its influence on the non-stationarity of the drop concentration field. The dry drift gives the expected value of the drop concentration at a point, around which random fluctuations may occur. The result of the subtraction will thus be a detrended drop concentration field, containing only the fluctuations. This field has an expectation of zero everywhere.

Like rain rates, drop concentrations have highly skewed distributions with many more small than large values. A possible approach to deal with this is to use a log transformation. We use a modified log transformation, such that

$$\tilde{N}(D_k, z) = \ln\left[N(D_k, z) + 1\right]. \tag{3.1}$$

Note that one is added to the drop concentrations before taking the log; this means that zeros, an important part of the DSD spectrum, are kept. The (log) DSD concentration field is assumed to be a non-stationary random function in space and time, and we decompose the transformed DSD concentrations into:

1. The dry drift per diameter class, and
2. The detrended drop concentration fields.

External to the DSD is a rain occurrence field $I(z)$ that informs the dry drift by defining the dry regions (see Section 3.3.1 for an example in which a radar field is used). The rainfall occurrence is a binary field indicating whether rain is measured at a point z:

$$I(z) = \begin{cases} 1 & \text{if } R(z) > 0 \\ 0 & \text{if } R = 0, \end{cases} \tag{3.2}$$

where $R(z)$ [mm h^{-1}] is the instantaneous rain rate at point z. Let $d(z)$ be a function that returns the Euclidean distance between a point z and the nearest dry region defined by I (Schleiss et al., 2014a). The dry drift is then a function of $d(z)$, and describes the functional

relationship between the distance of a point from the nearest dry region and the expected log-transformed drop concentration at that point.

To decompose each drop concentration, we assume that the expected transformed drop concentration value at a point is a function of the point's distance to a dry region in space or time. Thus

$$\mathbb{E}[\tilde{N}(D_k, z)] = f_k[d(z)], \tag{3.3}$$

where f_k is the deterministic dry drift function for diameter class D_k. With f_k defined, we subtract the dry drift from the observed drop concentration field to obtain the detrended drop concentrations $N^\dagger$:

$$N^\dagger(D_k, z) = \tilde{N}(D_k, z) - f_k[d(z)]. \tag{3.4}$$

The detrended process $N^\dagger$ has several useful properties (Schleiss et al., 2014a). It represents random fluctuations of the drop concentration fields around the expected drop concentrations given by the dry drift. By definition from Equations 3.3 and 3.4 its expectation is zero everywhere. The log-transformation on the input data helps to stabilise its variances (see Section 3.3.1). The assumption made here is that the dry drift is the only reason that the drop concentration fields are non-stationary. Although there can be other contributions to their non-stationarity, such as orographic or coastal effects, Schleiss et al. (2014a) showed that intermittency is a major factor in the non-stationarity of rainfall intensity fields at high temporal resolutions. We assume that the resulting detrended fields are second-order stationary. Experimental determination of the functional form of the dry drift $f_k[d(z)]$ is addressed in Section 3.3.1.

3.2.2 Principal component analysis

We now have a set of detrended drop concentration fields $N^\dagger(D_k, z)$, with one field per diameter class D_k. The different drop concentration fields are not independent or uncorrelated. In order to transform them into a set of orthogonal components that are uncorrelated, principal component analysis (PCA) (e.g. Jolliffe, 2002) is used. Until now only drop concentrations at a single time have been considered. To perform PCA we consider all available observations, across all locations in the event or time period for which we want to interpolate the DSD. There is no particular number of samples or time steps required in order to perform PCA, because the components are always guaranteed to be uncorrelated. However, the geostatistics that follow require consistency in the PCA transform used, in that the same orthogonal components must

be used for all analysed time steps. Both PCA and geostatistics are performed on a per-event basis.

PCA is a widely used statistical method, so in this chapter our description of it will remain informal. For a more complete description, the reader is invited to refer to, for example, Jolliffe (2002). Let there be U drop diameter classes available across L spatial locations for T separate observation times, and let $N^{\dagger}(D_k, z, t)$ be the log-transformed and detrended volumetric drop concentration of the kth class at location z and time t. Let us construct a matrix of measurements $\mathbf{M}$ in which each row corresponds to one observation, while each column corresponds to one class of drop diameters:

$$\mathbf{M} = \begin{bmatrix} N^{\dagger}(D_1, z_1, 1) & \cdots & N^{\dagger}(D_U, z_1, 1) \\ \vdots & \ddots & \vdots \\ N^{\dagger}(D_1, z_L, 1) & \cdots & N^{\dagger}(D_U, z_L, 1) \\ \vdots & \ddots & \vdots \\ N^{\dagger}(D_1, z_L, T) & \cdots & N^{\dagger}(D_U, z_L, T) \end{bmatrix}. \tag{3.5}$$

To ensure that each diameter class is treated equally by the PCA algorithm, the class values are scaled and shifted so that the diameter class mean is zero and its standard deviation is one. This process is to normalise the different variabilities and different widths of the input classes. Each class is a column in $\mathbf{M}$, and we have for the kth class:

$$\tilde{\mathbf{M}}_k = \frac{\mathbf{M}_k - \overline{\mathbf{M}_k}}{\sigma(\mathbf{M}_k)}, \tag{3.6}$$

where $\overline{\mathbf{M}_k}$ represents the mean and $\sigma(\mathbf{M}_k)$ represents the standard deviation of $\mathbf{M}_k$. To understand PCA, view the matrix $\tilde{\mathbf{M}}$ as a collection of TL points in U-dimensional space. PCA reprojects the points in $\tilde{\mathbf{M}}$ into a new coordinate system, in which each successive dimension explains the largest possible variance in the data set. The first dimension, then, is the longest axis through the U-dimensional point cloud. The second is the longest that is orthogonal to the first, and so on until all variability in the points is explained. The result of the PCA is a $U \times S$ matrix $\mathbf{W}$, with $S \leq U$. The matrix $\mathbf{W}$ transforms the original data set into a $TL \times S$ components matrix $\mathbf{C}$, such that $\mathbf{C} = \tilde{\mathbf{M}}\mathbf{W}$. Each component has a corresponding contribution to the total variance, Q_s for the sth component.

$\mathbf{C}$ has one row per observation and one column per component. To determine how many components are returned, a threshold can be applied so that only components that explain a certain amount of the variance (those with Q_s above some threshold) are kept. Because the least important components are essentially noise, we could use this tolerance to perform

dimensionality reduction and remove noise to a chosen level. In this work, however, we keep all PCA components in order to perfectly reconstruct measured DSD spectra.

PCA outputs are orthogonal (Jolliffe, 2002). If the inputs happen to be multivariate normal, then the outputs will be both orthogonal and independent. In our case, our inputs are the detrended drop concentrations, which are the random fluctuations of the process around the dry drift. Although the concentration distributions are unimodal and almost symmetric, their tails are generally heavier than normal distribution tails. The components are therefore not strictly multivariate normal. As a result, we cannot assume that the PCA outputs are independent, but they are guaranteed to be orthogonal and uncorrelated.

PCA has two main advantages for this work. First, the S components in $\mathbf{C}$ are uncorrelated, so univariate geostatistics can be used on each one independently. Second, PCA offers precise control over how much noise is kept in the interpolation, through optional removal of less important components.

3.2.3 Variograms of components

Having translated the measured DSD spectra into assumed second-order stationary, uncorrelated variables, our task is now to use geostatistics to predict the component values at unmeasured locations. We view each PCA component as a sample of a stochastic process X_s, where the values in X_s are found in the sth column of the matrix $\mathbf{C}$. The variogram of a random process is used to characterise its spatial structure (Matheron, 1971). The variogram is a measure of the (dis)similarity of the process at points separated by a certain distance l, and is defined for an intrinsically stationary process Z as (Cressie, 1993)

$$\gamma_Z(l) = \frac{1}{2}\mathrm{Var}\left[Z(z+l) - Z(z)\right]. \tag{3.7}$$

We have a finite number of samples, so we must use the sample variogram, in which the expectation is approximated using the sample mean. The mean is easily affected by outliers, so by using the sample variogram we are assuming that the input distributions are close to symmetric. This is generally the case with our component distributions (see Section 3.3.2), but we use the Cressie robust variogram (Cressie, 1993) instead of the standard variogram to ensure that any outliers are well handled.

If there are N_l sampled pairs of points for a distance lag l, then the Cressie sample variogram

for each component is (Cressie, 1993)

$$2\tilde{\gamma}_s(l) = \frac{\left\{\frac{1}{|N_l|}\sum_{r=1}^{N_l}|X_s(z_r+l)-X_s(z_r)|^{\frac{1}{2}}\right\}^4}{\left(0.457+\frac{0.494}{|N_l|}\right)}. \tag{3.8}$$

For interpolation, we need to be able to know the variogram at any given distance lag. To allow this, a variogram model is fitted to each sample variogram (Chilès and Delfiner, 1999).

3.2.4 Kriging of components

Kriging estimates the value of a process at an unmeasured point. The new value is a weighted combination of the measured values; the task of the kriging process is to determine the optimal weights, informed by the variogram that defines the spatial structure of the process (Chilès and Delfiner, 1999). We use ordinary kriging, in which the mean of the studied process is assumed to be constant but unknown. Our processes are X_s, the principal components of the DSD.

Estimations are made for a single time step. The estimated value of the sth PCA component at an unmeasured location z_0 at time t is

$$X_s^*(z_0,t) = \sum_{r=1}^{L} w_r X_s(z_r,t), \tag{3.9}$$

where the weights w are determined by the ordinary kriging algorithm and there are L observation locations at time t. We use the $*$ symbol to indicate estimated, rather than observed, values. The kriging process uses a Lagrange multiplier to find the weights, under the constraints that the bias of X_s^* is zero and the estimation variance is minimised. Kriging is an exact interpolator, meaning that the estimation at a measured location is the measurement itself, and for Gaussian random functions kriging is an optimal solution (Chilès and Delfiner, 1999).

3.2.5 Back-transformation of the components

Back-transformation of the estimated component values to an estimated DSD spectrum is simply a matter of reversing the transformations that were applied to the measured DSDs, in reverse order. Starting with the kriged value for the sth component, X_s^*, we first back-transform the estimated components into detrended DSD concentrations, by placing the

estimations for time t into a vector $\mathbf{C}^*(z_0,t)$:

$$\mathbf{C}^*(z_0,t) = \left[X_1^*(z_0,t),\ldots,X_S^*(z_0,t)\right], \tag{3.10}$$

which can be rotated to return to DSD diameter classes:

$$\tilde{\mathbf{M}}^*(z_0,t) = \mathbf{C}^*(z_0,t)\mathbf{W}^T. \tag{3.11}$$

To obtain the estimated detrended DSD, we must rescale and recentre $\tilde{\mathbf{M}}^*$:

$$\mathbf{M}^*(z_0,t)_k = \tilde{\mathbf{M}}^*(z_0,t)_k \cdot \sigma(\mathbf{M}_k) + \overline{\mathbf{M}_k}. \tag{3.12}$$

The components of $\mathbf{M}^*(z_0,t)$ are the estimated detrended per-diameter-class drop concentrations for location z_0 at time t, such that

$$\mathbf{M}^*(z_0,t) = \left[N^{\dagger *}(D_1,z_0,t),\ldots,N^{\dagger *}(D_U,z_0,t)\right]. \tag{3.13}$$

It remains to re-add the dry drift using

$$\tilde{N}^*(D_k,z_0,t) = N^{\dagger *}(D_k,z_0,t) + f_k[d(z_0)], \tag{3.14}$$

and finally to back-transform the original logarithmic transformation:

$$N^*(D_k,z_0,t) = \exp[\tilde{N}^*(D_k,z_0,t)] - 1. \tag{3.15}$$

$N^*(D_k,z_0,t)$, $k \in \{1,U\}$ is the desired result: the estimated drop size distribution spectrum, across K classes, for time t at a location z_0.

3.2.6 Estimation uncertainty

One of the advantages of using geostatistics for interpolation is that as well as an estimation of the required value at a given point, it provides an estimate of the associated error, called the estimation variance. Thus for each kriged value for the sth component, X_s^*, the estimation variance V_s is provided. The estimated value and estimation variance can be seen as respectively the mean and variance of a normal distribution that represents the distribution of possible values. Because each estimate of per-class drop concentration is a weighted linear combination of the uncorrelated PCA components, the estimation variances can easily be analytically back-transformed into variances for each individual log-transformed drop concentration class.

However, analytically computing the estimation variance on bulk variables would involve mixing the variances of the non-independent drop classes. Another possible way to quantify the estimation uncertainty on bulk variables is a probabilistic approach. In this case, realisations of possible values can be drawn from the estimation distributions for each PCA component. Back-transformation of these realisations provides sets of possible DSDs from which sets of bulk variables can be computed. Statistics can then be performed on these sets of bulk variable realisations, to find, for example, their inter-quartile ranges. Through this process, the kriging estimation variance of each component can be used to derive a measure of the uncertainty in the estimated bulk variable outputs. A note of caution: kriging estimation variances must be handled with care when dealing with non-Gaussian random functions, and as we mentioned in Section 3.2.2, our component distributions are close to but not strictly Gaussian. We consider that the computed estimation variances, if not exact, provide a reasonable order of magnitude of the true error of each component.

3.3 Application to HyMeX data

We tested the DSD interpolation technique by applying it to Parsivel (both first-generation and Parsivel^2) data collected by the HyMeX instrument network introduced in Section 2.2. The Parsivel data were all corrected using the technique described in the previous chapter. The resulting processed data take the form of drop concentrations [$\mathrm{mm}^{-1}\ \mathrm{m}^{-3}$] in 32 unevenly-sized classes of equivolume drop diameter per time step. The radar (MXPol, see Schneebeli et al., 2013 for the full instrument description) was a dual-polarimetric X-Band Doppler weather radar, scanning at preset elevation angles above the disdrometer network about every five minutes.

For these analyses, we resampled the Parsivel data into lower temporal resolutions of one minute, five minutes, and 10 minutes. Resampling was performed for each Parsivel station, by finding the mean DSD over each new time period. Parsivels record a precipitation type flag which indicates whether the measured precipitation was solid or liquid (see Löffler-Mang and Joss, 2000, Fig. 9), and a quality-control flag that indicates if there were any problems with the laser beam. These flags were used to select time steps that could be assumed to represent

good-quality measurements of liquid precipitation. For each time resolution, we subset the resampled time steps for each station to those in which none of the input high-resolution measurements registered solid precipitation, no problems with the laser were detected, the resulting derived rain rate was greater than 0.1 mm h^{-1}, and a radar scan was available.

There were three sets of collocated stations in the network, at Pradel, Pradel Grainage, and Villeneuve-de-Berg. We made the assumption that collocated instruments sampled the same precipitation volume over integration time. Any differences between the stations were therefore assumed to be due to measurement error and not due to spatial effects. Although the collocated Parsivels performed similarly, there were of course measurement differences. To avoid such measurement errors being falsely classified as spatial effects we used only the best-performing station from each group, judged by comparison with collocated rain gauges at five-minute resolution. Thus of the collocated stations we used Pradel 1, Pradel Grainage (first-generation), and Villeneuve-de-Berg as interpolation input. Pradel 2 data was included until variogram models were fitted, so it could be used to determine nugget values (see Section 3.3.3), after which it was not included as interpolation input.

Events were defined using disdrometer data at one-minute temporal resolution, while the rest of the testing and analyses were performed using five-minute temporal resolution. An event was defined as a period in which no more than one hour of completely dry time was observed, and for which sample variograms could be found (at five-minute resolution; see section 3.3.3). Defined events are shown in Table 3.1. Using these criteria, 15 events were defined, all with at least five disdrometer stations reporting. The shortest event was 1.8 hours long, while the longest lasted for 23.7 hours. Each event was analysed separately. Per-station quantities of rainfall recorded within these events are shown in Table 3.2. In the following sections, examples are shown from event 13. This event lasted for 5.2 hours, during which the maximum per-station five-minute amount was 2.1 mm and the maximum five-minute rain rate was 25.1 mm h^{-1}. We chose event 13 because it was a good example case containing moderate rain rate and frequent intermittency. Note that this event is used purely for illustrative purposes and results from all events are shown in Section 3.4.

3.3.1 The dry drifts of drop concentrations

The calculation of the dry drift for drop concentrations depends upon an occurrence field (Equation 3.2). We used radar data to determine this occurrence process for our field area. In the specific case of our radar, a (horizontal polarisation) signal to noise ratio (SNR) of greater than or equal to 5 dB meant that the radar was measuring a signal of sufficiently good quality to use. To ensure this signal was indeed rainfall, we set a threshold for radar reflectivity at 10

	Start	H	Peak R	Stns	1 min	5 min	10 min
1	2012-09-18 19:56:00	6.1	30.9	9	58	52	27
2	2012-09-25 05:55:00	8.6	65.1	9	56	51	28
3	2012-10-25 22:14:00	15.5	15.2	9	129	130	66
4	2012-10-26 15:09:00	11.8	40.8	9	118	115	67
5	2012-10-31 10:23:00	6.8	6.6	10	53	55	28
6	2012-11-03 16:30:00	3.9	15.9	9	32	32	20
7	2012-11-09 19:09:00	22.4	30.4	6	152	150	79
8	2012-11-26 04:54:00	23.7	24.7	5	195	190	101
9	2013-09-28 17:05:00	1.8	39.1	10		13	
10	2013-10-04 15:33:00	4.8	33.4	9	30	29	
11	2013-10-12 15:39:00	7.5	6.6	10	65	64	37
12	2013-10-15 03:44:00	7.5	25.0	10	68	68	39
13	2013-10-20 02:32:00	5.2	56.8	10	43	43	26
14	2013-10-27 03:22:00	5.5	60.1	10	46	46	28
15	2013-11-02 19:11:00	4.3	12.1	10	28	29	17

Table 3.1 – HyMeX event times, showing the start time (UTC, end of first time step), the event length in hours (H), the peak one-minute rain rate [mm h^{-1}], the number of stations that measured DSDs during the event, and the number of within-event time steps suitable for geostatistical analysis per time resolution. For each time resolution, events were included if it was possible to estimate variograms with at least 30 pairs of observations per distance lag class (Note: this table was first published in Raupach and Berne (2016) (Chapter 4 of this thesis). It contains the same event definitions as the table of events originally published with the rest of this chapter. It also contains the number of available time steps, and lists one-minute instead of five-minute peak R.

dBZ. We defined the occurrence process as

$$I_R(z) = \begin{cases} 1 & \text{if SNR}(z) > 5 \text{ and } Z_H > 10 \\ 0 & \text{otherwise.} \end{cases} \tag{3.16}$$

Calculations assuming typical noise levels showed that over the region of our network, the 5 dB cut-off occurred at a radar reflectivity of between -11 and +12 dBZ. The dry/rain threshold could therefore appear at values of between 10 dBZ and 12 dBZ, which, using the Z-R relationship of Marshall et al. (1955), translate to an approximate rain rate of 0.2 mm h^{-1}. Note that it was possible for the SNR and Z_H values in the radar fields to be undefined, and there were areas that were not observed by the radar. The distance between a point and the closest dry region was calculated using the Euclidean distance, such that

$$d(z) = \min_{y \in \Omega} \|z - y\|, \tag{3.17}$$

where $\Omega = \{y \,|\, I_R(y) = 0\}$ is the set of dry points defined by the occurrence map. Occurrence maps were calculated for each considered five-minute time step, using plan position indicator

Station	Hours [H]	Amount [mm]
Mirabel	93	240
Lussas	75	220
Lavilledieu	92	256
Les Blaches	90	249
St-Germain	90	252
Pradel 1	92	254
Pradel 2	91	253
Montbrun	19	52
Pradel Grainage (first-generation)	24	81
Saint Etienne de Fontbellon	26	82
Villeneuve de Berg	55	129
Pradel-Vignes	3	7
Mont-Redon	47	112

Table 3.2 – Station information, showing the station name, the number of hours, and the total amount the station recorded during event times (calculated using one-minute resolution data and including all within-event time steps). This information was originally published in Raupach and Berne (2016) (Chapter 4 of this thesis).

(PPI) radar scans, in which the elevation is fixed while the radar changes azimuth. We used the lowest available elevation angle that did not produce significant clutter and beam-blocking, which was four degrees above horizontal. The range resolution was 75 m and the inner angular resolution (antenna beam-width) was 1.45°, with data collected with an angular spacing of about 1°. The occurrence maps covered a region extending 10 km from all sides of the disdrometer network. For time steps in which more than one radar scan was present, only points at which no rain was observed in any radar scan were taken to be dry. The radar value for a grid point was taken to be the mean value of the radar observations that, when their coverage areas were projected to the horizontal plane, overlapped the grid point. The occurrence maps were found at grids of 25 m resolution that captured fine detail, then aggregated to 100 m resolution. Points which were outside the radar coverage area, or for which SNR or Z_H were undefined, were considered to have an undefined occurrence value. When determining the dry distance $d(z)$ for a given point, if the occurrence map border or a point with undefined occurrence was closer than the nearest dry region, then the dry distance was also undefined. Points with undefined dry distance could not be used to train dry drift models.

Similar to results of previous studies on rain rate and DSD model parameters (Schleiss et al., 2012, 2014a), our results show that there is a relationship between drop concentrations and the dry distance. Figure 3.1 shows drop concentration versus distance to dry region for within-event five-minute time steps, for the first equivolume drop diameter class that contained drops. The distributions of drop concentrations are shown both before and after the log transformation (Equation 3.1). Schleiss et al. (2014a) showed that an ordinary log transformation results in stabilised distributions of values per dry distance class. After the transformation, the variability was more uniform across dry distance classes. There was a similar stabilising effect in our results, even though the resulting drop concentration distributions were not normal due to the fact that we keep zeros in the DSDs (see Section 3.2.1).

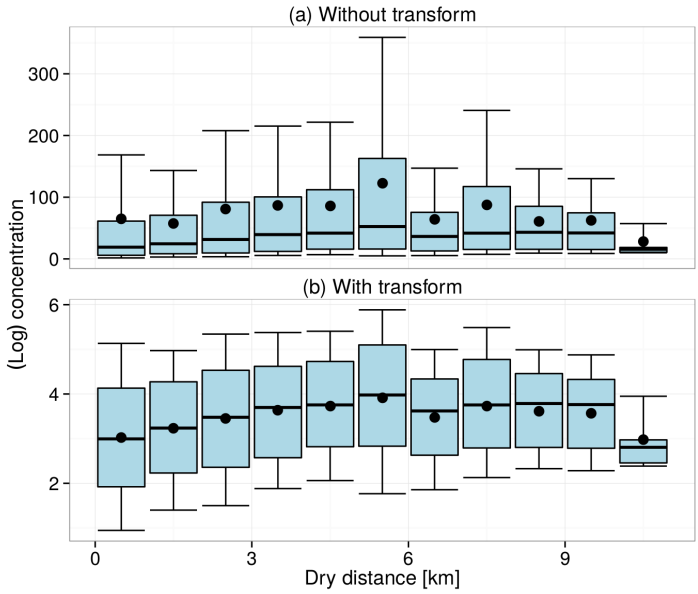


Figure 3.1 – Experimental dry drifts across all events, for disdrometer class 3 ([0.2495, 0.3745) mm). Shown are drop concentration distributions (a) with no log transformation, and (b) with log transformation. Distance class width is 1 km. Boxes show class inter-quartile range, whiskers show 10th to 90th percentile range, bold horizontal lines indicate medians, means are indicated by dots. The log transformation stabilised the variances of the drop concentrations by distance class.

The effect of the dry drift is most apparent for small drops, and becomes less important for large drops, which occur less frequently. Figure 3.2 shows the dry drift over all available time steps for the 17th Parsivel drop size class (3 to 3.5 mm). The drop concentrations were low both close to and far from a dry region. Close to dry regions (under about 3 km), these large drops were so rare that the inter-quartile ranges of concentrations were zero. The log transformation had only a small stabilising effect.

A functional form was fitted to the dry drift for each drop concentration class. The shape of the functional form was well fitted by one of two models. The first model was a spherical plus nugget model, like the type that is often fitted to variograms. This model has three parameters:

1. The range a [m] represents the distance after which the drop concentration is no longer expected to change.

2. The nugget c_0 [–] represents the expected transformed drop concentration at the boundary of the wet/dry regions.

3. The partial sill c_d [–] defines the expected transformed drop concentration ($c_0 + c_d$) at the range distance (when $d = a$).

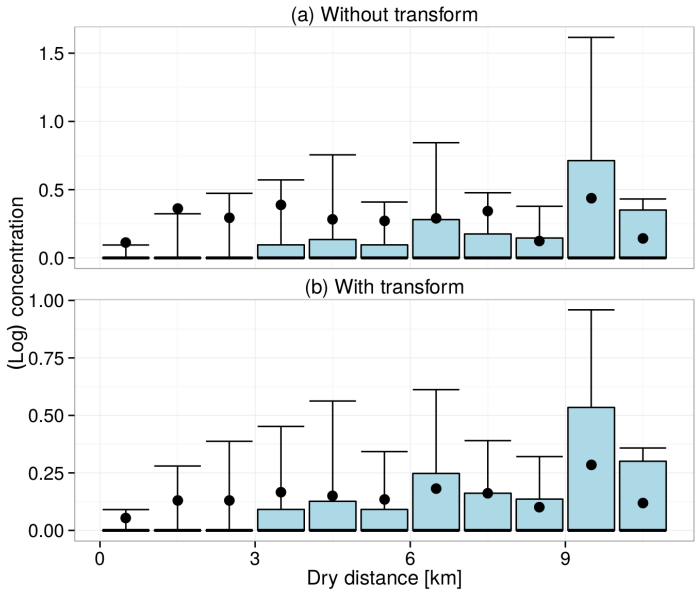


Figure 3.2 – Experimental dry drifts across all events, for disdrometer class 17 ([3, 3.5) mm). Shown are drop concentration distributions (a) with no log transformation, and (b) with log transformation. Distance class width is 1 km. Boxes show class inter-quartile range, whiskers show 10th to 90th percentile range, bold horizontal lines indicate medians, means are indicated by dots.

This spherical model is defined as (Chilès and Delfiner, 1999)

$$f_S(d) = \begin{cases} c_0 + c_d \left[\frac{3d}{2a} - \frac{1}{2} \left(\frac{d}{a} \right)^3 \right] & \text{if } d \leq a \\ c_0 + c_d & \text{if } d > a. \end{cases} \tag{3.18}$$

The second model allows for the drop concentration to remain zero for a distance from the dry region, before it slowly increases. This is useful for concentrations of larger drop sizes, which tend to appear only at some distance from the dry region. The second model is a modified Gaussian model, with three parameters:

1. A scale parameter a [m]; the drop concentration is not expected to change after the pseudorange, when $d \geq \sim 1.73a + d_0$ (Chilès and Delfiner, 1999).

2. The sill c_d [–] represents the maximum expected transformed drop concentration.

3. The zero distance d_0 [–] represents the distance below which the expected transformed drop concentration is zero.

This Gaussian model is defined as

$$f_G(d) = \begin{cases} c_d \left[1 - e^{-\frac{(d-d_0)^2}{a^2}} \right] & \text{if } d \geq d_0 \\ 0 & \text{if } d < d_0. \end{cases} \tag{3.19}$$

For each drop diameter class, the model f_S or f_G that best fitted all the data points (using non-linear weighted least squares) was used. Most drop concentration classes were fitted best by the spherical plus nugget model f_S. The Gaussian model f_G was the best fit for some classes for larger drops. Large drops are rare and therefore it makes sense that when rainfall is lighter, as it is expected to be near dry regions, the expected drop concentration for large drops may be zero. These spherical and Gaussian models are used for the dry drift functions simply because they fit the data, not because of their mathematical resemblance to variogram model functions. In fact, any model, however simple, that satisfactorily fits the dry drift data can be used.

It is expected that different rainfall events will have different dry drift properties. For each event, we found sample values by dry distance and fitted functional forms. Means and standard deviations of model parameters over the 15 events are shown in Table 3.3. It is clear from the large standard deviations for the range parameter a that it is advisable to fit the dry drift models to data per event, and that for events in which an abrupt change occurs (for example a convective front to stratiform pattern) it would be better to split the event into the two phases and fit dry drift models to each phase separately. Figure 3.3 shows examples of the dry drift relationships between distance to dry region and drop concentration, for two drop diameter classes in event 13, and Figure 3.4 shows the fitted models for these same classes, including the use of a Gaussian model for the larger drop class.

During interpolation, the dry distance for every grid point was calculated, and the dry drift models were used to determine the expected value of each PCA component at that point. Since the dry drift models have a range after which the expected value is constant, this range could be used for some points for which the dry distance was undefined. Specifically, for points with undefined dry distance that were further from a border or undefined occurrence value than the largest dry drift model (pseudo)range, that maximum range was used as the point's dry distance during interpolation. No interpolation was performed at points for which the occurrence was undefined. For each drop diameter class, the concentrations $\tilde{N}(D_k, z)$ were adjusted to remove the effect of the dry drift (Equation 3.4), leaving the detrended process $N^{\dagger}(D_k, z)$.

Class (k)	D_k	δ_k	Model	Events	$\overline{c_d}$	$\sigma(c_d)$	$\overline{c_0}$	$\sigma(c_0)$	$\overline{d_0}$	$\sigma(d_0)$	$\overline{a}$	$\sigma(a)$
3	0.25	0.12	Spherical	14	1.21	0.69	2.54	0.59			5.96	6.23
4	0.37	0.12	Spherical	14	1.13	0.77	3.42	0.63			5.47	5.28
5	0.50	0.13	Spherical	14	0.89	0.70	4.48	0.73			5.88	6.78
6	0.62	0.12	Spherical	14	1.01	0.69	4.48	0.63			4.68	5.15
7	0.75	0.12	Spherical	15	1.12	0.56	4.48	0.48			4.50	4.58
8	0.87	0.12	Spherical	15	1.51	0.61	4.00	0.31			4.25	4.66
9	1.00	0.12	Spherical	15	1.93	0.69	3.34	0.41			4.65	4.08
10	1.12	0.13	Spherical	15	2.24	0.66	2.60	0.47			4.32	3.33
11	1.25	0.25	Spherical	15	2.52	0.71	1.77	0.52			4.53	3.53
12	1.50	0.25	Spherical	15	2.60	0.83	0.93	0.47			5.52	4.61
13	1.75	0.25	Spherical	15	2.31	0.77	0.48	0.37			6.60	5.07
14	2.00	0.25	Spherical	15	1.94	0.78	0.26	0.27			7.93	5.54
15	2.25	0.25	Spherical	12	1.45	0.60	0.17	0.20			6.87	5.47
16	2.50	0.50	Spherical	12	0.97	0.54	0.08	0.11			8.04	7.06
17	3.00	0.50	Spherical	10	0.36	0.27	0.03	0.07			7.80	7.21
18	3.50	0.50	Spherical	9	0.28	0.50	0.02	0.04			8.25	7.57
19	4.00	0.50	Spherical	7	0.06	0.09	0.00	0.00			2.16	1.27
20	4.50	0.50	Gaussian	9	0.01	0.03			0.12	0.20	4.41	6.69
21	5.00	1.00	Gaussian	9	0.01	0.02			0.15	0.15	2.95	4.88
22	6.00	1.00	Gaussian	8	0.02	0.03			0.10	0.11	6.01	6.61

Table 3.3 – Mean (overline) and standard deviation (σ) of dry drift model parameters for the example events. The model shown is the one that was the best fit greater than 50% of the time (spherical model parameters are shown if there was a tie). k is the class number. D is the class centre diameter, and δ_k is the class width; both are in mm. Parameters are sill (c_d), nugget (c_0), distance lag (d_0) and range (a). Events shows the number of events included in the calculation. Note that larger drop classes were not present in all events.

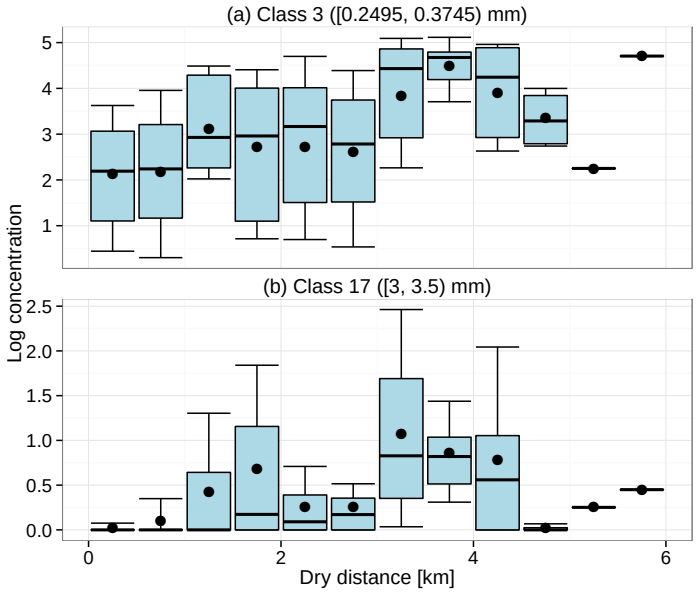


Figure 3.3 – Experimental dry drifts for event 13 for (a) class 3 ([0.2495, 0.3745) mm) and (b) class 17 ([3, 3.5) mm). Distance class width is 500 m. Boxes show class inter-quartile range, whiskers show 10th to 90th percentile range, horizontal lines indicate medians, means are indicated by dots.

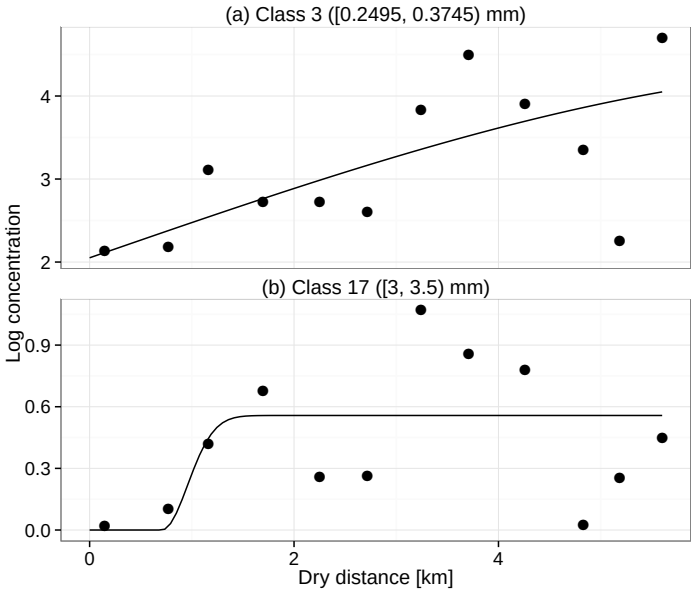


Figure 3.4 – Fitted dry drift models for event 13 for (a) class 3 ([0.2495, 0.3745) mm) and (b) class 17 ([3, 3.5) mm). The spherical model was the best fit for class 3, while the Gaussian model was the best fit for class 17. The models (lines) were fitted to all data, but to improve readability only the mean concentration values per 500 m distance class (dots) are shown.

3.3.2 Principal components

PCA was performed to find the principal components of the detrended drop concentrations $N^{\dagger}(D_k, z)$. Parsivel drop diameter classes 3 to 22 recorded drops during event 13, and PCA produced 20 orthogonal principal components. The properties of the first eight components are shown in Table 3.4. In this example, these first eight components explained 97% of the variance in the input data, with the first two components explaining more than two thirds of the variance. The least important components affect the accuracy of the higher-order moments of the reconstructed DSDs. Preliminary tests showed that in the majority of cases, good reconstructions of DSDs were possible with as few as ten components. In this work, however, we kept all principal components to preserve all available information.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Std. dev.	3.07	2.04	1.57	1.12	0.97	0.75	0.58	0.45
Prop. var.	0.47	0.21	0.12	0.06	0.05	0.03	0.02	0.01
Cumu. prop.	0.47	0.68	0.80	0.86	0.91	0.94	0.96	0.97

Table 3.4 – Properties of the first eight PCA components for the example event. Std. dev. is standard deviation, Prop. var. is proportion of variance, Cumu. prop. is cumulative proportion of variance.

In general, the resulting component distributions were close to symmetrical. Figure 3.5 shows the distributions of relative difference between mean and median of each component distribution. In this plot the difference between mean and median is relative to the 10th to 90th percentile range of each component distribution. The relative differences were low,

with no difference between median and mean exceeding 9%. Our use of Cressie's robust estimator (Equation 3.8) ensures that possible outlier influence is minimised in the variogram calculations.

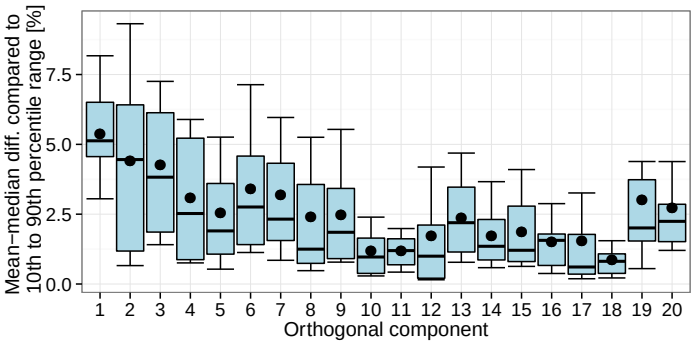


Figure 3.5 – Distributions of mean to median difference as a percentage of the 10th to 90th percentile range, for each PCA component, across all events. Boxes show inter-quartile range, dots indicate means, bold bars show medians, whiskers display 10th to 90th percentile range of mean to median differences.

3.3.3 Fitting variograms

After calculating principal component values, the next step is to characterise their spatial structures by finding component variograms. The sample semivariances for each of our components were found at various distance lags, and a variogram model was fitted to each sample variogram. Variogram estimation, model fitting, and kriging were performed using the R Gstat package (Pebesma, 2004).

To accurately estimate the variogram of a process, a minimum of about 30 pairs of samples of the process is recommended (Cressie, 1993). In most studies involving networks of disdrometers, limitations on the number of instruments mean that this is not possible, and our study is no exception. To increase the number of samples available for the variogram, we assumed that measurements at different time steps represented realisations of the same event-scale process. For each event, we grouped together all measurements, such that in Equation 3.8, N_l was the total number of unique data pairs for distance lag l, across the whole event. Only points measured during the same time step could be paired. The distance class size was set to 100 m, then distance classes were dynamically joined so that at least 30 pairs of points appeared per distance class. The events analysed (and shown in Table 3.1) were those in which there were at least six distance classes containing at least 30 point pairs.

Per-component sample variograms were fitted with a spherical model (Equation 3.18). For the component variograms, the range a represents the decorrelation distance of the process, the sill c_d represents the approximate variance of the process, and the nugget c_0 represents the variogram at a distance lag of zero, which is made up of measurement error and process microstructure (Chilès and Delfiner, 1999). In this case, we used collocated Parsivel stations (Pradel 1 and Pradel 2) to estimate the variogram nugget by calculating the mean

per-component variogram with pairs taken from the collocated station data. For the rest of the analyses, only DSD values at Pradel 1 were used for that location. If the fitted range converged to a value less than the mean distance of pairs in the second distance class, the model had only one sample variogram point to fit to and the fit could not be properly determined. In this case, we forced the range to be equal to the mean distance of pairs in the second distance class, and the variogram model represented noise after this range.

In our network of disdrometers, there were not enough data to determine whether the component processes were iso- or anisotropic. We therefore assumed isotropy. If more data points were available, this assumption could be tested, and if anisotropy was present it could be taken into account through the geostatistical framework. Figure 3.6 shows the sample variograms and fitted variogram models for the first eight transformed PCA components for the example event. Even though the least important components had little spatial structure, we kept all components so that measured values could be perfectly reconstructed.

3.3.4 Kriging

Using the variograms trained for the example event, we were able to reproduce the DSD spectrum at any given point near the measurement stations. To demonstrate the technique, we created a grid at 100 × 100 m^2 resolution, with a buffer of 1 km added around the locations that provided observations, and estimated the DSD spectrum at each point for which the occurrence mask indicated there was rain. The minimum allowed value for any DSD bin was the smallest non-zero observed concentration, and smaller estimated concentrations were set to zero. For each DSD in the grid, we calculated bulk variables: the rain rate R, total drop concentration N_t, mass-weighted mean drop diameter D_m, and radar reflectivity Z_H. We used the model of Andsager et al. (1999) for drop axis ratios, and T-matrix code (Mishchenko and Travis, 1998) to calculate back-scattering cross sections. Inter-quartile ranges (IQRs) of estimations were were calculated as described in Section 3.2.6, using 1000 samples drawn from each component estimation distribution. As expected, the estimation variances were lower close to observation locations.

3.3.5 Example gridded interpolations

Here we show examples for a single five-minute time step during event 13, ending at 2013-10-20 03:00 UTC. In Figures 3.7, 3.8, 3.9, 3.13 and 3.14, disdrometer stations are marked as black triangles. Example interpolations and IQRs are shown in Figures 3.7, 3.8, and 3.9 for rain rate R, mass-weighted mean drop diameter D_m, and radar reflectivity Z_H respectively. Dry regions according to the occurrence mask are shown in white. These plots show some interesting properties of the DSD interpolation technique. The effect of the dry drift calculations can be seen as structure related to the occurrence process. Note that the different bulk variables show different dry drift properties. Also note that since the dry drift models provide expected values, and fluctuations around these are modelled by the geostatistics, it is possible, but

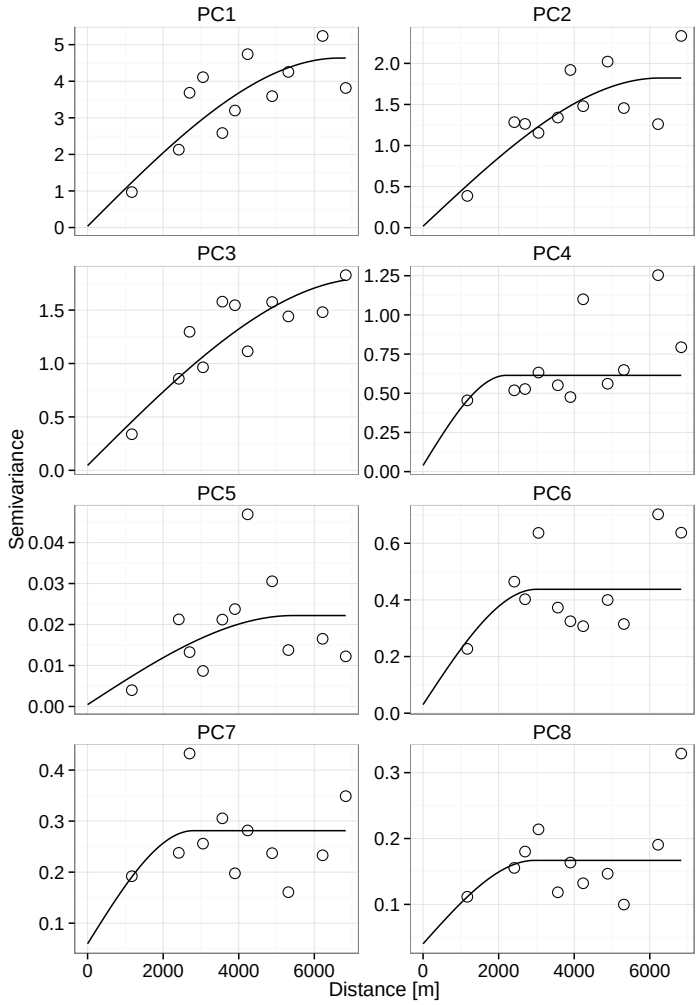


Figure 3.6 – Example variograms for the first eight PCA components in event 13. Sample variograms are shown as dots, while the lines are the fitted variogram models. The PCA components are ordered by importance. The less important components have less spatial structure; indeed the least important components contain essentially only noise.

less likely, for large values to appear close to dry regions. These interpolation results show realistic rainfall structure due to the fact that the dry drift was taken into account. Kriging has a smoothing effect which is evident in these results. The interpolated field of radar reflectivity (Figure 3.9 (a)) compares well with the measured radar reflectivity (Figure 3.14 (a)) in terms of broad rainfall pattern.

3.4 Leave-one-out testing

Leave-one-out testing was conducted for all events individually. For each event, we performed the DSD interpolation process. Dry drift estimation, principal components, and variograms were calculated using all non-collocated stations in the event. Then, within each event, each

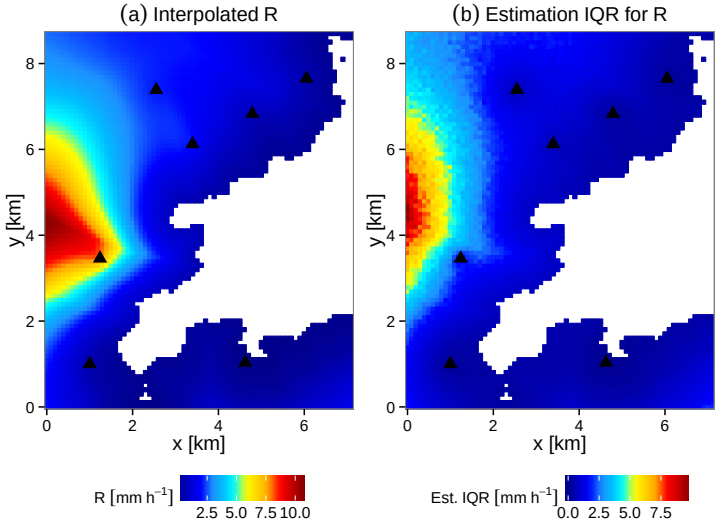


Figure 3.7 – Example interpolation results, for (a) rain rate R, and (b) the estimation inter-quartile range for R, for an example time step (2013-10-20 03:00 UTC) during event 13.

station was left out in turn, and the interpolation method was used to estimate the DSD spectra and through it bulk variables at the left-out station. The bulk variables were the total drop concentration N_t, rain rate R, mass-weighted mean drop diameter D_m, and radar reflectivity Z_H. Performance statistics were calculated per DSD equivolume drop diameter class, and per bulk variable, on every time step and station combination in all events. There were 4795 such combinations. The error was calculated as interpolated value minus measured value. The statistics used were bias (mean error) and relative error. Relative error for the bulk variables was calculated by taking error as a percentage of absolute measured value, then taking the median of these relative errors. The drop concentrations per drop size class often had a measured value of zero, so standard relative error (undefined for measured values of zero) was inappropriate. We therefore used a relative error in which the difference was calculated as a proportion of the event 10th to 90th percentile range, for a given drop concentration class. This relative error is undefined when the 10th to 90th percentile range is zero. Median relative error was used to reduce the impact of outliers in the distributions of relative error.

Figures 3.10 and 3.11 show bias distributions and relative error distributions for drop concentrations per equivolume diameter class. For drop concentrations, the bias and relative error distribution inter-quartile ranges were roughly centred around zero, indicating that much of the time the interpolation technique was able to reproduce the DSD at the left-out station in an unbiased way. While spreads of absolute errors were higher for drop classes with the highest proportion of concentrations (0.4 to 1.1 mm), spreads of relative errors on these classes were similar. The absolute value of the median relative error was less than 0.5% for all drop diameter classes, showing that the error is only a small fraction of the range of possible values of drop concentration in a given class.

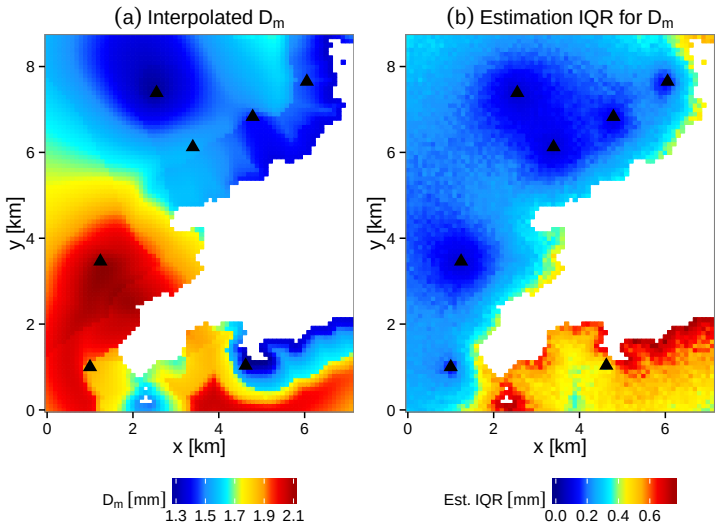


Figure 3.8 – Example interpolation results: (a) mass-weighted mean drop diameter D_m, and (b) the estimation inter-quartile range for D_m, for an example time step (2013-10-20 03:00 UTC) during event 13.

Results for bulk variables are shown in terms of distributions of relative error in Figure 3.12. The inter-quartile ranges fell around zero and were positively skewed. The error ranges were similar for N_t and R, and considerably lower for D_m and Z_H. D_m is the ratio of the 4th to the 3rd moments, which we have observed display similar variability to other DSD moments. We hypothesise that the fact that D_m is a ratio has a stabilising effect and reduces the range of its errors. Similarly, the fact that we use Z_H in dBZ instead of linear units (mm^6 m^{-3}) has the effect of reducing the influence of very large outliers and thus the spread of error. We chose to use dBZ because it is the most common unit in which to express radar reflectivity. The absolute values of median relative errors were low (less than 2.5%) for all bulk variables. Overall, rain rate R had the worst performance and mass-weighted mean drop diameter D_m was reproduced best. Leave-one-out performance results per bulk variable are shown in Table 3.5.

	N_t [%]	R [%]	D_m [%]	Z_H [%]
Mean	34.52	39.13	2.46	4.35
10th	-52.58	-61.55	-15.28	-21.60
25th	-30.73	-35.01	-7.40	-10.63
Median	-2.30	-1.11	0.31	0.32
75th	39.54	51.14	9.27	13.31
90th	124.27	159.48	21.28	34.61

Table 3.5 – Leave-one-out errors per bulk variable, showing mean, 10th, 25th, 50th (Median), 75th and 90th percentiles of errors.

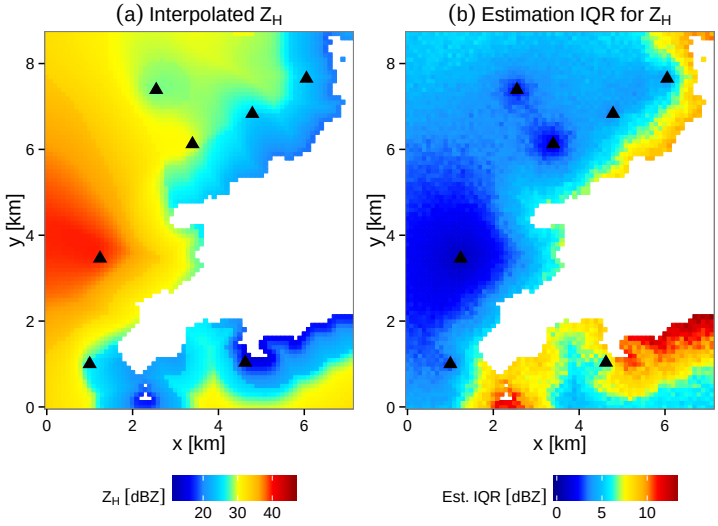


Figure 3.9 – Example interpolation results: (a) radar reflectivity Z_H, and (b) the estimation inter-quartile range for Z_H for an example time step (2013-10-20 03:00 UTC) during event 13.

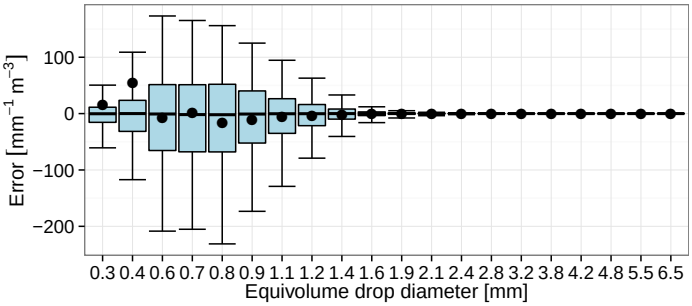


Figure 3.10 – Distributions of bias per DSD equivolume diameter class (centres shown) from leave-one-out testing. Boxes show inter-quartile range, whiskers show 10th to 90th percentile range, dots show means, bold horizontal bars show medians.

3.4.1 Effect of neglecting the dry drift

Taking the dry drift into account is important to satisfy the stationarity requirements of the geostatistics techniques we use. However, rainfall occurrence maps are not always available or easy to estimate. When only point rainfall measurements are available, it is possible to estimate the dry drift in space from temporal dry drifts, using time series information and an estimate of storm advection, as shown in Schleiss et al. (2014a). For those who may want to use our technique without a rain occurrence map, it is worth mentioning that the technique can operate without the dry drift steps if the fields are (incorrectly) assumed to be second-order stationary. Testing using the same leave-one-out station/time combinations showed that not taking the dry drift into account affected the median relative leave-one-out errors by less than five percentage points, and did not greatly affect the 10th to 90th percentile spread of relative errors. Leave-one-out testing on only station points within 0.5 km of a dry region,

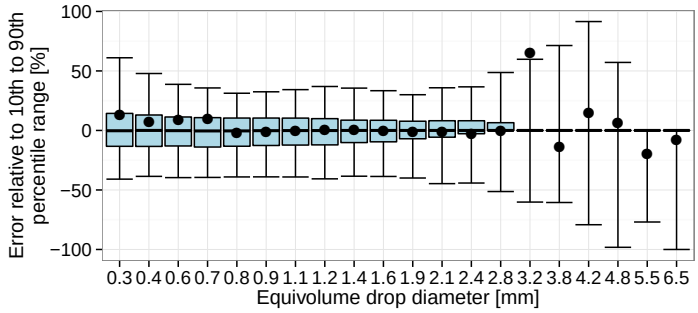


Figure 3.11 – Distributions of leave-one-out relative error (error proportional to per-event 10th to 90th percentile range) per DSD equivolume diameter class (centres shown). Boxes show inter-quartile range, whiskers show 10th to 90th percentile range, dots show means, bold horizontal bars show medians.

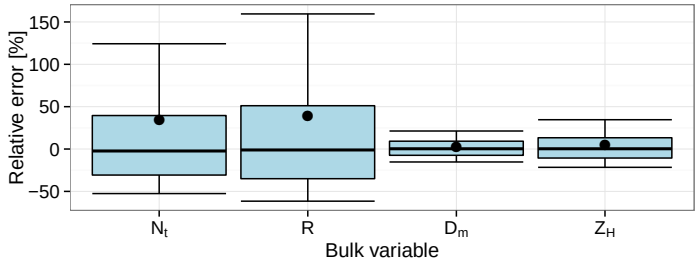


Figure 3.12 – Distributions of leave-one-out relative error (error proportional to non-zero measured value) per bulk variable. Boxes show inter-quartile range, whiskers show 10th to 50th percentile range, dots show means, bold horizontal bars show medians.

however, showed that extreme errors were worse without the dry drift steps, with the 10th to 90th percentile range of relative errors on these points significantly larger for N_t and R when dry drift was neglected. Hence, not taking the dry drift into account will not lead to significant overall bias, but to larger possible errors (and therefore uncertainty) on estimations at points close to dry regions.

3.5 Stochastic simulation of the DSD

While this chapter is primarily about interpolation of the measured DSD spectrum, it is worth mentioning that the concepts presented here can equally be applied to perform stochastic simulation of DSD spectra. By replacing the kriging process with a sequential conditional Gaussian simulation approach (see e.g. Pebesma, 2004; Schleiss et al., 2014b), multiple similar realisations of the DSD can be simulated. In contrast to interpolation, in which the most likely value is found for each point, stochastic simulation produces many equally-likely realisations of the field, all of which have identical spatial properties (variograms). The mean of the stochastic fields approaches the kriging result as the number of realisations increases. While kriging has a smoothing effect, individual stochastic realisations are more realistic fields that are not as smooth as kriging outputs. When stochastic simulation is used with our

DSD estimation technique, the principal components of the detrended DSD processes are simulated at points that are indicated by the occurrence process to contain precipitation. Once values for each realisation are obtained, the same process outlined in Section 3.2.5 is used to back-transform the components into simulated DSDs.

We used conditional sequential Gaussian simulation from the R Gstat package (Pebesma, 2004) to calculate 100 simulated realisations of the DSD for the same grid, region, and example time step used in Section 3.3. For computational efficiency, the simulation algorithm was restricted to using at most 500 nearest neighbours at each iteration. The same occurrence field was used for all simulations. Figure 3.13 shows two rain intensity fields derived from DSDs simulated using this technique, while Figure 3.14 shows measured and simulated radar reflectivity. The radar reflectivities are derived from the same simulated DSDs as the rain rate in Figure 3.13 (a). These simulation results are equally-likely realisations for the same time step. The realisations each have the same spatial properties and the same intermittency, but the values of individual points are different. The per-location mean of the simulated DSDs across all realisations would converge to the same values given by the kriging method, as the number of realisations increased. These individual simulated fields are obviously less smooth than the kriged results and may contain extreme values. Stochastic simulation offers a useful probabilistic approach to generate ensembles of realistic DSD fields, and to examine the variability of the DSD while taking into account possible extreme values that would be "smoothed out" by the interpolation technique.

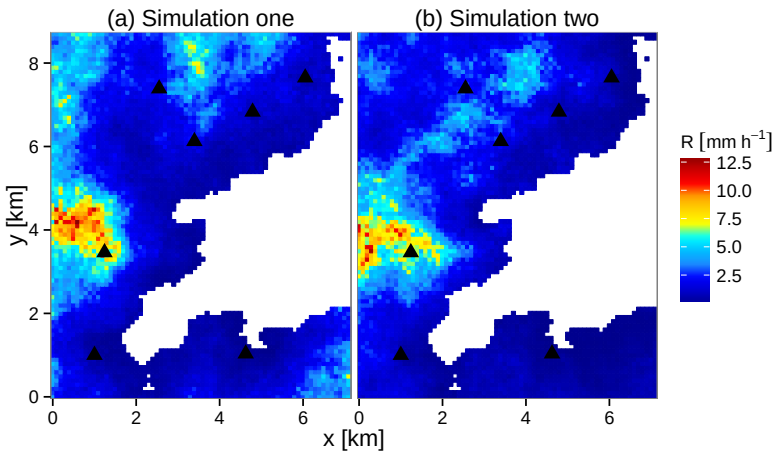


Figure 3.13 – Two example simulated fields for rain rate R, for an example time step (2013-10-20 03:00 UTC) during event 13. Note that the fields differ and are not smooth.

3.6 Conclusions

We have presented a new approach for the interpolation of experimental raindrop size distribution spectra. Using the technique, non-parametric drop size distribution spectra can be estimated at unmeasured locations. We showed that raindrop concentrations per diameter

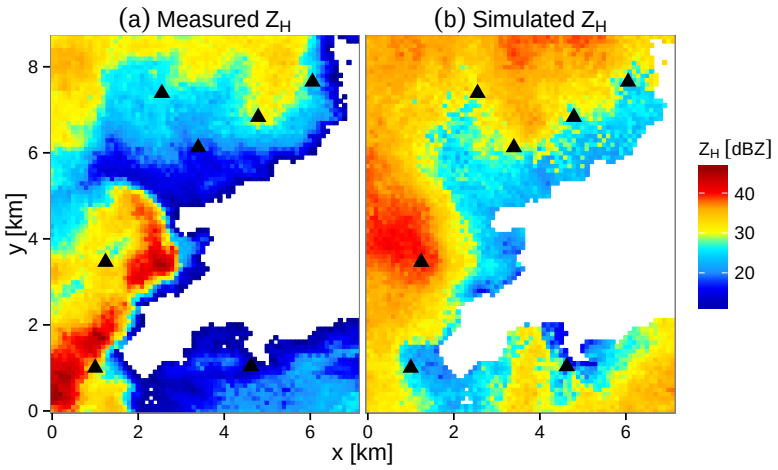


Figure 3.14 – Comparison to measured values for an example time step (2013-10-20 03:00 UTC) during event 13: (a) the measured radar reflectivity (attenuation corrected horizontal polarisation Z_H [dBZ] at 100 m resolution, SNR > 5 dB and Z_H > 10), and (b) an example simulated field for Z_H. The simulated field is less smooth and more realistic than the interpolated field shown in Figure 3.9.

class are subject to a dry drift. Given a rainfall occurrence field, the dry drift can be subtracted from the DSD concentration fields to obtain detrended fields that are not affected by this source of non-stationarity. The DSD interpolation technique works on these detrended fields, and uses geostatistical methods to characterise and interpolate principal components of the detrended DSD. We applied the technique to disdrometer network data from HyMeX campaigns in Ardèche, France. The method is equally applicable to data from other sensors. We used radar data to determine occurrence fields which were used to calculate the dry drifts. If dedicated radar data were not available, the occurrence field could be calculated from other sources such as numerical weather prediction models or operational weather radar products, or indeed estimated from disdrometer time series (Schleiss et al., 2014a). The presented technique will also operate effectively, although not as accurately, without any consideration of the dry drift.

Leave-one-out testing using HyMeX data demonstrated that the new technique is able to estimate the DSD spectrum with low bias at unmeasured locations. Since the method is geostatistical, an associated measure of estimation uncertainty is provided for every estimation. Stochastic simulation of the DSD is possible through simple modification of the technique, and could be used to estimate the probability densities of DSD values at points in space, or to examine possible extreme values of the DSD. All bulk rainfall variables can be calculated from the DSD, so the DSD interpolation method is effectively able to interpolate or simulate all bulk variables simultaneously. However, its main utility comes from the fact that it interpolates or simulates the non-parametric DSD spectrum assuming no prescribed functional form. The method is useful for studies that use networks of disdrometers to investigate the small-scale variability of the drop size distribution and its associated precipitation variables. Possible extensions to the technique include the use of kriging with external drift to take topographical

effects into account and the use of climatic dry drifts for specific regions.

4 Small-scale variability of the raindrop size distribution and its effect on areal rainfall retrieval

This chapter is adapted from:

- T. H. Raupach and A. Berne. Small-scale variability of the raindrop size distribution and its effect on areal rainfall retrieval. *Journal of Hydrometeorology*, 17:2077–2104, July 2016. doi: 10.1175/JHM-D-15-0214.1.

This work was completed by T. Raupach under the supervision of A. Berne. Research, analyses, and writing are by T. Raupach. For data acknowledgements, see Appendix A.

4.1 Introduction

An inherent difficulty in measurement of precipitation is that of change of support. The support of a measurement may be a point or an areal region, and at small scales it may be tempting to assume the two are equivalent. In this chapter we study the small-scale sub-grid variability of the DSD in order to answer two questions. First, what error is introduced by assuming that a point measurement of precipitation represents an areal region? Second, if an estimate of the DSD is derived from areal precipitation measurements, how representative is it of the underlying sub-grid precipitation process?

Sub-grid variability of the DSD and non-linearities between rainfall variables imply that any assumption of equivalence between areal and point precipitation measurements introduces error. Take for example a weather radar that measures electromagnetic radiation reflected off hydrometeors within a particular volume of air. This radar reflectivity Z [mm^6 m^{-3}] is related to rain intensity R [mm h^{-1}] via the DSD (e.g. Marshall and Palmer, 1948; Uijlenhoet, 2001), and this relationship is known to be scale dependent (Verrier et al., 2013; Sassi et al., 2014). To calculate R from Z, one of two assumptions regarding scale is usually made. The first assumption is that a point measurement can represent a pixel. For example, a point measurement of the DSD is used to relate R to Z, or a point measurement of R is related to an areal measurement of Z. In the second assumption, an areal measurement is assumed to be representative of the sub-grid process. For example, the radar measurements are used to infer properties of a DSD model that is assumed to describe the areal DSD from which R can be calculated. This approach makes the assumption that the retrieved DSD model is representative at the pixel scale. None of these approaches takes sub-grid variability of the DSD into account.

Previous studies of DSD variability have involved the use of single disdrometers and investigation of DSD variability over time-series measurements (e.g. Uijlenhoet et al., 2003b; Lee and Zawadzki, 2005; Chapon et al., 2008). Jameson (2015) provided a technique for up-scaling single disdrometer measurements, which produces gridded simulations of non-parametric point DSDs that honour the statistical properties of the observations. The resulting gridded fields are useful for statistical characterisation of the rainfall field, but are not appropriate for direct comparison to fields of observations. Lee et al. (2007) derived the spatial and temporal distributions the DSD, using radar data, a time series of point DSD measurements, and a double-moment normalised DSD model. They found that with two DSD moments (measured Z and simulated R) they were able to capture the bulk but not all of the variability in the DSD.

Other studies have used networks of disdrometers to examine DSD variability in space. Miriovsky et al. (2004) used a network of four disdrometers of different types within a 1 km^2 region in Iowa, United States (US), and reported that radar reflectivity was highly variable within this area. However, large instrumental differences meant they were unable to determine quantitative variability, and the study was focused on the Z-R relationship. Lee et al. (2009) used four disdrometers with inter-spacings up to about 33 km, combined with radar data,

to study four stratiform rain events in Montreal, Canada. Variability in DSD moments and bulk variables was found even over small distance lags, and lower order DSD moments were found to be slightly more correlated than moments of higher order, indicating that large drops play a larger role than small drops in the variability of the DSD. This study was limited by the network that was not specifically designed for DSD variability research. Tokay and Bashor (2010) used a network of three disdrometers to quantify variability in DSD model parameters and bulk variables along a 1.7 km linear transect on Wallops Island, US, with the goal of studying DSD variability within a typical ground-based radar pixel (2×2 km^2). This network was limited by the small number of instruments. Tapiador et al. (2010) set up a network of 16 disdrometers at eight locations over a 4×4 km^2 area near Ciudad Real, Spain. They found large spatial variability at kilometre scale, but the analysis focused on bulk variables and not the DSD itself. Jaffrain and Berne (2012b) used a network of 16 disdrometers over a 1 km^2 region (described in Jaffrain et al., 2011) to observe the DSD over the area of a typical operational weather radar pixel in Lausanne, Switzerland. Using stochastic simulation of characteristic drop diameter, total drop concentration, and rain rate, they found that the assumption that a point measurement represents an areal estimate introduced close to Gaussian errors with zero mean and significant standard deviations. For example, at 60-seconds resolution and over the whole network, the error standard deviation was up to 25% for rain rate in convective cases. Noise in the data overtook natural DSD variability for time steps longer than 30 minutes. Using the same network of disdrometers, Jaffrain and Berne (2012a) studied the sub-grid variability of the power laws used for rain-rate estimation from radar. When the power laws were trained using point measurements, then applied at 1 km^2 scale, an error in rain rate estimation of between -2% and $+15$% was observed. More recently, Jameson et al. (2015a) and Jameson et al. (2015b) studied DSD variability using a network of 21 logarithmically-spaced disdrometers over a very small area (distance lags of 1 to 100 m) in South Carolina, US. Jameson et al. (2015a) found that spatial and temporal clustering structures of the DSD were significantly different from each other and that they differed across drop size classes.

Capturing the full variability of the DSD is non-trivial. In an instrument network, it is a challenge to place enough instruments to measure the full DSD variability. For example, Tapiador et al. (2010) found that at least six disdrometers would be required to accurately capture the full spatial variability of the DSD at kilometre scale. Efforts to address this issue have included stochastic simulation techniques to estimate areal DSD gamma model parameters (e.g. Schleiss et al., 2012). Stochastic simulation was used by Jaffrain and Berne (2012b) to quantify DSD model parameter variability. Jameson et al. (2015b) highlighted that it is inadequate to evaluate DSD variability using only integral variables, because different DSDs can produce the same integral variable values. Studies using disdrometer networks have drawn some broad conclusions: that DSD variability within an event can be larger than that between events (Tapiador et al., 2010; Jaffrain and Berne, 2012b), and that variability increases with greater domain size (Jaffrain and Berne, 2012b; Jameson et al., 2015b) and greater drop size (Jameson et al., 2015b) and decreases with temporal integration (Jaffrain and Berne, 2012b; Tokay and Bashor, 2010).

In this chapter, we focus on DSD volumetric drop concentrations in their measured drop size classes, without assuming any functional form of the DSD. This is in contrast to previous studies that have, for the most part, focused on bulk variables or parameters of a DSD model. Jameson (2015) also simulated non-parametric DSDs; their approach differed from ours in that they used a Bayesian up-sizing approach and only a single disdrometer. We present results that used the geostatistical approach shown in Chapter 3 to perform stochastic simulation of measured DSDs on a regular, high-resolution grid of points. Simulations were constrained by measured experimental DSDs from a network of instruments. These gridded simulations allowed us to estimate the DSD at both grid and sub-grid scales, and thus make comparisons between these scales.

We considered the variability of the DSD over a range of scales from 500×500 m^2 to 7.5×7.5 km^2. Special focus was put on two scales specifically chosen to correspond to real-world DSD applications. The first was 5×5 km^2, about the size of the ground footprint of the Global Precipitation Measurement (GPM) space-borne weather radar (Hou et al., 2014). The second was 2.8×2.8 km^2, the operational pixel size of the COSMO atmospheric model as used in Germany (Baldauf et al., 2011). Simulated grids of DSDs allowed us to quantify the error introduced by assuming that a point measurement of the DSD represents an areal region at various scales. The results were generalised by normalising the scale by the decorrelation distance of rainfall intensity. The errors found should be taken into account, for example, in ground-validation of radar and model outputs using point measurements. Further, using the stochastic simulation gridded output, we simulated the way in which GPM and COSMO would observe pixel-scale DSD processes, and tested how representative these areal estimates were of the underlying sub-grid processes. The use of fitted model parameters allowed us to determine the primary sources of errors in the areal DSD retrievals.

The rest of this chapter is structured as follows. In Section 4.2 we outline two DSD models and their parameters. In Section 4.3 the data used in this study are described. In Section 4.4 we explain the stochastic simulation approach, the methods used to calculate variabilities and errors, and the GPM-style and COSMO-style retrievals of the DSD at pixel scale. Results are presented in Section 4.5, with subsections provided for results concerning raw DSD concentrations (4.5.1), bulk variables (4.5.2), GPM-style retrieval of rain rates (4.5.4) and COSMO-style retrieval of rain rates (4.5.5). A brief discussion in Section 4.6 puts the results into perspective. Conclusions are drawn in Section 4.7.

4.2 Models of the DSD

This chapter presents an investigation into the variability of the DSD and the resulting variability of some commonly used bulk variables. These are defined from the weighted moments of the DSD, as explained in Chapter 1. In this section it is explained how the DSD is often modelled, and how parameters for a DSD model are found by GPM and COSMO. Because in this chapter we deal with more than one wavelength, bulk variables that rely on wavelength

are written with the wavelength λ specified.

It is often convenient, although not necessarily always correct, to model the DSD using a functional form. The first proposed model was the exponential DSD model of Marshall and Palmer (1948). While this model provides a good estimation of DSDs integrated over time, a model that performs better for instantaneous DSDs is the more general gamma DSD model (Ulbrich, 1983), written

$$N(D) = N_0 D^{\mu} \exp(-\Lambda D). \tag{4.1}$$

The gamma DSD model has three parameters: the shape parameter μ [–], the slope parameter Λ [mm^{-1}], and the intercept N_0 [m^{-3} mm$^{-1-\mu}$]. When $\mu = 0$ the model reduces to the Marshall and Palmer exponential form. For the rest of this section, we assume integration over drop sizes from 0 to ∞. While measured DSD data are necessarily truncated at some minimum and maximum observable $D_{\min}$ and $D_{\max}$, the effects of DSD truncation have been investigated and found to be negligible when the range of diameters is sufficiently large around the median volume drop diameter (e.g. Willis, 1984; Ulbrich, 1985; Vivekanandan et al., 2004). Assuming integration over all drop sizes, Λ and μ are related via D_m:

$$\Lambda D_m = 4 + \mu. \tag{4.2}$$

This relationship is used in another common DSD model that is based on the normalised DSD of Willis (1984). In this chapter we refer to this model as the "normalised DSD model". In contrast to the gamma DSD model in which the unit of N_0 depends on μ, all variables in the normalised DSD model have conventional units. It is formulated (using D_m as in Seto et al., 2013)

$$N(D) = N_w f_N(\mu) \left(\frac{D}{D_m}\right)^{\mu} \exp\left[-(4+\mu)\frac{D}{D_m}\right], \tag{4.3}$$

where N_w [mm^{-1} m^{-3}] is a scaling factor, and $f_N(\mu)$ is defined using the gamma function Γ:

$$f_N(\mu) = \frac{6}{(4)^4} \frac{(4+\mu)^{\mu+4}}{\Gamma(\mu+4)}. \tag{4.4}$$

Following the derivation for N_w in Bringi and Chandrasekar (2001), but replacing median vol-

ume diameter with mass-weighted mean diameter D_m and using the relationship in Equation 4.2, N_w is related to the liquid water content W by

$$N_w = \frac{(4)^4}{\pi \rho_\omega} \left(\frac{10^3 W}{D_m^4} \right). \tag{4.5}$$

When DSD measurements are available, DSD model parameters can be fitted to the observed data. However, it is often the case that the DSD itself is not measured, and the DSD model parameters for an areal region must be estimated from observations of other rainfall variables. In this section we outline how DSD model parameters are retrieved in two separate systems that correspond to the two scales we focus on in this chapter. First, we describe algorithms proposed for use by the Dual-Frequency Precipitation Radar (DPR) on board the GPM Core Observatory satellite (Hou et al., 2014). These algorithms infer the DSD from radar reflectivities. Second, we show how the COSMO atmospheric model (e.g. Baldauf et al., 2011) infers DSD parameters from the modelled total mass fraction of liquid water.

GPM-style DSD retrieval

The GPM DPR measures radar reflectivity at Ku- (13.6 GHz) and Ka-band (35.55 GHz) (Hou et al., 2014). For each pixel, a measurement is made in either single-frequency (Ku- or Ka-band) or dual-frequency (Ku- and Ka-band) mode. From the returned radar reflectivities, the parameters of a DSD model are estimated. Here we briefly review the single- and dual-frequency methods described in Seto et al. (2013) and Liao et al. (2014), in which the DSD model used is the normalised DSD model shown in Equation 4.3. In this chapter, we assume an idealised system in which radar reflectivities are known perfectly, which is to say that attenuation is assumed to have been estimated and corrected (for discussion of an attenuation-correction algorithm used with this DSD retrieval method, see Seto et al., 2013). Further, we assume a uniform distribution of energy within the radar beam. The specific attenuation, k [dB km^{-1}], is the amount of radar signal attenuated by the precipitation per kilometre of distance travelled, and is calculated from the DSD as

$$k(\lambda) = \frac{1}{\ln(10)} \times \int_{D_{\min}}^{D_{\max}} \sigma_e(D, \lambda) N(D) dD, \tag{4.6}$$

where $\sigma_e(D, \lambda)$ [cm^2] is the extinction cross-section for a drop with equivolume diameter D mm at wavelength λ cm.

It is convenient to use the N_w-normalised reflectivity I_b [dB] and specific attenuation I_e [dB

km^{-1} mm m^3], in which N_w is assumed to equal 1 mm^{-1} m^{-3} (Liao et al., 2014). Using this assumption, and substituting the DSD model defined in Equation 4.3 into Equations 1.7 and 4.6, the normalised reflectivity and normalised specific attenuation are obtained (Liao et al., 2014). Normalised reflectivity I_b is given by

$$I_b(\lambda, D_m) = 10\log_{10}\left[\frac{10^6\lambda^4}{\pi^5|K_\omega|^2}\int_0^\infty f_N(\mu)\left(\frac{D}{D_m}\right)^\mu \exp(-\Lambda D)\sigma_b(D,\lambda)dD\right], \tag{4.7}$$

and normalised specific attenuation I_e is given by

$$I_e(\lambda, D_m) = \frac{1}{\ln(10)} \times \int_0^\infty f_N(\mu)\left(\frac{D}{D_m}\right)^\mu \exp(-\Lambda D)\sigma_e(D,\lambda)dD. \tag{4.8}$$

In dual-frequency mode, the dual-frequency ratio (DFR) is used to determine DSD model parameters (Seto et al., 2013; Liao et al., 2014). DFR [dB] is defined as

$$\mathrm{DFR} = Z(\lambda_{\mathrm{Ku}}) - Z(\lambda_{\mathrm{Ka}}), \tag{4.9}$$

where λ_{Ku} and λ_{Ka} are the wavelengths for Ku- and Ka-band respectively, and measured reflectivities Z are in dBZ. DFR is independent of N_w, since $Z = 10\log_{10}(N_w) + I_b$, and therefore

$$\mathrm{DFR} = I_b(\lambda_{\mathrm{Ku}}, D_m) - I_b(\lambda_{\mathrm{Ka}}, D_m). \tag{4.10}$$

In the single-frequency case, the ratio between specific attenuation k and radar reflectivity in linear units $Z_l = 10^{(Z/10)}$ [mm^6 m^{-3}] is used to determine DSD model parameters (Seto et al., 2013). Dividing $k = N_w I_e$ by $Z_l = N_w 10^{(I_b/10)}$, we have

$$\frac{k}{Z_l} = \frac{I_e}{10^{(I_b/10)}}, \tag{4.11}$$

where k/Z_l is in units of dB km^{-1} mm^{-6} m^3. By assuming a constant value of μ, look-up

tables can be pre-calculated to determine the values of DFR and k/Z_l for a range of values of D_m. These look-up tables are then used to determine D_m given the measured DFR (in the dual-frequency case) or k/Z_l (in the single-frequency case) (Seto et al., 2013; Liao et al., 2014). With μ assumed and D_m estimated from a look-up table, N_w is then calculated using (Liao et al., 2014)

$$10\log_{10}(N_w) = Z(\lambda) - I_b(\lambda, D_m). \tag{4.12}$$

The retrieved values of the parameters N_w and D_m and assumed value of the parameter μ can then be used with the normalised DSD model (Equation 4.3). The model can be substituted for $N(D)$ in bulk variable equations to calculate other variables of interest.

COSMO-style DSD retrieval

The COSMO atmospheric model uses the gamma DSD model (Equation 4.1), with fixed values of shape parameter μ and intercept N_0. The remaining parameter Λ, for slope, is determined from the modelled total mass fraction of water, q [–] (Doms et al., 2011). The total mass fraction of water is defined from the DSD as

$$q = \frac{\pi 10^{-9}\rho_\omega}{6\rho} \int_{D_{\min}}^{D_{\max}} D^3 N(D) dD, \tag{4.13}$$

where $\rho = \rho_\omega + \rho_a$ [g cm^{-3}] is the total density of the air/water mixture and ρ_a [g cm^{-3}] is the air density. Note that the liquid water content $W = 10^6 \times \rho q$ (see Equation 1.3). Using the gamma DSD model and assuming integration over drop sizes from 0 to ∞, the slope parameter Λ is found using

$$\Lambda = \left(\frac{\pi 10^{-3}\rho_\omega N_0 \Gamma(4+\mu)}{6W} \right)^{\frac{1}{4+\mu}}. \tag{4.14}$$

Before version 4.21 of the COSMO model (released 2011), the exponential DSD model was used, with parameters $\mu = 0$ and $N_0 = 8\times10^6$ m^{-4} = 8000 m^{-3} mm^{-1} for rain. Since version 4.21 the gamma model has been in use, and N_0 has been made to vary depending on the assumed μ (Seifert et al., 2011). The current operational code uses a modified Ulbrich (1983)

N_0 and μ relationship, such that for N_0 in $\mathrm{m}^{-4-\mu}$,

$$N_0 = 8 \times 10^6 \exp(3.2\mu). \tag{4.15}$$

The value of μ in the current operational code is 0.5 and thus $N_0 = 1253\ \mathrm{m}^{-3}\ \mathrm{mm}^{-3/2}$. Once parameters μ and N_0 have been assumed and Λ has been found, the gamma DSD model (Equation 4.1) can be substituted for $N(D)$ in bulk variable equations to calculate other variables of interest. This is a one-moment microphysical scheme, because a single moment (W) of the DSD is used to describe the DSD. Two-moment microphysical schemes, in which, for example, particle concentration is also predicted and used, have been considered (e.g. Baldauf et al., 2011; Van Weverberg et al., 2014), however Baldauf et al. (2011) found that in the specific case of current COSMO operational use, the benefits of a two-moment scheme did not outweigh the increased computational cost. As a result, the one-moment scheme is currently in operational use, and in this chapter we focus on that scheme.

4.3 Data

The DSD data used in this chapter were measured during the autumns of 2012 and 2013, using the HyMeX network of Parsivel disdrometers in Ardèche, France. The instrument network is described in Section 2.2, and the meteorological properties of the region are introduced in Section 1.4. The data used here were measured by eleven (nine in 2012) first-generation Parsivels and two Parsivel^2 disdrometers. Figure 2.1 shows a map of the network; station locations are shown in 2.1, and Table 3.2 shows quantities of recorded data. Stations were deployed over both autumn campaigns, with the exceptions of Montbrun and Pradel Grainage (2013 only). Collocated stations must be handled with care for geostatistical analysis, because measurement differences should not be interpreted as spatial effects. To avoid these problems, we used the best-performing Parsivel at each collocated station, judged by comparison of rain rates with collocated tipping bucket rain gauges at five-minute resolution. Thus three more collocated Parsivel^2 stations (two at Villeneuve and one at Pradel Grainage) were excluded from this analysis. The collocated stations Pradel 1 and Pradel 2 were nevertheless used to determine variogram nuggets (see Chapter 3), but of these stations, only data from Pradel 1 were used in the generation of stochastic simulations. The maximum difference in disdrometer to rain gauge relative error between two collocated disdrometers was 13%. DSDs recorded by each disdrometer were corrected using the technique presented in Chapter 2. Up to 100 nearest neighbours were used for each point. The equivolume diameter classes in which drops appeared in the corrected Parsivel data are shown in Table 2.3.

In both campaigns, a transportable dual-polarimetric X-band Doppler weather radar called MXPol (for full instrument details see Schneebeli et al., 2013) was located to the north-east of the disdrometer network. Among other scans, the radar performed horizontal plan position

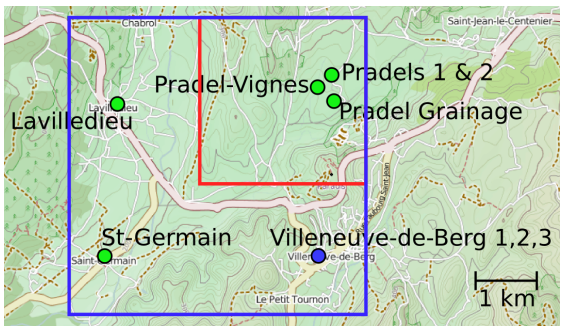


Figure 4.1 – The studied regions of interest, on a subset of the station map (Figure 2.1). The radar was positioned to the north-east at Montbrun. Squares indicate the approximate placement of the chosen regions of interest, in blue for the 5×5 km^2 area and in red for the 2.8×2.8 km^2 area. The top left corner in the 5×5 km^2 region is at UTM zone 31 coordinates 614524.8 m, 4938550 m, while the top-left corner in the 2.8×2.8 km^2 region is at UTM zone 31 coordinates 616724.8 m, 4938550 m. Maps ©Thunderforest (CC BY-SA, `http://www.thunderforest.org/`), map data ©OpenStreetMap (ODbL, `http://www.openstreetmap.org/copyright`).

indicator (PPI) scans at various elevations above the disdrometer network. We used the lowest reliable scan elevation, which was four degrees above horizontal. These scans were made about every five minutes.

Raw Parsivel measurements were made using an integration time of either 30 s or 60 s. These measurements were resampled to one, five, and 10-minute resolution, by averaging the drop concentrations per equivolume diameter bin for each time step. Time steps at each resolution were only considered if no solid precipitation was observed, the Parsivel indicated no error condition, the rain intensity observed was greater than 0.1 mm h^{-1}, and there was at least one radar scan made within the Parsivel integration time. Events were defined using one-minute resolution DSD data, as periods of rain in which there was no more than one hour of completely dry time across the network, and for which at least one station recorded an amount greater than 1 mm. The use of geostatistical analysis restricted the events that could be used to those with a sufficient number of observation pairs (see Chapter 3). Definitions of the analysed events are shown in Table 3.1.

The equations in Section 4.2 require the assumption of integration over drop sizes from 0 to ∞. Disdrometers necessarily have minimum and maximum drop sizes that they measure. The processed DSD data used in this chapter were truncated at minimum and maximum drop sizes of 0.2495 mm and 7 mm respectively. Willis (1984) showed that for the gamma DSD model, as long as the maximum drop diameter $D_{\max}$ is greater than 2.5 times the median drop diameter D_0 [mm], the error introduced by truncation of the DSD at $D_{\max}$ is negligible. This was case for 99.8% of the measured one-minute DSDs on which we based our analyses. Vivekanandan et al. (2004) determined that if $D_{\min} < D_0/2$ and $D_{\max} > 4D_0$, truncation causes less than 5% error on resulting bulk variables. The one-minute measured DSDs used here satisfied this constraint 93.7% of the time. The measured DSDs that did not satisfy this constraint were for moderate to heavy rain, with a mean (median) intensity of 8.2 (4.2) mm h^{-1}. The effect of truncation on the measured and simulated DSDs was hence assumed to be negligible.

4.4 Methods

In this section we describe the methods that were used. First, the two typical pixel areas on which most analyses were performed are introduced. Second, the method for generation of gridded DSD estimates from measured DSD data is explained. Third, we show the method used to compare point- to pixel-scale DSDs and bulk variables. Fourth, the steps involved in estimating and comparing GPM- and COSMO-style pixel-scale outputs to sub-grid values are outlined. Finally, we briefly describe how we calculated accumulated rain amounts per method.

4.4.1 Regions of interest for typical scales

DSD stochastic simulation was used to estimate DSDs on regular grids. There were two main regions of interest (ROIs). The ROI areas were 5×5 km^2 and 2.8×2.8 km^2, so as to represent respectively the typical size of a space-borne weather radar footprint, and the horizontal pixel size of a high-resolution NWP model. The locations of the ROIs were judiciously chosen so as to have a reasonable trade-off between maximising the number of disdrometer stations within each one, and the ROIs being as far as possible from the edge of the radar coverage area. In this way the DSD fields simulated in the chosen areas were able to be well constrained by ground and radar measurements. The resulting ROIs are shown in Figure 4.1. The smaller ROI is contained entirely within the larger ROI. Grids were made at 100 m resolution using coordinates in the Universal Transverse Mercator (UTM) zone 31 projection, including ROI edge points.

4.4.2 Estimation of gridded DSDs

The DSD simulation method presented in Chapter 3 was used to estimate the DSD at a regularly spaced grid of locations covering the area of the disdrometer network. Each estimate in the grid represents the DSD at the point scale at a given location, and the grid as a whole is conditioned by the measurements from the network of disdrometers.

DSD simulation was performed for integration times of one, five, and 10 minutes. One-minute resolution was the highest temporal resolution available at all stations, and therefore provided the closest possible approximation of DSD variability within an instantaneous radar scan. We produced 100 stochastic realisations for every available time step. Each time step was treated independently, with the simulations conditional on the disdrometer measurements for each. Bulk variables were calculated from the estimated DSDs. Rain rate calculations require the use of a drop velocity model, which in turn requires the altitude and latitude of the point in question. The altitude was taken as the mean station altitude within the ROI, which was 275 m. Digital elevation data (Jarvis et al., 2008) showed that the large ROI has a mean elevation of 260 m with a standard deviation of 44 m. The mean latitude of grid points in the large ROI, 44.56826°N, was used as the latitude. Raindrop terminal fall velocities were found using the

technique of Beard (1976), assuming a sea-level temperature of 15° C and relative humidity of 0.95. This assumed sea-level temperature led to a temperature for the large region of about 13° C, based on a standard atmospheric temperature lapse rate of $-6.5°\text{km}^{-1}$ (Wallace and Hobbs, 2006). Radar reflectivities were calculated using vertical incidence. Drop shapes were calculated using the model of Andsager et al. (1999), and back-scattering coefficients were found using the T-matrix code of Mishchenko and Travis (1998).

Because the stochastic simulation method we use assumes normally distributed variables, unrealistically large drop concentrations were occasionally predicted by the stochastic simulation algorithm. To remove these values, we set a threshold on rain rate: any simulated DSDs for which the corresponding rain rate was greater than 200 mm h^{-1} were removed and not counted. This threshold removed less than 0.05% of simulated DSDs. The largest measured one-minute rain rate from the disdrometer network was 73 mm h^{-1}.

Example grids of simulated fields of radar reflectivity are shown in Figure 4.2, where it can be seen that the stochastic simulation results for Z are consistent with radar observations. The stochastic simulation results are less smooth than interpolated fields (using kriging). This is because the result of interpolation by kriging is equivalent to the mean of all possible stochastic outputs. An interpolated output would present the single most likely grid of point values. In contrast, the 100 stochastically simulated realisations per time step are different but equally likely sets of values, each of which honours the spatial properties and values of the measured inputs. We used simulation instead of interpolation for two reasons. First, it allows for inclusion of more extreme values that would be "smoothed out" by interpolation using kriging. Second, stochastic simulation generates many realisations of the same process and thus creates a larger sample set to analyse.

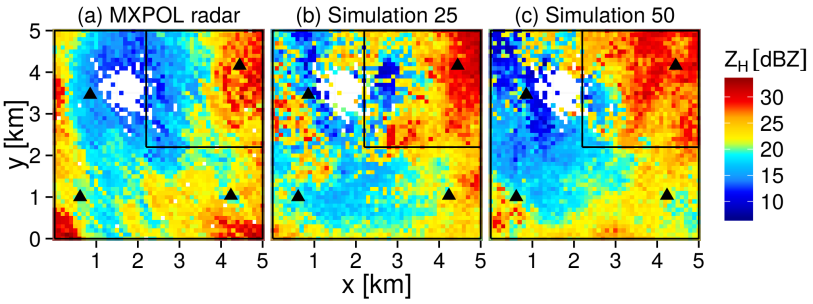


Figure 4.2 – Example gridded output for horizontal radar reflectivity $Z_h(\lambda_X)$ [dBZ] at X-band (9.4 GHz) and four degrees elevation, for a one-minute time step at 2012-09-18 21:28 UTC. Shown is the radar reflectivity recorded by the X-Band radar MXPol (a), and two example stochastic simulation realisations (b) and (c). Triangles show station locations (that reported during that time step), and the black lines trace the outermost points in the two regions of interest.

4.4.3 Comparison of point to areal estimates

Areal DSDs were calculated by taking the mean drop concentration per equivolume drop diameter class for every point in each available grid. Areal bulk variables were then calculated

from the mean DSD. These areal DSDs could then be compared to point DSDs. To ensure that any DSD estimation error resulting from the simulation technique was not mis-classified as error resulting from assuming a point measurement to be areal, we compared simulated areal estimates to individual simulated point estimates. Comparisons were made for those points that were within 250 m of the centre of each region of interest and that had rain intensity greater than or equal to 0.1 mm h^{-1}. Points with rain intensity less than 0.1 mm h^{-1} were considered to contain zero rain.

The error introduced when a point measurement of the DSD is assumed to be an areal measurement, for the kth drop equivolume diameter class D_k, is

$$E_A(D_k) = N_P(D_k) - N_A(D_k) \tag{4.16}$$

where $N_P(D_k)$ [mm^{-1} m^{-3}] and $N_A(D_k)$ [mm^{-1} m^{-3}] are point and areal simulated estimates of drop concentration for the kth drop diameter class, which contains drops with equivolume diameter in the range $[D_k, D_k + \delta_k)$ mm. The relative error is given as a percentage by

$$E_R(D_k) = \frac{E_A(D_k)}{|N_A(D_k)|} \times 100 \tag{4.17}$$

which is only defined when the areal drop concentration is non-zero. For bulk variables and model parameters, we used relative difference E_B, defined as

$$E_B = \frac{B_P - B_A}{|B_A|} \times 100 \tag{4.18}$$

where B_P is a bulk variable defined on the point scale, and B_A is the bulk variable calculated from the areal mean DSD. E_B is a percentage of the areal value. Again, this relative difference is only defined when the areal value B_A is non-zero. To make the relative difference calculation fair, we calculated them only for times when both the point and areal values were non-zero. Due to rainfall intermittency, sampled point values were not always non-zero. We noted when a point measurement was zero and therefore missed rain on the areal scale, but we performed detailed comparisons only for points at which the estimated drop concentrations were non-zero.

4.4.4 Simulation of GPM- and COSMO-style DSD retrieval

For each 5×5 km^2 stochastic realisation, the way in which a satellite-based system such as GPM might estimate the DSD for the region was simulated. To do so, radar reflectivities $Z(\lambda_{\mathrm{Ku}})$ and $Z(\lambda_{\mathrm{Ka}})$ and specific attenuation $k(\lambda_{\mathrm{Ku}})$ were calculated using the simulated areal DSD for each realisation. We assumed that $\mu = 3$ (as in Seto et al., 2013), and calculated a look-up table for D_m in terms of both DFR and k/Z_l, using Equations 4.10 and 4.11 and values of D_m from 0.001 mm to 7 mm in intervals of 0.001 mm. All calculations were performed using Parsivel drop size classes. The resulting relationships closely match those shown in Seto et al. (2013) and Liao et al. (2014), and are shown in Figure 4.3.

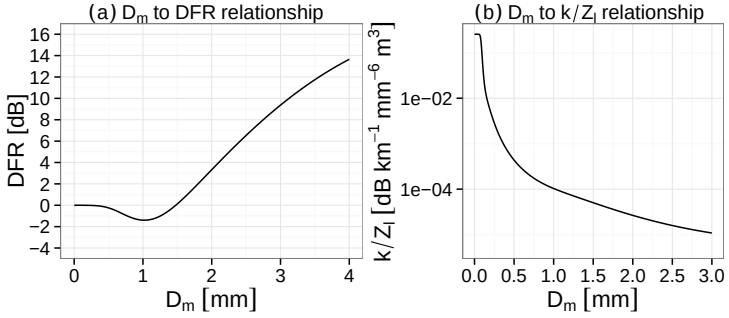


Figure 4.3 – Look-up table results for determining D_m. In (a), Dual-frequency ratio (DFR) [dB] and corresponding mass-weighted mean drop diameter D_m [mm] is shown. In (b), the relationship between specific attenuation $k(\lambda_{\mathrm{Ku}})$ and radar reflectivity $Z_l(\lambda_{\mathrm{Ku}})$ is shown with the corresponding value of D_m.

As is well-known (e.g. Seto et al., 2013), in the dual-frequency case, low values of DFR correspond to two values of D_m. In the computed look-up table, the minimum value of DFR occurred when D_m was equal to 1.02 mm. To reduce the error introduced by the retrieval of D_m using DFR, we followed the approach of Liao et al. (2014) and only considered areal DSDs for which D_m was greater than or equal to 1.02 mm. The range of DFR (D_m) values in the look-up table was -1.4 dB (1.02 mm) to 18 dB (7 mm). Using the look-up table and the DSD-derived values of $Z(\lambda_{\mathrm{Ku}})$, $Z(\lambda_{\mathrm{Ka}})$, and $k(\lambda_{\mathrm{Ku}})$, we simulated the way that a GPM-style algorithm would retrieve D_m and N_w in both the single- and dual-frequency cases. It is important to note that the specific attenuation k was calculated directly from the DSD and was thus assumed to be perfectly known. In reality, k must be estimated (e.g. Seto et al., 2013). In this work, in which the GPM algorithm is simulated, no beam effects were taken into account. To more accurately simulate GPM-like results, only records for which the radar reflectivities were greater than or equal to the DPR sensitivities of 18 dBZ for Ku-band and 12 dBZ for Ka-band (Hou et al., 2014) were analysed. Records for which the DFR value was outside the range of DFR values in the look-up table were not analysed.

In a similar approach, for each 2.8×2.8 km^2 stochastic realisation of gridded DSDs, the way that a NWP model such as COSMO might estimate the DSD was simulated. The liquid water content W for the mean DSD for each realisation was calculated, and using these values of W the value of Λ was found using Equation 4.14. These simulations were performed for both the exponential model constants quoted in Doms et al. (2011) ($\mu = 0$, $N_0 = 8000$ m^{-3} mm^{-1}) and

the newer gamma model constants quoted in Baldauf et al. (2011) ($\mu = 0.5$, $N_0 = 1253$ m^{-3} mm$^{-3/2}$). For both GPM-style and COSMO-style DSD model retrieval, the retrieved parameters were used in the corresponding DSD models to find drop concentrations for each Parsivel drop size class. To make comparisons fair, the resulting modelled DSDs were truncated to the same range of drop sizes (0.2495 to 7 mm) as the measured and simulated DSDs. Areal bulk variables were then calculated for the retrieved modelled DSDs.

4.4.5 Total rain amount

The accumulated rain amount is of primary interest to hydrologists. To see how accumulated rain amount is affected by the scale on which the measurement is taken, we calculated the sum of all one-minute rain amounts, for time steps on which all methods estimated a DSD. Simulation realisations were grouped together, so each method provided 100 accumulated rain amounts. The point simulations used were chosen randomly, one per realisation, from the sampled 2.8×2.8 km^2 region's previously sampled points. Note that these rain accumulations include only one-minute time steps for which radar data was available, which at their most frequent occurred about every five minutes.

4.5 Results

In this section, the results are shown. The error introduced by the assumption that a point measurement represents an areal measurement is addressed, for both DSD concentrations and bulk variables. We then show how precipitation variables simulated using GPM-style and COSMO-style retrieval of the DSD relate to the sub-grid distributions of corresponding values.

4.5.1 Drop diameter class concentrations

The difference $E(D_k)$ was calculated for each diameter class D_k, comparing simulated point measurements to corresponding simulated areal measurements. Per realisation, 21 grid points were within 250 m of the centre. At one-minute temporal resolution, of points that were sampled, 76% contained rainfall greater than 0.1 mm h^{-1}, so intermittency caused point estimates to miss areal rainfall about one-quarter of the time. Again for one-minute resolution, 2785674 non-zero point estimations of rainfall were compared to areal estimations. Errors introduced by the assumption that a point measurement represents an areal measurement are shown in absolute and relative terms for a 2.8×2.8 km^2 area in Tables 4.1 and 4.2 respectively, and for a 5×5 km^2 area in Tables 4.3 and 4.4 respectively. Figure 4.4 shows absolute error by equivolume drop diameter class, by time resolution and estimation type. In these tables and plots, we show only those drop classes for which the 10th and 90th percentile absolute differences were not both zero at one-minute resolution. At one-minute resolution, the median errors show negative bias. Median relative error increased for larger drop classes (greater than 2 mm), due to the fact that large drops occur infrequently and thus their areal concentrations

were often small. The greatest range of errors occurred for drops with equivolume diameters between about 0.3 and 2 mm.

	1 min			5 min			10 min		
D	q25	q50	q75	q25	q50	q75	q25	q50	q75
0.31	−13.1	−1.7	8.3	−8.2	−0.6	8.8	−6.8	−0.4	7.0
0.44	−24.0	−2.7	19.0	−16.5	−0.8	18.4	−12.9	−0.4	15.9
0.56	−61.5	−7.3	44.1	−30.0	0.4	48.6	−35.0	−1.8	35.9
0.69	−48.9	−3.4	54.5	−29.3	1.4	50.4	−20.6	2.4	41.3
0.81	−48.6	−2.4	61.1	−27.3	3.3	54.1	−21.1	3.3	43.3
0.94	−41.1	−2.7	43.4	−23.0	1.9	39.3	−16.8	1.9	31.0
1.06	−36.1	−4.3	25.0	−18.6	0.0	24.1	−12.8	0.6	19.3
1.19	−24.3	−3.5	12.4	−12.7	−0.5	13.5	−8.8	0.0	11.0
1.38	−10.9	−1.5	5.5	−6.7	−0.4	6.2	−4.6	−0.1	5.1
1.62	−3.5	−0.5	1.9	−2.3	−0.2	2.2	−1.6	−0.1	1.7
1.88	−1.2	−0.2	0.9	−0.8	−0.1	0.9	−0.6	0.0	0.7
2.12	−0.5	−0.1	0.4	−0.3	0.0	0.5	−0.3	0.0	0.4
2.38	−0.2	0.0	0.1	−0.2	0.0	0.3	−0.1	0.0	0.2
2.75	−0.1	0.0	0.0	−0.1	0.0	0.1	−0.1	0.0	0.1
3.25	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3.75	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table 4.1 – Absolute error statistics comparing mean DSDs at 2.8×2.8 km^2 pixel scale to point scale. D [mm] is the centre of each class to two decimal places. q25, q50, and q75 are respectively the 25th, 50th, and 75th percentile of the differences, expressed in mm^{-1} m^{-3}.

Figure 4.5 shows the interquartile range (IQR) of relative error by drop diameter, temporal resolution, and scale. The spread of relative error was always larger in the larger area than the smaller area. Error spreads in five and 10-minute temporal resolution results were similar, but the results for one-minute temporal resolution show a larger spread of relative errors for drops up to about 2.5 mm in equivolume diameter. The spread of relative errors decreased from the minimum drop size to about 0.75 mm, then increased with drop size. One-minute resolution error spread decreased for drops larger than about 2.25 mm. We conclude that in general terms, the variability of the drop concentration per diameter class increases with drop size and areal scale, and decreases with temporal resolution. This conclusion agrees with that of Jameson et al. (2015b), who gave a physical explanation: as the network size increases, the probability of sampling rarer parts of the DSD spectrum increases, and the variability of the DSD increases. This effect is greater for larger (and rarer) drop sizes.

For larger drops, areal concentrations become smaller with greater integration time, which leads to outlier values of relative error and increased IQRs. To decrease the effect of outlier relative errors, we divided the drops as evenly as possible into three diameter classes: the small third of drops in the range [0.2495, 0.6245) mm (Parsivel classes 3 to 5), the medium third of drops in the range [0.6245, 0.8745) mm (classes 6 and 7), and the large third of drops in the range [0.8745, 7) mm (classes 8 to 22). In the measured Parsivel data from all stations, at one-minute resolution, these classes contained 29%, 32%, and 39% of the drops respectively. Figure 4.6 shows, for these classes, the distributions of relative errors introduced by the assumption that a point measurement represents an areal region, by temporal resolution and spatial scale.

	1 min			5 min			10 min		
D	q25	q50	q75	q25	q50	q75	q25	q50	q75
0.31	−47.9	−12.6	42.7	−35.0	−4.1	38.8	−32.2	−3.4	35.6
0.44	−42.7	−9.5	39.6	−31.5	−2.8	36.4	−28.7	−1.6	35.1
0.56	−43.7	−10.2	37.7	−25.1	0.8	36.4	−29.8	−2.6	32.3
0.69	−35.3	−4.4	40.2	−23.5	1.7	36.1	−19.5	3.2	35.2
0.81	−32.9	−2.7	41.5	−20.6	3.3	36.4	−18.0	3.9	34.4
0.94	−34.7	−4.0	40.6	−21.7	2.5	35.9	−18.9	3.1	33.8
1.06	−43.4	−10.3	39.9	−25.8	0.2	35.6	−21.8	1.8	34.1
1.19	−50.1	−15.6	38.4	−30.0	−2.2	35.7	−25.6	0.2	35.2
1.38	−48.5	−13.6	41.4	−33.1	−3.9	36.4	−29.1	−0.9	36.8
1.62	−44.8	−6.4	55.8	−34.0	−2.1	42.6	−31.6	−0.3	42.4
1.88	−36.0	5.5	73.2	−30.6	3.3	51.5	−29.9	3.9	51.1
2.12	−25.4	18.8	90.8	−24.8	10.5	62.5	−25.3	10.3	61.8
2.38	−18.7	25.5	98.8	−19.3	16.9	72.3	−19.6	17.8	74.5
2.75	−9.6	32.6	103.1	−11.8	24.4	85.1	−12.0	27.1	91.8
3.25	−4.0	34.2	101.8	−5.5	31.3	98.7	−3.7	38.8	116.4
3.75	−0.9	37.7	120.9	−2.3	32.3	107.0	−1.5	37.2	110.4

Table 4.2 – Relative error statistics comparing mean DSDs at 2.8×2.8 km^2 pixel scale to point scale. D [mm] is the centre of each class to two decimal places. q25, q50, and q75 are respectively the 25th, 50th, and 75th percentile of the relative differences, expressed as percentages of areal concentrations.

The quantile statistics for these errors are shown in Table 4.5.

The range of relative error was largest for the class of small drops. Error range decreased and median error became more positive with decreasing temporal resolution. For high-resolution one-minute data, the median error was always negative, meaning that more than half of the time the point measurement underestimated the areal measurement. At its largest (one-minute, large area) this median error was -10%. The distributions were positively skewed. This means that the largest errors occurred when the point value overestimated the areal value. These results make intuitive sense. Light rainfall is more common than heavy rainfall, so given an areal region that contains both light and heavy rain, a point measurement is more likely to sample the light rain and thus underestimate the areal amount. In the rarer cases when the point does sample heavy rain, however, the error is likely to be a large overestimation: a point measurement inside a storm will overestimate an areal measurement in which there are many dry or light rain regions. As the integration time increases, severe underestimation becomes less likely. We note that the IQR of relative errors did not increase with increasing drop size when the classes contained similar numbers of drops. This observation supports the physical explanation mentioned above, that increased variability of large drops is due to their appearance in a rarer part of the DSD spectrum.

The resulting error distributions were grouped by rain rate estimated at each simulated point measurement. The classes of rain rate were chosen to each contain roughly one third of the rain rates at compared points for one-minute resolution, and were (0.1, 0.5] mm h^{-1}, (0.5, 2] mm h^{-1}, and (2, 200] mm h^{-1}. This grouping highlights how common light rainfall was, with almost two-thirds of the estimated rain rates being below 2 mm h^{-1}. The results of the

	1 min			5 min			10 min		
D	q25	q50	q75	q25	q50	q75	q25	q50	q75
0.31	−17.9	−2.4	9.8	−11.0	−1.0	10.0	−7.8	−0.2	9.9
0.44	−34.9	−4.0	22.7	−22.9	−1.4	22.0	−15.6	0.2	23.1
0.56	−84.6	−10.6	55.3	−38.2	0.3	60.4	−38.6	0.7	54.0
0.69	−63.3	−4.7	69.7	−37.5	1.5	62.7	−23.0	5.8	57.5
0.81	−63.5	−4.0	74.2	−36.0	3.8	65.6	−22.7	7.7	61.3
0.94	−54.8	−4.7	52.5	−30.5	2.2	49.8	−18.4	5.2	45.6
1.06	−47.9	−6.3	31.2	−25.3	−0.1	32.1	−15.0	1.8	29.0
1.19	−33.3	−5.3	16.2	−17.8	−0.9	18.9	−10.7	0.3	16.6
1.38	−15.7	−2.2	7.1	−9.6	−0.7	8.6	−5.7	−0.1	8.1
1.62	−5.3	−0.7	2.3	−3.4	−0.3	2.8	−2.1	−0.1	2.8
1.88	−1.7	−0.3	1.0	−1.2	−0.1	1.1	−0.7	0.0	1.1
2.12	−0.6	−0.1	0.5	−0.5	−0.1	0.6	−0.3	0.0	0.5
2.38	−0.3	0.0	0.0	−0.2	0.0	0.3	−0.2	0.0	0.3
2.75	−0.1	0.0	0.0	−0.1	0.0	0.1	−0.1	0.0	0.1
3.25	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3.75	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table 4.3 – Absolute error statistics comparing mean DSDs at 5×5 km^5 pixel scale to point scale. D [mm] is the centre of each class to two decimal places. q25, q50, and q75 are respectively the 25th, 50th, and 75th percentile of the differences, expressed in mm^{-1} m^{-3}.

grouped analyses are shown in Figure 4.7, and show that there is a strong influence of the rain intensity on the observed difference between point and areal DSDs. For very light rain, the median errors were always negative, while for heavier rain there were often positive median errors, and the distributions were more strongly positively skewed. These results confirm the intuitive reasoning given for the overall differences observed.

The results in this section show that the assumption that a point measurement of the DSD represents an areal measurement can introduce significant error. The distributions of relative errors for the chosen drop size classes had median values that showed bias between -10% and +8%. The distributions had large IQRs and were positively skewed. Very light rainfall produced negatively biased point DSD estimates, while simulated point measurements of heavier rainfall produced generally positive bias and some large over-estimations of the areal value.

4.5.2 Bulk variables

Variability in the DSD implies, of course, variability in precipitation bulk variables. We quantified the error introduced by assuming that a point measurement represents an areal region, for total drop concentration N_t [m^{-3}], rain rate R [mm h^{-1}], liquid water content W [g m^{-3}], mass-weighted mean drop diameter D_m [mm], and radar reflectivity in Ka-band $Z(\lambda_{\mathrm{Ka}})$ [dBZ] and Ku-band $Z(\lambda_{\mathrm{Ku}})$ [dBZ]. Relative error distributions are shown in Figure 4.8, and quantile statistics are given in Table 4.6 (4.7) for absolute (relative) errors.

The results show trends that are similar to those of the drop concentration relative errors. In

	1 min			5 min			10 min		
D	q25	q50	q75	q25	q50	q75	q25	q50	q75
0.31	−54.6	−16.3	47.8	−42.0	−7.0	44.7	−36.1	−1.6	46.2
0.44	−49.6	−12.9	45.1	−37.6	−4.6	42.8	−31.8	0.7	45.7
0.56	−50.1	−13.2	43.9	−29.4	0.8	43.8	−30.8	1.3	44.3
0.69	−40.3	−5.6	47.3	−27.4	2.1	43.7	−21.0	6.9	45.9
0.81	−38.6	−4.4	47.9	−24.8	3.8	44.0	−19.0	8.0	45.6
0.94	−41.0	−6.3	46.5	−26.3	2.9	44.3	−20.1	7.4	45.2
1.06	−50.2	−13.4	45.6	−31.6	−0.1	44.3	−24.8	4.6	44.8
1.19	−56.9	−19.4	43.9	−37.0	−3.4	44.6	−29.8	1.8	45.0
1.38	−55.5	−17.1	47.6	−40.8	−5.5	45.8	−33.8	−0.2	46.8
1.62	−51.6	−10.0	62.5	−41.2	−3.5	52.4	−35.7	1.6	54.4
1.88	−42.3	3.5	83.4	−36.2	3.6	64.3	−32.3	7.8	66.2
2.12	−32.2	17.3	104.9	−29.6	11.8	77.5	−26.3	15.8	78.7
2.38	−25.9	22.9	111.6	−23.4	19.2	90.3	−20.2	24.0	94.0
2.75	−16.0	30.9	115.4	−14.2	28.9	108.8	−10.9	35.6	117.3
3.25	−9.6	32.8	116.0	−5.4	39.1	130.3	−1.5	48.9	156.1
3.75	−6.2	41.5	138.6	0.7	45.5	135.5	0.5	45.3	161.1

Table 4.4 – Relative error statistics comparing mean DSDs at 5×5 km^2 pixel scale to point scale. D [mm] is the centre of each class to two decimal places. q25, q50, and q75 are respectively the 25th, 50th, and 75th percentile of the relative differences, expressed as percentages of areal concentrations.

		1 min			5 min			10 min		
Class	SL	q25	q50	q75	q25	q50	q75	q25	q50	q75
Small	2.8	−38.8	−7.2	36.8	−26.1	−0.2	35.1	−28.2	−1.9	32.1
	5	−45.8	−9.9	43.8	−31.2	−0.8	42.5	−30.0	1.5	43.9
Medium	2.8	−30.1	−1.4	39.3	−20.7	2.9	35.2	−17.6	4.0	34.2
	5	−35.6	−2.4	47.0	−24.8	3.5	43.1	−18.7	8.0	45.3
Large	2.8	−30.9	−2.4	38.0	−21.1	2.3	34.1	−18.8	2.9	33.0
	5	−37.9	−4.3	45.4	−26.6	2.7	43.4	−21.1	6.6	44.1

Table 4.5 – Relative error statistics by drop class (Small = $[0.2495, 0.6245)$ mm, Medium = $[0.6245, 0.8745)$ mm, Large = $[0.8745, 7)$ mm), and scale side length (SL [km]), comparing DSDs at pixel scale to point scale. q25, q50, and q75 are respectively the 25th, 50th, and 75th percentile of the relative differences, expressed as percentages of areal concentrations.

both small and large areas, the spread of relative errors for all variables decreased when the temporal resolution was decreased from one to five minutes. IQRs were larger for the larger areal scale. Using the IQR as an indication of how much a variable could be affected by an assumption that a point measurement is areal, the least affected variable was D_m, followed by $Z(\lambda_{\mathrm{Ka}})$, $Z(\lambda_{\mathrm{Ku}})$, N_t, W, and R. We hypothesise that the fact that the radar reflectivities are expressed in logarithmic units (dBZ), and the fact that D_m is a ratio of two DSD moments, leads to stabilised error distributions and reduced error ranges for these variables (this effect was also seen in Chapter 3). The distributions of errors were positively skewed for lower-order DSD moments N_t, W, R, and D_m, which were most likely to be overestimated. The errors were negatively skewed for the higher-order DSD moments, with the two radar reflectivities more likely to be underestimated.

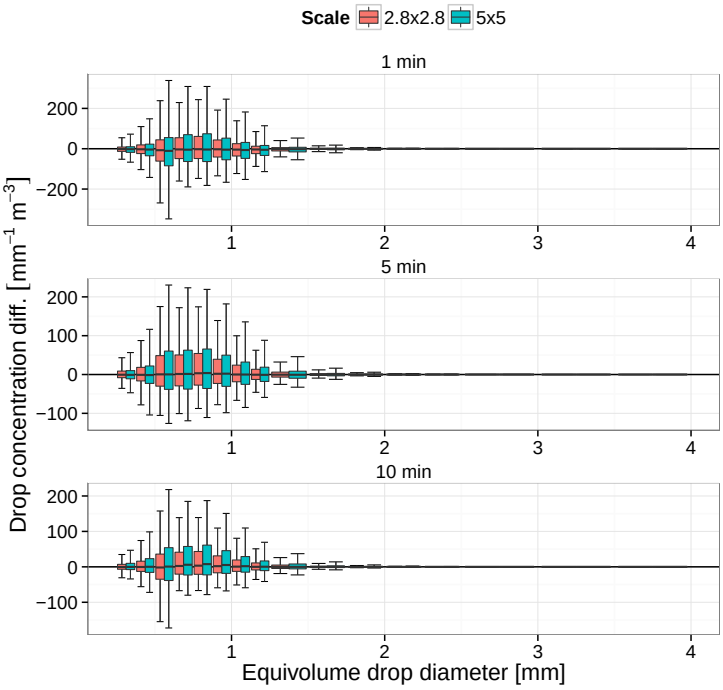


Figure 4.4 – Box-and-whisker plots showing absolute error distributions per time resolution and areal size, for simulated DSDs. The horizontal black bar indicates the median error, boxes show the interquartile range, and whiskers show the 10th to 90th percentile range.

The same analysis as for drop concentrations, of splitting the differences into classes of simulated point measurement rain rate, was performed. These results are shown in Figure 4.9. Again there was a strong influence of the rain rate at the point compared: median errors were negative for all variables for very light rain, and were all positive for rain with intensities greater than 2 mm h^{-1}. For N_t, W, and R, the error distributions were more positively skewed for the heaviest class of precipitation than for the other classes.

4.5.3 Generalised comparisons

This study focuses on areal regions that correspond in size to real-world areal sizes used by the COSMO model and GPM satellite. The results of the point to areal comparisons can be generalised somewhat by normalising the areal size (side length of the square area) by some characteristic distance or scale, here taken to be the decorrelation distance of the rainfall process. We found the decorrelation distance for each event using experimental variograms on rain rate from the disdrometer network. Each time step was taken as a separate realisation of the process, a log-transform and Cressie's robust variogram estimator (Cressie, 1993) were used, and experimental variograms were fitted with a spherical model (spherical variogram models have previously been successfully used in rainfall studies, e.g. Schuurmans et al., 2007; van de Beek et al., 2011). Across events for which the variogram model range was within the observation distance, the mean decorrelation distance was 5.5 km. We used this

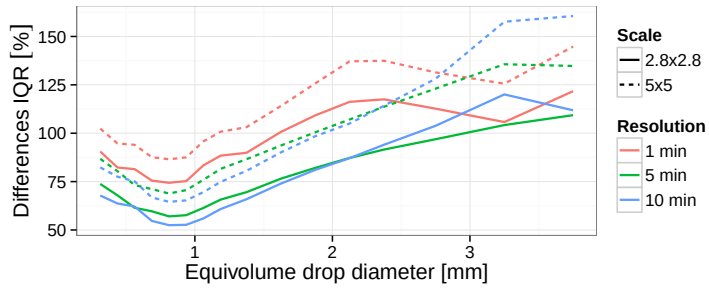


Figure 4.5 – Interquartile ranges (IQRs) of relative errors in drop concentrations introduced by the assumption that a point represents an area. The IQRs are displayed by drop diameter, time resolution, and scale.

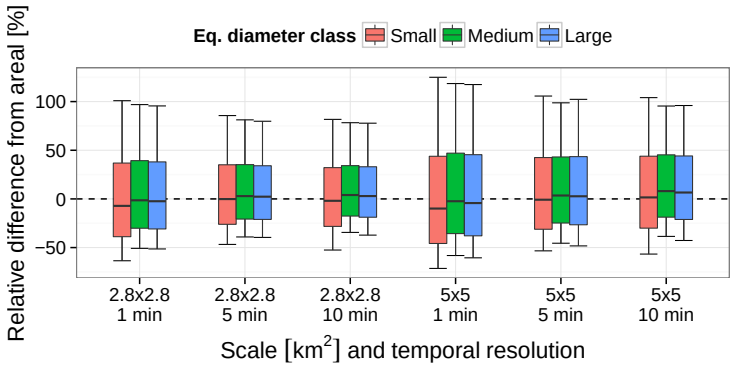


Figure 4.6 – Distributions of relative errors per class of drop size. The horizontal black bar indicates the median relative error, boxes show the interquartile range, and whiskers show the 10th to 90th percentile range. Drops are classed by equivolume diameter into the small ([0.2495, 0.6245) mm), medium ([0.6245, 0.8745) mm), and large ([0.8745, 7) mm) third of drops.

as a climatological value for the studied region. Note that this also means that even in the GPM-size pixel, the full variability of the rainfall process was not always sampled.

The same areal to point value comparisons as described above were performed for areal regions around the centre of a region formed by adding 2500 m to the north and west sides of the large ROI. The regions were square with side lengths from 500 m to 7500 m in 500 m increments. The interquartile range of relative errors was used as an indication of the error spread for each areal size. Figures 4.10 and 4.11 show the interquartile range of relative errors for drop concentration classes and bulk variables respectively, by normalised areal size. The spread of error increases almost linearly with the size of the areal region, and again the error spread is largest for the smaller third of drops and decreases with temporal integration. We expect that the rate of increase in error range would slow towards some finite limit as the ratio of areal side length to decorrelation distance becomes larger than one, but our network size did not allow us to simulate such large areal measurements. Linear model intercepts and slopes are given for normalised areal sizes less than one in Table 4.8. These results provide an estimate of expected error spread that can be applied to other regions with similar climatologies, if the decorrelation distance of rain rate is known.

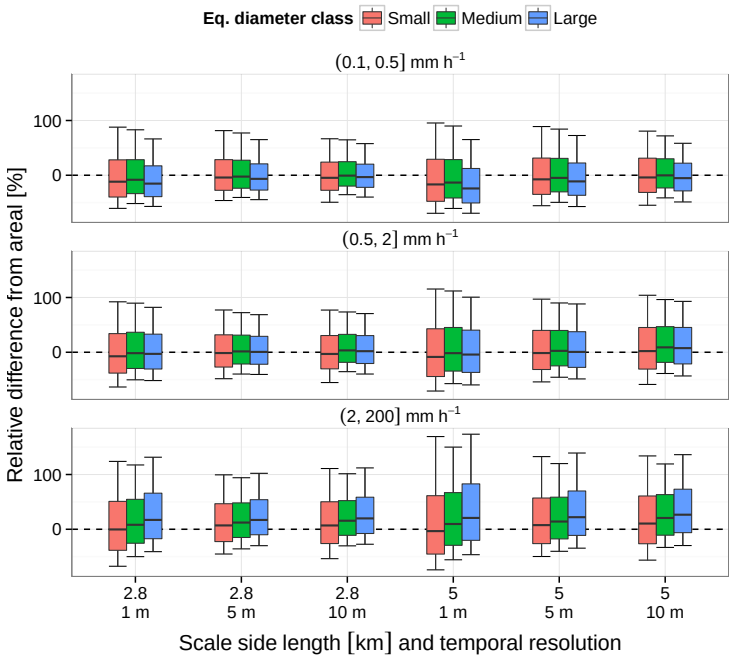


Figure 4.7 – Relative error distributions on classes of drop sizes grouped by estimated R at the point location. The horizontal black bar indicates the median error, boxes show the interquartile range, and whiskers show the 10th to 90th percentile range. Drops are classed by equivolume diameter into the small ([0.2495, 0.6245) mm), medium ([0.6245, 0.8745) mm), and large ([0.8745, 7) mm) third of drops.

4.5.4 GPM-style retrieval of bulk variables

The previous sections showed that if a point DSD is assumed to represent an area, an error is introduced. This error, in general, increases with the size of the area and with drop size. We now turn to the inverse case in which a retrieved areal DSD is assumed to be representative of the sub-grid process. We first address GPM-style DSD retrieval. The GPM DPR makes areal measurements of Ku- and/or Ka-band radar reflectivity at a resolution of about 5×5 km^2. In this section we show how simulated GPM-inferred bulk variables compared to sub-grid values in the observed area.

Recall from Section 4.4.4 that using the mean-DSD-derived $Z(\lambda_{\mathrm{Ku}})$ and $Z(\lambda_{\mathrm{Ka}})$ for each stochastic realisation, we simulated single- and dual-frequency GPM-derived D_m values using look-up tables. We used these values of D_m to retrieve DSD model parameters and thus to calculate the GPM-derived values for areal total drop concentration N_t [m^{-3}], rain rate R [mm h^{-1}], and sixth-moment radar reflectivity for each realisation. A total of about 38% of the areal DSDs were excluded from these analyses: 27% for having D_m less than 1.02 mm and a further 11% for having radar reflectivity values below the GPM DPR detectable range. These values were excluded in order to reduce the impact of errors in the prediction of D_m from DFR. Some DSDs had DFR values outside the look-up table range, at most 2.2 (0.9) dB above (below) the look-up table extents, and are hypothesised to reflect the influence of temperature on the values in the look-up table. These values were assigned the closest value of D_m in the look-up

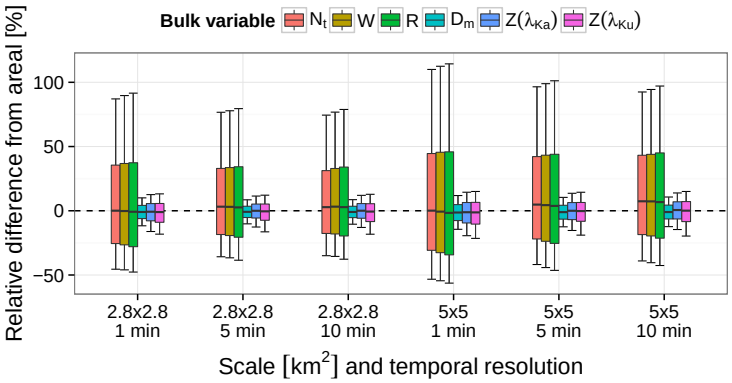


Figure 4.8 – Relative error distributions for bulk variables. The horizontal black bar indicates the median error, boxes show the interquartile range, and whiskers show the 10th to 90th percentile range.

		1 min			5 min			10 min		
Var	SL	q25	q50	q75	q25	q50	q75	q25	q50	q75
N_t	2.8	−30.60	0.05	42.54	−19.41	2.50	37.90	−16.06	1.84	29.28
	5	−42.09	0.08	57.03	−25.08	4.06	49.87	−17.11	5.52	44.85
W	2.8	−0.01	0.00	0.02	−0.01	0.00	0.02	−0.01	0.00	0.01
	5	−0.02	0.00	0.02	−0.01	0.00	0.02	−0.01	0.00	0.02
R	2.8	−0.24	0.00	0.30	−0.16	0.01	0.29	−0.12	0.01	0.22
	5	−0.36	−0.01	0.40	−0.23	0.02	0.38	−0.15	0.03	0.35
D_m	2.8	−0.07	−0.01	0.05	−0.06	−0.01	0.04	−0.06	−0.01	0.04
	5	−0.09	−0.02	0.06	−0.08	−0.01	0.05	−0.08	−0.01	0.06
$Z(\lambda_{\mathrm{Ka}})$	2.8	−1.92	−0.19	1.43	−1.41	−0.01	1.35	−1.40	0.01	1.39
	5	−2.43	−0.30	1.67	−1.78	0.00	1.66	−1.57	0.15	1.74
$Z(\lambda_{\mathrm{Ku}})$	2.8	−2.09	−0.25	1.43	−1.79	−0.12	1.39	−1.96	−0.17	1.44
	5	−2.53	−0.33	1.66	−2.05	−0.07	1.71	−2.04	0.02	1.83

Table 4.6 – Absolute error statistics by bulk variable and scale side length (SL [km]), comparing DSDs at pixel scale to point scale. q25, q50, and q75 are respectively the 25th, 50th, and 75th percentile of the relative differences. Units are m^{-3} for N_t, $\mathrm{g\,m}^{-3}$ for W, $\mathrm{mm\,h}^{-1}$ for R, mm for D_m, and dBZ for $Z(\lambda_{\mathrm{Ka}})$ and $Z(\lambda_{\mathrm{Ku}})$.

table. 66477 areal DSDs were analysed for the results in this section.

Figure 4.12 shows GPM-derived rain rates compared to DSD-derived rain rates for the dual- and single-frequency cases. The results show that assuming an accurate estimation of the specific attenuation k, both dual and single-frequency methods for estimation of rain rate produced good matches to DSD-derived rain rates. We note again that these are idealised results that ignore any beam effects or error in the estimation of specific attenuation. To test whether the resulting areal values were representative of the sub-grid values, we found the percentile of each GPM-simulated variable within the respective sub-grid distribution of values, in order to look at where within the distribution of sub-grid values each GPM-simulated value appeared. Densities of percentiles across all realisations are shown for N_t, N_w, D_m, R, and Z in Figure 4.13, for the two GPM-style retrieval methods, as well as for values obtained using the mean areal DSD. These plots show, graphically, where each areal value appeared most often within the sub-grid distributions of values.

		1 min			5 min			10 min		
Var	SL	q25	q50	q75	q25	q50	q75	q25	q50	q75
N_t	2.8	−25.5	0.1	35.5	−18.5	3.2	33.0	−17.7	2.8	31.2
	5	−30.8	0.1	44.6	−22.0	4.8	42.2	−18.5	7.4	43.2
W	2.8	−26.5	−0.2	36.8	−19.3	3.1	33.6	−18.1	3.2	32.9
	5	−32.6	−0.7	45.6	−23.8	4.4	43.3	−19.6	7.3	43.9
R	2.8	−28.0	−0.8	37.4	−20.6	2.6	34.3	−19.6	2.8	34.0
	5	−34.3	−1.6	45.9	−25.4	3.8	44.0	−21.3	6.7	45.0
D_m	2.8	−6.1	−0.9	4.2	−5.3	−0.8	3.5	−5.4	−0.9	3.5
	5	−7.6	−1.4	4.8	−6.5	−1.1	4.3	−6.4	−1.0	4.4
$Z(\lambda_{\mathrm{Ka}})$	2.8	−7.8	−0.7	5.6	−5.6	0.0	5.2	−5.8	0.0	5.6
	5	−9.5	−1.1	6.4	−6.9	0.0	6.4	−6.4	0.6	7.0
$Z(\lambda_{\mathrm{Ku}})$	2.8	−9.0	−1.0	5.7	−7.4	−0.4	5.3	−8.4	−0.6	5.5
	5	−10.4	−1.3	6.5	−8.2	−0.2	6.5	−8.4	0.1	7.1

Table 4.7 – Relative error statistics by bulk variable and scale side length (SL [km]), comparing DSDs at pixel scale to point scale. q25, q50, and q75 are respectively the 25th, 50th, and 75th percentile of the relative differences expressed as percentages of areal amounts.

The mean DSD provided a good approximation of all values for the area. From these results it is clear that extreme values affected the mean DSD, and therefore the mean-derived values were usually larger than the median sub-grid values. The small sampling size of the simulated point measurements implies that the sub-grid median values are underestimates of the real population values, with this effect increasing with the moment order of the bulk variable (Uijlenhoet et al., 2006). The GPM-style methods were able to produce values of R and Z that were generally as representative of the sub-grid values as the mean DSD for the area. However, they both produced values of lower-order moments N_t, N_w and D_m that indicated that the retrieved DSD models were not always representative of the sub-grid DSDs.

In the case of dual-frequency retrieval, GPM-simulated values of low DSD moments N_t and N_w were most likely to be above the median of sub-grid values, with 8% of N_t and 4% of N_w values above the sub-grid range. D_m tended to be overestimated or underestimated, but was only above (below) the range of sub-grid values 2% (5%) of the time. R and Z were well estimated, although like the results for the mean DSDs they produced density peaks above the median sub-grid value. For single-frequency retrieval, GPM-simulated variables were always within the sub-grid range, but lower-order moments tended to be overestimated, with density peaks greater than the median sub-grid value.

Error in GPM-style retrieval of the DSD is a combination of error caused by inexact estimation of D_m, and error in the fit to the DSD of the normalised gamma model with fixed μ. We fitted μ to each observed DSD using the technique of Johnson et al. (2014), (modified such that $\Lambda = (\mu+4)/D_m$). This maximum likelihood technique is appropriate for truncated DSDs recorded in binned drop size classes. Of our areal DSDs, 4% could not be fitted by this algorithm. Figure 4.14 shows how the assumed value of μ used by the GPM algorithms compared to the fitted gamma model parameters for the areal DSDs. The fixed value of $\mu = 3$ used by the GPM models was the 4th percentile of the distribution of μ fitted to experimental

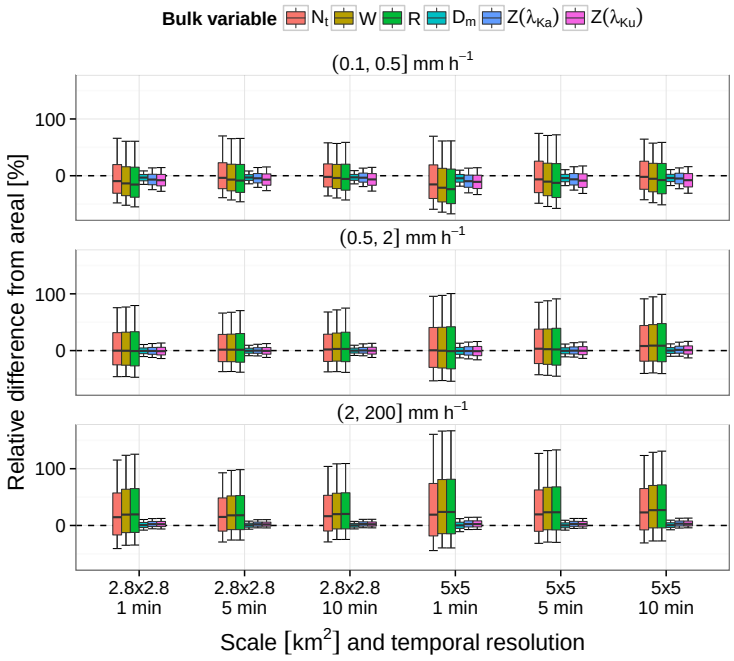


Figure 4.9 – Relative error distributions for bulk variables, classed by the simulated point estimation of rain rate. The horizontal black bar indicates the median error, boxes show the interquartile range, and whiskers show the 10th to 90th percentile range.

DSDs.

We tested the GPM DSD retrieval using (a) DSD-derived values of D_m and the original fixed value of μ, thus eliminating the retrieval of D_m from radar reflectivities, (b) DSD-derived D_m and fitted values of μ, and (c) GPM-derived D_m (look-up table with $\mu = 3$) and fitted values of μ. The performance was improved for (a), with lower-order moments more representative of the sub-grid values, but N_t still showing a probability peak above the median sub-grid value. In (b), the performance of N_t was better, but N_w, R, and Z were slightly degraded with respect to their proximity to the median sub-grid values, though they were still representative of the sub-grid distributions. In (c) the performance was similar to the performance when $\mu = 3$, with Z and R slightly degraded. We conclude that in GPM-style retrieval, N_t and N_w represent the sub-grid processes more poorly than the other tested variables primarily because of error in the estimation of D_m from radar reflectivities.

4.5.5 COSMO-style retrieval of bulk variables

A numerical weather model such as COSMO must also infer the parameters of a DSD model. In the case of COSMO, the parameters are inferred from an areal estimation of total mass fraction of water q. Recall from Section 4.4.4 that the COSMO areal DSD retrieval was simulated for each stochastic realisation on the 2.8×2.8 km^2 scale region. COSMO retrievals were simulated using both the exponential DSD model and updated gamma DSD model parameters. COSMO-

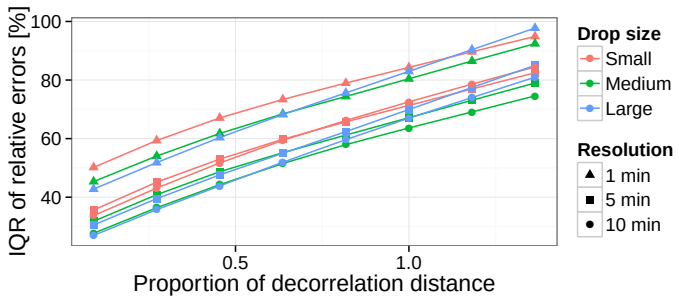


Figure 4.10 – Interquartile ranges (IQRs) of relative errors on drop concentration by classes containing roughly one-third of drops (small [0.2495, 0.6245) mm, medium [0.6245, 0.8745) mm, and large [0.8745, 7) mm) introduced by using a point as an areal value, by the areal side length's proportion of the climatological decorrelation distance.

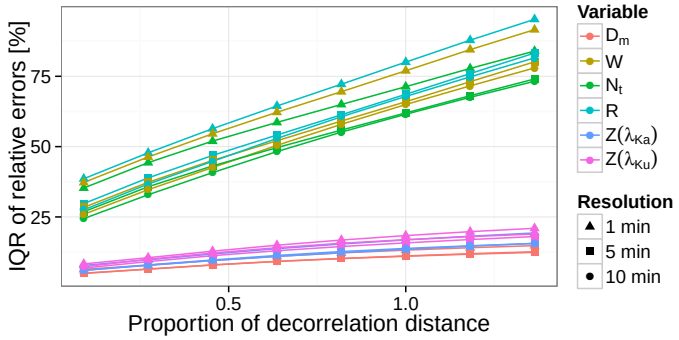


Figure 4.11 – Interquartile ranges (IQRs) of relative errors on bulk variables introduced by using a point as an areal value, by the areal side length's proportion of the climatological decorrelation distance.

inferred values of N_t, D_m, R, and sixth-moment Z_l (in linear units) were compared to sub-grid simulated values to see whether they were representative of the underlying processes.

A comparison between the COSMO-simulated rain rates and the mean-DSD rain rates for the $2.8 \times 2.8\,\mathrm{km}^2$ region is shown in Figure 4.15 for the gamma and exponential models. The gamma DSD model reproduced R slightly more accurately than the exponential DSD model. Just as for the GPM-style retrieval analysis presented above, the percentiles of the bulk variables within the corresponding sub-grid distributions were determined for each realisation. Density plots for these percentiles are shown in Figure 4.16. The two DSD models both had a tendency to overestimate N_t and Z, and to produce values of D_m outside the sub-grid range. R values were always within the range of sub-grid values. The exponential model had N_t values above the sub-grid range 53% of the time and D_m values outside the range 67% of the time (61% below, 6% above), but performed well for R and Z. The gamma model had D_m values above (below) the sub-grid range 39% (14%) of the time and Z above the sub-grid range 21% of the time. Although they returned reasonable values of R the DSDs found using the COSMO-derived exponential and gamma models were often unrealistic in terms of N_t and D_m and tended to overestimate Z.

The accuracy of COSMO-style retrieval is affected by the choice of DSD model parameters μ

Res.	Var.	Intercept	Slope	r^2
1 min	N_t	33	41	0.99
1 min	R	35	46	1.00
1 min	W	34	44	1.00
1 min	D_m	6	8	1.00
1 min	$Z(\lambda_{\mathrm{Ka}})$	7	11	1.00
1 min	$Z(\lambda_{\mathrm{Ku}})$	7	12	1.00
5 min	N_t	24	40	0.99
5 min	R	27	43	1.00
5 min	W	25	42	1.00
5 min	D_m	4	7	1.00
5 min	$Z(\lambda_{\mathrm{Ka}})$	5	9	1.00
5 min	$Z(\lambda_{\mathrm{Ku}})$	6	11	0.99
10 min	N_t	21	42	1.00
10 min	R	24	45	1.00
10 min	W	22	44	1.00
10 min	D_m	4	7	0.99
10 min	$Z(\lambda_{\mathrm{Ka}})$	5	9	1.00
10 min	$Z(\lambda_{\mathrm{Ku}})$	7	11	0.99

Table 4.8 – Linear model intercepts and slopes, to provide IQR by areal side length proportion of decorrelation distance of R. Linear models are fitted to proportions less than one. Results are organised by temporal resolution (Res.) and bulk variable (Var.), and the r^2 value is provided for each linear model fit.

and N_0. Recall that in COSMO-style one-moment retrieval, μ and N_0 remain static at all times, while Λ varies as a function of liquid water content. To investigate whether the parameter values used by COSMO were appropriate for the simulated areal DSDs, we used the technique of Johnson et al. (2014) to fit gamma model parameters to the areal DSDs. Less than 1% of areal DSDs could not have parameters fitted. Figure 4.17 shows how the COSMO DSD model parameters compared to the fitted parameters for simulated areal DSDs; in all cases, the COSMO parameters were at the low end of the fitted parameter distributions. The fixed values of μ were both at the 0th percentile of fitted values. The fixed value of N_0 with the gamma (exponential) model was the 4th (8th) percentile of fitted values, and the mean COSMO Λ value for the gamma (exponential) model was the 3rd (12th) percentile. The large ranges of the fitted parameters speak to the variability of the parameters over time.

We reran the COSMO retrievals, using the fitted values for μ and N_0. The fitted values of Λ were close to the values determined from fitted μ and N_0 using Equation 4.14, with a median relative difference of less than 1.5%. The fitted parameters produced a marked improvement in the performance of the COSMO DSD retrieval, especially for N_t. R, and Z were always within the sub-grid range. Some over-estimation of D_m remained. We conclude that error in COSMO-style DSD retrieval is largely affected by the fixed values of DSD model parameters μ and N_0.

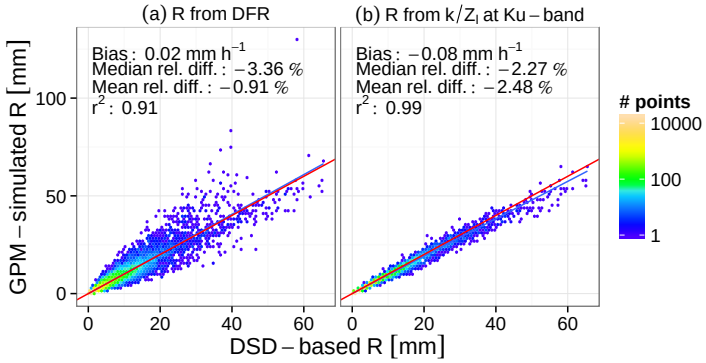


Figure 4.12 – Scatter plots showing the match between GPM-derived rain rate and DSD-derived rain rate, for mean stochastic realisation DSDs, using (a) dual-frequency DFR-based retrieval and (b) single-frequency k/Z_l-based retrieval of R. Red lines are 1:1 lines, while blue lines show the line of best fit. Note that the GPM-derived rain rates benefit from idealised assumptions, including that attenuation is perfectly known. We also assume an equal distribution of energy within the beam.

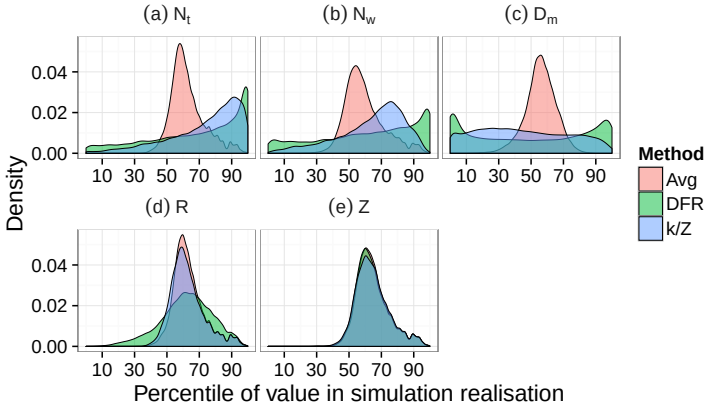


Figure 4.13 – Densities of percentiles of simulated areal (a) N_t, (b) N_w, (c) D_m, (d) R, and (e) sixth-moment Z (linear units) within sub-grid realisation values of each variable. Areal values are for dual-frequency DFR-based retrieval (DFR), single-frequency k/Z_l-based retrieval (k/Z), and for the mean DSD (Avg). These plots show where areal values most often fell within the distributions of sub-grid values. "Dry" sub-grid values are counted as zeros, except for D_m for which they are not included.

4.5.6 Total rain amount

Figure 4.18 shows the distributions of total rain amounts by method. Rain amounts were calculated per simulation number, so there are 100 values in each distribution shown in this plot. The difference between the accumulations calculated using point estimates and those calculated using an area are clear. Recall that we use the areal mean DSD as the reference amount. Simulated point measurements typically overestimated the total rain amount. COSMO with the gamma model also shows an overestimation, which agrees with earlier results. The accumulation amounts for the larger areal region were often smaller than those for the small area, which we hypothesise is due to the larger region containing more low or zero values. The density of values of R inside the 5×5 km^2 region is shown in Figure 4.19; it is clear from this plot that the distributions of R in the sub-grid regions are indeed highly skewed. Also shown

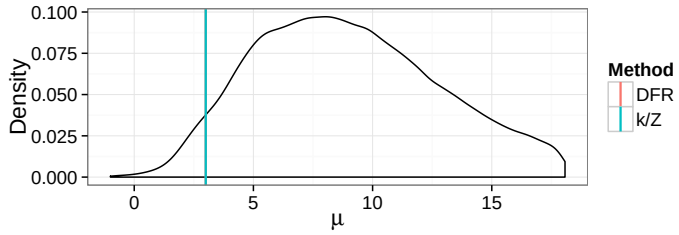


Figure 4.14 – The density of gamma model parameter μ, fitted to areal DSDs for the 5×5 km^2 region (black line). The parameter distribution has a long tail and for display purposes it has been truncated at its 90th percentile. This density is compared to the fixed GPM parameter value (vertical line).

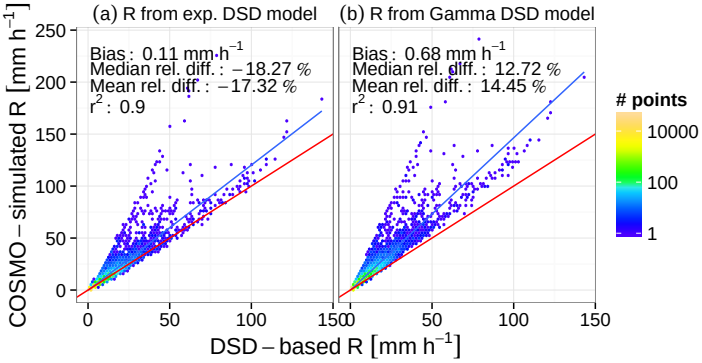


Figure 4.15 – Scatter plots showing the match between rain rates derived using COSMO-style retrieval to mean-DSD rain rates, using (a) the exponential DSD model, and (b) the gamma DSD model. Red lines are 1:1 lines, while blue lines show the line of best fit.

are the range of accumulated amounts for the idealised GPM case in which D_m was calculated directly from the DSD but the fixed value of μ was used, and for the COSMO cases in which fitted DSD model parameters were used. In both cases, the idealised versions are closer to the areal reference values, which supports our previous hypotheses that the sources of error are primarily the estimation of D_m for GPM and the fixed model parameters for COSMO.

4.6 Discussion

The results in this chapter are primarily symptomatic of the large variability that exists in DSDs, even over relatively small regions such as those studied here. In this section we put the results into perspective by discussing their meaning in a broader context. While the use of a point measurement to represent an areal region is common practice, our results show that even if the bias is low, the error introduced by such an assumption has wide variability, both for drop concentrations and bulk variables. Especially at high temporal resolution, great caution should be used if a point measurement of the DSD or a related variable is assumed to represent an area. As shown also by other studies (e.g. Jameson et al., 2015b), variability on the DSD increases with areal size and drop size and decreases with temporal integration. The variability of the mass-weighted mean drop diameter D_m is notably lower than the variability in other

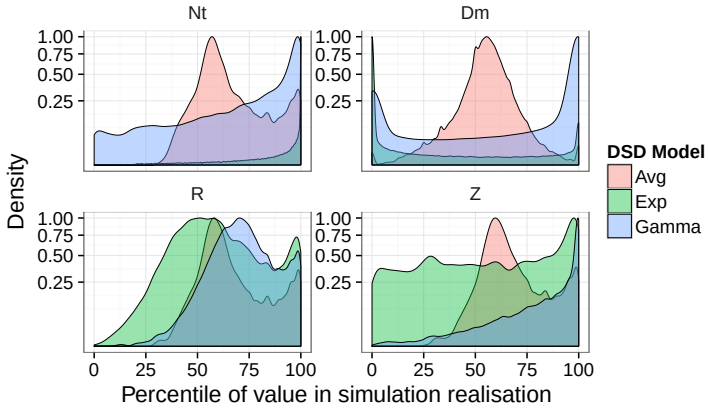


Figure 4.16 – Densities of percentiles within sub-grid realisation values for simulated areal (a) N_t, (b) D_m, (c) R, and (d) Z. Areal values were calculated using the COSMO-style algorithm and either an exponential (exp) or gamma DSD model. To better show differences between the densities, all densities have been scaled so as to have a peak of one, and there is a square-root scale on the y-axis. "Dry" sub-grid values are counted as zeros, except for D_m for which they are not included.

bulk variables; this may be explained by Jameson and Kostinski (1999), who showed that in clustered rain the variability of rain rate is due more to variability in drop concentrations than in drop sizes. This would lead to moment ratios having less variability since concentration factors would "cancel out" (Thurai et al., 2012). The results can be generalised by normalising the areal size by the decorrelation distance of the rain rate process. These results, shown in Section 4.5.3, could be applied to similar climatological regions for which the decorrelation distance of rain rate is known.

The variability in the DSD poses challenges for the estimation of areal DSDs. Due to the skewed nature of drop concentration distributions, even simply taking the mean DSD over dense measurements within the area produces an areal DSD that is likely to represent the higher-valued portion of the sub-grid distribution. Algorithms that attempt to derive an areal DSD from other areal measurements have a harder task. Our results show that, at least in the case of idealised GPM and COSMO algorithms, rainfall rate and radar reflectivity are estimated more reliably than lower order moments of the DSD. Low order moments from estimated areal DSDs should therefore be treated with more caution than higher order moments.

It is normal to use a DSD model to represent an areal (or point) DSD. These models generally do an excellent job of capturing the DSD and its bulk variables. That being said, our results have shown large variability in the DSD parameters that are appropriate for DSDs measured at high temporal resolution. In the case of the COSMO algorithm, errors in the bulk variable outputs were largely due to the fixed DSD model shape and slope parameters, which we showed were often different to values fitted to the areal DSDs. On the other hand, errors in the GPM-style algorithm were largely due to error in the estimation of the characteristic drop diameter D_m from radar reflectivities.

Our results are based on stochastic simulation of the rainfall process over typical areas that

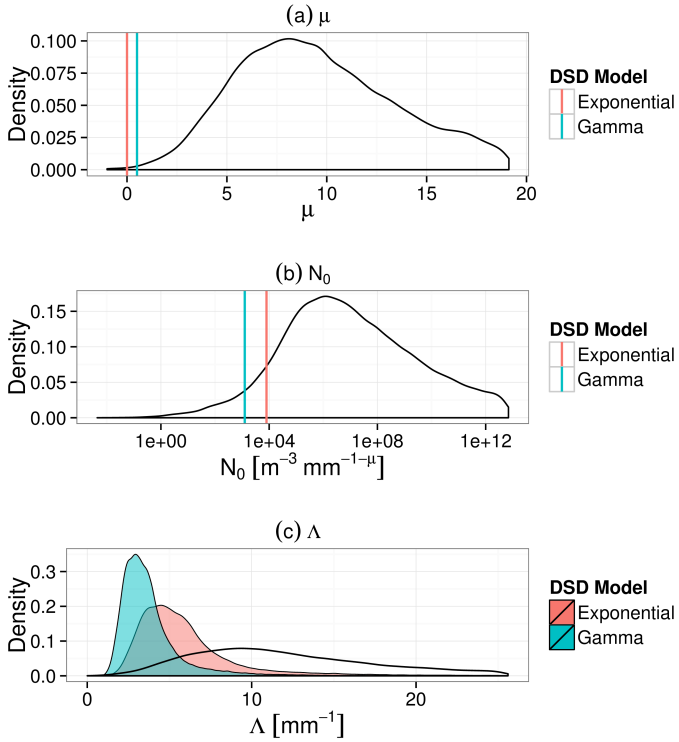


Figure 4.17 – Densities of gamma model parameters (a) μ, (b) N_0, and (c) Λ, fitted to areal DSDs for the 2.8×2.8 km^2 region (black lines). The parameter distributions have long tails and for display purposes they have been truncated at their 90th percentile. These densities are compared to COSMO parameter values (colour).

correspond to pixel sizes in real rainfall products (e.g. GPM and COSMO). Simple experimental variograms on R recorded by the Parsivel network can give an approximation of the decorrelation distances in the rain process in time and space for each event. In our case, the large region was large enough to cover the decorrelation distance for 80% of the events. The small areal region, however, was never large enough to observe the full variability of the rainfall process. In time, the decorrelation distance was larger than any integration time we used for all but one event. It should not be assumed, therefore, that our simulated domains captured the full variability of the rainfall process.

Finally, it should be noted that these results are for a specific area in Ardèche, France. This is a mid-latitude region subject to Mediterranean rainfall, and we studied events that occurred during the autumn months, when rainfall is known to be higher in these regions (see Section 4.3). The results shown here are specific to this region. It is our hypothesis, however, that because these results depend primarily on the variability in the DSD, the broad conclusions just outlined are valid wherever there is large DSD variability present.

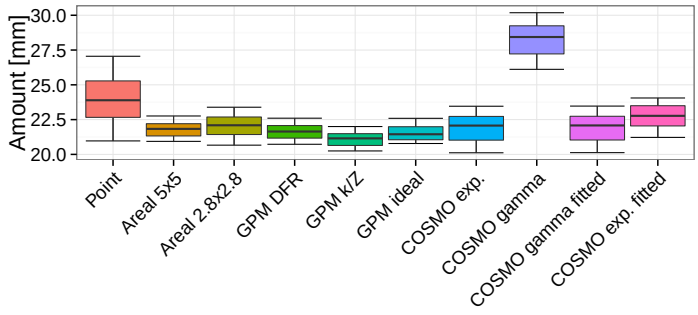


Figure 4.18 – The sum of all available one-minute rain amounts, with distributions over 100 realisations per method. Point amounts are randomly sampled from near the centre of simulated grids. Only those time steps for which all methods produced a DSD were included. "GPM ideal" refers to the case in which D_m was derived from the DSD, but a fixed value of μ was used. COSMO "fitted" cases refer to those in which fitted values of the model parameters were used.

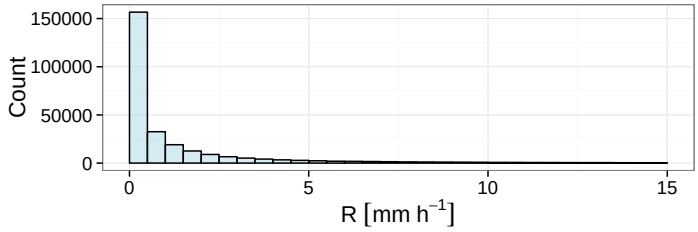


Figure 4.19 – The histogram of rain rate R, calculated for a sample. We used R values from one realisation of the large ROI from each time step, then randomly sampled 10% of these collected values, to give a total sample size of 279087 points. The x-axis is truncated at 20 mm h^{-1} and the bin width is 0.5 mm h^{-1}.

4.7 Conclusions

In this chapter we have shown results of analyses designed to quantify the effects of DSD variability on areal measurements of precipitation. We focused on two typical scales: 5×5 km^2, corresponding to the footprint of the GPM satellite-borne weather radar, and 2.8×2.8 km^2, a high-resolution horizontal pixel size of the COSMO NWP model. The analyses were based on high resolution grids of simulated experimental DSDs that were found using the DSD interpolation and simulation technique presented in Chapter 3, constrained using radar and disdrometer network data from the HyMeX 2012 and 2013 autumn campaigns in Ardèche, France. The mean DSD over the simulated grid was taken as the areal DSD. These areal DSD estimations were then compared to simulated DSDs on the point scale, and used to investigate whether GPM- and COSMO-style areal DSD retrievals were representative of sub-grid values.

The effect of assuming that a point measurement of the DSD represents an areal measurement was tested first. The results showed that this assumption introduces an error that, in general, is greater for larger drop sizes, increases with the size of the areal measurement, and decreases when measurements are integrated in time. Although these errors are most often small to moderate, with median errors on equally distributed drop size classes between −10% and

8%, the distributions of errors on both DSD concentrations and resulting bulk variables had wide interquartile ranges. Thus for any one application of this assumption it is possible for a large error to occur. While point estimates generally underestimated areal estimates for very light rainfall, they overestimated areal values for heavier rainfall, and the error distributions were mostly positively skewed. Generally, bias was more positive and the error distributions had larger spread for rainfall of greater than 2 mm h^{-1}. Bulk variables also exhibited large interquartile ranges of errors. Of the bulk variables, rain rate, liquid water content and drop concentration were all more affected than radar reflectivities and mass-weighted median drop diameter. To generalise the results, we provided linear relationships between the expected interquartile range of point-to-areal error, and areal side-length normalised by a reference distance, which we took to be the climatological decorrelation distance of rain intensity. For a similar region in which this climatological information is known, these relationships could be used to give an estimate of the point-to-areal error range for common bulk rainfall variables.

Secondly, the case in which retrieved areal DSDs are assumed to represent underlying sub-grid processes was investigated. We simulated GPM-style retrieval of the DSD on the pixel scale with specific attenuation assumed perfectly known, and COSMO-style retrieval on a smaller pixel scale. Both single- and dual-frequency GPM-style retrieval methods were able to reproduce rain rate and radar reflectivity in a representative way, but the retrieved lower-order moment values were not always properly representative of the sub-grid DSDs. Similarly, COSMO-style retrieval performed well for rain rate and radar reflectivity retrieval but produced values of drop concentration and characteristic drop diameter that were often unrealistic in comparison to sub-grid ranges. Assumed and derived DSD model parameters used by GPM- and COSMO-style retrieval were compared to parameters fitted to the simulated areal DSDs. The results showed the large variability in appropriate DSD model parameters. The GPM-style retrieval error was largely due to error in estimation of the characteristic drop size from radar reflectivities. In contrast, error in COSMO-style estimation of lower-order DSD moments was largely due to the use of fixed DSD model parameters. In this chapter we focused only on horizontal variability of the DSD. Future work will turn towards investigating vertical variability of the DSD and associated precipitation variables.

5 Invariance of the double-normalised raindrop size distribution through 3D spatial displacement

This chapter is adapted from:

- T. H. Raupach and Alexis Berne. Invariance of the double-normalised raindrop size distribution through 3D spatial displacement. Submitted to *Journal of Applied Meteorology and Climatology*.

This work was completed by T. Raupach under the supervision of A. Berne. Research, analyses, and writing are by T. Raupach. For data acknowledgements, see Appendix A.

5.1 Introduction

DSD normalisation, in which the DSD is represented as a combination of its statistical moment(s) and a "generic" normalised DSD, allows the DSD to be summarised in a compact form. Normalisation can also be used to study DSD variability, with the idea being that the normalised DSD is constant and any variability is described by variation of the chosen moment(s) alone. Sempere-Torres et al. (1994) presented a single-moment DSD normalisation approach that describes the DSD in terms of one DSD moment of arbitrary order and an arbitrary normalised DSD function. This approach unified previous normalisation approaches that used specific scaling functions and moment orders (e.g. Sekhon and Srivastava, 1971; Willis, 1984), and has been experimentally verified (Sempere-Torres et al., 1998). The method has been used to study DSD variability (Uijlenhoet et al., 2003b; Chapon et al., 2008) and the microstructure of extreme precipitation (Uijlenhoet et al., 2003a). Testud et al. (2001) proposed a two-moment normalisation approach that uses the third- and fourth-order moments of the DSD. The techniques of Testud et al. (2001) and Sempere-Torres et al. (1994) were unified by Lee et al. (2004), who proposed a double-moment scaling normalisation that does not require any assumption on the functional form of the normalised DSD. In this chapter we present an analysis of the invariance of the double-moment DSD normalisation of Lee et al. (2004) through horizontal and vertical displacement in space.

Whereas the single-moment normalised DSD function has been shown to vary by rainfall type (e.g. Sempere-Torres et al., 2000), the double-moment normalised DSD was found to be similar across different rainfall types (Lee et al., 2004). The double-moment technique is therefore more effective at capturing the variability of precipitation and provides a convenient and compact representation of the DSD. Double-moment normalisation has been used to study precipitation microstructure (Moumouni et al., 2008) and DSD variability (Lee et al., 2007, 2009). Berne et al. (2012) compared single- and double-moment normalisation of the DSD at typical operational radar pixel scale (1×1 km^2). They showed that two moments captured more variability than one, and that with double-moment normalisation, point measurements of the DSD could be used to provide pixel-scale estimates of DSD moments up to order six. At temporal resolutions finer than 10 minutes, neither technique captured all the variability in the DSD.

When double-moment normalisation of the DSD is used, the normalised DSD is usually estimated from observations and then assumed to be the same at unsampled locations or times. Lee et al. (2007) showed low scatter around the average two-moment normalised DSD for time series measurements of the DSD at one location. However, it has not yet been tested whether there are significant changes to the normalised DSD in the vertical, nor in horizontal domains larger than the 1×1 km^2 region used by Berne et al. (2012). In this study we present an analysis of the consistency of the double-moment normalisation of Lee et al. (2004), across vertical displacement of up to 2 km and horizontal displacement of up to 100 km, and between climatic regions. If the assumption of an invariant double-normalised DSD can be made, then double-moment normalisation not only provides a compact DSD representation but could be

useful in a wide range of applications, including DSD retrieval.

The rest of this chapter is organised as follows: in Section 5.2 the double-moment DSD normalisation technique of Lee et al. (2004) is described. We describe the data used in Section 5.3. The invariance of the double-moment normalised DSD is examined in Section 5.4. The effect of using a normalised DSD model trained in one location on data from another location is shown in Section 5.5. The performance of normalised DSD models, by combination of input moment orders, is shown in Section 5.6. The significance of the invariance of the double-moment normalised DSD is discussed in Section 5.7, and conclusions are made in Section 5.8.

5.2 Double-moment normalisation of the DSD

The nth-order moment of the DSD, M_n [mm^n m^{-3}], is written

$$M_n = \int_0^\infty N(D) D^n dD. \tag{5.1}$$

In the double-moment normalisation method of Lee et al. (2004), a DSD moment is written as a combination of two others as

$$M_n = C_n M_j^{(n-i)/(j-i)} M_i^{(j-n)/(j-i)}, \tag{5.2}$$

where C_n [-] is the nth-order moment of the normalised DSD h [-], which appears in the double-normalised expression for the DSD (Lee et al., 2004):

$$N(D) = M_i^{(j+1)/(j-i)} M_j^{(i+1)/(i-j)} h(x). \tag{5.3}$$

$x = D M_i^{1/(j-i)} M_j^{-1/(j-i)}$ [–] is the second-normalised diameter (Lee et al., 2004). Setting $n = i$ and $n = j$, self-consistency constraints are found:

$$C_i = 1 = \int_0^\infty h(x)x^i\,dx, \text{ and}$$
$$C_j = 1 = \int_0^\infty h(x)x^j\,dx. \tag{5.4}$$

While h can be any function, Lee et al. (2004) suggest the use of the generalised gamma function for its ability to describe all observed DSD shapes. This function is used in this study. The double-normalised form of the DSD used here can be summarised as follows (Lee et al., 2004):

$$\hat{N}(D) = M_i^{(j+1)/(j-i)} M_j^{(i+1)/(i-j)} \hat{h}(x), \tag{5.5}$$
$$\hat{h}(x) = c\Gamma_i^{(j+c\mu)/(i-j)} \Gamma_j^{(-i-c\mu)/(i-j)} x^{c\mu-1} \times \exp\left[-\left(\frac{\Gamma_i}{\Gamma_j}\right)^{c/(i-j)} x^c\right], \tag{5.6}$$

where $\Gamma_i = \Gamma(\mu + i/c)$ and $\Gamma_j = \Gamma(\mu + j/c)$, Γ is the gamma function, and c [–] and μ [–] are fitted parameters. Any DSD can then be described by Equation 5.6 with moment orders i and j.

While the integral properties defined here use drop diameters from zero to infinity, raindrops have finite size, and measurement limits mean that the distributions of drop sizes are necessarily truncated. As noted in Section 1.1, in this case become sums over classes of D. Likewise, integrals over x become definite integrals over the available classes of x. Several studies have considered the effect of DSD truncation on the calculation of bulk rain variables (e.g. Willis, 1984; Ulbrich, 1985; Vivekanandan et al., 2004). In this study the truncation limits depended on the instrument used. We ensured that any DSDs compared were truncated to the same limits, and statistical moments calculated from truncated DSDs were used qualitatively.

5.3 Data

Data from three instrument networks were used in this work. The first data set, which we refer to as the HyMeX data set, was from the disdrometer network deployed in Ardèche, France, during the autumns of 2012 and 2013, described in Section 2.2.1. Data used were from OTT Parsivel (Löffler-Mang and Joss, 2000) disdrometers, a 2D-video-disdrometer (2DVD, Schönhuber et al., 2008), two vertically-pointing METEK GmbH micro rain radars (MRRs,

Peters et al., 2002, 2005; Tridon et al., 2011), a mobile, dual-polarisation Doppler weather radar (MXPol, for instrument details see Schneebeli et al., 2013), and a Vaisala WXT520 weather station. Some Parsivels were collocated, and we used only the best-performing Parsivel station at each site, judged by comparison with collocated tipping bucket rain gauges. The two MRRs collected profiles with a vertical resolution of 100 m and an integration time of 10 seconds. MXPol made vertical measurements every 3-5 minutes with a vertical resolution of 75 m and an integration time of 21-27 seconds.

The vertical profiles in the HyMeX data set often sampled solid precipitation. For this reason we identified rainfall events in the HyMeX data set and estimated the freezing level for each one. Rainfall events were identified as time periods containing no more than an hour of dry time in the Parsivel data, and including measurements from MXPol and at least one of the MRRs. Freezing levels were estimated per time step using temperature data from the weather station (at Pradel Grainage), and an atmospheric lapse rate of $-6.5°$ km^{-1} (Wallace and Hobbs, 2006). Per event, the cutoff height was the minimum estimated freezing level minus 300 m. Visual inspection of radar scans showed that when a bright band was present, it was above the cutoff height. MRR and MXPol vertical profile data were subset to data from these event times, and altitudes greater than the cutoff height per event were excluded from this study. All valid Parsivel and 2DVD data from HyMeX in 2012 and 2013 were used. The six defined events and their estimated freezing levels are shown in Table 5.1.

Event	Date	Start	End	Freezing level [m]
1	09/28	17:05	18:56	2586
2	10/12	15:39	23:11	1171
3	10/15	03:44	11:16	1956
4	10/20	02:32	07:47	2156
5	10/27	03:22	08:49	2463
6	11/02	19:11	23:27	2063

Table 5.1 – Details of HyMeX campaign rain events, with their associated freezing level height cutoffs used to avoid the freezing level.

The second data set, which we call the Payerne data set, was collected in Payerne, Switzerland, by five first-generation Parsivel disdrometers with 30-second integration time, from February to July 2014. The third data set, referred to here as the Iowa data set, was collected as part of the National Aeronautics and Space Administration (NASA) Iowa Flood Studies (IFloodS) GPM ground validation campaign, from April to June 2013 in Iowa, United States. Data used were from 14 Parsivel^2 disdrometers (Petersen et al., 2014) with one-minute measurement integration time. A brief comparison of the climates of the HyMeX, Payerne and Iowa data set regions is given in Wolfensberger et al. (2015). Instrument station coordinates, and the number of rainy hours recorded by each station, are shown in Table 5.2. No events were defined for the Payerne and Iowa datasets, since only Parsivel data were used.

All integrated measurements were resampled to one-minute resolution. Parsivel data for HyMeX and Payerne were corrected using the technique explained in Chapter 2 (correction

Network	Station	Lat. °N	Lon. °E	Alt. [m]	Hrs
Payerne	HARAS Avenches	46.8872	7.0141	435	169.2
	Military Airport Payerne	46.8425	6.9184	451	197.1
	Morat Airport	46.9783	7.1300	433	210.5
	Payerne MCH Roof	46.8133	6.9428	489	215.0
	Station SwissMetNet Payerne	46.8115	6.9424	489	163.7
HyMeX	Lavilledieu	44.5772	4.4532	227	201.2
	Les Blaches	44.6008	4.4810	429	194.8
	Lussas	44.6123	4.4706	289	168.0
	Mirabel	44.6069	4.4987	496	213.6
	Mont-Redon	44.6141	4.5148	636	163.2
	Montbrun	44.6141	4.5461	602	56.8
	Pradel 1	44.5829	4.4987	278	200.7
	Pradel Grainage	44.5790	4.5011	271	79.0
	Pradel-Vignes	44.5801	4.4950	256	40.0
	Saint-Etienne-de-Fontbellon	44.6000	4.3826	302	58.1
	St-Germain	44.5551	4.4497	204	179.6
	Villeneuve-de-Berg	44.5548	4.4953	301	136.8
Iowa	apu01	42.2388	-92.4637	284	125.5
	apu02	42.1823	-92.3654	293	147.9
	apu03	42.1260	-92.2817	283	123.7
	apu04	42.1224	-92.2806	280	101.2
	apu05	41.9927	-92.0602	286	137.3
	apu06	41.9782	-92.0757	274	142.1
	apu07	41.9926	-92.0914	272	167.4
	apu08	41.9927	-92.0708	282	100.1
	apu09	41.8614	-91.8853	240	97.3
	apu10	41.8605	-91.8737	255	93.4
	apu11	41.8471	-91.8603	259	92.4
	apu12	41.8474	-91.8458	258	81.3
	apu13	41.6406	-91.5418	197	123.5
	apu14	41.6406	-91.5416	197	98.2

Table 5.2 – Station locations and information. Shown are the network, the station name, its WGS84 coordinates and altitude, and the number of precipitation hours recorded by Parsivel that were included in this study (Hrs). For HyMeX, the 2DVD was at Pradel Grainage, the MRRs were at Montbrun and Pradel Grainage, and MXPol was at Montbrun.

factors used for both were those trained using HyMeX data) and subset to time steps for which no solid precipitation and no instrument errors were detected. Parsivel data from Iowa were already quality-controlled and were used as provided. The 2DVD data were also processed as in Chapter 2 and subset to liquid rain steps indicated by the collocated Parsivel (available in 2013 only). Attenuation correction was performed for the MXPol vertical profile data using the Hitschfeld and Bordan method (Hitschfeld and Bordan, 1954). The vertical range of the profiles was limited to a few kilometres, and since at these ranges attenuation was in the order of a few dB at most, the Hitschfeld and Bordan method could be considered stable (Delrieu et al., 1999). MRR data were provided with attenuation correction performed (METEK, 2010) using the method of Peters et al. (2010). For the vertically-pointing radars, DSDs for each altitude were retrieved using the Doppler spectrum and the MRR retrieval method (Strauch, 1976; Peters et al., 2002), in which vertical wind and turbulence are ignored (Peters et al., 2002).

Negative drop concentrations returned by the MRR (METEK, 2013) were set to zero.

All DSD data were truncated. For Parsivel and 2DVD data the truncation was from 0.2495 mm to 7 mm (see Chapter 2) and classes were defined as in Table 2.3. For MRR- and MXPol-derived DSDs class widths were variable. The maximum drop size considered for MRR and MXPol-derived DSDs was 5.8 mm, corresponding to the maximum possible drop size in the fall velocity to drop diameter relationship of Atlas et al. (1973), which is used by the Doppler DSD retrieval algorithm (Peters et al., 2005). This retrieval algorithm suffers from instabilities as the drop size D approaches zero (Peters et al., 2005). Peters et al. (2005) used a minimum drop size of 0.246 mm. In our MRR data sets, however, very large numbers of small drops were returned for drop sizes larger than this minimum. To choose the minimum drop diameter to use for Doppler-profile-derived DSDs, we compared DSDs retrieved for the third range gate of MXPol and MRR radars at Montbrun, resampled into Parsivel drop diameter classes, to those measured by the Parsivel on the ground at Montbrun. To account for the fall time of the sampled volumes, the comparisons were made at 10-minute resolution. The results showed overestimation in comparison with Parsivel of drop concentrations returned by the DSD retrieval algorithm for classes of drops with diameters below 0.5 mm. Comparison to Parsivel-derived total drop concentration N_t [m^{-3}], liquid water content W [$g\,m^{-3}$], and rain rate R [$mm\,h^{-1}$] with the truncated DSDs showed the best overall agreement occurred using a minimum drop size of 0.6 mm for MRR and 0.5 mm for MXPol. The DSDs retrieved from Doppler information were truncated using these limits.

DSD records were subset to those records for which the rain intensity was greater than 0.1 mm h^{-1}. Despite the DSD truncation applied, the DSD retrieval algorithm used for the MRR and MXPol data sometimes returned unrealistic concentrations of small drops, which lead to very large values of rain rate and liquid water content. We excluded these records by excluding any DSD from the vertical profiles which had an associated rain rate greater than 150 mm h^{-1} (in contrast, the largest corresponding intensity for HyMeX was 62 mm h^{-1} for Parsivel and 94 mm h^{-1} for the weather station). This constraint led to the exclusion of 0.5% of MRR DSDs and 0.7% of MXPol DSDs. Finally, to avoid sampling effects, only Parsivel stations and profiler altitudes for which more than 100 within-event records were available were included; this led to the exclusion of altitudes greater than 2470 m for MRR, and altitudes greater than 2080 m for the MXPol vertical scans. The maximum altitude was lower for MXPol because this radar measured vertical profiles less often than the MRRs and therefore sampled less data.

5.4 Invariance of the double-normalised DSD

In this section we present a study on the invariance of the double-normalised DSD through horizontal and vertical displacement in space. If the assumption that the normalised DSD is invariant is acceptable, double-moment normalisation of the DSD is a powerful tool for the investigation of three-dimensional variability of the DSD. To test the invariance we compared normalised DSDs at each location to those at other locations from the same region and for the

same one-minute time step. Differences in moments of the normalised DSD were quantified by displacement distance. Horizontal displacement was tested via comparisons between Parsivel stations in each disdrometer network. The maximum altitude difference between two stations was 432 m (in the HyMeX network) and the maximum horizontal distance between stations was about 100 km (in the Iowa network). Vertical displacement was tested using the three vertically-pointing radars in the HyMeX network. Performance statistics used were median and interquartile range of relative bias (difference divided by measured value, shown as a percentage), and squared Pearson correlation coefficient (r^2). Examples are shown for moment orders three and four, a combination that is commonly used to represent the DSD (e.g. Testud et al., 2001).

The double-normalised DSD was estimated for every individual time step at each individual station and altitude, for all combinations of input moment orders between zero and seven. For each location and each one-minute time step, values of $h(x)$ were calculated using Equation 5.3. To quantify the similarity of the normalised DSDs at different points, we compared integral moments of $h(x)$, of orders zero to seven. For a pair of locations, the moments of the two normalised DSDs were compared through calculations of the absolute values of relative differences. Examples of such differences are shown in Figures 5.1 and 5.2 for horizontal and vertical displacement respectively. For comparison, these figures also show differences in the moments of $N(D)$, the un-normalised, measured DSD for the same distance lags. The pattern of difference in the DSD by distance is similar for all moments, and the median values of relative bias are always much larger than those for the normalised DSD. For readability, the figures do not include the interquartile ranges; the IQRs on moments of $h(x)$ were relatively constant with distance, while those for the DSD moments were much larger and increased in magnitude with displacement distance. These results confirm previous findings that the double-moment normalised DSD is highly effective at collapsing the DSD into a generic shape (Lee et al., 2004, 2007). The slope of the regression line on median relative bias by displacement distance is a useful indication of the invariance of the normalised DSD through displacement. An effect of distance on the similarity of the empirical normalised DSDs at one time step is visible in both horizontal and vertical cases, and it is clear that the effect is more pronounced for moment orders zero and seven, which are far from the input moment orders in this example. On the other hand, the self-consistency constraints (Equations 5.4) guarantee that moment orders i and j of $h(x)$ are always 1, and thus there is no change by displacement distance for these moments of $h(x)$.

The displacement effect is shown more strongly in the vertical results than the horizontal. Over all tested input moment combinations, the maximum slope on the best-fit lines to median relative differences in normalised DSD moments through horizontal displacement was 1.9 percentage points per kilometer (% pts km $^{-1}$), for input moments zero and one, producing change in normalized DSD moment seven, in the HyMeX data set. The median change in normalised DSD moment through horizontal displacement was 0.07 % pts km^{-1} with an IQR of 0.19 % pts km^{-1}. The maximum slope of change in any normalised DSD moment through vertical displacement was 45.8 % pts km^{-1}, for the MRR at Montbrun, with input moments zero

and one, producing change in normalised DSD moment seven. Through vertical displacement the median slope of the change was 4.84 % pts km^{-1} with an IQR of 10.27 % pts km^{-1}. While moderate changes are shown for specific moments of the normalised DSD that are of orders far from the chosen input moment orders, the normalised DSD shows only minor changes with horizontal and vertical displacement on average.

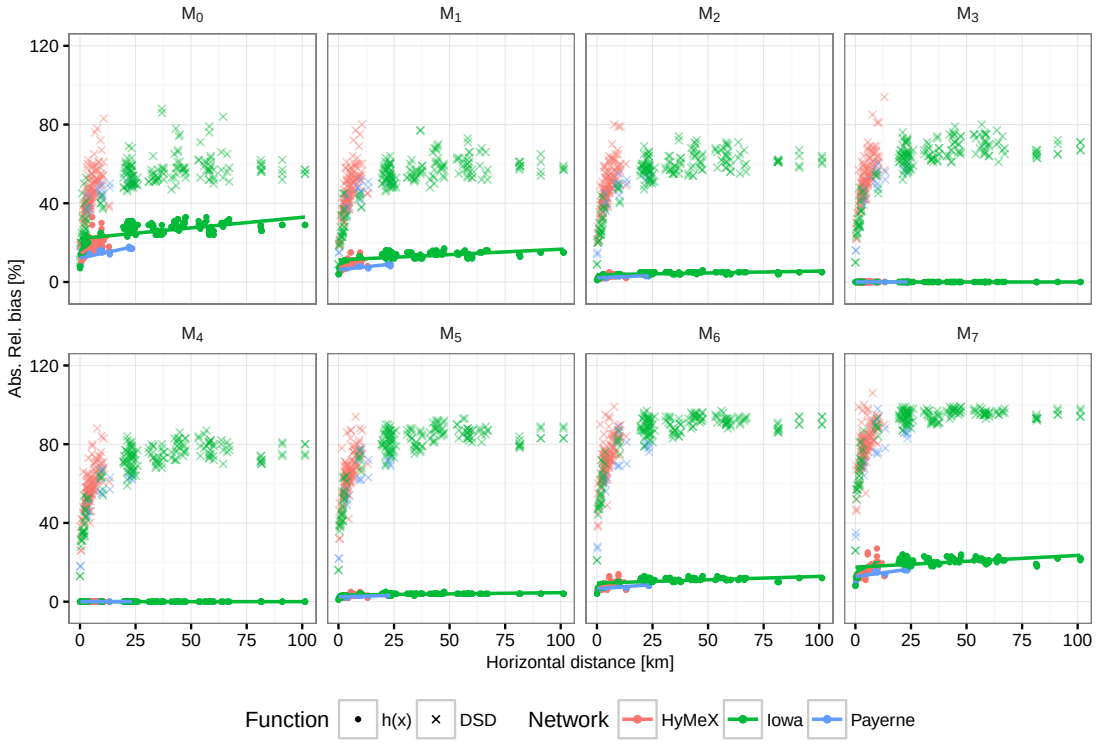


Figure 5.1 – Absolute values of median relative bias between integral moments of normalised DSDs ($h(x)$) by horizontal displacement, for $i = 3$ and $j = 4$, with regression lines. Also shown are median points of difference distributions between measured, un-normalised DSDs by displacement distance. Coloured lines show the regression lines on medians for moments of $h(x)$. The slopes of these lines were used to quantify the change in $h(x)$ by displacement distance. For readability the y-axis is cut at 120%.

5.5 Displacement effect with normalised DSD model

It has been shown that the empirical normalised DSD is effective at collapsing the variability in the DSD into a mean shape that is similar between spatial locations at the same time. When the normalised DSD is used in practice it is often modelled using a functional form such as the generalised gamma distribution shown in Equation 5.6. This DSD model was fitted to the measured normalised DSDs, using all data at each location, and its performance in reconstructing each one-minute DSD at other locations was tested.

To model the normalised DSDs, parameters μ and c of the generalised gamma model were found by using least squares fitting to minimise differences between $\log(h(x))$ and $\log(\hat{h}(x))$

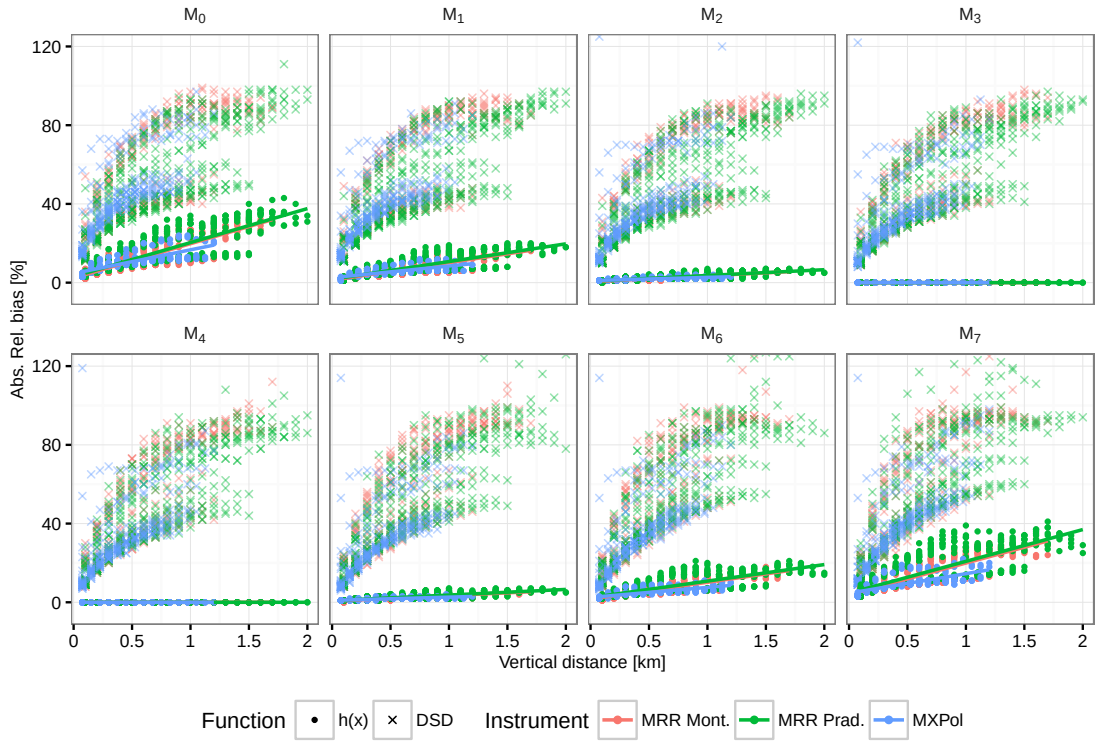


Figure 5.2 – Absolute values of median relative bias between integral moments of normalised DSDs ($h(x)$) by vertical displacement, for $i = 3$ and $j = 4$, with regression lines. Also shown are median points of difference distributions between measured, un-normalised DSDs by displacement distance. The different instrument altitudes and requirement for at least 100 DSDs per height account for the different vertical distances sampled. Symbols as for Figure 5.1. "Mont." stands for Montbrun, "Prad." stands for Pradel Grainage. For readability the y-axis is cut at 120%.

(Lee et al., 2004). The fit was performed only on non-zero values of $h(x)$, and to reduce the influence of outliers it was performed in each case only for values of x between the 10th and 90th percentiles of observed x, inclusive. This constraint meant that values of x near the edges of its observed range, for which there were often fewer data points, were not counted. Minimum allowed parameter values were zero for μ and 0.5 for c; this value of c was chosen to stop the fitted model from producing unreasonably high small-drop concentrations. Examples of generalised gamma models fitted to normalised DSDs are shown in Figure 5.3 for disdrometers and Figure 5.4 for MXPol. Differences in normalised DSD shape between the disdrometers and Doppler-based instrument types are hypothesised to be due to the different DSD truncations and significant differences in measurement principles.

The performance of the model trained at one location was tested by reconstructing the DSD using Equation 5.6 and moment data at another location; the reconstructed DSDs were then compared to measured DSDs. For disdrometer data, the DSDs were reconstructed in the same truncated Parsivel diameter classes. For radar data, the DSDs were reconstructed in the same truncated diameter classes as the DSDs used to calculate the input moments.

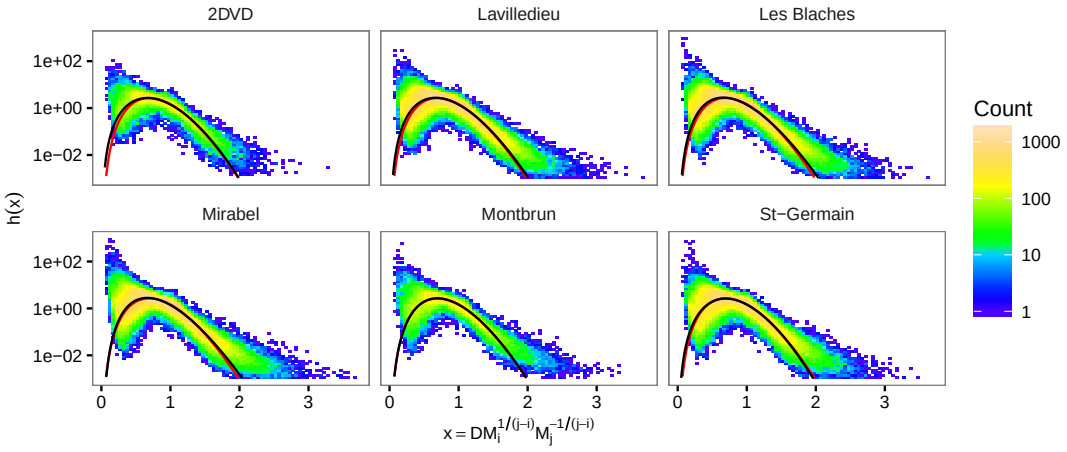


Figure 5.3 – Normalised DSDs ($i = 3,\ j = 4$) for disdrometer data for a subset of disdrometer stations. Black lines show the generalised gamma model fitted to data for each station. In red is the model fitted to data from the station at Montbrun. For visual comparison, only points with $h(x) \geq 10^{-3}$ are shown.

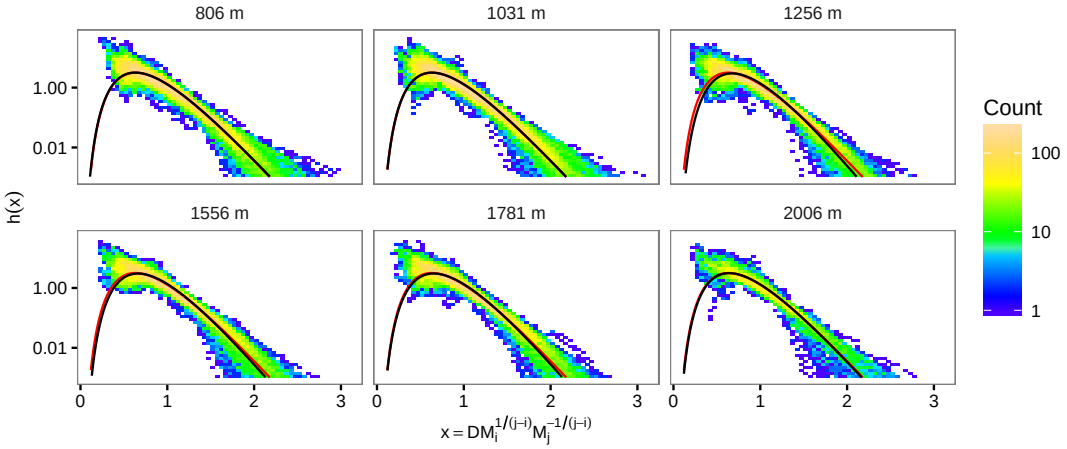


Figure 5.4 – Normalised DSDs ($i = 3,\ j = 4$) for MXPol vertical data for a subset of altitudes. Black lines show the generalised gamma model fitted to data for each altitude. In red is the model fitted to data for altitude 881 m. For visual comparison, only points with $h(x) \geq 10^{-3}$ are shown.

Measured and reconstructed DSD moments were compared per instrument. Figures 5.5 and 5.6 show examples of relative bias as a function of displacement distance in the horizontal and vertical, respectively. Compared to Figures 5.1 and 5.2, in which one-minute measured normalised DSDs were compared, these figures show comparisons of DSDs reconstructed using location-specific climatological normalised DSD models, at displaced locations. In these examples, the regression lines on median relative bias show no obvious trend on performance across horizontal displacement, while in the vertical, performance decreased with distance for moment orders six and seven, and was stable for the other DSD moment orders. IQRs of relative bias are similar across displacement distances.

Tests with all moment combinations showed that the reproduced moments with the greatest slope, and therefore greatest change with distance, were those moments orders far from the input orders, meaning that often the greatest slope was for moment zero or seven. For Parsivel data, the slopes ranged from -1.15 to 0.01 percentage points change in median relative bias per kilometer. These slopes were usually small negative values, but since performance is unlikely to increase with distance, and there were less data available with large distances, we consider the negative slopes to represent negligable changes affected by sampling effects. Bias values were larger for the Iowa network than for Payerne and HyMeX, which may be partly due to the fact that these data were not provided in raw form, and were therefore processed differently to the other two sets. For vertical displacement with MXPol, all moment combinations showed change of under 2.1 % pts km^{-1}. In the MRR data, the change was less than 6.2 % pts km^{-1} for all moment combinations. The fact that results from MXPol show little performance change through vertical displacement means that the small performance changes seen at times in the MRR data may be due to instrument effects rather than variance of the normalised DSD. The overall mean slope for horizontal displacement was -0.04 % pts km^{-1} and for vertical it was 0.57 % pts km^{-1}. This analysis shows that climatological normalised DSDs modelled at one location using a generalised gamma function can feasibly be applied to other locations, with almost no change in DSD-reconstruction performance with horizontal displacement, and acceptable loss of performance with vertical displacement.

5.6 Performance of normalised DSD models

All HyMeX data were used to fit normalised DSD models per instrument type and combination of moment orders. The model parameters found are shown in Table 5.3; these parameters are those used in the rest of this chapter. We turn now to the evaluation of the performance of these models in the different regions studied. For each combination of moment orders, all DSDs were reconstructed using the model appropriate for the instrument, fitted using HyMeX data. Figure 5.7 shows the median relative bias with the largest magnitude per input moment combination and region. The best performance is shown when the two moment orders are not close to each other; for example moments one and six make a good combination. Importantly, the Parsivel model trained in France is able to be applied to measurements from Switzerland and the USA, without large performance loss. The figure shows a performance loss for the

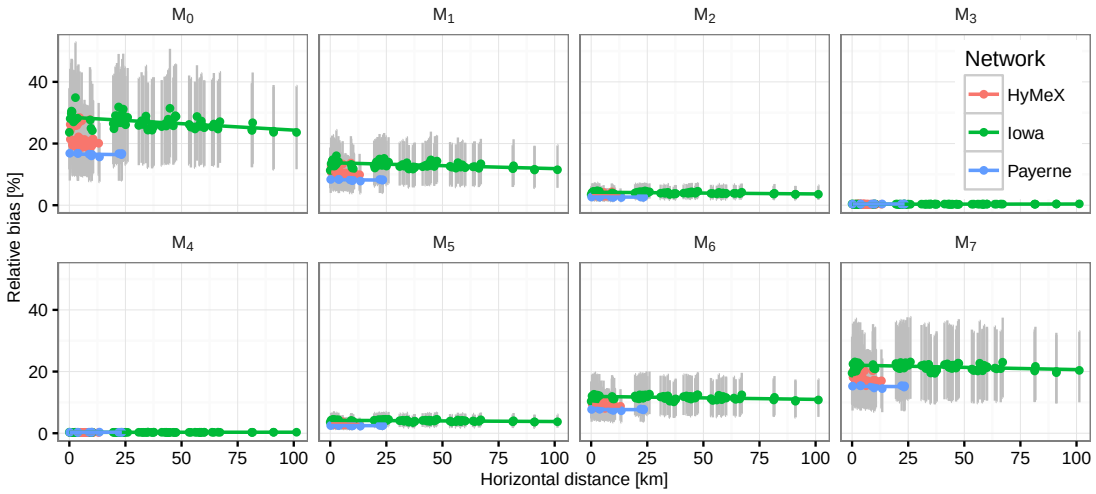


Figure 5.5 – Absolute values of relative bias on DSD moments for the generalised gamma model trained at one location and used at another, by horizontal displacement, for input moment orders $i = 3$ and $j = 4$. Lines and points as for Figure 5.1.

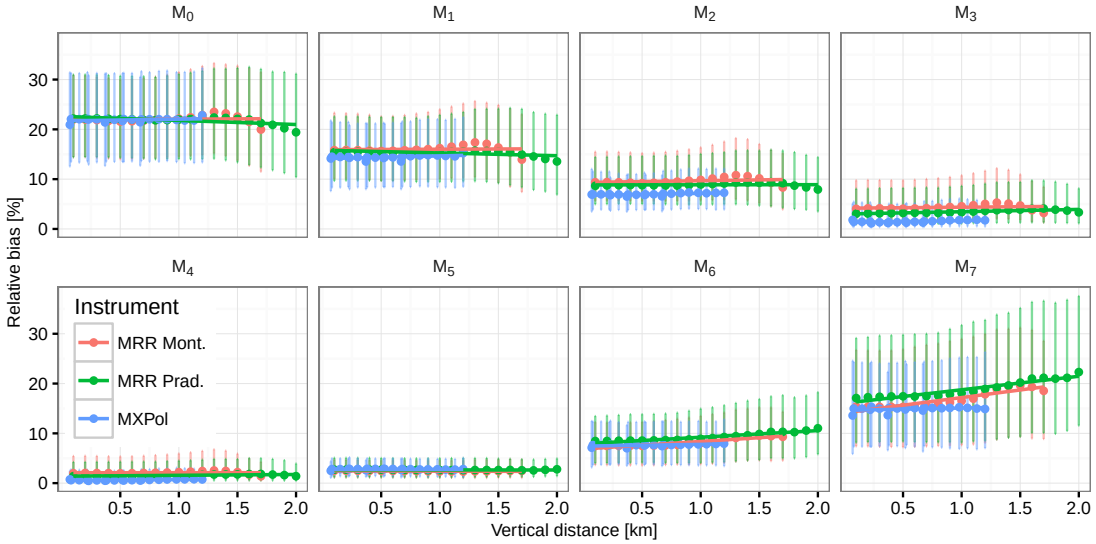


Figure 5.6 – Absolute values of relative bias on DSD moments for the generalised gamma model trained at one location and used at another, by vertical displacement, for input moment orders $i = 3$ and $j = 4$. Lines and points as for 5.1, abbreviations as for Figure 5.2.

Iowa data set, and worse performance for the vertically pointing radars, but we note that those shown in the figure are the worst-case outcomes which usually occurred for moments far from the input moments (for example, the worst relative bias for the Iowa data set was -22% for reconstruction of moment seven with input moments zero and one, but over all tested reconstructions in the Iowa data set the average absolute value of median relative bias was -3%). For the vertically-pointing radars, the worst bias ranged from −28% to 43%, and was usually for moment zero (and the other few times for moment seven), while the average

value of median relative bias over all tested combinations was −5% for the MRRs and −6% for MXPol.

Orders	2DVD			MRR			MXPol			Parsivel (HyMeX)		
	μ	c	RSE	μ	c	RSE	μ	c	RSE	μ	c	RSE
0, 1	22.15	0.63	0.60	21.81	0.66	1.68	22.02	0.64	1.55	22.27	0.62	0.60
0, 2	30.10	0.55	0.58	31.06	0.52	1.64	30.28	0.55	1.44	30.42	0.55	0.59
0, 3	11.59	0.88	0.58	36.19	0.52	1.43	36.31	0.51	1.33	36.45	0.51	0.59
0, 4	7.33	1.09	0.59	38.82	0.50	1.30	38.81	0.50	1.22	16.78	0.74	0.61
0, 5	6.09	1.17	0.61	39.06	0.50	1.19	39.73	0.50	1.11	11.44	0.88	0.64
0, 6	4.62	1.32	0.63	42.54	0.50	1.11	22.50	0.65	1.02	9.33	0.96	0.67
0, 7	3.49	1.49	0.65	43.40	0.50	1.04	9.73	0.92	0.94	7.18	1.07	0.71
1, 2	10.74	0.92	0.58	20.24	0.66	1.47	20.30	0.66	1.36	20.07	0.68	0.60
1, 3	5.63	1.25	0.59	28.18	0.57	1.35	28.16	0.57	1.26	14.37	0.80	0.61
1, 4	4.65	1.35	0.60	34.02	0.53	1.23	33.90	0.53	1.15	8.18	1.05	0.64
1, 5	3.95	1.44	0.62	38.67	0.50	1.12	38.44	0.51	1.05	6.31	1.17	0.67
1, 6	3.51	1.51	0.64	39.33	0.50	1.03	16.62	0.74	0.96	5.40	1.25	0.71
1, 7	2.95	1.63	0.67	40.01	0.50	0.96	7.61	1.01	0.88	4.53	1.34	0.75
2, 3	3.74	1.51	0.61	18.61	0.68	1.26	18.54	0.69	1.18	7.14	1.12	0.65
2, 4	3.08	1.63	0.63	26.21	0.59	1.15	26.03	0.60	1.08	4.85	1.33	0.69
2, 5	2.73	1.70	0.66	31.87	0.54	1.05	25.18	0.62	0.98	3.91	1.45	0.73
2, 6	2.39	1.80	0.68	36.31	0.52	0.96	11.25	0.87	0.90	3.33	1.54	0.77
2, 7	2.02	1.92	0.71	40.09	0.50	0.89	5.60	1.13	0.82	2.81	1.64	0.81
3, 4	2.39	1.81	0.67	17.00	0.70	1.07	16.86	0.73	1.00	3.39	1.55	0.75
3, 5	2.01	1.94	0.70	24.26	0.61	0.97	13.12	0.82	0.91	2.76	1.67	0.80
3, 6	1.66	2.08	0.74	29.64	0.57	0.89	6.49	1.07	0.83	2.28	1.80	0.85
3, 7	1.35	2.25	0.77	34.00	0.54	0.82	3.46	1.33	0.77	1.90	1.92	0.89
4, 5	1.47	2.19	0.76	15.46	0.74	0.90	6.43	1.08	0.84	2.06	1.87	0.87
4, 6	1.15	2.40	0.80	22.42	0.64	0.83	22.81	0.59	0.86	1.67	2.01	0.92
4, 7	0.98	2.54	0.83	15.23	0.75	0.77	1.92	1.57	0.72	1.38	2.15	0.97
5, 6	0.90	2.60	0.85	14.04	0.77	0.78	1.84	1.59	0.72	1.29	2.19	1.00
5, 7	0.76	2.75	0.89	6.39	1.02	0.73	1.08	1.79	0.67	1.06	2.33	1.05
6, 7	0.61	2.94	0.94	3.15	1.25	0.71	0.60	1.96	0.64	0.85	2.49	1.13

Table 5.3 – Generalised gamma model parameters μ and c per moment combination and instrument type, and showing residual standard error (RSE) of the model fit.

Figure 5.8 shows the maximum percentage of variance unexplained (PVU = $(1 - r^2) \times 100$) for each combination of input moments. Again, we see that models fitted on data in France can be applied to the other data sets without significant performance loss. The moment with the most unexplained variance was most often either moment zero or seven. Using the best-performing input moment combination, it was possible that more than 85% of the variance was explained for every reconstructed moment order. The best-performing moment pair differed by instrument type. For example, for the Parsivel data, the best-performing combination was moments one and five, which produced a maximum PVU of 12% across all three regions. For the vertical radars, the best combination was moments three and five, which produced a maximum PVU of 6% over the three instruments. The worst values of PVU were lower for MXPol than the two MRRs. Over all locations, moment combinations, and instruments, the mean PVU was 8%.

As a further check, we tested the rain rate calculated using the reconstructed DSDs against rain

rates calculated from measured DSDs. For reconstructed DSDs in the HyMeX and Payerne Parsivel data sets, the median relative bias was always within five percentage points of zero. In Iowa it was always within nine percentage points of zero and usually within five percentage points of zero. For the vertical radars, median relative bias was usually within five percentage points of zero, and was within 10 percentage points of zero in all cases except for the MRR at Montbrun with moments zero and one, which produced a median relative bias of 12%. IQRs on the median relative bias were often low: for example for input moments three or four combined with orders four, five or six, the IQR was never larger than ten percentage points for any instrument in any region. The IQR rose to a maximum of 48% (for Iowa with M_0 and M_1).

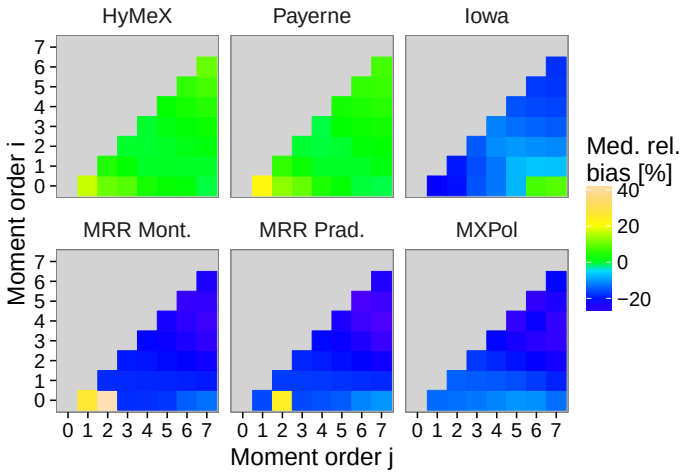


Figure 5.7 – The "worst" median relative bias [%] of all reconstructed versus measured moments, by combination of input DSD moments i and j. The top row shows the Parsivel data sets in the three tested regions, "Mont." indicates Montbrun, "Prad." indicates Pradel Grainage.

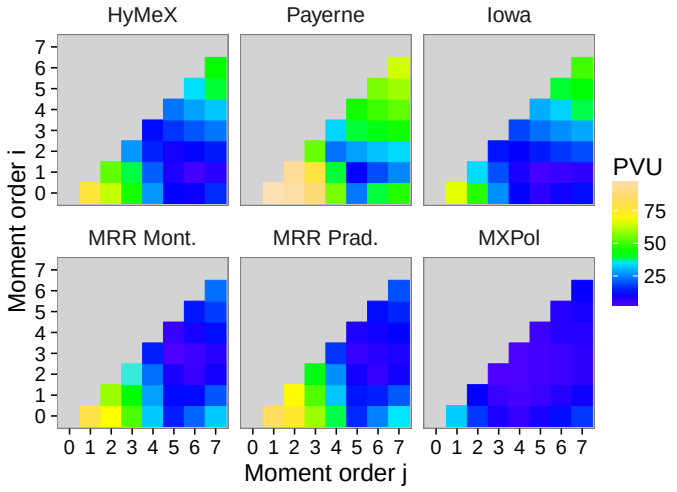


Figure 5.8 – The percentage of variance unexplained (PVU) of reconstructed moments, by combination of input DSD moments i and j. Sub-plots as in Figure 5.7.

The results shown in this section speak to both the best choice of input DSD moments and the invariance of the normalised DSD. The input moments used have a clear effect on the performance of reconstructed moments, with moment orders far from the input moments reconstructed the worst. Input moment orders should thus be chosen to be close to the moment order requiring the best accuracy. The results show that with horizontal displacement, the performance of the double-moment normalised DSD is close to invariant with any choice of input moments, with almost no change in bias on reconstructed DSD moments, even over 100 km horizontal displacement. With vertical displacement, the results show differences in the normalised DSD by altitude for MRR but not MXPol data. When the normalised DSD is used to reconstruct the DSD, however, the performance loss for reconstruction of DSD moments per km of height displacement is usually only a few percentage points. The relative invariance of the normalised DSD is such that it is possible to fit a generalised gamma model to data from France, and use it with data from a different climatic region more than 7000 km away in the USA, with comparable performance. We conclude that for practical use, when input moments are well-chosen, the double-moment normalisation technique is stable enough that the normalised DSD can be considered invariant.

5.7 Discussion

The main conclusion of this work is that the double-moment normalised DSD of Lee et al. (2004) is close to invariant through spatial displacement. We have shown that using a generalised gamma model and the assumption of a normalised DSD, a model trained in one location can be applied in another without significant performance loss. Further, the choice of input moment affects the amount of variability captured, and we quantified this effect through presentation of the worst-case bias and percentage of variance unexplained per moment combination. These results have implications for applications of the double-moment normalised DSD, which include DSD variability studies, DSD-retrieval algorithms, and DSD models used in rainfall retrieval. In this section we provide a short discussion of these implications.

A close to invariant normalised DSD implies that most of the variability of the DSD can be captured by the double-moment normalisation technique. While previous studies have shown that most but not all variability is captured (e.g. Berne et al., 2012; Lee et al., 2007), we have quantified the amount of variability captured by combination of input moments, and have shown that it is possible to explain more than 85% of DSD variability with the double-moment technique, with more than 90% explained on average. This is true even when the normalised DSD model has been trained in one location and applied in another. It is possible to envision, then, a functional mapping between variability in two bulk variables of rainfall and whole-DSD variability. Such a system would allow for studies of DSD variability in areas in which limited data are available.

The assumption that the normalised DSD is invariant could be used in any application in which the DSD must be reconstructed from limited information, such as in DSD retrieval.

Retrieval of the DSD from polarimetric data has been a long-standing research goal. Since it was shown that radar reflectivity corresponds to the sixth moment of the DSD (Marshall and Palmer, 1948), the advent of polarimetric radar has allowed other microphysical properties to be linked to radar measurements (Seliga and Bringi, 1976; Jameson, 1983, 1985; Bringi and Chandrasekar, 2001). As suggested by Lee et al. (2004), radar measurements could be related to moments of the DSD to apply the double-moment normalisation to radar remote sensing. We have shown that this idea is indeed feasible, because the normalised DSD can be assumed invariant for practical purposes. Radar variables correspond to higher moments of the DSD Bringi and Chandrasekar (2001), and the work in this chapter shows that it should be expected, then, that the largest errors in such an application would be for the zeroth DSD moment.

In areal rainfall retrieval, it is already common for two moments of the DSD to be used, but DSD models in use may be less flexible and able to capture less DSD variability than the double-normalised DSD. For example, the operational one-moment DSD scheme in the consortium for small-scale modelling (COSMO) numerical weather prediction model uses a gamma DSD (Ulbrich, 1983) with two fixed parameters (Baldauf et al., 2011). It has been shown, in Chapter 4 of this thesis, that this algorithm is susceptible to errors in estimation of DSD properties, which are caused primarily by the fixed values of DSD model shape and intercept parameters. While a sophisticated two-moment scheme is available for COSMO (Seifert and Beheng, 2006), it has been found to have limited performance benefit, and higher computation cost than the one-moment scheme (Baldauf et al., 2011). The assumption of an invariant normalised DSD allows for further research to test whether the scheme is of use to such numerical weather models, as well as areal rainfall retrieval algorithms similar to those used by the global precipitation measurement (GPM) mission (Hou et al., 2014).

Establishing links between microphysical processes in precipitation and changes in the shape of the double-normalised DSD remains an open research topic (Lee et al., 2004). While our work has shown that changes in $h(x)$ due to spatial displacement are small and can be quantified, it is likely that some of the observed changes in $h(x)$ in the vertical are due to microphysical processes that occur as the rain falls. Further studies on rainfall classified by the dominant microphysical process are required to further analyse the causes of these differences.

5.8 Conclusions

The double-moment normalisation of Lee et al. (2004) provides a convenient way to summarise the DSD in a compact form. The assumption made is that the normalised DSD is the same everywhere; if this is the case then the variability of the DSD can be explained through variability of only two of its moments. In this work we tested this assumption. Data from networks of disdrometers and radars in three different climatic regions – Ardèche in France, Payerne in Switzerland, and Iowa in the USA – were used. The disdrometer networks were used to test the invariance of the normalised DSD through (close to) horizontal displacement, and

vertically profiling radars in one data set were used to test its invariance in the vertical. The normalised DSD was tested with all combinations of moment orders between one and seven. The normalised DSD was found to be almost invariant through horizontal displacement, and to show moderate changes through vertical displacement.

Normalised DSDs found at each location were tested against data at other locations. Testing of reconstructed DSDs using displaced normalised DSDs showed little or no performance loss across horizontal displacement, and performance loss of only a few percentage points per kilometre across vertical displacement. We conclude that for practical use, when the input moments are well chosen, the double-moment normalised DSD can be considered invariant. Double-moment normalisation of the DSD is a technique with wide applicability. The results we have shown demonstrate the relative invariance of the normalisation of Lee et al. (2004), across horizontal and vertical displacement. Normalised DSD models trained in France were able to be applied in Switzerland and the USA with relatively good performance. The normalisation is thus well suited for use in studies of DSD variability, DSD retrieval, and areal rainfall estimation. Future work should investigate the small amount of remaining variability in the double-moment normalised DSD and its links to the microphysical processes of precipitation.

6 Retrieval of the raindrop size distribution from polarimetric radar data using double-moment normalisation

This chapter is adapted from:

- T. H. Raupach and Alexis Berne. Retrieval of the raindrop size distribution from polarimetric radar data using double-moment normalisation. Submitted to *Atmospheric Measurement Techniques*.

This work was completed by T. Raupach under the supervision of A. Berne. Research, analyses, and writing are by T. Raupach. For data acknowledgements, see Appendix A.

6.1 Introduction

For applications such as numerical weather prediction (e.g. Baldauf et al., 2011) or radar remote sensing (e.g. Bringi and Chandrasekar, 2001) it is often necessary to know the areal DSD at the pixel scale. In other cases, such as studies of the microphysics of precipitation (Pruppacher and Klett, 2000), it would be useful to be able to remotely infer the DSD aloft or in remote locations. For these reasons, retrieval of the DSD from radar data has been a long-standing goal. Polarimetric weather radars are particularly useful for remote retrieval of the DSD, because differences between vertically and horizontally polarised electromagnetic waves reflected off hydrometeors in the atmosphere provide information on the particles' concentration, size, and shape. In rainfall, radar reflectivity in horizontal (Z_H [dBZ]) or vertical (Z_V [dBZ]) polarisation primarily relates to drop concentration and size. Differential reflectivity (Z_{DR} [dB]) reflects drop shape, and specific differential phase shift (K_{dp} [° km^{-1}]) relates to both the concentration and shape of the drops (Bringi and Chandrasekar, 2001). In this chapter we present a new technique for DSD retrieval from polarimetric radar data, which is based on the double-moment normalisation technique of Lee et al. (2004).

Seliga and Bringi (1976) showed that Z_{DR} can be linked to the median volume drop diameter, a microphysical property of rain. Since then, many methods for DSD retrieval from radar variables have been proposed. Zhang et al. (2001) introduced the "constrained gamma" method, in which the shape and slope parameters of a gamma DSD model (Ulbrich, 1983) are assumed dependent. This assumption is subject to debate (e.g. Zhang et al., 2003; Atlas and Ulbrich, 2006; Moisseev and Chandrasekar, 2007; Cao and Zhang, 2009). The technique, modified by Brandes et al. (2003), can provide useful DSD information (Brandes et al., 2004b). In the "beta" method (Gorgucci et al., 2002), the effective slope of the drop axis ratio to diameter relationship is retrieved. The slope is used to find parameter values for the normalised gamma model of Willis (1984), which has advantages for use with polarimetric observations (Illingworth and Blackman, 2002). Retrieval of the gamma model shape parameter with the beta method is subject to high uncertainty (Gorgucci et al., 2002; Anagnostou et al., 2008). To deal with noisy Z_{DR} and K_{dp} data at low rain rates, Bringi et al. (2002, 2003) used the beta method for heavy rain and disdrometer-based regressions on Z_H and Z_{DR} for light rain. Brandes et al. (2004a) found that the constrained gamma method was in better agreement with disdrometer data than the beta method, while Anagnostou et al. (2008) reported similar performance from the two techniques, and both studies noted that the beta method is sensitive to errors in K_{dp}. Vulpiani et al. (2006) developed a neural-network DSD-retrieval technique, and spatial correlations of DSD model parameters have been retrieved from radar data (Thurai et al., 2012; Bringi et al., 2015).

X-band polarimetric weather radars are popular due to their portability, small size, and high resolution and sensitivity, but measurements at X-band suffer from attenuation by heavy rain (Kalogiros et al., 2013; Anagnostou et al., 2013) and must be corrected (Matrosov et al., 2005; Park et al., 2005a). Several DSD-retrieval algorithms have been developed for X-band (e.g. Park et al., 2005b; Gorgucci et al., 2008; Kalogiros et al., 2013; Anagnostou et al., 2013),

including some with integrated attenuation correction (e.g. Testud et al., 2000; Yoshikawa et al., 2014). The self-consistent with optional parameterization attenuation correction and microphysics estimation (SCOP-ME) algorithm, developed through studies by Anagnostou et al. (2009, 2010) and Kalogiros et al. (2013), uses relationships calculated for the Rayleigh limit, corrected for Mie scattering at X-band. It performs well compared to contemporary algorithms and disdrometer observations (Anagnostou et al., 2013). In this chapter we present a new method for DSD retrieval that uses the double-moment DSD normalisation of Lee et al. (2004), and compare it to SCOP-ME.

The rest of this manuscript is organised as follows: two polarimetric variables used in this chapter are introduced in Section 6.2. Data used are presented in Section 6.3. In Section 6.4 we propose a new DSD-retrieval method that uses double-moment normalisation to retrieve the DSD from polarimetric radar data. Its performance is compared to that of SCOP-ME using radar variables simulated from DSD measurements in Section 6.5. In Section 6.6 we introduce a new method to reduce the effects of noise in radar measurements. Using this method, the DSD-retrieval algorithms are compared using radar data in Section 6.7. Conclusions are made in Section 6.8.

6.2 Polarimetric variables

Recall from Chapter 1 that all bulk variables of rainfall can be derived from the DSD. These include polarimetric radar variables such as the radar reflectivity (Equation 1.7). In this section some more useful polarimetric variables are introduced (a detailed review is provided by Bringi and Chandrasekar, 2001). Differential reflectivity in linear units, ξ_{dr} [–], defined as Z_h/Z_v, has been shown to relate to the reflectivity-weighted mean drop axis ratio r_z (Jameson, 1983). r_z is defined as

$$r_z = \frac{\int_0^\infty r(D) D^6 N(D) dD}{\int_0^\infty D^6 N(D) dD}, \tag{6.1}$$

where $r(D)$ is the vertical to horizontal axis ratio of a drop of equivolume diameter D. The relationship found by Jameson (1983) is

$$r_z \sim (\xi_{\mathrm{dr}})^{-\frac{3}{7}}, \tag{6.2}$$

which is valid for narrow distributions of raindrop axis ratio (Bringi and Chandrasekar, 2001).

Dual-polarisation radars measure specific differential phase shift (on propagation), K_{dp} [° km^{-1}], which is the difference in phase change between horizontally and vertically polarised waves that pass through one kilometre of rain. It is defined as (Bringi and Chandrasekar, 2001)

$$K_{\mathrm{dp}} = \frac{180\lambda}{\pi} 10^{-1} \int_0^\infty \mathrm{Re}\left[f_{hh}(D) - f_{vv}(D)\right] N(D) dD, \tag{6.3}$$

where Re represents the real part of a complex number and Re(f_{hh}) [cm] and Re(f_{vv}) [cm] are the real parts of the forward scattering amplitudes for horizontal and vertical polarisation respectively. K_{dp} can also be defined from the DSD as (Jameson, 1985)

$$K_{\mathrm{dp}} = \left(\frac{180}{\lambda}\right) 10^{-1} CW(1 - r_m), \tag{6.4}$$

where the dimensionless value $C \sim 3.75$ (Bringi and Chandrasekar, 2001). r_m is the mass-weighted mean raindrop axis ratio, defined as (Jameson, 1985)

$$r_m = \frac{\int_0^\infty r(D) D^3 N(D) dD}{\int_0^\infty D^3 N(D) dD}. \tag{6.5}$$

Various axis ratio functions are available (e.g. Pruppacher and Beard, 1970; Andsager et al., 1999; Brandes et al., 2002; Thurai and Bringi, 2005). Where unspecified, the ratio function used in this chapter was that of Thurai and Bringi (2005). We return to the question of axis ratios and K_{dp} in Section 6.4. Just as previously in Sections 1.1 and 5.2, the integrals in this section are idealised because the range of drop sizes is written from zero to infinity, and in practice a sum over discrete classes of D or x is used. We used the same truncation limits for compared quantities. When polarimetric variables were calculated from DSDs, the T-matrix codes of Mishchenko and Travis (1998) were used. Unless specified otherwise, the codes were used with an assumed temperature of 12.5° C, a Gaussian distribution of raindrop canting angles with zero mean and a standard deviation of 6° (stated as reasonable by Bringi and Chandrasekar, 2001), and a radar frequency of 9.4 GHz.

6.3 Data

To test the new method, data from the three networks of OTT Parsivel (Löffler-Mang and Joss, 2000) disdrometers introduced in Section 5.3 were used. Each network had a nearby X-band weather radar that scanned above the disdrometers. From the HyMeX network (introduced in Section 2.2.1), the test data included measurements from 11 of the disdrometers, the MRR at Pradel Grainage, and MXPol. The MRR provided vertical profiles of DSD estimations recorded with 100 m vertical resolution and 10 s integration time. MXPol was located to the north-east of the disdrometer network (Figure 2.1). In 2013, MXPol recorded "stacked" plan position indicator (PPI) scans above the Parsivel network at elevations of four, five, six, eight, 10, 12, 14, 16, and 20 degrees above horizontal, with a return time of about six minutes. For the MRR test data we used the events identified in Section 5.3 and shown in Table 5.1. As before, only those MRR data below the estimated freezing level were used for each event.

The second network used was the Payerne data set, and included the MXPol radar as well as Parsivel disdrometers. We used the MXPol PPI scan at five degrees above horizontal, which had a return time of about five minutes. The third data set used was from ten Parsivel^2 disdrometers (Petersen et al., 2014) in Iowa data set. Overlooking this network was the University of Iowa's X-band radar XPOL5 (Mishra et al., 2016). We used PPI data recorded at three degrees above horizontal, with a return time of about eight minutes, for three days of heavy rainfall, the 25th, 26th, and 27th May 2013. The three networks were in regions with different climatologies (as described in Wolfensberger et al., 2015). Table 6.1 shows the instrument stations used here, the distance of each station to the PPI radar volumes used, and the number of radar scans that overlapped with one-minute observations. Table 5.2 shows the spatial coordinates for the stations used here.

Disdrometer and MRR DSD data were truncated and treated as shown in Section 5.3. All available disdrometer and PPI data were used, while MRR data were subset to HyMeX event times so that likely solid precipitation was not considered. PPI radar reflectivities were compared to measurements from disdrometers (and the MRR in HyMeX), and bias in Z_H was corrected for each of the three locations. Two days of radar data from Payerne (2014-03-22 and 2014-04-08) exhibited higher radar bias and were not included in this study. Attenuation correction for the PPI data was performed using the ZPHI algorithm (Testud et al., 2000), and K_{dp} was estimated using the method of Schneebeli et al. (2014). PPI scan data were sampled by taking the mean values for radar volumes that overlapped horizontally the latitude and longitude of the point in question within the instrument's corresponding one-minute integration period. To discount noise, PPI values were subset to those for which Z_H was greater than or equal to 10 dBZ and the signal to noise ratio in horizontal polarisation was greater than or equal to 5 dB.

Network	Station	S	H (ground)	H (a.s.l)	D	MI
Payerne	HARAS Avenches	481	910	1350	9.8	15.2
	Military Airport Payerne	406	370	820	3.7	16.3
	Morat Airport	326	2090	2520	23.2	16.6
HyMeX	Lavilledieu	1159	970	1190	8.4	53.4
	Les Blaches	1218	550	980	5.4	60.1
	Lussas	1234	730	1020	6.0	65.1
	Mirabel	1222	370	870	3.8	56.8
	Mont-Redon	1226	140	780	2.5	15.9
	Pradel 1	1202	680	960	5.1	39.1
	Pradel Grainage	1179	700	970	5.3	42.8
	Pradel-Vignes	1185	730	990	5.5	22.0
	Saint-Etienne-de-Fontbellon	1008	1210	1520	13.1	40.7
	St-Germain	1089	1100	1310	10.1	73.4
	Villeneuve-de-Berg	1108	840	1140	7.7	62.0
Iowa	apu05	85	1520	1810	29.5	44.5
	apu06	84	1570	1840	30.1	52.0
	apu07	80	1660	1930	31.9	38.1
	apu08	87	1570	1850	30.3	59.7
	apu09	158	700	940	12.9	40.0
	apu10	164	640	890	12.0	22.9
	apu11	155	600	860	11.4	20.8
	apu12	154	540	801	10.3	27.8
	apu13	96	1730	1920	31.7	60.1
	apu14	95	1730	1920	31.7	65.2

Table 6.1 – Instrument stations with corresponding PPI volumes, with the number of scans for that volume (S), and the volume centre's height above the ground (H (ground) [m], to nearest 10 m), the volume centre's height above sea level (H (a.s.l) [m], to nearest 10 m), and its horizontal range from the radar (D [km]). Also shown is the maximum one-minute rain intensity (MI [mm h^{-1}]) recorded by each Parsivel at a radar scan time. The radar locations were 44.6141° N, 4.5461° E, 602 m altitude for HyMeX, 46.8133° N, 6.9428° E, 489 m altitude for Payerne, and 41.8870° N, 91.7341° W, 263 m altitude for Iowa. The MRR was at Pradel Grainage. Military Airport is short for Military Airport Payerne.

6.4 DSD retrieval from polarimetric radar data

In Chapter 5 it was shown that for practical purposes, the double-normalised DSD can be assumed to be invariant across spatial displacement. Lee et al. (2007) showed that $h(x)$ derived from time series measurements at one location had low scatter around the average normalised DSD. In this chapter, we make the assumption that the double-moment normalised DSD function $\hat{h}(x)$ is invariant in space and time over the typical domain of interest. Once again we use the generalised gamma DSD model and the double-moment normalised DSD shown in Equation 5.6. Using this assumption, the DSD can be reconstructed at a point in space using polarimetric radar data. Given a known normalised DSD, the task of DSD reconstruction becomes that of determining from radar information the values of two DSD moments. In this section we present a new DSD-retrieval method that uses this idea. The aim of the proposed DSD-retrieval technique is to retrieve two DSD moments, using only polarimetric radar data. We used disdrometer data to simulate radar variables and estimate their relationships to DSD moments. The disdrometer data used were the Parsivel data from the HyMeX network (101494

one-minute DSDs). Z_H, K_{dp} and Z_{DR} were calculated for these DSDs for the MXPol stacked PPI incidence angles, temperatures of five, ten, and 15° C, and each of four drop axis ratio functions: those of Andsager et al. (1999), Brandes et al. (2002), Thurai and Bringi (2005), and that of Beard and Chuang (1987) in the form shown in Kalogiros et al. (2013).

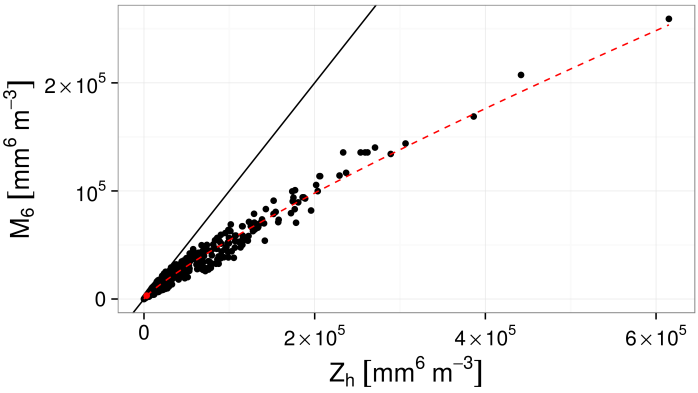


Figure 6.1 – A sample of 20,000 points from the training set of Parsivel data, showing the relationship between radar reflectivity in linear units and moment six of the DSD. Each black point represents a one-minute DSD measurement. The one-to-one line is shown in black, and the red dashed line shows the fitted relationship of Equation 6.6. The low-high Z_h threshold of 35 dBZ is shown with a red point.

Radar reflectivity in linear units, Z_h [mm^6 m^{-3}], is the sixth moment of the DSD in the Rayleigh scattering regime (Bringi and Chandrasekar, 2001). At X-band frequencies, larger drops enter into the Mie scattering regime and differences appear between M_6 and Z_h. We use the observation that Z_h departs from M_6 for heavier rain, and assume that this departure occurs when Z_H is greater than some value. This threshold was determined through comparison of M_6 and Z_h for DSDs with Z_H in classes of width 1 dBZ between 30 dBZ and 40 dBZ, and was set to 35 dBZ. For larger reflectivity values, a power law relationship was found using least-squares. The resulting relationship is

$$\widehat{M_6} = \begin{cases} Z_h & \text{if } 10\log_{10}(Z_h) \leq 35 \\ 3.2\, Z_h^{0.85} & \text{if } 10\log_{10}(Z_h) > 35. \end{cases} \tag{6.6}$$

On the training set, relative bias between $\widehat{M_6}$ and M_6 was −2.7%, the interquartile range (IQR) of relative bias was 2.9 percentage points, and the r^2 value was 0.98. The fitted relationship is shown on samples of training data in Figure 6.1. Retrieving a second, lower-order DSD moment is more difficult than estimating M_6, because radar variables are more closely linked to the higher-order moments of the DSD. Using theoretical relationships as much as possible, we present here a method to retrieve the third moment of the DSD from polarimetric data.

As shown in Equation 6.2, the reflectivity-weighted mean drop axis ratio, r_z, is related to a negative power of the differential reflectivity in linear units, ξ_{dr}. In Kalogiros et al. (2013), the

reflectivity-weighted and mass-weighted drop axis ratios were assumed to be the same and differences were dealt with through least-squares fitting of qualitative relationships between radar variables. A similar approach is taken here. Since r_z and the mass-weighted mean drop axis ratio r_m are both weighted mean drop axis ratios, we assume that r_m can also be approximated through a negative power of ξ_{dr}, such that $r_m \sim \xi_{\mathrm{dr}}^{-\beta_M}$. Recall from Equation 1.3 that W is the third moment of the DSD multiplied by a constant. Replacing W, r_m, and constants in Equation 6.4 with M_3 (to be estimated), $\xi_{\mathrm{dr}}^{-\beta_M}$, and a constant α_M respectively, we arrive at the expression

$$\hat{M}_3 = \frac{\alpha_M K_{\mathrm{dp}}}{(1-\xi_{\mathrm{dr}}^{-\beta_M})}, \tag{6.7}$$

where α_M and β_M are parameters to be found. K_{dp} is sensitive to the raindrop axis ratio (e.g. Bringi and Chandrasekar, 2001), so this relationship was found per axis ratio function using least-squares fitting. The same Z_H threshold was used to divide the data into "Rayleigh-like" and "Mie-like" sets. The results and their performance statistics are shown in Table 6.2.

	Low Z_H		High Z_H				
Ratio	α_M	β_M	α_M	β_M	RB	IQR	r^2
Thurai	1524	6.8	730	2.5	2.5	13	0.97
Brandes	1332	4.3	684	2.0	1.4	23	0.96
Andsager	1623	6.1	803	2.5	0.6	20	0.96
Beard	1131	3.9	672	2.1	0.7	21	0.97

Table 6.2 – Fitted values of α_M and β_M for DSDs with $Z_H \leq 35$ dBZ (Low Z_H) and those with $Z_H > 35$ dBZ (High Z_H), by drop axis ratio function (Ratio). Also shown are the resulting median relative bias (RB [%]), IQR of relative bias (IQR [% pts]), and r^2 on the training data.

The proposed DSD retrieval technique can be summarised as follows: the double-normalised DSD $\hat{h}(x)$ and its parameters c and μ are assumed trained from data and known. Then, given K_{dp}, ξ_{dr} and Z_h, (1) DSD moment six is estimated using Equation 6.6; (2) DSD moment three is estimated using Equation 6.7 and parameter values from Table 6.2. The DSD is then retrieved using Equation 5.6 with $i = 3$ and $j = 6$.

6.5 Comparison to an existing DSD-retrieval method

The new DSD retrieval method was compared to the DSD-retrieval method SCOP-ME (Anagnostou et al., 2009, 2010; Kalogiros et al., 2013). We implemented SCOP-ME using its description in Anagnostou et al. (2013). SCOP-ME was developed for X-band using T-matrix simulations, and in Anagnostou et al. (2013) it is shown to outperform the algorithms of Anagnostou et al. (2008) and Park et al. (2005a). The DSD model used by SCOP-ME is based on the normalised DSD of Willis (1984) (see also Bringi and Chandrasekar, 2001). Kalogiros

et al. (2013) provided an explicit expression for rain rate using polarimetric variables, but since we are interested in the whole DSD, in the following sections we compare R computed from reconstructed DSDs. The comparison of the two methods is first shown using Parsivel data in which the radar values were simulated using T-matrix codes and were therefore free of radar measurement noise.

Comparisons of the two techniques (both in this section and Section 6.7) were made using the three Parsivel data sets from HyMeX, Payerne, and Iowa. Comparison statistics were computed with difference defined as retrieved minus measured value, for DSD moments zero to seven, D_m, and R. For each one-minute DSD record, Z_h, K_{dp} and Z_{DR} were calculated using T-matrix codes, for an elevation angle of 4° above horizontal, and using each of the four drop axis ratio functions. SCOP-ME and the double-moment method were used to retrieve the DSD concentrations $N(D)$ for D in the class centres of the truncated Parsivel diameter classes. For the double-moment technique, the generalised gamma model parameters for the normalised DSD $\hat{h}(x)$ (Equation 5.6) for Parsivel data with $i = 3$ and $j = 6$ were used. As found in Chapter 5, these values were $c = 1.8$ and $\mu = 2.28$. The HyMeX data set is used as an example: measured and retrieved rain rates are shown for one event in Figure 6.2, for the drop axis ratio model of Beard and Chuang (1987), with which SCOP-ME was trained. Scatter plots of measured and reconstructed values for the Beard axis ratio model are shown in Figure 6.3. Comparisons of distributions of relative errors are shown in Figure 6.4.

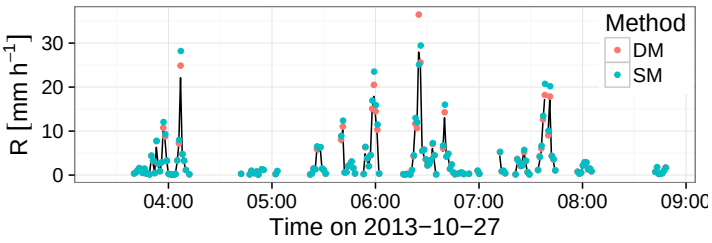


Figure 6.2 – A time series plot showing retrieved versus measured rain rate for double-moment (DM) and SCOP-ME (SM) methods, on for the fourth tested HyMeX event, using the axis ratio function of Beard and Chuang (1987). Measured rain rates by Parsivel are shown as a black line, retrieved values as coloured points.

Performance results are shown in detail for the HyMeX region in Table 6.3. Differences in performance statistic are shown by ratio function, variable, and region in Table 6.4. These differences are shown visually in Figure 6.5. In over half of the tested region and variable combinations (moments one to seven, R, and D_m), the double-moment technique produced a better median relative bias than the SCOP-ME technique; on average the double-moment technique produced a median relative bias that was 2 percentage points better. IQR of relative bias was usually slightly higher for the double-moment technique, with a mean difference of 2.4 percentage points. Correlation coefficients and scatter plot slopes were usually very similar between the two techniques. On average, the double-moment results produced an r^2 value and slope that were respectively 0.02 and 0.03 further from one than the SCOP-ME values.

The average differences across the three tested regions and four tested raindrop axis ratio

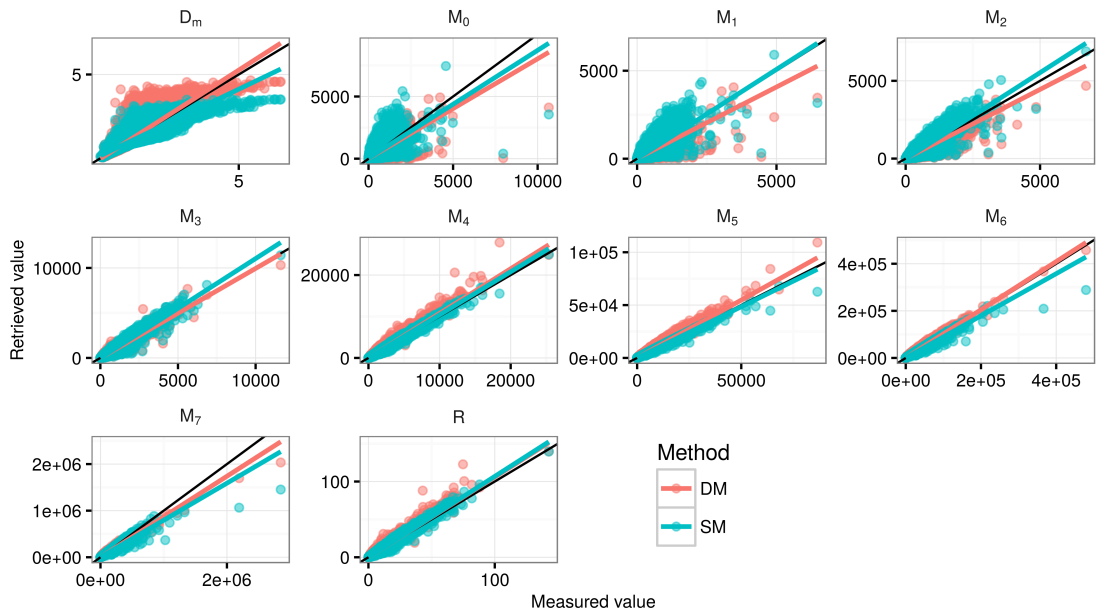


Figure 6.3 – Scatter plots showing retrieved versus measured moments M_n [mm^n m^{-3}], R [mm h^{-1}], and D_m [mm], for double-moment (DM) and SCOP-ME (SM) methods, on HyMeX data only, using the axis ratio function of Beard and Chuang (1987). One-to-one lines are shown in black, lines of best fit in colour.

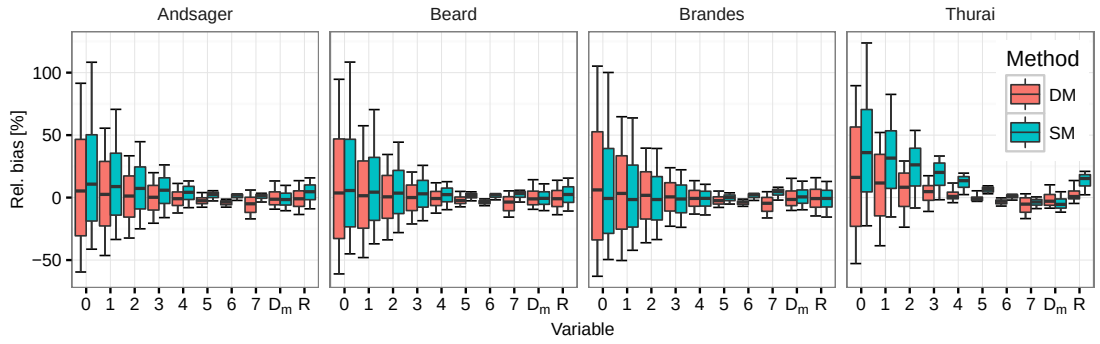


Figure 6.4 – Relative bias distributions from the double-moment and SCOP-ME DSD-retrieval methods, by drop axis ratio function, for HyMeX data. Variables are moment order n [mm^n m^{-3}], D_m [mm], and R [mm h^{-1}]. Bold bars show medians, boxes show IQRs, whiskers show 10th to 90th percentile ranges.

		Double-moment				SCOP-ME				Difference			
Ratio	Var	RB	IQR	r^2	S	RB	IQR	r^2	S	RB	IQR	r^2	S
Ands.	D_m	−1	10	0.85	1.05	−2	9	0.87	0.82	−0	1	0.02	−0.13
	M_0	5	77	0.67	0.80	11	69	0.65	0.89	−5	8	−0.01	0.08
	M_1	3	52	0.77	0.81	9	50	0.79	1.01	−6	2	0.02	0.18
	M_2	1	33	0.89	0.88	7	34	0.90	1.07	−6	−1	0.01	0.05
	M_3	0	20	0.96	0.98	6	21	0.96	1.05	−6	−1	0.00	−0.03
	M_4	−1	11	0.98	1.07	4	11	0.99	0.98	−3	0	0.01	0.05
	M_5	−3	5	0.98	1.09	3	4	0.99	0.93	0	0	0.02	0.02
	M_6	−4	3	0.98	1.02	2	3	0.97	0.87	2	0	−0.01	−0.11
	M_7	−5	12	0.98	0.86	2	3	0.93	0.79	3	9	−0.05	−0.07
	R	−1	12	0.98	1.05	5	12	0.98	1.01	−4	1	0.01	0.04
Thur.	D_m	−3	9	0.88	1.07	−5	7	0.89	0.84	−3	1	0.01	−0.10
	M_0	16	80	0.70	0.97	36	66	0.74	1.15	−20	14	0.04	−0.11
	M_1	12	49	0.80	0.89	32	46	0.84	1.21	−20	3	0.04	−0.10
	M_2	8	27	0.91	0.91	26	30	0.92	1.22	−18	−4	0.01	−0.14
	M_3	5	12	0.98	1.00	20	18	0.97	1.17	−15	−6	0.00	−0.17
	M_4	1	5	0.99	1.07	14	9	0.99	1.08	−13	−3	0.01	−0.01
	M_5	−1	3	0.98	1.07	7	4	0.99	0.99	−5	−1	0.01	0.06
	M_6	−3	3	0.99	0.98	2	2	0.97	0.90	1	1	−0.01	−0.08
	M_7	−5	12	0.98	0.82	−4	5	0.93	0.79	2	7	−0.05	−0.03
	R	1	6	0.99	1.05	15	9	0.99	1.12	−14	−3	0.00	−0.06
Bran.	D_m	−1	12	0.83	1.06	1	11	0.86	0.81	1	1	0.02	−0.13
	M_0	6	87	0.63	0.79	−1	68	0.59	0.80	5	19	−0.03	0.00
	M_1	3	59	0.74	0.79	−2	50	0.75	0.93	2	9	0.01	0.14
	M_2	2	38	0.87	0.86	−2	35	0.87	1.02	0	3	0.00	0.11
	M_3	1	23	0.96	0.97	−1	22	0.95	1.03	−1	0	0.00	0.00
	M_4	−1	13	0.98	1.07	−1	12	0.99	0.98	0	1	0.01	0.05
	M_5	−2	6	0.98	1.11	−0	5	0.99	0.92	2	1	0.01	0.03
	M_6	−4	3	0.98	1.05	2	3	0.97	0.87	1	−0	−0.01	−0.08
	M_7	−5	11	0.98	0.91	5	5	0.92	0.78	−0	7	−0.06	−0.13
	R	−1	14	0.98	1.04	−1	13	0.98	1.00	0	1	0.00	0.04
Beard	D_m	−1	11	0.85	1.06	−1	10	0.87	0.79	0	1	0.02	−0.15
	M_0	4	80	0.66	0.80	6	70	0.61	0.87	−2	10	−0.05	0.07
	M_1	1	54	0.77	0.81	4	50	0.76	1.01	−3	3	−0.01	0.17
	M_2	1	34	0.89	0.89	4	35	0.89	1.11	−3	−0	0.00	0.00
	M_3	0	20	0.97	0.99	3	21	0.96	1.11	−3	−1	0.00	−0.10
	M_4	−1	11	0.98	1.08	2	11	0.99	1.04	−2	0	0.01	0.04
	M_5	−2	5	0.98	1.10	2	4	0.99	0.97	1	2	0.01	0.06
	M_6	−3	2	0.98	1.02	2	3	0.97	0.89	1	−0	−0.01	−0.08
	M_7	−4	11	0.98	0.87	3	5	0.93	0.79	0	7	−0.05	−0.07
	R	−1	13	0.98	1.06	2	12	0.99	1.07	−2	0	0.00	−0.01

Table 6.3 – Comparison of double-moment method to SCOP-ME results on Parsivel data from five rainfall events in the HyMeX data set by axis ratio function (Ratio). Med. RB is median relative bias [%], RB IQR is interquartile range of relative bias [% points], r^2 is squared correlation coefficient. Slope is slope of best fit line on measured vs. reconstructed plot. Difference is difference in absolute values for RB and IQR, and difference in distance from 1 for r^2 and slope. A negative difference shows that the double-moment method improved on SCOP-ME's performance.

		Payerne				Iowa			
Ratio	order	RB	IQR	r^2	S	RB	IQR	r^2	S
Ands.	D_m	0	1	0.02	0.08	1	2	0.04	−0.03
	M_0	1	18	0.16	0.14	13	0	−0.04	0.12
	M_1	−0	6	0.08	0.16	10	−0	0.05	0.33
	M_2	−1	1	0.01	0.12	8	−1	0.03	0.06
	M_3	−2	−1	0.00	0.00	4	−0	0.04	0.07
	M_4	−2	0	0.02	0.09	0	2	0.03	0.18
	M_5	−0	1	0.01	−0.01	1	2	0.01	0.06
	M_6	1	0	0.01	−0.16	2	0	−0.01	−0.14
	M_7	3	8	0.03	−0.01	2	12	−0.02	0.01
	R	−2	0	0.02	0.08	0	2	0.04	0.21
Thur.	D_m	−2	1	0.02	0.11	−3	2	0.03	0.04
	M_0	−11	15	0.19	0.25	−5	9	0.07	−0.26
	M_1	−15	3	0.09	0.19	−12	3	0.11	−0.06
	M_2	−15	−3	0.01	0.04	−17	−3	0.03	−0.10
	M_3	−14	−5	0.00	−0.09	−15	−6	0.03	−0.08
	M_4	−11	−4	0.01	0.08	−10	−3	0.02	0.12
	M_5	−4	−2	0.00	0.02	−4	0	0.00	0.07
	M_6	1	1	0.01	−0.06	1	1	−0.01	−0.07
	M_7	1	5	0.03	0.06	−0	6	−0.02	0.04
	R	−12	−5	0.02	0.01	−12	−3	0.03	0.08
Bran.	D_m	−3	2	0.02	0.09	−1	2	0.03	0.01
	M_0	−13	26	0.17	0.10	4	13	−0.02	0.02
	M_1	−8	13	0.10	0.11	2	8	0.07	0.29
	M_2	−5	5	0.02	0.08	0	3	0.03	0.09
	M_3	−2	1	0.00	−0.02	−1	0	0.04	0.06
	M_4	−0	0	0.02	0.07	−0	2	0.03	0.19
	M_5	3	1	0.01	0.02	2	3	0.01	0.10
	M_6	1	−0	0.01	−0.20	2	−1	−0.01	−0.12
	M_7	−2	5	0.03	−0.06	−2	7	−0.02	−0.03
	R	−0	−0	0.03	0.04	−0	2	0.04	0.21
Beard	D_m	−1	1	0.02	0.07	1	1	0.03	−0.04
	M_0	−5	16	0.16	0.16	10	1	−0.07	0.10
	M_1	−2	7	0.08	0.17	8	−1	0.03	0.32
	M_2	1	2	0.01	0.14	6	−2	0.02	−0.01
	M_3	2	−0	0.00	0.03	4	−2	0.03	−0.02
	M_4	3	1	0.02	0.09	1	2	0.02	0.18
	M_5	4	1	0.00	0.00	1	3	0.01	0.09
	M_6	0	−0	0.01	−0.14	1	−0	−0.01	−0.12
	M_7	−3	5	0.02	−0.02	−1	8	−0.02	0.00
	R	3	0	0.02	0.08	1	1	0.03	0.13

Table 6.4 – Differences in performance by variable and region, for DSDs retrieved from Parsivel data using the double-moment technique and SCOP-ME. Differences are defined as for Table 6.3, so a negative difference shows that the double-moment method improved on SCOP-ME's performance. Note that differences for HyMeX are shown in Table 6.3.

functions are shown in Table 6.5. On average, the double-moment technique produced better median relative bias than SCOP-ME on D_m, R, and moments zero to five. IQRs were similar on average with the exception of moments zero and seven for which SCOP-ME produced smaller IQRs. The double-moment technique produced better scatter plot slope on D_m and moments three, six and seven. As is shown in tables 6.3 and 6.4, the results were different for different drop axis ratio functions. For example, when the Thurai function was used, the double-moment technique performed better overall. In contrast, the performances of the two methods were more similar when the Brandes function was used, and SCOP-ME outperformed the double-moment technique for moments zero and one in the Iowa data set with the Andsager and Beard axis ratio functions. The double-moment technique's performance changes by axis ratio function relate to the accuracy of the prediction of DSD moment three from K_{dp} and Z_{DR}. As shown in Table 6.2, moment three is recovered most precisely when the Thurai axis ratio function is used. While differences exist between the results for the different regions, the inter-region differences in comparative performance of the two techniques were generally small. We now move to testing the two techniques on measured radar data, in which noise is a problem that must be dealt with.

Variable	RB	IQR	r^2	Slope
D_m	-0.75	1.34	0.02	-0.01
M_0	-2.31	12.32	0.05	0.06
M_1	-3.71	4.78	0.06	0.16
M_2	-4.22	-0.08	0.01	0.04
M_3	-4.05	-1.76	0.01	-0.03
M_4	-3.03	-0.14	0.02	0.09
M_5	-0.02	0.96	0.01	0.04
M_6	1.22	0.09	-0.00	-0.11
M_7	0.27	7.07	-0.02	-0.03
R	-3.35	-0.25	0.02	0.07

Table 6.5 – Average differences between double-moment and SCOP-ME techniques, on Parsivel data, over three regions and four raindrop axis ratios.

6.6 Reducing the effects of noise

Radar data is noisy at light rain rates, particularly for K_{dp} and Z_{DR} (e.g. Bringi et al., 2002; Schneebeli et al., 2014). We propose here a method to deal with this noise for the current application of DSD retrieval. Regressions on Z_H and ξ_{dr} are used to determine "expected" values for these variables, which can be used when the measured values are likely to be noisy. We found that Z_{DR} can be reasonably predicted from Z_h using

$$\hat{Z}_{\mathrm{DR}} \sim \alpha_Z Z_h^{\beta_Z}, \tag{6.8}$$

and K_{dp} can be predicted from Z_h and ξ_{dr} using

$$\hat{K}_{\mathrm{dp}} \sim \alpha_K Z_h^{\beta_{K1}} \xi_{\mathrm{dr}}^{\beta_{K2}}. \tag{6.9}$$

with parameters α_Z, β_Z, α_K, β_{K1} and β_{K2}. Least-squares fitting in log-log space, using the training data set described in Section 6.4, was used to find best-fitting parameter values per raindrop axis ratio function. The resulting values and performance statistics are shown in Table 6.6.

			Z_{DR} performance					K_{dp} performance	
Ratio	α_Z	β_Z	RB [%]	IQR [% pts]	α_K	β_{K1}	β_{K2}	RB [%]	IQR [% pts]
Thurai	0.10	0.27	0	45	0.0004	0.90	-3.78	4	31
Brandes	0.03	0.41	-3	74	0.0001	1.02	-2.50	-1	12
Andsager	0.05	0.35	-2	59	0.0002	0.97	-3.13	-0	15
Beard	0.06	0.35	-2	63	0.0002	1.00	-3.14	-2	12

Table 6.6 – Fitted coefficients and the performance of the fits on disdrometer data with simulated polarimetric variables, for Equations 6.8 and 6.9. Performance is shown in terms of median relative bias (RB) and the IQR of relative bias (IQR).

We use threshold values based on those of Bringi et al. (2002) to determine when K_{dp} and Z_{DR} may be noisy. To reduce the effects of noise, then, if $Z_H \leq 35$ dBZ or $Z_{\mathrm{DR}} \leq 0.2$ dB, measured Z_{DR} is replaced by the the expected value $\hat{Z}_{\mathrm{DR}}$ and ξ_{dr} is replaced by $10^{(\hat{Z}_{\mathrm{DR}}/10)}$. Likewise, if $Z_H \leq 35$ dBZ or $K_{\mathrm{dp}} \leq 0.3\ ^\circ\ \mathrm{km}^{-1}$, K_{dp} is replaced by $\hat{K}_{\mathrm{dp}}$ (calculated with $\hat{\xi_{\mathrm{dr}}}$ if ξ_{dr} was replaced). This method for treating radar data allows radar data with negative values of K_{dp} or Z_{DR} to be used. The data treatment improved DSD-retrieval performance for both the double-moment and SCOP-ME techniques. For example, when retrieved DSDs were matched to measured MRR data, the median relative bias was reduced by an average (across variables) of ~10 percentage points for SCOP-ME and by ~18 percentage points for the double-moment technique, while IQRs were reduced more; for example on the MRR data the IQRs were reduced by ~92 (95) percentage points for the SCOP-ME (double-moment) method. When retrieved DSDs were compared to Parsivel data, the noise in the radar data contributed to errors to such an amount that for both techniques the proposed method for dealing with K_{dp} and Z_{DR} reduced both relative bias and IQR on relative bias on moments of orders three and lower by hundreds of percentage points. PPI data used in the following section were treated using this technique. We note that because most values of Z_H recorded in the PPIs analysed here were lower than 35 dBZ, the majority of radar records were corrected this way.

6.7 Comparison using radar data

The DSD-retrieval techniques were applied to noise-corrected PPI radar data from the three locations, so the retrieval techniques were evaluated independently of the noise-correction

method. We used the elevation angles of the stacked PPIs for HyMeX, 5° for Payerne, and 3° for Iowa. Measured radar variables Z_H, K_{dp} and Z_{DR} were recovered for volumes corresponding to instrument locations. DSD retrieval was applied using these values, and the resulting DSDs compared to those that were measured by other instruments. Because the axis ratio of Thurai and Bringi (2005) produced the best results for the double-moment technique on the Parsivel data, the double-moment technique was used with parameters for the Thurai axis ratio function. Note that the assumption of axis ratio function affects only parameters of the double-moment technique, because the radar data used in this section are measured, not simulated, and the SCOP-ME technique is used as provided in Anagnostou et al. (2013). In the HyMeX campaign, the lowest available PPI elevation angle (4°) was used to compare results to Parsivels, but there was also an MRR at Pradel Grainage which retrieved estimates of the DSD aloft. MRR-derived DSDs were compared at multiple altitudes using the MXPol stacked PPIs above the HyMeX instrument network. All comparisons using PPI data involved a difference in measurement volume – a change-of-support problem that we expect will lead to greater error spread (see Chapter 4). We first address the comparisons with MRR for HyMeX, then move to the comparisons with the Parsivel networks in all three regions.

MXPol volume centre altitudes were projected into MRR altitude classes for comparison. The double-moment DSD-retrieval algorithm was used with the generalised gamma model $\hat{h}$ parameters (Equation 5.6) for MRR data and $i = 3$ and $j = 6$ found in Chapter 5, of $c = 0.57$ and $\mu = 29.6$. The reconstructed DSDs were made using classes of drop diameter from 0.6 to 5.8 mm with a class width of 0.1 mm, so that the truncation matched that of the MRR data. PPI values from eight 100 m altitude classes between about 900 and 2100 m above sea level, from 496 PPI scans, were compared to MRR estimates of the DSD aloft. Results of comparisons between MRR- and PPI-derived DSDs are shown for three example altitudes in Figure 6.6. There was good agreement between the recorded radar reflectivity recorded by both instruments, with a median relative bias of −2%, an IQR on relative bias of 15 percentage points, and a value of r^2 of 0.66. Both techniques overestimated DSD moment orders zero to four and underestimated orders six and seven. Rain rate was recovered with a median relative bias of 6% (IQR 92 % pts) by the double-moment technique and 21% (IQR 106 % pts) by SCOP-ME. The double-moment technique showed lower median relative bias than SCOP-ME on all variables except moments five and six, and smaller IQRs on all variables except D_m. Correlation coefficients were low for both techniques (the maximum r^2 was 0.32, by SCOP-ME for D_m), but the double-moment technique had a slightly higher value of r^2 in the majority of cases. High best-fit slopes (values from 1.6 to 4.6) were observed for moments five, six, and seven, and show the effect of a few outlier points in these cases, which appeared in results for both techniques. Performance differences between the two techniques are shown in Table 6.7. Overall, the double-moment technique for DSD-retrieval out-performed SCOP-ME for the retrieval of DSD parameters and rain rate measured aloft by the MRR.

DSDs retrieved from polarimetric radar data were also compared to those recorded by ground-based Parsivels in the three climatic regions we studied. It should be noted that there were, at times, significant vertical distances between the radar volume and the instruments used

Variable		RB	IQR	r^2	S
D_m	MRR	−2	1	0.03	−0.09
	HyMeX	−3	2	0.03	−0.12
	Payerne	−2	0	0.00	−0.04
	Iowa	−5	2	0.06	−0.17
M_0	MRR	−19	−39	0.00	0.00
	HyMeX	13	−13	0.03	0.08
	Payerne	−3	−5	0.02	−0.02
	Iowa	−0	−25	0.11	0.18
M_1	MRR	−20	−35	0.00	0.01
	HyMeX	17	−19	0.04	0.10
	Payerne	−11	−20	0.02	0.02
	Iowa	−1	−21	0.12	0.11
M_2	MRR	−22	−19	−0.01	0.01
	HyMeX	16	−18	0.03	0.10
	Payerne	−9	−17	0.02	0.04
	Iowa	9	−16	0.12	0.06
M_3	MRR	−18	−17	−0.01	0.02
	HyMeX	12	−13	0.00	0.07
	Payerne	5	−15	0.00	0.06
	Iowa	11	−10	0.06	0.03
M_4	MRR	−14	−9	0.00	−0.01
	HyMeX	8	−9	0.01	0.04
	Payerne	12	−12	−0.01	0.07
	Iowa	4	−10	0.02	0.02
M_5	MRR	1	−8	−0.03	0.35
	HyMeX	5	−6	0.09	0.02
	Payerne	7	−10	−0.02	0.06
	Iowa	2	−5	0.06	0.02
M_6	MRR	5	−3	−0.08	1.12
	HyMeX	3	−5	0.11	0.01
	Payerne	3	−5	−0.02	0.02
	Iowa	2	−5	0.12	0.02
M_7	MRR	−1	−1	−0.10	1.56
	HyMeX	2	−3	0.09	0.01
	Payerne	−0	−6	−0.01	0.00
	Iowa	1	−4	0.13	0.02
R	MRR	−16	−13	−0.01	0.02
	HyMeX	9	−10	0.00	0.05
	Payerne	10	−13	−0.01	0.07
	Iowa	6	−8	0.01	0.02

Table 6.7 – Differences in performance by variable and region, for DSDs retrieved from PPI data using the double-moment technique and SCOP-ME, compared to the MRR at Pradel Grainage (MRR) and Parsivels by region (HyMeX, Payerne, and Iowa). Differences are defined as for Table 6.4.

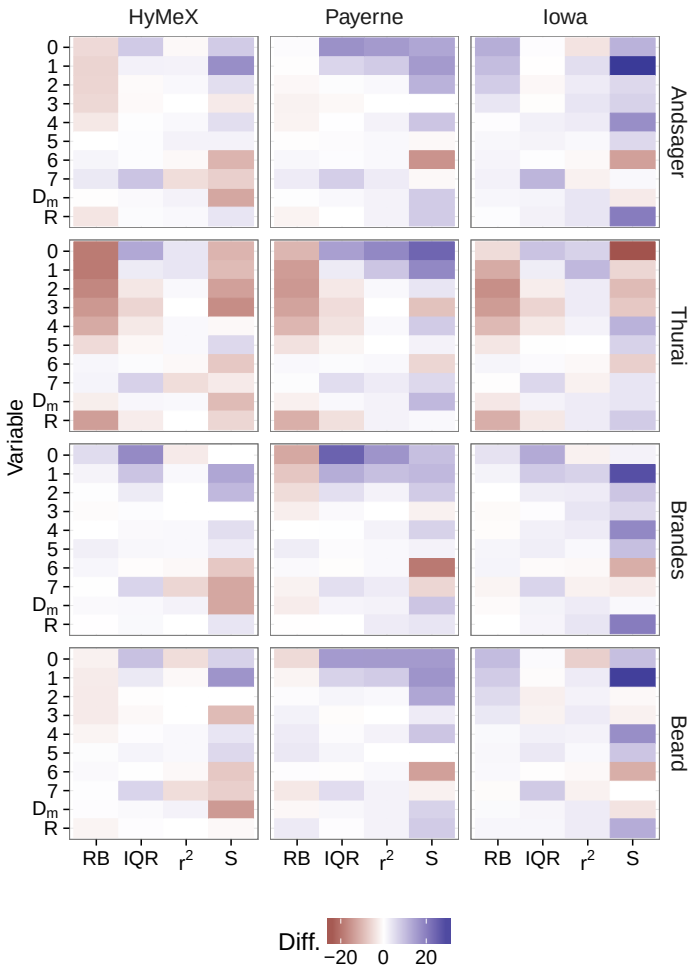


Figure 6.5 – Differences in performance between the double-moment technique and SCOP-ME, using radar variables simulated from Parsivel data, by region and drop axis ratio function (differences in Tables 6.3 and 6.4). Reds indicate negative differences, where the double-moment technique outperformed SCOP-ME. Variables are moment order n [mm^n m^{-3}], D_m [mm], and R [mm h^{-1}]. Differences are shown for median relative bias (RB [% pts]), IQR of relative bias (IQR [% pts]), squared correlation coefficient (r^2, difference from one multiplied by 100 for display on the same scale), and scatter plot slope (S, difference from one multiplied by 100).

in these comparisons (see Table 6.1). The DSDs were retrieved in truncated Parsivel drop diameter classes, using the Parsivel generalised gamma model parameters quoted in Section 6.5. Figure 6.7 shows distributions of DSD-retrieval relative error for each region.

In the majority of the tested cases, the double-moment technique produced smaller ranges of relative bias than SCOP-ME, for all variables except D_m. Where the double-moment technique produced better median relative bias, the mean per-case difference was −4 percentage points, while in cases where SCOP-ME performed better, the mean per-case difference was 8 percentage points. Values of r^2 and scatter plot slope were similar between the two techniques, with the majority of cases showing differences less than 0.05 for both variables. Differences in performance between the two techniques are shown in Figure 6.8 and Table 6.7.

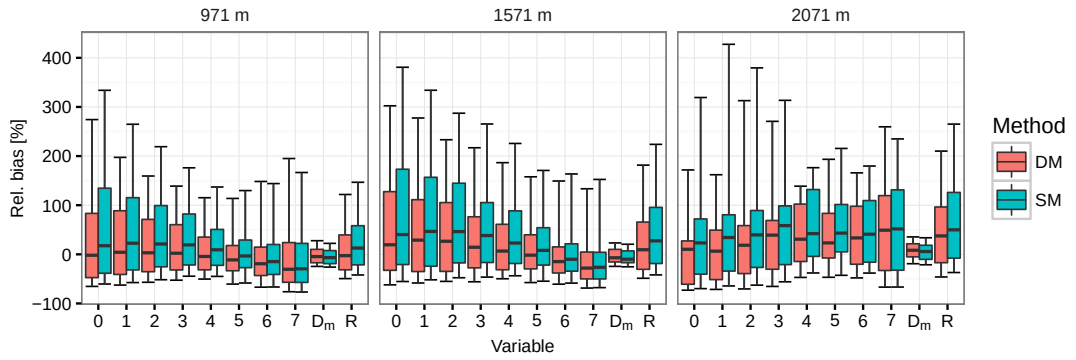


Figure 6.6 – Distributions of relative bias on DSD moments orders between zero and seven, comparing DSDs retrieved using PPI data, and those measured by MRR at Pradel Grainage. The results are classed by altitude for a selection of three altitudes across the compared range. Symbols as for Figure 6.4.

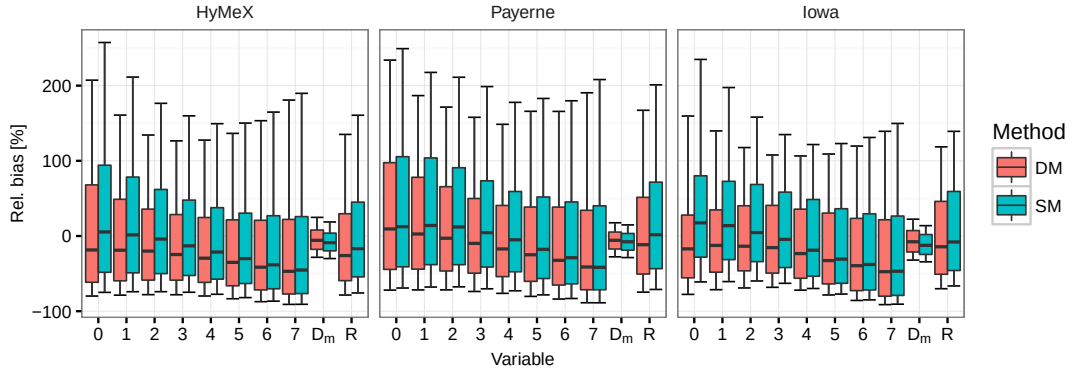


Figure 6.7 – Distributions of relative bias on DSD moments orders between zero and seven, comparing DSDs retrieved using PPI data, and those measured by the Parsivel networks. Symbols as for Figure 6.4.

The performance of the double-moment DSD-retrieval technique is reliant on how accurately two DSD moments can be extracted from radar data, and in turn on how accurate the radar data are. Both retrieval techniques appear to be similarly affected by radar inaccuracies such as bias in Z_H, and experiments with different bias corrections showed similar patterns of results. It is worth noting again that these comparisons were performed using data for which the noisy values of K_{dp} and Z_{DR} had been treated using the method proposed in Section 6.6, which significantly improved the performance of both techniques with real radar data. The proposed DSD-retrieval technique was applied using the normalised DSD fitted to data in Ardèche, France, to the regions of Payerne in Switzerland and Iowa in the USA, without significant performance loss. This supports our finding in Chapter 5 that for practical use, the double-normalised DSD can be considered invariant.

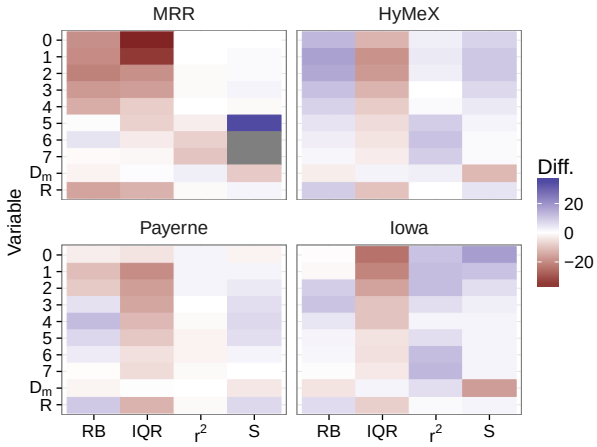


Figure 6.8 – Differences in performance between the double-moment technique and SCOP-ME using noise-corrected radar data, for MRR and for Parsivels by region (differences in Table 6.7). Variables and performance statistics as for Figure 6.5. For display, difference in r^2 and slope are multiplied by 100. Red indicates that the double-moment technique outperformed SCOP-ME. Grey indicates an r^2 difference greater than 100 on this scale; these points were affected by scatter plot outliers.

6.8 Conclusions

Given the assumption of an invariant normalised DSD, and an estimate of that function, the DSD can be predicted using only two of its moments. Two DSD moments are available from polarimetric radar data. At X-band, radar reflectivity can be used to accurately predict the sixth moment of the DSD, and moment three can be retrieved relatively accurately using K_{dp} and ξ_{dr}. We showed that by estimating these two moments from radar data, the DSD for a radar volume can be predicted using the double-moment technique. Tests on disdrometer data from three networks in different climatic regions showed that DSD-retrieval using this new technique produced similar or slightly better performance than the SCOP-ME DSD-retrieval technique of Kalogiros et al. (2013). The proposed method is also more flexible, because there is no prescribed functional form for the double-normalised DSD, and even a non-parametric $\hat{h}(x)$ could be used. Nor is there a prescribed method of DSD moment extraction, which means that the moments used could be tailored to the intended purpose.

A new method for treatment of radar data with possibly noisy values of K_{dp} and Z_{DR} was proposed. The method is based on predicting the expected values of these variables from radar reflectivity, and considerably improved the performance of both the DSD-retrieval techniques. Using noise-corrected radar data, DSDs were predicted from polarimetric variables in PPI scans measured by X-band radars in each of the three regions. A test of the retrieved DSDs against MRR data for DSDs aloft in the HyMeX region in France, and comparisons of radar-retrieved DSDs against disdrometer data from the three regions, showed reasonable agreement but large error spread for both methods. The double-moment technique predicted DSD moments measured by ground-based disdrometers with lower error spread than SCOP-ME.

Compared to DSDs measured aloft by the MRR, the DSDs retrieved by the double-moment technique outperformed those of SCOP-ME. This study provides a proof-of-concept for DSD-retrieval using noise-corrected radar data, the double-moment normalisation method of (Lee et al., 2004), and a generalised gamma model for the normalised DSD. Future work will address more precise prediction of low-order DSD moments from polarimetric radar data.

7 Multifractal analysis of snowfall recorded using a 2D-video-disdrometer

This chapter is adapted from:

- T. H. Raupach, A. Gires, I. Tchiguirinskaia, D. Schertzer, and A. Berne. Multifractal analysis of snowfall recorded using a 2D-video-disdrometer, submitted to *Journal of Hydrometeorology*.

This work was completed by T. Raupach under the supervision of A. Gires and A. Berne. I. Tchiguirinskaia and D. Schertzer provided expert advice on multifractal techniques. Research, analyses, and writing are by T. Raupach, using Python code for universal multifractals written by A. Gires. For data acknowledgements, see Appendix A. This collaboration was made possible by financial support from the Partenariat Hubert Curien – Germaine de Staël (Projet 32709UK), which allowed for T. Raupach to travel to Paris, and A. Gires to travel to Lausanne, to collaborate.

Some symbols in common use in the universal multifractal literature also appear in the radar and precipitation microphysics literature. To avoid introducing confusion by using these symbols multiple times in this thesis, the following commonly used symbols have been replaced by different symbols in this chapter:

- The resolution, normally λ, has been replaced with κ.
- A singularity, normally γ, has been replaced with ψ.
- The fractal co-dimension function, normally c, has been replaced with c_D.
- A moment order, normally written q, has been written as e.
- The wave number, normally k, has been replaced with w.

7.1 Introduction

Multifractal analysis is a useful way to study the properties of precipitation fields. It can provide insights into their intermittency and extremes, and whether they are invariant to scale or exhibit one or more different scaling regimes in time or space. The framework of Universal Multifractals (UM) (Schertzer and Lovejoy, 1987a, 1997) has been used to analyse rainfall over a wide variety of scales (see e.g. reviews by Lovejoy and Schertzer, 1995; Schertzer et al., 2010). Few studies have examined the multifractal properties of snowfall. In this chapter we report on multifractal analyses of snowfall recorded using a 2D-video-disdrometer (2DVD, Kruger and Krajewski, 2002; Schönhuber et al., 2008) situated in the Swiss Alps.

A wide variety of data sources have been used to study the multifractal properties of precipitation. At the large and medium scales, courser than a few kilometres in space or a few minutes in time, satellite products (Tessier et al., 1993; Lovejoy et al., 2008), climate simulations (Royer et al., 2008), weather model outputs (Gires et al., 2011), weather radars (Nykanen and Harris, 2003; Verrier et al., 2010; Gires et al., 2011), and rain gauge data (e.g. Fraedrich and Larnder, 1993; Olsson, 1995; Tessier et al., 1996; De Lima and Grasman, 1999; Molnar and Burlando, 2008; De Lima and De Lima, 2009) have been used. At smaller scales, studies have used lidar (Mandapaka et al., 2009), and disdrometer measurements (De Montera et al., 2009; Gires et al., 2014, 2016).

Previous multifractal work has mostly focused on liquid precipitation. Molnar and Burlando (2008) studied gauged precipitation in Switzerland, and found stronger multifractality in summer precipitation than winter precipitation, which was close to monofractal. They posited that intra-season variability in scaling parameters could be caused by mixed rain and snow at some locations, and recommended a study of snow-only scaling parameters. Shook and Pomeroy (2010) studied time series of daily snowfall on the Canadian prairies, and found weak multifractality for scales shorter than one month. Spatial distributions of snow accumulations have been shown to demonstrate fractal behaviour (e.g. Shook and Gray, 1996, 1997; Granger et al., 2002; Iggy Litaor et al., 2002; Deems et al., 2006). Fabry (1996) used spectral analysis to study high resolution time series (down to 0.1 s) of precipitation, including snow, captured using a sonic gauge. They performed spectral analysis and observed scaling in the snowfall time series with a break at about 20 s, after which little scaling was observed. They attributed the difference in scaling regimes to mixing-like processes caused by the differing fall speeds of the hydrometeors.

In this study we used similar methodology to analyses that were previously made on liquid precipitation by Gires et al. (2015), applied to measurements of solid precipitation. Gires et al. (2015) used data collected using a 2DVD to analyse raindrop distributions at high resolution in space and time. During the heaviest portions of their studied rainfall events, they found scaling behaviour in vertical rainfall structure at ranges from 0.5 to 36 m, and homogeneously distributed raindrops at smaller scales. In this study we used the same instrument and performed the same analyses on data recorded in the Swiss Alps in winter, to test whether

multifractality is evident in snowfall at high spatial and temporal resolutions. This study is a first application of multifractal analysis to snow data for horizontal accumulations of individual snowflakes and reconstructed vertical columns of snowflake concentration and mass, and to the best of our knowledge this is the first time that full universal multifractal analysis has been used on snowfall-only data at temporal resolutions finer than one day.

The rest of this chapter is organised as follows. A brief review of universal multifractal analysis is given in Section 7.2. The data used are presented in Section 7.3. In Section 7.4 the methods used to treat the data are shown. Results are shown in Section 7.5 and conclusions are drawn in Section 7.6.

7.2 Universal multifractal analysis

In this section we briefly explain multifractal analysis with the aim of showing the meaning behind the variables we analyse in the following sections. For more detail, the reader is encouraged to refer to the review of Schertzer and Lovejoy (2011). Let ϵ_κ be a field at resolution κ. This field has a spatial dimension d_ϵ. An example of a one-dimensional field is a time series of rain rate measurements, and an example of a two-dimensional field is the positions of snowflakes on a plane. The resolution of the field is $\kappa = L_\epsilon / l_\epsilon$, where L_ϵ is the outer scale – the length of the time series, or the size of the whole plane – and l_ϵ is the observation scale. The field contains a process which is assumed to have a mean of one over its outer scale. In practice, data are normalised by dividing all measured values by the overall ensemble mean.

Let A be the set of points where the process is active. In a fractal process, the number of non-overlapping d_ϵ-dimensional boxes with side length l_ϵ required to cover all the process points in the field, $N_{\kappa,A}$, is related to the resolution via the fractal dimension D_F:

$$N_{\kappa,A} \approx \kappa^{D_F}. \tag{7.1}$$

A sparse set will be covered by fewer boxes, so the fractal dimension is a measure of sparseness of the process within its outer scale. Given any single one of these boxes, the probability that it intersects with the process set A is

$$\Pr(\epsilon_\kappa) = \frac{N_{\kappa,A}}{N_\kappa} \approx \frac{\kappa^{D_F}}{\kappa^{d_\epsilon}} = \kappa^{-c_F}, \tag{7.2}$$

where N_κ is the number of boxes required to cover the entire field, and $c_F = d_\epsilon - D_F$ is called the fractal co-dimension of the process.

The sparseness of the process and the fractal dimension will change if the process is thresholded by a value (e.g. Lovejoy et al., 1987). To characterise not just the process occurrence but also its values, then, a scale-dependent threshold κ^{ψ} can be used. The probability that a single box intercepts the thresholded field is then

$$\Pr(\epsilon_\kappa \geq \kappa^{\psi}) \approx \kappa^{-c_D(\psi)}, \tag{7.3}$$

where ψ is a called a singularity and $c_D(\psi)$ is called the fractal co-dimension function (Schertzer and Lovejoy, 1987a). It can be shown (Schertzer and Lovejoy, 1987a) that Equation 7.3 is equivalent to a relationship between the statistical moments of the field and its resolution, such that

$$\langle {\epsilon_\kappa}^e \rangle \approx \kappa^{K(e)}, \tag{7.4}$$

where e is a moment order and $K(e)$ is called the moment scaling function. $K(e)$ is related to $c_D(\psi)$ via the Legendre transform, and for any moment e there is an associated singularity ψ. Knowledge of $K(e)$ thus allows for the field to be characterised at any resolution. In the UM framework, $K(e)$ can be written as an expression depending on only three parameters (Schertzer and Lovejoy, 1987a, 1997):

$$K(e) = \frac{C_1}{\alpha - 1}(e^{\alpha} - e) + He. \tag{7.5}$$

The parameters are as follows:

1. H, the degree of non-conservation. H measures the scale-dependence of the mean (of the field), such that $\langle \epsilon_\kappa \rangle \approx \kappa^{-H}$. $H = 0$ for a conservative field in which $\langle \epsilon_\kappa \rangle = 1$ at all resolutions. H is also the order of fractional integration (if H is negative) or derivation (if H is positive) required to transform the observed field into a conservative field (Schertzer and Lovejoy, 1991; Tessier et al., 1993).

2. C_1, the mean intermittency. C_1 is the fractal co-dimension of the mean field. The mean field is the field thresholded by the singularity associated with $e = 1$. This singularity is also equal to C_1 if the field is conservative (Schertzer and Lovejoy, 2011). If $C_1 = 0$ then the mean field is homogeneous. A larger C_1 indicates a sparser mean field. For non-null fields, $0 \leq C_1 \leq d_\epsilon$.

3. α, the multifractality index. α indicates how quickly the intermittency of the field

changes as the considered singularity moves away from that associated with $e = 1$, the singularity of the mean field. It thus measures where the process sits on a scale between monofractality (for which $\alpha = 0$ and the intermittency is the same for all thresholds) and log-normality ($\alpha = 2$).

Spectral analysis is used to find H and provides a first estimate of the scaling behaviour of the field. If the field is multifractal then its power spectra E can be written

$$E(w) \approx w^{-\beta} \tag{7.6}$$

where w is the wave number and β is called the spectral slope. H is then estimated as (Tessier et al., 1993)

$$\beta = 1 + 2H - K(2). \tag{7.7}$$

with H taken as zero in Equation 7.5 (i.e. considering only the conservative part). For a given value of e, trace moment (TM) analysis (e.g. Schertzer and Lovejoy, 1987b) can be used to find $K(q)$: on a log-log plot of κ on the x-axis vs. $\langle {\epsilon_\kappa}^e \rangle$ on the y-axis, the slope of the line of best fit is $K(e)$.

The values of C_1 and α for a given field can be found directly using double trace moment (DTM) analysis, in which the normalised η-power of the field is used (Lavallée et al., 1993). The normalised η-power is written

$$\epsilon_\kappa^{(\eta)} = \frac{\epsilon_\kappa^\eta}{\langle \epsilon_\kappa^\eta \rangle}. \tag{7.8}$$

The eth statistical moment of the η-power field also scales with resolution as described by a moment scaling function, $K(e,\eta)$, such that

$$\langle {\epsilon_\kappa}^{(\eta)e} \rangle \approx \kappa^{K(e,\eta)}, \tag{7.9}$$

and in the case of the UM framework,

$$K(e,\eta) = \eta^{\alpha} K(e). \tag{7.10}$$

DTM analysis uses a log-log plot of η on the x-axis vs. $K(e,\eta)$ on the y-axis; on the linear part of the plot, the slope is equal to α and the intercept equal to C_1.

TM and DTM analysis rely on the assumption that the field is conservative. When $H \neq 0$, the field is considered non-conservative and an additional fractional integration is required before TM and DTM analyses are used to estimate α and C_1 (Schertzer and Lovejoy, 1991; Tessier et al., 1993; De Montera et al., 2009). In this work we used a fractional integration in which each one-dimensional realisation was transformed individually (see Schertzer and Lovejoy, 1991, appendix B2). For a one-dimensional field ϵ, let $I = \epsilon - \langle\epsilon\rangle$, and let $\mathscr{F}(I)$ be the Fourier transform of I, i the complex unit, and w the wave number. The fractionally integrated field $\hat{\epsilon}$ was calculated as

$$\mathscr{F}(\tilde{\epsilon}) = i w^{H} \mathscr{F}(I) \tag{7.11}$$

$$\hat{\epsilon} = \frac{\tilde{\epsilon} - \min(\tilde{\epsilon})}{\langle \tilde{\epsilon} - \min(\tilde{\epsilon}) \rangle} \langle\epsilon\rangle . \tag{7.12}$$

Note that an inverse Fourier transform is used to convert $\mathscr{F}(\tilde{\epsilon})$ to $\tilde{\epsilon}$, and that as well as subtracting the mean of ϵ, the zeroth Fourier component was set to zero to force the mean of each field to be zero (e.g. Tessier et al., 1993). This fractional transformation corresponds to integration when $H < 0$ and derivation when $H > 0$.

Finite sample sizes mean that there is a maximum singularity ψ_s for which reliable estimates of the co-dimension and moment scaling functions are possible (e.g. Hubert et al., 1993). ψ_s has an associated maximum moment order e_s. To calculate the field at a given resolution, the measured data (at high resolution) are up-scaled through an iterative process, in which at each step the observation scale l_e is doubled to create a lower-resolution field by averaging adjacent pixels. For this reason, the outer scale of the field must be a power of two.

7.3 Data

The data used in this study were collected using a 2DVD (of low-profile type) located at Weissfluhjoch Versuchsfeld (9.8096° E, 46.8301° N) near Davos, Switzerland. The site is in an alpine environment 2540 m above sea level. The 2DVD was placed on a platform about

two metres above the ground, between the two fences of a Double Fence Intercomparison Reference (DFIR, e.g. Goodison et al., 1998) structure. Data used in this study were collected during the northern winter of 2014/2015. Figure 7.1 shows the installed instrument.

Figure 7.1 – The Double Fence Intercomparison Reference (DFIR) structure (left) and the 2DVD instrument installed on a platform between the two fences (right) at 2545 m near Davos, Switzerland.

The 2DVD contains two perpendicularly-facing line-scan cameras, each with an opposing light source, which detect the shadow, from two angles, of any particle that falls through its collection area. The imaging resolution is finer than 0.2 mm. The nominal collection area is 10 × 10 cm^2. The two camera planes are vertically offset so that particle velocity is measured directly, through the difference in particle detection time in each camera. The shape of each particle can be reconstructed using the two views (Schönhuber et al., 2008). The 2DVD can not be considered an absolute reference for snow measurements (Battaglia et al., 2010), and for liquid precipitation it has been shown to underestimate raindrops under 0.3 mm in equivolume diameter (Tokay et al., 2013). This being said, it is able to provide data on individual snowflakes at high resolutions, so it remains useful for this study.

There were collocated meteorological instruments at the 2DVD location, and we used data from these instruments to select three one-hour periods during which there was snowfall and calm wind conditions. Wind speed was measured outside the DFIR by a Young 05103 Wind Monitor, temperature was measured outside the DFIR by a Thygan VTP 6 instrument, and precipitation was measured inside the DFIR by a Thies Laser Precipitation Monitor disdrometer (see Roulet et al., 2014, for information on the meteorological instruments). We found, using the meteorological data, one-hour periods in which there was solid precipitation without hail, zero mean wind speed, and a temperature of less than -10° C. Of these, the hour-long time periods with maximum solid precipitation for each of the months of December 2014, January 2015, and February 2015 were analysed. The resulting time periods (henceforth referred to by event number) are summarised in Table 7.1.

The 2DVD data contained, for each particle, the time stamp, velocity [m s^{-1}], the particle

E	End of hour	Int [mm h^{-1}]	Temp [° C]	Num
1	2014-12-28 02:00:00	2.36	-13.2	7000
2	2015-01-27 18:00:00	4.16	-13.5	13131
3	2015-02-24 08:00:00	2.56	-10.7	17378

Table 7.1 – Information on the studied hours of precipitation by event number (E), showing the studied time, mean solid precipitation intensity measured by Thies laser disdrometer (Int), mean temperature (Temp), and the number of particles measured by the 2DVD and included in this study from the event (Num).

width and height recorded by each camera [mm], and the pixel number of the centre of the particle in lines from each camera. These pixel numbers were converted to x and y positions inside the 2DVD's measurement window, in millimetres, by multiplying the pixel numbers by the pixel width which depends on the particle's distance from the camera. During the three selected one-hour periods, 49117 particles were recorded. One particle had a velocity of zero and was excluded. The 2DVD appeared to capture fewer particles on two particular edges of the measurement area, which we hypothesise was due to edge effects and the positioning of the instrument. To deal with this we subset the particles to those that appeared within 100 mm of the least-affected side of both camera lines. With this constraint, 3779 particles (7.7% of the total number) were removed. To ensure we sampled snowfall only, particles were only kept if their velocities were between 0.2 and 8 m s^{-1}; this constraint removed only one particle.

7.4 Data treatment

In this section, the ways that the 2DVD snow particle data were treated before applying multifractal analysis are presented. The mass of each solid precipitation particle was estimated. These masses were then formed into reconstructed ballistic columns (Section 7.4.2) and time series (Section 7.4.3). Two-dimensional particle accumulation maps (Section 7.4.4) were also calculated.

7.4.1 Estimation of particle masses

The method of Huang et al. (2015) was used to estimate the mass of each recorded particle from the 2DVD data. In their method, the mass m [g] of a particle is calculated as

$$m = \frac{\pi \eta_{\text{air}}^2 X}{8 g \rho_a} \left(\frac{A_e}{A_c} \right)^{1/4} . \tag{7.13}$$

where η_{air} [kg m^{-1} s^{-1}] is the air viscosity, g [m s^{-2}] is acceleration due to gravity, A_e [m^2] is the area of pixels covered by the particle in one camera view, and A_c [m^2] is the area of the minimum ellipse or circle that covers the particle in the same camera view. In this equation the units of the air density ρ_a are g cm^{-3}. We have added the multiplier of 10^3 to convert the

mass to grams. X [–] is the Davies number, calculated as

$$X = \left\{ \frac{\left[\left(\frac{Re}{8.5} \right)^{(1/2)} + 1 \right]^2 - 1}{0.1519} \right\}^2 , \tag{7.14}$$

and Re [–] is the Reynold's number, which depends on the particle velocity V_f [m s^{-1}]:

$$Re = \frac{2 \times 10^3 \rho_a V_f}{\eta_{\text{air}}} \left(\frac{A_c}{\pi} \right)^{1/2} . \tag{7.15}$$

The viscosity and air density were calculated per-event, based on the altitude and temperature at the 2DVD site. The average viscosity was 1.662 × 10^{-5} kg m^{-1} s^{-1} and the average air density was 0.979 kg m^{-1}. Particle areas were calculated using data from one of the 2DVD's two cameras (camera A). For a given particle, A_e was calculated as the area of the ellipse with major and minor axes equal to the major and minor axes of the bounding rectangle of the particle. Following Huang et al. (2015), if A_e / A_c was greater than one, A_c was replaced by the area of the smallest circle containing the bounding rectangle. Densities of estimated mass are shown by event in Figure 7.2.

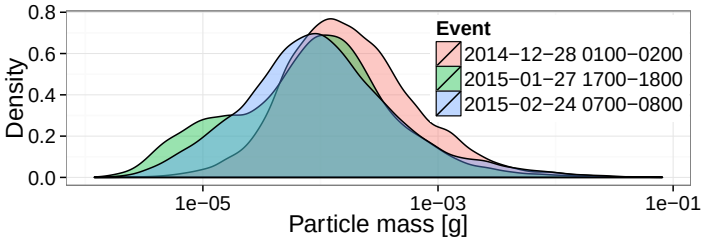


Figure 7.2 – Densities of estimated particle mass by event. The x-axis has a log scale.

7.4.2 Reconstruction of vertical columns

Just as in Gires et al. (2015), we reconstructed vertical columns above the 2DVD. Each reconstructed column had dimensions of 0.1 × 0.1 × 35 m, and we reconstructed one column per second of the studied precipitation. For each second, the height above the instrument for every particle was calculated using the particle's velocity (as measured at the ground), and its centre point, and those within 35 m of the instrument were kept. Each column was divided vertically into 512 boxes, so that each box had a height of about 68 mm. This reconstruction of vertical columns relies on the assumptions that there was no wind, that particles fell vertically at constant velocity, and that there were no microphysical processes such as aggregation or

breakup of snowflakes. The grouping of particles into vertical boxes goes some way towards reducing the effect of ignoring horizontal movement of the snowflakes, but these remain very coarse assumptions. Our aim in this study is to gain some preliminary insights into the vertical structure of falling snow, in the absence of measurements of the true vertical snow column, so we continue in spite of these obvious limitations.

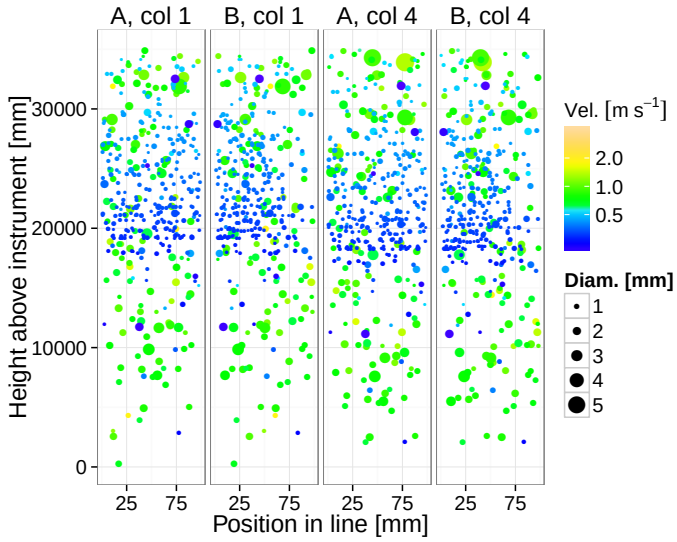


Figure 7.3 – Examples of two of the reconstructed columns. "Col 1" and "col 4" correspond to respectively the first and fourth second of the studied hour of precipitation on the 27th January 2015. "A" and "B" correspond to the two 2DVD cameras. Points show particle locations in the column, point size shows recorded particle (maximum) width, colour shows velocity.

The quantity analysed was calculated per time and per box as

$$X_p = \sum_{n}^{P} {m_n}^p, \tag{7.16}$$

where the sum was taken over all P particles in the box. By varying p, different quantities are recovered; X_0 provides the number of particles per box, while X_1 equals the total estimated snow mass per box. Each set of 512 values for one second of precipitation was assumed to represent one realisation of the solid precipitation process in the vertical column, and multifractal analysis was performed on these realisations. As a control data set, we also randomly redistributed the particle centre points uniformly in the vertical, and performed the same multifractal analysis on these homogeneously distributed fields. Figure 7.4 shows the difference between the reconstructed columns and the homogeneously distributed ones. Before analyses were performed, all values were normalised by the mean of all data across all realisations.

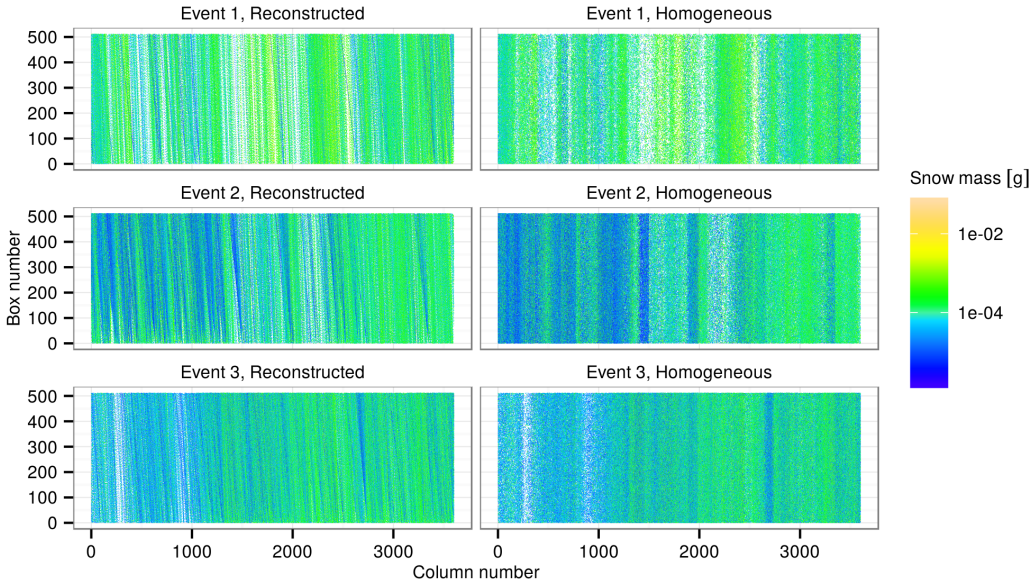


Figure 7.4 – Snow mass per box and per column, per event, and showing the difference between reconstructed and homogeneously distributed particles. In the reconstructed fields, the particles falling towards the 2DVD over time are visible as oblique patterns.

7.4.3 Time series

Because the 2DVD provides information about individual particles, we could construct time series at a chosen time resolution. We used the following formula to estimate precipitation intensity:

$$\tilde{R}_p = \sum_{n}^{P} \left(\frac{10^3 \, m_n}{S_n (\Delta t / 3600)} \right)^p \tag{7.17}$$

where S_n [mm^2] is the sampling area, Δt [s] is the temporal resolution, P is the total number of particles over which the sum is taken (i.e. the number of particles that fell through the collection area during Δt), and we assume that 1 g represents exactly 1000 mm^3 of water. The sampling area S_n is calculated for each particle's diameter D [mm] as $S_n = (100 - D_n)^2$, because particles that were not fully inside the sampling area were removed, thus reducing the effective sampling area (e.g. Schönhuber et al., 2008; Battaglia et al., 2010). p is again a power, such that when $p = 0$, the quantity produced is the number of particles observed per time step, and when $p = 1$ the intensity is given in mm h^{-1} so that it is equivalent to rain rate if the precipitation was liquid.

For each event, we calculated a time series for a temporal resolution of 0.1 s, and cut the time series so that it contained a power-of-two number of time steps. The subset chosen

contained the largest sum of $\tilde{R}_p$ values. Figure 7.5 shows the time series, with the analysed portions shown in red. Individual data points were normalised by the overall ensemble mean before analyses were made. For comparison with the recorded time series, homogeneous time series were created by randomly sampling without replacement from the measured values to randomise their order.

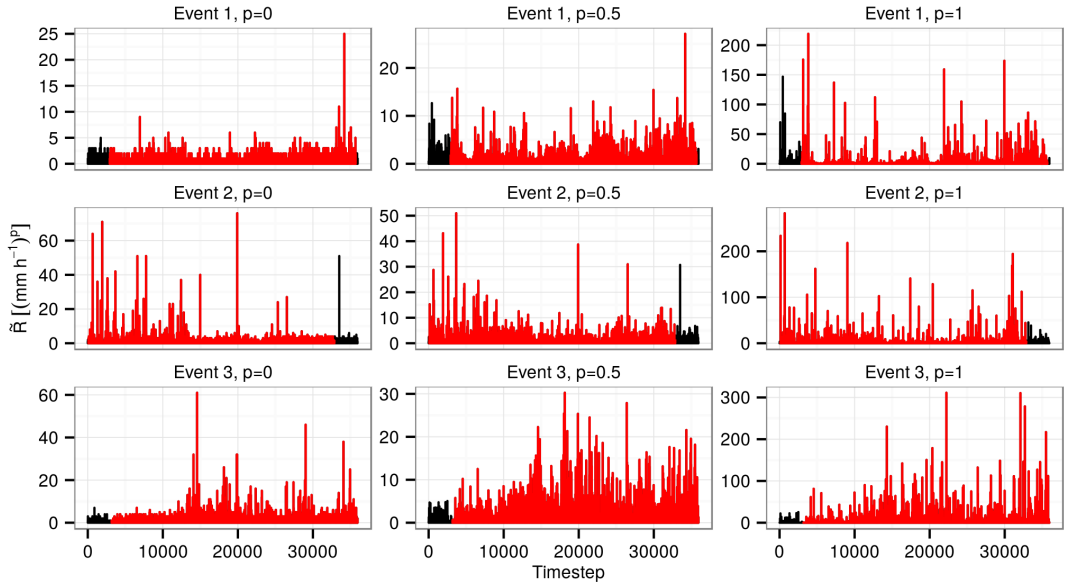


Figure 7.5 – Time series at 100 ms, showing estimated solid precipitation liquid water intensity $\tilde{R}_p$ $[(\mathrm{mm\,h^{-1}})^p]$ per event and by power p.

7.4.4 Particle accumulations

Particle accumulation maps were calculated using particle positions and widths in mm from each camera, projected onto a grid of pixels with a pixel side length of 0.1 mm. The sampling area was cropped so that it had a power-of-two number of pixels per side; the resulting maps are at a resolution of 512 × 512 pixels. The area in the centre of the sampling area was used, to reduce any potential edge effects. Accumulation maps were calculated by number of accumulated particles per map, for 10, 50, and 100 particles, up to a maximum of 200 maps that were each separate in time. For each map, each pixel contained the number of snowflakes that covered that pixel when they landed. Figure 7.6 shows examples of particle maps. The rectangular shapes of the particles are due to the fact that the particle is assumed to completely cover the area defined by the widths from each camera. These maps were considered to contain realisations of a two-dimensional field. The number of maps of snowflake accumulations calculated for each event are shown in Table 7.2. Just as with the other "cuts" we used, these data were processed in sets (by number of accumulated particles) and for each set the data were normalised by the ensemble mean of the set data.

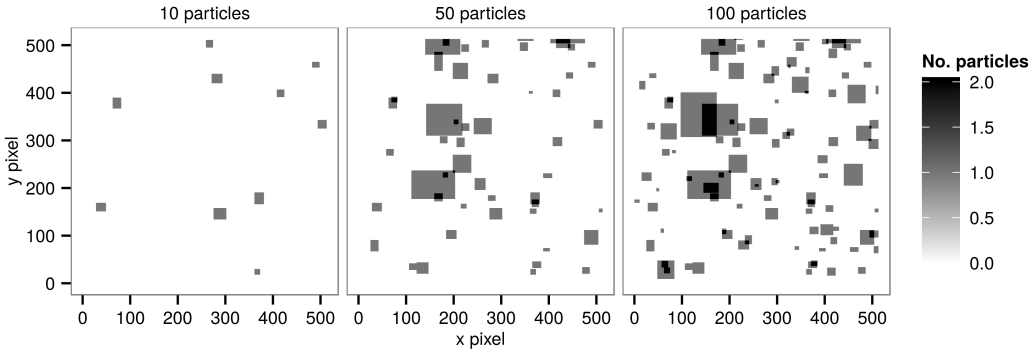


Figure 7.6 – Examples of accumulation maps in the studied hour from 2015-01-27, showing the first time steps of particle accumulation series for 10, 50, and 100 particles.

Event	Particles	No. maps
1	10	200
1	50	53
1	100	26
2	10	200
2	50	97
2	100	48
3	10	200
3	50	130
3	100	65

Table 7.2 – The number of maps found per number of particles per event.

7.4.5 Multifractal analysis

For each type of data analysed (vertical columns, time series, and accumulation maps), we performed multifractal analysis for values of p of 0 (particle concentration), 0.5, and 1 (particle mass). First, spectral analysis and TM analysis were used together to identify scaling regimes; the scaling regimes were manually identified, with the criteria being to a) observe scaling in the spectral analysis (the mean of the field, represented by resolution at $w = 1$, was ignored), and b) to ensure that the TM correlation coefficient r^2 was maximised. DTM analysis was performed for $e = 1.5$. Spectral analysis was used to determine whether any additional fractional integration was required before TM and DTM were used. As a first guess of H for the fractional integration, we used $\tilde{H} = (\beta - 1)/2$. If $\tilde{H}$ was less than zero, fractional integration of order $\tilde{H}$ was used before TM and DTM analyses were applied. Informed by the identified scaling regimes, DTM analysis was used to estimate values of α, C_1, ψ_s, e_s, and H. To calculate α and C_1, the slope of the DTM line on a neighbourhood of six points on the $\log_{10}(\eta)$ axis was used (in linear space, this axis comprises a sequence of 34 evenly-spaced values of η between -2 and 1). This line was centred on $\log(\eta) = 0$ if fractional integration was performed, and on $\log(\eta) = 0.15$ otherwise.

7.5 Results

In this section we show results and interpretations of universal multifractal analysis of the three spatio-temporal "cuts" in turn.

7.5.1 Vertical columns

Exactly the same analyses were performed on both reconstructed vertical columns and on the columns in which the particles were distributed randomly. This was to ensure that potential scaling results found on the reconstructed columns were not artefacts of the analysis methods. The first step in all our multifractal analyses is to look at spectral analyses of the fields, to see whether scaling exists and if so which ranges of scales display multifractal properties. Example plots for a single hour of precipitation (event three) are shown in Figure 7.7. Here we compare the spectral analysis of reconstructed vertical column data, and of vertical columns in which the particles were randomly distributed in a uniform way.

The spectral analysis of fields with homogeneous particle distributions shows, as expected, a slope of zero. On the other hand, the reconstructed columns show a section in which scaling is observed, over scales down to about two metres. All but one of these fields had negative values of H, and thus on those fields, a single fractional integration of order $\tilde{H}$ was performed before conducting the TM and DTM analyses. TM and DTM analyses were carried out for resolutions of 1 to 16, corresponding to observational scales of 35 to 2.19 m. The scaling regimes shown in the TM analyses often extended more towards the small scale, as can be seen in Figure 7.8, in which the scaling regime seems to extend to $\kappa = 64$ or about 0.5 m. For consistency with the spectral analysis results, however, we analysed the same scale for all fields. Results for the analyses of vertical columns are shown in Table 7.3. TM and DTM analysis plots for the vertical columns in the example hour are shown in Figure 7.8.

Event	p	β	r^2 (spec)	r^2 (TM)	α	C_1	ψ_s	e_s	H
1	0.0	0.41	0.80	1.00	1.63	0.02	0.22	10.12	-0.27
1	0.5	0.34	0.75	1.00	1.77	0.02	0.25	8.10	-0.31
1	1.0	0.13	0.66	1.00	1.99	0.03	0.29	6.32	-0.41
2	0.0	1.09	0.99	1.00	2.02	0.07	0.47	3.67	0.12
2	0.5	0.60	0.88	1.00	1.90	0.03	0.28	6.82	-0.17
2	1.0	0.10	0.62	1.00	1.98	0.03	0.30	6.13	-0.42
3	0.0	0.76	0.94	1.00	1.84	0.02	0.25	7.95	-0.10
3	0.5	0.39	0.78	1.00	1.82	0.02	0.23	8.60	-0.28
3	1.0	0.10	0.53	1.00	2.16	0.02	0.31	5.61	-0.43

Table 7.3 – UM results for the reconstructed vertical columns, by event E, for scales from 35 m to 2.19 m (β calculated for 17.5 m to 2.19 m). r^2 (spec) is squared correlation coefficient for the spectral slope excluding the largest scale, and r^2 (TM) is the squared correlation coefficient for the TM analysis over the selected scales. The correlation scores for the TM analyses are only 1 after the use of fractional integration, which introduces correlation.

The spectral analyses for vertical columns show scaling for particle concentrations ($p = 0$)

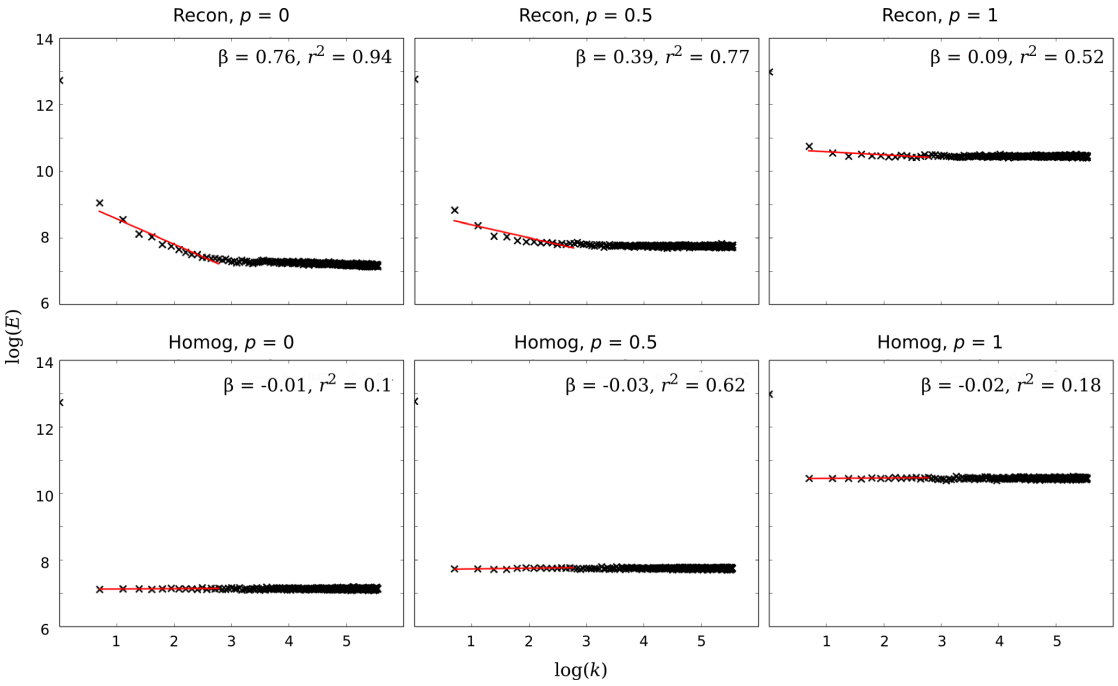


Figure 7.7 – Spectral analyses for vertical columns, for the analysed hour of precipitation on the 24th February 2015. Shown are results for $p = 0$, $p = 0.5$ and $p = 1$ for both reconstructed particle positions ("Recon", first row) and homogeneously assigned ("Homog", second row) particle positions.

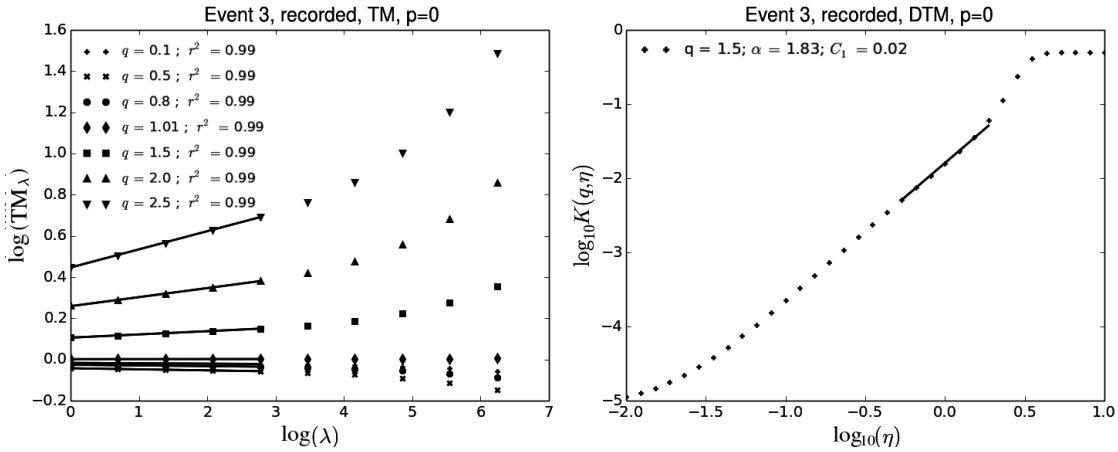


Figure 7.8 – TM and DTM analyses for vertical columns in the analysed hour of precipitation on the 24th February 2015, for values of p of 0 and 1. In this plot, $TM_K = \langle {\epsilon_K}^e \rangle$ for a given moment order e, and η is the power to which the field is raised then normalised.

Event	p	β	r^2 (spec)	r^2 (TM)	α	C_1	ψ_s	e_s	H
1	0.0	-0.00	0.01	0.96	1.62	0.01	0.17	13.45	-0.28
1	0.5	0.01	0.19	0.97	1.87	0.02	0.22	9.02	-0.32
1	1.0	0.01	0.06	1.00	2.09	0.02	0.28	6.37	-0.46
2	0.0	-0.00	0.00	0.85	2.15	0.00	0.13	14.53	0.05
2	0.5	-0.01	0.06	0.94	1.97	0.01	0.20	9.35	-0.19
2	1.0	0.01	0.01	0.99	2.07	0.02	0.27	6.80	-0.45
3	0.0	-0.01	0.11	0.91	1.38	0.01	0.10	30.65	-0.10
3	0.5	-0.02	0.63	0.96	1.87	0.01	0.19	10.56	-0.28
3	1.0	-0.01	0.19	0.99	2.21	0.02	0.27	6.34	-0.43

Table 7.4 – UM results for vertical columns with homogeneous particle positions, fractionally integrated using values of β in Table 7.3. Results are shown by event E, for scales from 35 m to 2.19 m (β calculated for 17.5 m to 2.19 m). r^2 (spec) and r^2 (TM) defined as for Table 7.3.

for scales down to about two metres. As the power p increases, however, the quality of the scaling, as measured by β and its associated correlation coefficient, reduces. This is visible in Figure 7.7, and is likely due to the fact that increased power implies a larger impact of heavier snowflakes. Heavier flakes are more rare, and therefore do not show scaling over the 35 m column that we studied. We hypothesise that larger (and thus heavier) snowflakes decorrelate from atmospheric turbulence at larger scale; but the studied scales were not large enough to confirm this. We focus on the case where good scaling was observed, i.e $p = 0$ corresponding to columns containing the vertical distribution of particle concentrations.

TM and DTM analysis were also performed for the homogeneously distributed particle fields, after fractional integration using the values of β for the reconstructed fields. These results are shown in Table 7.4. They show that with homogeneously distributed particle positions, β was close to zero. Where less or no scaling is observed, for example for event three and when $p = 1$, the DTM results (α and C_1) are similar between the two sets, indicating that the reconstructed columns contained no clustering and little scaling, and that in those cases the results of the DTM analysis were rather an artefact of the fractional integration that was applied. When scaling is evident, however, there are differences shown between the homogeneous and reconstructed results, which shows that the obtained values are representative of the underlying physical process and not an artefact of the analysis technique. For example, in events two and three with $p = 0$, there are clear differences in the results between the homogeneous and reconstructed cases. The differences are particularly clear for the maximum singularity ψ_s in these cases.

The TM and DTM analyses show values of α that are outside normal limits (i.e. greater than 2) in cases where no or little scaling was observed, and in Event 2 in which we hypothesize that estimation error has pushed α just over 2. In all cases, analyses were carried out on scales ranging from 35 m to ~2 m, meaning that the results are valid only on this range. Further investigations with would be needed to extend these results to larger scales. For the analyzed scales of 35 m to about 2 m and $p = 0$, the DTM analysis shows a fractal co-dimension C_1 of 0.02 for two events and 0.07 for the other, indicating that the mean field

shows low intermittency at these scales. The multifractality index α is high, which means that the field is highly multifractal, and that its intermittency changes rapidly as it is thresholded at higher singularities.

These results confirm that indeed the reconstructed distributions of particle numbers show scaling properties in the vertical columns. The UM parameters obtained for scales between 35 m and two metres open the possibility of numerical simulations of vertical snowfall fields on this range of scales, using UM. However, noise in the data, the small data set, and associated estimation errors mean that complete simulations would require further investigation to precisely determine UM parameters on larger scales and for values of p not equal to zero. The differences between reconstructed and homogeneously distributed particle concentrations offers evidence that scaling is, at times, present in vertical snowflake concentrations for scales between about 35 m and two metres, and possibly down to 0.5 m.

7.5.2 Time series

An example of spectral analysis results for the time series is shown in Figure 7.9. The spectral analyses results were noisy and had low correlation coefficients for spectral slope, indicating that the quality of the scaling is limited. This is common behaviour when only one field realisation is available, as was the case for each time series. It was not possible to distinguish different scaling regimes from these analyses alone. The spectral slope was found for scales from 27 minutes to 100 ms. The values of β produced by spectral analysis of the time series data showed a small spectral slope and therefore some scaling, which was confirmed by comparisons with time series in which the data order was randomised. In these randomised cases, no scaling was observed. The values of H returned by the spectral analyses were negative, so a fractional integration of order $\tilde{H}$ was performed on each time series. TM analyses on the resulting fields further confirm that scaling exists, at least on the larger scaling regimes from 54 minutes to ~3.4 or ~1.7 minutes. Examples of TM and DTM plots for the time series analysis are shown in Figure 7.10. UM analysis results for the time series are shown in Table 7.5.

Just as for the vertical columns, we compared the results to those using exactly the same fractional integration and processing on randomised fields. In the case of the time series, ψ_s was significantly different for every event and value of p, so we consider that the scaling shown by these results is not introduced by the data processing. While the low amount of data means that we should not put too much emphasis on individual results, we conclude that scaling does exist in the time series of snowfall from the tested outer scale of about 54 minutes down to about two minutes. We note that at the average particle fall velocity (1.1 m s^{-1}), the lower limit of the scaling regime observed in vertical columns (~ 2 m) equates to a time series scale of about 2 s. While in some cases, the TM analysis showed a second scaling regime at higher resolutions, given the available data and the poor quality of the time series scaling results on this range of scales, we are not able to state definitively whether scaling exists in the time series for scaling regimes from about two minutes to two seconds.

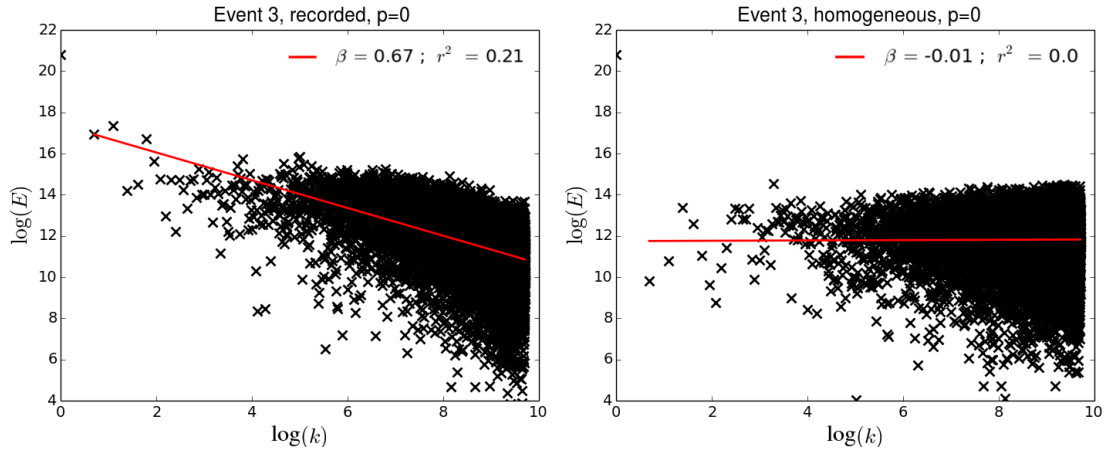


Figure 7.9 – Spectral analyses for the analysed time series of precipitation on the 24th February 2015, for recorded and homogeneous time series.

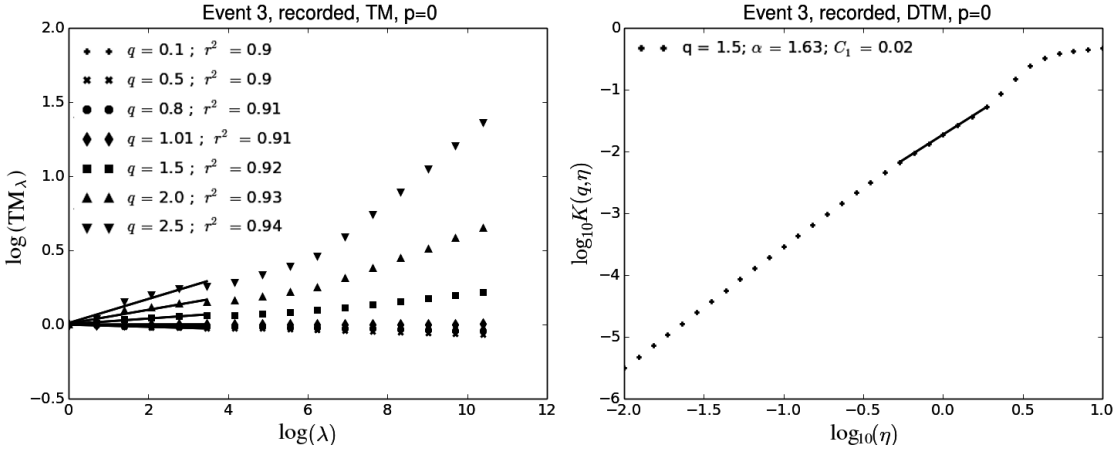


Figure 7.10 – TM analyses for the analysed time series of precipitation on the 24th February 2015.

7.5.3 Particle accumulations

Results of spectral analysis for accumulation maps with 100 particles, for measured and randomly distributed particle fields, are shown as an example in Figure 7.11. The results of spectral analysis using uniformly randomly distributed particles are very similar to those for the measured distributions in space, as is shown in Table 7.6. This is evidence that no scaling is shown in the measured distributions. For all values of p, the values of β were significantly lower than two, the dimension of the field. We conclude that at the analysed scale of 5.12×5.12 cm^2, the snow particles exhibit the same behaviour as if they were distributed homogeneously – that is, as if their positions were drawn from a uniform distribution. This conclusion agrees with the analysis of the vertical columns, which showed no scaling for particle concentrations at scales below ~0.5-2 m.

The exceedance probabilities of the time required to observe 10, 50, and 100 snowflakes were calculated (as shown for liquid precipitation in Gires et al., 2015). These results are shown

Event	p	β	r^2 (spec)	r^2 (TM)	α	C_1	ψ_s	e_s	H
1	0.0	0.38	0.08	0.99	1.47	0.04	0.25	9.61	-0.28
1	0.5	0.27	0.04	0.99	1.50	0.04	0.28	8.28	-0.33
1	1.0	0.09	0.01	0.96	1.87	0.02	0.22	8.98	-0.44
2	0.0	0.91	0.33	0.96	3.36	0.01	0.31	4.48	-0.03
2	0.5	0.63	0.19	0.95	1.93	0.01	0.22	8.83	-0.17
2	1.0	0.15	0.01	0.84	1.88	0.01	0.21	9.62	-0.41
3	0.0	0.67	0.21	0.92	1.64	0.03	0.24	9.10	-0.14
3	0.5	0.43	0.10	0.93	1.47	0.03	0.21	11.40	-0.26
3	1.0	0.13	0.01	0.96	1.62	0.02	0.21	10.90	-0.42

Table 7.5 – UM analysis results for recorded snowfall time series, for scales from ~ 54 min to ~ 1.7 min (β calculated for scales from 27 min to 100 ms). r^2 (spec) and r^2 (TM) defined as for Table 7.3.

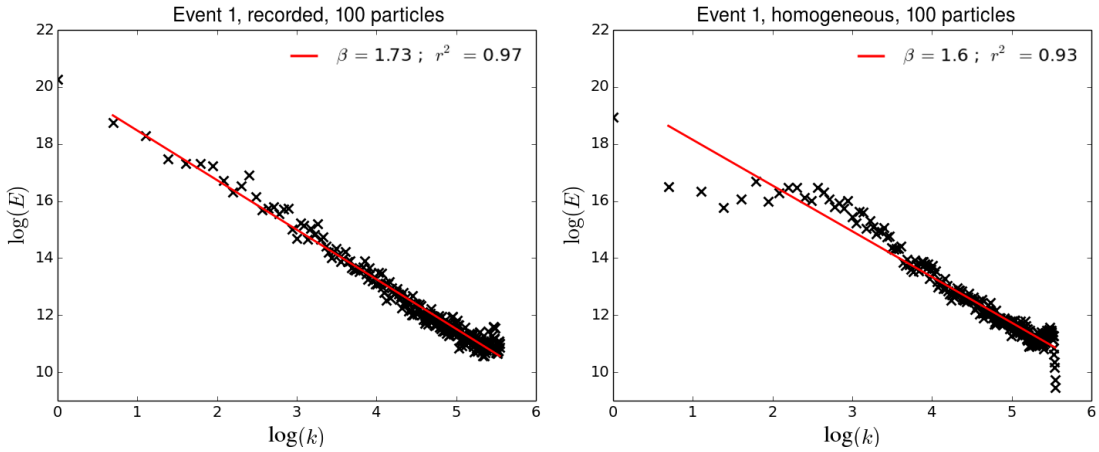


Figure 7.11 – Spectral analyses for the analysed snow accumulation maps on the 24th February 2015.

in Figure 7.12. There is power-law behaviour (a straight line in the log-log plot) shown for some temporal scales, for example for accumulations of 100 snow particles between time periods of about 15 to 100 seconds. At the mean observed snowflake velocity in our three events (1.1 m s^{-1}), this range of times corresponds to spatial scales of about 16 to 110 m. The power-law behaviour thus supports the idea that there is some scaling behaviour in snow particle distributions over scales in the order of metres to tens of metres in this context, as we showed in Section 7.5.1.

7.6 Conclusions

In this chapter we have shown the results of a universal multifractal analyses of snowfall at high spatial and temporal resolutions. Methodologies similar to Gires et al. (2015) were used to analyse snowfall data. Three one-hour-long periods of snowfall in calm conditions were studied. The data were collected by a 2DVD in the Swiss Alps in the northern winter of 2014/2015. Multifractal analyses were performed on ballistically reconstructed vertical columns and high resolution time series of quantities derived from particle concentration and

E	Num P	β (meas)	r^2 (meas)	β (hom)	r^2 (hom)
1	10	1.76	0.96	1.81	0.95
1	50	1.78	0.97	1.75	0.97
1	100	1.78	0.97	1.76	0.97
2	10	1.58	0.95	1.46	0.89
2	50	1.75	0.97	1.64	0.95
2	100	1.78	0.97	1.73	0.96
3	10	1.79	0.97	1.65	0.95
3	50	1.75	0.98	1.65	0.95
3	100	1.74	0.97	1.60	0.94

Table 7.6 – Results of spectral analysis on accumulated particle maps, for the number of particles “Num P”, for measured (meas) and homogeneously (hom) distributed particle maps, per event (E).

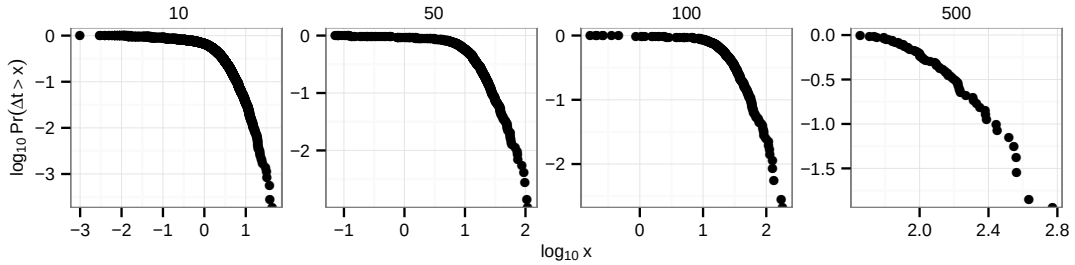


Figure 7.12 – Exceedance probabilities for the time t required to observe n snowflakes, where n is equal to 10, 50, 100 and 500.

mass, and two-dimensional snowfall accumulation maps within the sampling window of the 2DVD.

There were some difficulties in estimation of the multifractal parameters, due to the small range of scales it was possible to investigate, and the scaling observed was generally of poor quality. With the available data, scaling was observed in the vertical columns of particle concentrations between scales of about 35 m to 2 m (and sometimes down to 0.5 m). At smaller scales it was found that the vertical fields behaved as if they were uniformly distributed. The results suggested that scaling was present in the snowfall intensity time series, for scales of almost one hour to just over one minute. Analysis of two-dimensional snowflake accumulation maps over the small sampling window of the 2DVD showed no scaling, which agreed with the other results that showed little or no scaling at the scale of the 2DVD's sampling area. Further investigations on larger data sets are needed to extend the results to a wider range of scales, and to obtain more robust UM parameter estimates that could be used to fully simulate snowfall fields.

8 Conclusions and perspectives

8.1 Summary

Rainfall is a vital process on Earth, but its extremes such as droughts and floods put life and property at risk and have large societal and environmental impacts. Accurate measurements of rainfall are required in order to better understand its physical processes, and through this understanding to improve our ability to model and predict rain. Point measurements of rainfall using rain gauges are generally accurate, but suffer from a lack of spatial representativity. Weather radars offer the ability to observe rainfall over large regions, but their measurements are indirect and must be interpreted together with knowledge of the microstructure of the water particles in the atmosphere. The microstructure of rain, summarised statistically by the raindrop size distribution (DSD), is highly variable in time and space. This variability complicates measurement of rain both on the point and areal scale. All bulk variables of rain can be calculated from the DSD, and although the DSD describes the rainfall process at the scale of individual drops, it is fundamental to many larger-scale rainfall processes. Perhaps most importantly for the measurement of rainfall, the DSD describes the interactions of the electromagnetic waves emitted by radars with the raindrops in the atmosphere. The work presented in this thesis forms a contribution towards the better understanding and characterisation of the variability of the DSD across scales.

The techniques developed and results drawn were made using, for the most part, a network of disdrometers and radars that was located in Ardèche, France. This region experiences heavy Mediterranean precipitation. While the broad meteorological ingredients and processes that cause these precipitation events are well understood, forecasting of precise storm locations and expected rainfall amounts remains difficult. In this thesis, we presented new methods for the investigation of DSD variability, and applied them to data collected in this Mediterranean region. The work followed a logical arc, from collection of accurate measurements of the DSD, to methods for their study and interpretation, to the effects of DSD variability on algorithms used not only in the Mediterranean but world-wide.

In Chapter 2, a new correction method for DSDs measured using Parsivel disdrometers was

presented. The method uses the high-resolution 2DVD as a reference instrument, and corrects measured drop velocities and concentrations so that they match those of the 2DVD on average. After correction, DSD moments closely matched those measured by the 2DVD, and in the majority of cases the rain rate measured by a disdrometer was closer to that measured by a collocated rain gauge. In Chapter 3, a new geostatistical method for spatial interpolation and stochastic simulation of DSDs was proposed. The technique deals explicitly with the statistical requirements of kriging, and is able to interpolate or simulate the experimental, non-parametric DSD, meaning that it is a useful way to investigate the effects of DSD variability on different functional forms of the DSD. Leave-one-out testing showed that DSD interpolations were made with little bias.

In Chapter 4 the two proposed techniques were brought together, and corrected DSDs from the Parsivel network in the Mediterranean were used to produce many stochastically simulated gridded DSD fields. These fields were used to quantify the sub-grid variability of the DSD in two typically-sized regions that corresponded to real-world areal rainfall retrieval pixel sizes. The error introduced by assuming a point measurement of rainfall represents an area was quantified. The rainfall retrieval methods of the GPM space-borne radar and the COSMO numerical weather prediction model were tested, and it was shown that while they usually provide good estimates of rain intensity and radar reflectivity, other DSD properties such as the characteristic drop diameter were often not representative of the sub-grid DSDs.

Double-moment normalisation of the DSD is used to present the DSD in a compact form, using the assumption that two DSD moments can describe most DSD variability around an invariant normalised DSD function. In Chapter 5 the invariance of the double-moment normalised DSD of Lee et al. (2004) was tested using instrument networks in three regions. It was the first test of the invariance of this normalised DSD over large horizontal distances (up to 100 km), in the vertical rainfall profile, and between different climatic regions. It was shown that for practical purposes the double-moment normalised DSD can be assumed to be invariant. This assumption was used in Chapter 6, in which a new method for the retrieval of the DSD from polarimetric radar data was proposed. The new technique uses the double-moment normalisation of Lee et al. (2004) on two DSD moments retrieved from polarimetric data. A new method for treating noisy radar data was proposed and shown to improve DSD-retrieval performance with real radar data. The proposed technique was tested against a state-of-the-art DSD-retrieval method and was found to perform as well as or better than the existing technique for retrieving DSDs measured on the ground. The new technique outperformed the existing technique when its results were compared to DSDs indirectly measured aloft by an MRR.

In Chapter 7, an application of universal multifractal analysis on high-resolution snowfall data was shown. The snowfall was measured by a 2DVD in the Swiss Alps, and UM analyses were made on ballistically reconstructed columns and high-resolution (0.1 s) time series of snowfall concentration and mass, as well as on high-resolution snowfall accumulation maps. This was the first application of multifractal analysis to reconstructed vertical columns of snowflake

concentration and mass, the also to horizontal accumulations of individual snowflakes. It was the first application of full universal multifractal analysis to snowfall-only time series at such high temporal resolution. The results showed no scaling at very small scales (under one minute or two metres), indicating that at sub-metre and sub-minute scales snowfall appears to be distributed homogeneously. Scaling was observed in the vertical columns between about 35 m and 2 m, and weak scaling was observed in the time series measurements for scales of almost one hour to just over one minute.

8.2 Contributions

The main contributions of this thesis are as follows:

1. A method for the correction of classed disdrometer measurements of the DSD, in which a higher-resolution instrument such as the 2DVD is used as a reference, was proposed.

2. We developed a new geostatistical method for spatial interpolation and stochastic simulation of experimental DSD spectra, in which no functional form of the DSD is required. The method can be used to produce grids of full non-parametric DSDs, which are useful for testing DSD variability and algorithms that use different DSD models.

3. The stochastic simulation method was used to investigate the small-scale variability of the DSD in Mediterranean rainfall, and the change of support problem for rainfall estimation. We showed that assuming that a point represents an area introduces significant error on rainfall measurements, and quantified this error for typical domains. It was also shown that algorithms used by GPM and COSMO for areal rainfall retrieval provide microphysical values that are sometimes not representative of the sub-grid process. The primary reasons for these errors were identified.

4. The double-moment DSD normalisation technique of Lee et al. (2004) was shown to produce a normalised DSD that can, for practical use, be considered invariant through horizontal and vertical displacement. This means that double-moment DSD normalisation is a powerful tool for the investigation of the three-dimensional variability of the DSD.

5. We introduced a new DSD-retrieval technique that can estimate the DSD from polarimetric radar data using double-moment normalisation. The technique performs as well as or better than a contemporary DSD-retrieval technique, and is flexible because no specific model of the double-normalised DSD is prescribed.

6. Using universal multifractal analysis on high-resolution snowfall data, scaling properties were observed in reconstructed vertical columns of snowfall from two to 35 metres, and on high-resolution time series between one-minute and almost one hour. It was shown that snowfall at sub-metre or sub-minute scale displays properties implying homogeneous distribution of snowflakes.

8.3 Perspectives

The work presented in this thesis contributes to the understanding of small-scale variability of the DSD, and its effects on the measurement of precipitation. There are many avenues for future research to build upon what has been presented here, and many research questions pertinent to this field remain. First, the correction method for DSDs measured by Parsivel disdrometers, presented in Chapter 2, was trained and tested on disdrometer data from Ardèche in France and Payerne in Switzerland. It would be useful to test this method on data from collocated Parsivels and 2DVDs in other climatic regions and more distant locations.

Second, this thesis focused on the variability of the DSD more in the horizontal than in the vertical. Indeed, in this research field there is a tendency to look at precipitation processes using "slices" in space, whether they be horizontal or vertical. This is simply because it is easier to observe rainfall processes this way, for example with scanning radars. Interpolation and simulation offer ways to characterise the process between measured points. We have shown in this thesis the utility of simulating the non-parametric DSD without assumption of a functional form; namely that any fitting errors are avoided, and (more importantly) that it is then possible to test functional DSD models and the algorithms that use them. While some work on 3D simulation of DSD model parameters has been published (e.g. Schleiss and Smith, 2015), more work is required on full 3D interpolation and simulation of the non-parametric DSD. Such methods could be used to quantify DSD variability in three dimensions, and thus extend the kind of tests we made on horizontal domains into 3D volumes that more closely resemble actual radar volumes or 3D NWP model outputs. Such research would be of particular use for precipitation estimation with satellite-based radars such as GPM.

Third, to simulate or interpolate the DSD in the vertical, the vertical variability of the DSD would need to be characterised. In our horizontal simulation framework, the DSD is treated so as to remove most of the non-stationarity in the rainfall process. This is not as simple in the vertical, in which the process is likely to be highly non-stationary. More research is required into the vertical evolution of the DSD and ways that it could be modelled or characterised. The vertical variability of the DSD is heavily influenced by the microphysical processes that occur as rain falls (e.g. Pruppacher and Klett, 2000; Rosenfeld and Ulbrich, 2003), so any research into the characterisation of vertical DSD variability should explicitly take microphysical processes into account.

Fourth, the relative invariance of the double-moment normalised DSD is useful for a range of applications (see Section 5.7). It would be of research interest to determine the reasons – apart from instrument error – for the small amount of variability that does exist in the normalised DSD. It is expected that different microphysical processes would produce changes in the shape of the single- or double-normalised DSD (Lee et al., 2004), and further work using rainfall classified by dominant microphysical process is required to find and characterise links between normalised DSDs and physical processes. Finally, the performance of the DSD-retrieval method proposed in Chapter 6 is tied directly to how accurately DSD moments can be

retrieved from radar data. Further work is required to improve the performance of prediction of low-order DSD moments from polarimetric variables that may be noisy.

A Data acknowledgements

Data used in this thesis were collected from networks of disdrometers and radars. In addition to the papers cited when datasets are introduced, we make the following data acknowledgements:

HyMeX LTE Parsivels/2DVD, SOP2012: deployed by J. Grazioli, J. Jaffrain, M. Schleiss, D. Scipion, S. Studzinski, A. Berne, maintained by T. Raupach, M. Schleiss, J. Grazioli, S. Studzinski (EPFL LTE). Data from HyMeX program sponsored by grants MISTRALS/HyMeX, ANR-2011-BS56-027, FLOODSCALE project, EPFL LTE. 2DVD and Parsivel data processed by T. Raupach.

HyMeX LTE Parsivels/2DVD/weather station, SOP2013: deployed/maintained by J. Grazioli, T. Raupach, D. Wolfensberger, S. Studzinski, A. Berne (EPFL LTE). Sponsors as for SOP2012. Data processed by T. Raupach.

HyMeX HPicoNet Parsivels/rain gauges: deployed/maintained by B. Boudevillain, S. Gérard, G. Molinié (Laboratoire d'étude des Transferts en Hydrologie et Environnement, Grenoble University). Data from HyMeX program sponsored by grants MISTRALS/HyMeX, ANR-2011-BS56-027 FLOODSCALE project and OHMCV (Cevennes-Vivarais Mediterranean Hydrometeorological Observatory). Parsivel data processed by T. Raupach.

HyMeX Parsivel^2 at Pradel Grainage: deployed/maintained by Y. Pointin, J. Van Baelen, and colleages (Laboratoire de Météorologie Physique (LaMP), Université Blaise Pascal de Clermont-Ferrand). Data from HyMeX program sponsored by grants MISTRALS/HyMeX, ANR-2011-BS56-027 FLOODSCALE project, Laboratoire de Meteorologie Physique, Domaine Olivier de Serres, CERMOSEM and OHMCV (Cevennes-Vivarais Mediterranean Hydrometeorological Observatory). Parsivel data processed by T. Raupach.

HyMeX MXPol: Deployed by J. Grazioli, S. Studzinski, A. Berne (EPFL LTE). Maintained by the above and T. Raupach. Project sponsors as for SOP2012. Data processed by J. Grazioli.

Payerne: Parsivels deployed by J. Jaffrain (Meteosuisse), maintained by J. Jaffrain, T. Raupach. 2DVD deployed/maintained by T. Raupach. MXPol deployed/maintained by J. Grazioli, A. Berne, S. Studzinski (EPFL LTE). MXPol data processed by J. Grazioli, Parsivel and 2DVD data processed by T. Raupach.

Appendix A. Data acknowledgements

HyMeX MRRs: Pradel-Grainage, deployed/maintained by Y. Pointin, J. Van Baelen, and colleages (Laboratoire de Météorologie Physique (LaMP), Université Blaise Pascal de Clermont-Ferrand). Montbrun deployed/maintained by J. Grazioli, A. Berne (EPFL LTE). MRR data processed by J. Grazioli.

Iowa IFloodS: NASA Iowa Flood Studies (IFloodS) Global Precipitation Mission (GPM) ground validation campaign, University of Iowa's X-band radar XPOL5 (Mishra et al., 2016), Parsivel^2 disdrometers (Petersen et al., 2014). Radar data processing to subtract bias by J. Grazioli (EPFL LTE).

Davos Versuchsfeld: 2DVD deployed/maintained by T. Raupach, D. Wolfensberger, L. Habegger, J. Grazioli, C. Praz (EPFL LTE), and M. Ruesch (WSL Institute for Snow and Avalanche Research SLF). We thank SLF for hosting the instrument. The meteorological data (temperature, wind speed, and laser disdrometer solid precipitation intensity) were obtained from the Solid Precipiation Intercomparison Experiment (SPICE) and MeteoSwiss (see Roulet et al., 2014). Regarding only Chapter 7 in which SPICE data was used: results presented in this work were obtained as part of the Solid Precipitation InterComparison Experiment (SPICE), conducted on behalf of the World Meteorological Organization (WMO) Commission for Instruments and Methods of Observation (CIMO). The analysis and views described herein are those of the author(s) at this time, and do not necessarily represent the official outcome of WMO SPICE. Mention of commercial companies or products is solely for the purposes of information and assessment within the scope of the present work, and does not constitute an endorsement by the author(s) or WMO.

Bibliography

Anagnostou, M. N., E. N. Anagnostou, G. Vulpiani, M. Montopoli, F. S. Marzano, and J. Vivekanandan, 2008: Evaluation of X-band polarimetric-radar estimates of drop-size distributions from coincident S-band polarimetric estimates and measured raindrop spectra. *IEEE T. Geosci. Remote Sens.*, **46 (10)**, 3067–3075, doi:10.1109/TGRS.2008.2000757.

Anagnostou, M. N., J. Kalogiros, E. N. Anagnostou, and A. Papadopoulos, 2009: Experimental results on rainfall estimation in complex terrain with a mobile X-band polarimetric weather radar. *Atmos. Res.*, **94 (4)**, 579–595, doi:10.1016/j.atmosres.2009.07.009.

Anagnostou, M. N., J. Kalogiros, E. N. Anagnostou, M. Tarolli, A. Papadopoulos, and M. Borga, 2010: Performance evaluation of high-resolution rainfall estimation by X-band dual-polarization radar for flash flood applications in mountainous basins. *J. Hydrol.*, **394 (1)**, 4–16, doi:10.1016/j.jhydrol.2010.06.026.

Anagnostou, M. N., J. Kalogiros, F. S. Marzano, E. N. Anagnostou, M. Montopoli, and E. Piccioti, 2013: Performance evaluation of a new dual-polarization microphysical algorithm based on long-term X-band radar and disdrometer observations. *J. Hydrometeor.*, **14 (2)**, 560–576, doi:10.1175/JHM-D-12-057.1.

Andrieu, H., J. Creutin, G. Delrieu, and D. Faure, 1997: Use of a weather radar for the hydrology of a mountainous area. Part I: radar measurement interpretation. *J. Hydrol.*, **193 (1–4)**, 1–25, doi:10.1016/S0022-1694(96)03202-7.

Andronache, C., 2004: Precipitation removal of ultrafine aerosol particles from the atmospheric boundary layer. *J. Geophys. Res.: Atmos.*, **109 (D16)**, doi:10.1029/2003JD004050, URL http://dx.doi.org/10.1029/2003JD004050.

Andsager, K., K. V. Beard, and N. F. Laird, 1999: Laboratory measurements of axis ratios for large rain drops. *J. Atmos. Sci.*, **56 (15)**, 2673–2683, doi:10.1175/1520-0469(1999)056<2673:LMOARF>2.0.CO;2.

Atlas, D., R. Srivastava, and R. S. Sekhon, 1973: Doppler radar characteristics of precipitation at vertical incidence. *Rev. Geophys.*, **11 (1)**, 1–35, doi:10.1029/RG011i001p00001.

Atlas, D., and C. Ulbrich, 2006: Drop size spectra and integral remote sensing parameters in the transition from convective to stratiform rain. *Geophys. Res. Lett.*, **33 (16)**, doi:10.1029/2006GL026824.

Baldauf, M., A. Seifert, J. Förstner, D. Majewski, M. Raschendorfer, and T. Reinhardt, 2011: Operational convective-scale numerical weather prediction with the COSMO model: description and sensitivities. *Mon. Weather Rev.*, **139 (12)**, 3887–3905, doi:10.1175/MWR-D-10-05013.1.

Barancourt, C., J.-D. Creutin, and J. Rivoirard, 1992: A method for delineating and estimating rainfall fields. *Water Resour. Res.*, **28 (4)**, 1133–1144, doi:10.1029/91WR02896.

Battaglia, A., E. Rustemeier, A. Tokay, U. Blahak, , and C. Simmer, 2010: PARSIVEL snow observations: a critical assessment. *J. Atmos. Oceanic Technol.*, **27 (2)**, 333–344, doi:10.1175/2009JTECHA1332.1.

Beard, K. V., 1976: Terminal velocity and shape of cloud and precipitation drops aloft. *J. Atmos. Sci.*, **33 (5)**, 851–864, doi:10.1175/1520-0469(1976)033<0851:TVASOC>2.0.CO;2.

Beard, K. V., and C. Chuang, 1987: A new model for the equilibrium shape of raindrops. *J. Atmos. Sci.*, **44 (3)**, 1509–1524, doi:10.1175/1520-0469(1987)044<1509:ANMFTE>2.0.CO;2.

Berne, A., J. Jaffrain, and M. Schleiss, 2012: Scaling analysis of the variability of the rain drop size distribution at small scale. *Advances in Water Resources*, **45**, 2–12, doi:10.1016/j.advwatres.2011.12.016.

Berne, A., and W. Krajewski, 2013: Radar for hydrology: Unfulfilled promise or unrecognized potential? *Advances in Water Resources*, **51**, 357 – 366, doi:10.1016/j.advwatres.2012.05.005, 35th Year Anniversary Issue.

Berrocal, V. J., A. E. Raftery, and T. Gneiting, 2008: Probabilistic quantitative precipitation field forecasting using a two-stage spatial model. *Adv. Atmos. Sci.*, **2 (4)**, 1170–1193, doi:10.1214/08-AOAS2.

Blocken, B., and J. Carmeliet, 2004: A review of wind-driven rain research in building science. *Journal of Wind Engineering and Industrial Aerodynamics*, **92 (13)**, 1079–1130, doi:10.1016/j.jweia.2004.06.003.

Brandes, E., G. Zhang, and J. Vivekanandan, 2002: Experiments in rainfall estimation with a polarimetric radar in a subtropical environment. *J. Appl. Meteor.*, **41 (6)**, 674–685.

Brandes, E. A., G. Zhang, and J. Vivekanandan, 2003: An evaluation of a drop distribution-based polarimetric radar rainfall estimator. *Journal of Applied Meteorology*, **42 (5)**, 652–660, doi:10.1175/1520-0450(2003)042<0652:AEOADD>2.0.CO;2.

Brandes, E. A., G. Zhang, and J. Vivekanandan, 2004a: Comparison of polarimetric radar drop size distribution retrieval algorithms. *Journal of Atmospheric and Oceanic Technology*, **21 (4)**, 584–598, doi:10.1175/1520-0426(2004)021<0584:COPRDS>2.0.CO;2.

Brandes, E. A., G. Zhang, and J. Vivekanandan, 2004b: Drop size distribution retrieval with polarimetric radar: Model and application. *J. Appl. Meteor.*, **43 (3)**, 461–475, doi:10.1175/1520-0450(2004)043<0461:DSDRWP>2.0.CO;2.

Braud, I., P. Crochet, and J. D. Creutin, 1994: A method for estimating mean areal rainfall using moving trend functions of the intensities. *J. Appl. Meteor.*, **33**, 1551–1561, doi:10.1175/1520-0450(1994)033,1551:AMFEMA.2.0.CO;2.

Bringi, V., G.-J. Huang, V. Chandrasekar, and E. Gorgucci, 2002: A methodology for estimating the parameters of a gamma raindrop size distribution model from polarimetric radar data: Application to a squall-line event from the TRMM/Brazil campaign. *J. Atmos. Oceanic Technol.*, **19 (5)**, 633–645, doi:10.1175/1520-0426(2002)019<0633:AMFETP>2.0.CO;2.

Bringi, V., L. Tolstoy, M. Thurai, and W. Petersen, 2015: Estimation of spatial correlation of drop size distribution parameters and rain rate using NASA's S-band polarimetric radar and 2D video disdrometer network: Two case studies from MC3E. *Journal of Hydrometeorology*, **16 (3)**, 1207–1221, doi:10.1175/JHM-D-14-0204.1.

Bringi, V. N., and V. Chandrasekar, 2001: *Polarimetric Doppler weather radar*. Cambridge University Press, 662 pp.

Bringi, V. N., V. Chandrasekar, J. Hubbert, E. Gorgucci, W. L. Randeu, and M. Schoenhuber, 2003: Raindrop size distribution in different climatic regimes from disdrometer and dual-polarized radar analysis. *J. Atmos. Sci.*, **60 (2)**, 354–365, doi:10.1175/1520-0469(2003)060<0354:RSDIDC>2.0.CO;2.

Calder, I. R., 1986: A stochastic model of rainfall interception. *Journal of Hydrology*, **89 (1)**, 65–71.

Cao, Q., and G. Zhang, 2009: Errors in estimating raindrop size distribution parameters employing disdrometer and simulated raindrop spectra. *J. Appl. Meteor. Climate*, **48 (2)**, 406–425, doi:10.1175/2008JAMC2026.1.

Chapon, B., G. Delrieu, M. Gosset, and B. Boudevillain, 2008: Variability of rain drop size distribution and its effect on the $Z-R$ relationship: a case study for intense Mediterranean rainfall. *Atmos. Res.*, **87 (1)**, 52–65, doi:10.1016/j.atmosres.2007.07.003.

Chilès, J.-P., and P. Delfiner, 1999: *Geostatistics: Modeling spatial uncertainty*. Probability and statistics, Wiley, 695 pp.

Chua, S.-H., and R. L. Bras, 1982: Optimal estimators of mean areal precipitation in regions of orographic influence. *J. Hydrol.*, **57 (1-2)**, 23–48, doi:10.1016/0022-1694(82)90101-9.

Ciais, P., and Coauthors, 2005: Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature*, **437 (7058)**, 529–533, doi:10.1038/nature03972.

Crane, R. K., 1971: Propagation phenomena affecting satellite communication systems operating in the centimeter and millimeter wavelength bands. *Proceedings of the IEEE*, **59 (2)**, 173–188, doi:10.1109/PROC.1971.8123.

Cressie, N. A. C., 1993: *Statistics for spatial data*. Probability and mathematical statistics, Wiley, 900 pp.

Creutin, J.-D., and C. Obled, 1982: Objective analyses and mapping techniques for rainfall fields: an objective comparison. *Water Resour. Res.*, **18 (2)**, 413–431, doi:10.1029/WR018i002p00413.

De Lima, M., and J. De Lima, 2009: Investigating the multifractality of point precipitation in the Madeira archipelago. *Nonlinear Processes in Geophysics*, **16 (2)**, 299–311, doi:10.5194/npg-16-299-2009.

De Lima, M., and J. Grasman, 1999: Multifractal analysis of 15-min and daily rainfall from a semi-arid region in Portugal. *J. Hydrol.*, **220 (1)**, 1–11, doi:10.1016/S0022-1694(99)00053-0.

De Montera, L., L. Barthès, C. Mallet, and P. Golé, 2009: The effect of rain-no rain intermittency on the estimation of the universal multifractals model parameters. *J. Hydrometeor.*, **10 (2)**, 493–506, doi:10.1175/2008JHM1040.1.

De Oliveira, V., 2004: A simple model for spatial rainfall fields. *Stoch. Env. Res. Risk A.*, **18 (2)**, 131–140, doi:10.1007/s00477-003-0146-4.

Deems, J. S., S. R. Fassnacht, and K. J. Elder, 2006: Fractal distribution of snow depth from lidar data. *Journal of Hydrometeorology*, **7 (2)**, 285–297, doi:10.1175/JHM487.1.

Delrieu, G., L. Hucke, and J. D. Creutin, 1999: Attenuation in rain for X- and C-band weather radar systems: Sensitivity with respect to the drop size distribution. *Journal of Applied Meteorology*, **38 (1)**, 57–68, doi:10.1175/1520-0450(1999)038<0057:AIRFXA>2.0.CO;2.

Delrieu, G., and Coauthors, 2005: The catastrophic flash-flood event of 8-9 September 2002 in the Gard region, France: a first case study for the Cévennes-Vivarais Mediterranean Hydro-meteorological Observatory. *J. Hydrometeor.*, **6 (1)**, 34–52, doi:10.1175/JHM-400.1.

Doms, G., and Coauthors, 2011: A description of the nonhydrostatic regional COSMO model, Part II: Physical parameterization. URL http://www.cosmo-model.org/content/model/documentation/core/cosmoPhysParamtr.pdf, accessed 04.08.2015.

Drobinski, P., and Coauthors, 2014: HyMeX, a 10-year multidisciplinary program on the Mediterranean water cycle. *Bull. Amer. Meteor. Soc.*, **95**, doi:10.1175/BAMS-D-12-00242.1.

Ducrocq, V., and Coauthors, 2014: Hymex-sop1: The field campaign dedicated to heavy precipitation and flash flooding in the northwestern mediterranean. *Bull. Amer. Meteor. Soc.*, **95 (7)**, 1083–1100, doi:10.1175/BAMS-D-12-00244.1.

Easterling, D. R., G. A. Meehl, C. Parmesan, S. A. Changnon, T. R. Karl, and L. O. Mearns, 2000: Climate extremes: Observations, modeling, and impacts. *Science*, **289 (5487)**, 2068–2074, doi:10.1126/science.289.5487.2068.

Emmanuel, I., H. Andrieu, E. Leblois, and B. Flahaut, 2012: Temporal and spatial variability of rainfall at the urban hydrological scale. *J. Hydrol.*, **430-431 (0)**, 162–172, doi:10.1016/j.jhydrol.2012.02.013.

Fabry, F., 1996: On the determination of scale ranges for precipitation fields. *Journal of Geophysical Research: Atmospheres*, **101 (D8)**, 12 819–12 826, doi:10.1029/96JD00718.

Fraedrich, K., and C. Larnder, 1993: Scaling regimes of composite rainfall time series. *Tellus A*, **45 (4)**, 289–298, doi:10.1034/j.1600-0870.1993.t01-3-00004.x.

Frei, C., and C. Schär, 1998: A precipitation climatology of the Alps from high-resolution rain-gauge observations. *Int. J. Climatol.*, **18 (8)**, 873–900, doi:10.1002/(SICI)1097-0088(19980630)18:8<873::AID-JOC255>3.0.CO;2-9.

Gires, A., I. Tchiguirinskaia, and D. Schertzer, 2016: Multifractal comparison of the outputs of two optical disdrometers. *Hydrological Sciences Journal*, **61 (9)**, 1641–1651, doi:10.1080/02626667.2015.1055270.

Gires, A., I. Tchiguirinskaia, D. Schertzer, and A. Berne, 2015: 2DVD data revisited: Multifractal insights into cuts of the spatiotemporal rainfall process. *J. Hydrometeor.*, **16 (2)**, 548–562, doi:10.1175/JHM-D-14-0127.1.

Gires, A., I. Tchiguirinskaia, D. Schertzer, and S. Lovejoy, 2011: Analyses multifractales et spatio-temporelles des précipitations du modèle Méso-NH et des données radar. *Hydrol. Sci. J.*, **56 (3)**, 380–396, doi:10.1080/02626667.2011.564174.

Gires, A., I. Tchiguirinskaia, D. Schertzer, A. Schellart, A. Berne, and S. Lovejoy, 2014: Influence of small scale rainfall variability on standard comparison tools between radar and rain gauge data. *Atmos. Res.*, **138**, 125–138, doi:10.1016/j.atmosres.2013.11.008.

Godart, A., S. Anquetin, E. Leblois, and J. Creutin, 2011: The contribution of orographically driven banded precipitation to the rainfall climatology of a Mediterranean region. *J. Appl. Meteor. Climate*, **50 (11)**, 2235–2246, doi:10.1175/JAMC-D-10-05016.1.

Goodison, B., P. Louie, and D. Yang, 1998: WMO solid precipitation measurement intercomparison: Final report. Tech. rep., WMO, Geneva. No. 67, WMO/TD - No. 872.

Goovaerts, P., 2000: Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *J. Hydrol.*, **228 (1-2)**, 113–129.

Gorgucci, E., V. Chandrasekar, and L. Baldini, 2008: Microphysical retrievals from dual-polarization radar measurements at X band. *J. Atmos. Oceanic Technol.*, **25 (5)**, 729–741, doi:10.1175/2007JTECHA971.1.

Gorgucci, E., V. Chandrasekar, V. N. Bringi, and G. Scarchilli, 2002: Estimation of raindrop size distribution parameters from polarimetric radar measurements. *J. Atmos. Sci.*, **59 (15)**, 2373–2384, doi:10.1175/1520-0469(2002)059<2373:EORSDP>2.0.CO;2.

Granger, R. J., J. W. Pomeroy, and J. Parviainen, 2002: Boundary-layer integration approach to advection of sensible heat to a patchy snow cover. *Hydrol. Processes*, **16 (18)**, 3559–3569, doi:10.1002/hyp.1227.

Haberlandt, U., 2007: Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event. *J. Hydrol.*, **332 (1-2)**, 144 – 157, doi:10.1016/j.jhydrol.2006.06.028.

Hitschfeld, W., and J. Bordan, 1954: Errors inherent in the radar measurement of rainfall at attenuating wavelengths. *J. Meteor.*, **11 (1)**, 58–67, doi:10.1175/1520-0469(1954)011<0058:EIITRM>2.0.CO;2.

Hou, A., G. Skofronick-Jackson, C. Kummerow, and J. Shepherd, 2008: Global precipitation measurement. *Precipitation: Advances in Measurement, Estimation and Prediction*, S. Michaelides, Ed., Springer Berlin Heidelberg, 131–169, doi:10.1007/978-3-540-77655-0_6.

Hou, A. Y., and Coauthors, 2014: The Global Precipitation Measurement mission. *Bulletin of the American Meteorological Society*, **95 (5)**, 701–722, doi:10.1175/BAMS-D-13-00164.1.

Houze, R. A., 2012: Orographic effects on precipitating clouds. *Reviews of Geophysics*, **50 (1)**, doi:10.1029/2011RG000365.

Huang, G.-J., V. Bringi, D. Moisseev, W. Petersen, L. Bliven, and D. Hudak, 2015: Use of 2D-video disdrometer to derive mean density–size and Z_e–SR relations: Four snow cases from the light precipitation validation experiment. *Atmos. Res.*, **153**, 34–48, doi:10.1016/j.atmosres.2014.07.013.

Hubert, P., and Coauthors, 1993: Multifractals and extreme rainfall events. *Geophys. Res. Lett.*, **20 (10)**, 931–934.

Huffman, G. J., and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *Journal of Hydrometeorology*, **8 (1)**, 38–55, doi:10.1175/JHM560.1.

Iggy Litaor, M., T. Seastedt, and D. Walker, 2002: Spatial analysis of selected soil attributes across an alpine topographic/snow gradient. *Landscape Ecology*, **17 (1)**, 71–85, doi:10.1023/A:1015216400909.

Iguchi, T., T. Kozu, J. Kwiatkowski, R. Meneghini, J. Awaka, and K. Okamoto, 2009: Uncertainties in the rain profiling algorithm for the TRMM precipitation radar. *J. Meteorol. Soc. Jpn.*, **87**, 1–30, doi:10.2151/jmsj.87A.1.

Iguchi, T., T. Kozu, R. Meneghini, J. Awaka, and K. Okamoto, 2000: Rain-profiling algorithm for the TRMM precipitation radar. *Journal of Applied Meteorology*, **39 (12)**, 2038–2052, doi:10.1175/1520-0450(2001)040<2038:RPAFTT>2.0.CO;2.

Illingworth, A. J., and T. M. Blackman, 2002: The need to represent raindrop size spectra as normalized gamma distributions for the interpretation of polarization radar observations. *J. Appl. Meteor.*, **41 (3)**, 286–297, doi:10.1175/1520-0450(2002)041<0286:TNTRRS>2.0.CO;2.

Jaffrain, J., and A. Berne, 2011: Experimental quantification of the sampling uncertainty associated with measurements from Parsivel disdrometers. *J. Hydrometeor.*, **12 (3)**, doi:10.1175/2010JHM1244.1.

Jaffrain, J., and A. Berne, 2012a: Influence of the subgrid variability of the raindrop size distribution on radar rainfall estimators. *J. Appl. Meteor. Climate*, **51**, doi:10.1175/JAMC-D-11-0185.1.

Jaffrain, J., and A. Berne, 2012b: Quantification of the small-scale spatial structure of the raindrop size distribution from a network of disdrometers. *J. Appl. Meteor. Climate*, **51**, doi:10.1175/JAMC-D-11-0136.1.

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

Jameson, A., 2015: A bayesian method for upsizing single disdrometer drop size counts for rain physics studies and areal applications. *IEEE T. Geosci. Remote Sens.*, **53 (1)**, 335–343, doi:10.1109/TGRS.2014.2322092.

Jameson, A., and A. Kostinski, 1999: Fluctuation properties of precipitation. Part v: Distribution of rain rates-theory and observations in clustered rain. *J. Atmos. Sci.*, **56 (22)**, 3920–3932, doi:10.1175/1520-0469(1999)056<3920:FPOPPV>2.0.CO;2.

Jameson, A., M. Larsen, and A. Kostinski, 2015a: Disdrometer network observations of fine scale spatial-temporal clustering in rain. *J. Atmos. Sci.*, **72 (4)**, 1648–1666, doi:10.1175/JAS-D-14-0136.1.

Jameson, A., M. Larsen, and A. Kostinski, 2015b: On the variability of drop size distributions over areas. *J. Atmos. Sci.*, **72 (4)**, 1386–1397, doi:10.1175/JAS-D-14-0258.1.

Jameson, A. R., 1983: Microphysical interpretation of multi-parameter radar measurements in rain. Part I: Interpretation of polarization measurements and estimation of raindrop shapes. *Journal of the Atmospheric Sciences*, **40 (7)**, 1792–1802, doi:10.1175/1520-0469(1983)040<1792:MIOMPR>2.0.CO;2.

Jameson, A. R., 1985: Microphysical interpretation of multiparameter radar measurements in rain. Part III: Interpretation and measurement of propagation differential phase shift between orthogonal linear polarizations. *J. Atmos. Sci.*, **42 (6)**, 607–614, doi:10.1175/1520-0469(1985)042<0607:MIOMRM>2.0.CO;2.

Jameson, A. R., and A. B. Kostinski, 2001: What is a raindrop size distribution? *Bull. Amer. Meteor. Soc.*, **82 (6)**, 1169–1177, doi:10.1175/1520-0477(2001)082<1169:WIARSD>2.3.CO;2.

Jarvis, A., H. Reuter, A. Nelson, and E. Guevara, 2008: Hole-filled seamless SRTM data V4, International Centre for Tropical Agriculture (CIAT). URL http://srtm.csi.cgiar.org, accessed 23.03.2016.

Johnson, R. W., D. V. Kliche, and P. L. Smith, 2014: Maximum likelihood estimation of gamma parameters for coarsely binned and truncated raindrop size data. *Q. J. Roy. Meteor. Soc.*, **140 (681)**, 1245–1256, doi:10.1002/qj.2209.

Jolliffe, I., 2002: *Principal Component Analysis*. 2nd ed., Springer-Verlag New York, doi:10.1007/b98835.

Joss, J., and A. Waldvogel, 1967: Ein Spektrograph für Niederschlagstropfen mit automatischer Auswertung / A spectrograph for raindrops with automatic interpretation. *Pure Appl. Geophys.*, **68**, 240–246, doi:10.1007/BF00874898.

Kalogiros, J., M. N. Anagnostou, E. N. Anagnostou, M. Montopoli, E. Picciotti, and F. S. Marzano, 2013: Optimum estimation of rain microphysical parameters from X-band dual-polarization radar observables. *IEEE T. Geosci. Remote Sens.*, **51 (5)**, 3063–3076, doi:10.1109/TGRS.2012.2211606.

Kawanishi, T., and Coauthors, 2000: Remote sensing and applications: Earth, atmosphere and oceans TRMM precipitation radar. *Advances in Space Research*, **25 (5)**, 969 – 972, doi:10.1016/S0273-1177(99)00932-1.

Koster, R. D., and M. J. Suarez, 1995: Relative contributions of land and ocean processes to precipitation variability. *J. Geophys. Res: Atmos.*, **100 (D7)**, 13 775–13 790, doi:10.1029/95JD00176.

Krajewski, W., and J. Smith, 2002: Radar hydrology: rainfall estimation. *Adv. Water Resour.*, **25 (8)**, 1387–1394, doi:10.1016/S0309-1708(02)00062-3.

Krajewski, W. F., and Coauthors, 2006: DEVEX-disdrometer evaluation experiment: basic results and implications for hydrologic studies. *Adv. Water Resour.*, **29**, 807–814, doi:10.1016/j.advwatres.2005.03.018.

Kruger, A., and W. F. Krajewski, 2002: Two-dimensional video disdrometer: a description. *J. Atmos. Oceanic Technol.*, **19 (5)**, 602–617, doi:10.1175/1520-0426(2002)019<0602:TDVDAD>2.0.CO;2.

Kundu, P. K., and R. K. Siddani, 2011: Scale dependence of spatiotemporal intermittency of rain. *Water Resour. Res.*, **47 (W08522)**, doi:10.1029/2010WR010070.

Lavallée, D., S. Lovejoy, and P. Ladoy, 1993: Nonlinear variability and landscape topography: analysis and simulation. *Fractals in geography*, L. de Cola, and N. Lam, Eds., Prentice-Hall, 171–205.

Lee, C. K., G. W. Lee, I. Zawadzki, and K.-E. Kim, 2009: A preliminary analysis of spatial variability of raindrop size distributions during stratiform rain events. *J. Appl. Meteor. Climate*, **48 (2)**, 270–283, doi:10.1175/2008JAMC1877.1.

Lee, G., and I. Zawadzki, 2005: Variability of drop size distributions: time-scale dependence of the variability and its effects on rain estimation. *J. Appl. Meteor.*, **44 (2)**, 241–255, doi:10.1175/JAM2183.1.

Lee, G., I. Zawadzki, W. Szyrmer, D. Sempere-Torres, and R. Uijlenhoet, 2004: A general approach to double-moment normalization of drop size distributions. *J. Appl. Meteor.*, **43 (2)**, 264–281, doi:10.1175/1520-0450(2004)043<0264:AGATDN>2.0.CO;2.

Lee, G. W., A. W. Seed, and I. Zawadzki, 2007: Modeling the variability of drop size distributions in space and time. *Journal of applied meteorology and climatology*, **46 (6)**, 742–756, doi:10.1175/JAM2505.1.

Liao, L., R. Meneghini, and A. Tokay, 2014: Uncertainties of GPM DPR rain estimates caused by DSD parameterizations. *J. Appl. Meteor. Climate*, **53 (11)**, 2524–2537, doi:10.1175/JAMC-D-14-0003.1.

Lin, Y. L., S. Chiao, T. A. Wang, M. L. Kaplan, and R. Weglarz, 2001: Some common ingredients for heavy orographic rainfall. *Weather Forecast.*, **16 (6)**, 633–660, doi:10.1175/1520-0434(2001)016<0633:SCIFHO>2.0.CO;2.

Löffler-Mang, M., and J. Joss, 2000: An optical disdrometer for measuring size and velocity of hydrometeors. *J. Atmos. Oceanic Technol.*, **17 (2)**, 130–139, doi:10.1175/1520-0426(2000)017<0130:AODFMS>2.0.CO;2.

Lovejoy, S., and D. Schertzer, 1995: Multifractals and rain. *New Uncertainty Concepts in Hydrology and Water Resources*, Z. W. Kundzewicz, Ed., Cambridge University Press, 61–103.

Lovejoy, S., D. Schertzer, and V. Allaire, 2008: The remarkable wide range spatial scaling of TRMM precipitation. *Atmos. Res.*, **90 (1)**, 10–32, doi:10.1016/j.atmosres.2008.02.016.

Lovejoy, S., D. Schertzer, and A. Tsonis, 1987: Functional box-counting and multiple elliptical dimensions in rain. *Science*, **235 (4792)**, 1036–1038, doi:10.1126/science.235.4792.1036.

Mandapaka, P. V., P. Lewandowski, W. Eichinger, and W. F. Krajewski, 2009: Multiscaling analysis of high resolution space-time lidar-rainfall. *Nonlinear Proc. in Geophys.*, **16 (5)**, 579–586, doi:10.5194/npg-16-579-2009.

Marshall, J. S., W. Hitschfeld, and K. L. S. Gunn, 1955: Advances in radar weather. *Adv. Geophys.*, **2**, 1–56, doi:10.1016/S0065-2687(08)60310-6.

Marshall, J. S., R. C. Langille, and W. M. Palmer, 1947: Measurement of rainfall by radar. *J. Meteor.*, **4**, 186–192, doi:10.1175/1520-0469(1947)004<0186:MORBR>2.0.CO;2.

Marshall, J. S., and W. M. Palmer, 1948: The distribution of raindrops with size. *J. Meteor.*, **5**, 165–166, doi:10.1175/1520-0469(1948)005<0165:TDORWS>2.0.CO;2.

Masson, D., and C. Frei, 2014: Spatial analysis of precipitation in a high-mountain region: exploring methods with multi-scale topographic predictors and circulation types. *Hydrol. Earth Syst. Sci.*, **18 (11)**, 4543–4563, doi:10.5194/hess-18-4543-2014.

Matheron, G., 1971: *The Theory of Regionalized Variables and its Applications*. Ecole Nationale Supérieure des Mines de Paris, Paris, 218 pp.

Matrosov, S. Y., D. E. Kingsmill, B. E. Martner, and F. M. Ralph, 2005: The utility of X-band polarimetric radar for quantitative estimates of rainfall parameters. *Journal of Hydrometeorology*, **6 (3)**, 248–262, doi:10.1175/JHM424.1.

METEK, 2010: MRR physical basics. Tech. rep., METEK Metorologische Messtechnik GmbH, Elmshom, Germany. Updated 2012-03-13, valid for MRR Service Version ≥ 5.2.0.9.

METEK, 2013: MRR-2 Micro Rain Radar user manual. Tech. rep., METEK Metorologische Messtechnik GmbH, Elmshom, Germany. 2014-03-07 Valid for MRR Service Version 6.0.0.6 / Firmware 6.10.

MeteoFrance, 2014: Données climatiques de la station de Montélimar. URL http://www.meteofrance.com/climat/france/station/26198001/normales, accessed 2014-11-20.

MeteoSwiss, 2013: Federal office of meteorology and climatology MeteoSwiss, climate normals Payerne, reference period 1981-2010. URL http://meteosuisse.ch/files/kd/climsheet/en/PAY_norm8110.pdf, accessed 2014-11-19.

Miniscloux, F., J.-D. Creutin, and S. Anquetin, 2001: Geostatistical analysis of orographic rain bands. *J. Appl. Meteor.*, **40 (11)**, 1835–1854, doi:10.1175/1520-0450(2001)040<1835:GAOOR>2.0.CO;2.

Miriovsky, B., and Coauthors, 2004: An experimental study of small-scale variability of radar reflectivity using disdrometer observations. *J. Appl. Meteor.*, **43 (1)**, 106–118, doi:10.1175/1520-0450(2004)043<0106:AESOSV>2.0.CO;2.

Mishchenko, M. I., and L. D. Travis, 1998: Capabilities and limitations of a current FORTRAN implementation of the T-matrix method for randomly oriented, rotationally symmetric scatterers. *J. Quant. Spectrosc. Radiat. Transfer*, **60 (3)**, 309–324, doi:10.1016/S0022-4073(98)00008-9.

Mishra, K. V., and Coauthors, 2016: Deployment and performance analyses of high-resolution Iowa XPOL radar system during the NASA IFloodS campaign. *J. Hydrometeor.*, **17 (2)**, 455–479, doi:10.1175/JHM-D-15-0029.1.

Moisseev, D. N., and V. Chandrasekar, 2007: Examination of the μ-λ relation suggested for drop size distribution parameters. *J. Atmos. Oceanic Technol.*, **24 (5)**, 847–855, doi:10.1175/JTECH2010.1.

Molnar, P., and P. Burlando, 2008: Variability in the scale properties of high-resolution precipitation data in the alpine climate of Switzerland. *Water Resources Research*, **44 (10)**, doi:10.1029/2007WR006142.

Moumouni, S., M. Gosset, and E. Houngninou, 2008: Main features of rain drop size distributions observed in Benin, West Africa, with optical disdrometers. *Geophys. Res. Lett.*, **35 (23)**, doi:10.1029/2008GL035755.

Nešpor, V., W. F. Krajewski, and A. Kruger, 2000: Wind-induced error of raindrop size distribution measurement using a two-dimensional video disdrometer. *J. Atmos. Oceanic Technol.*, **17 (11)**, 1483–1492, doi:10.1175/1520-0426(2000)017<1483:WIEORS>2.0.CO;2.

Nuissier, O., V. Ducrocq, D. Ricard, C. Lebeaupin, and S. Anquetin, 2008: A numerical study of three catastrophic precipitating events over southern France. I: numerical framework and synoptic ingredients. *Q. J. Roy. Meteor. Soc.*, **134 (630)**, 111–130, doi:10.1002/qj.200.

Nuissier, O., B. Joly, A. Joly, V. Ducrocq, and P. Arbogast, 2011: A statistical downscaling to identify the large-scale circulation patterns associated with heavy precipitation events over southern France. *Q. J. Roy. Meteor. Soc.*, **137 (660)**, 1812–1827, doi:10.1002/qj.866.

Nykanen, D. K., and D. Harris, 2003: Orographic influences on the multiscale statistical properties of precipitation. *J. Geophys. Res: Atmos.*, **108 (D8)**, doi:10.1029/2001JD001518.

Olsson, J., 1995: Limits and characteristics of the multifractal behaviour of a high-resolution rainfall time series. *Nonlinear processes in Geophysics*, **2 (1)**, 23–29, doi:10.5194/npg-2-23-1995.

Park, S.-G., V. N. Bringi, V. Chandrasekar, M. Maki, and K. Iwanami, 2005a: Correction of radar reflectivity and differential reflectivity for rain attenuation at X band. Part I: Theoretical and empirical basis. *J. Atmos. Oceanic Technol.*, **22 (11)**, 1621–1632, doi:10.1175/JTECH1803.1.

Park, S.-G., M. Maki, K. Iwanami, V. N. Bringi, and V. Chandrasekar, 2005b: Correction of radar reflectivity and differential reflectivity for rain attenuation at X band. Part II: Evaluation and application. *J. Atmos. Oceanic Technol.*, **22 (11)**, 1633–1655, doi:10.1175/JTECH1804.1.

Pebesma, E. J., 2004: Multivariate geostatistics in S: the gstat package. *Comput. Geosci.*, **30 (7)**, 683–691, doi:10.1016/j.cageo.2004.03.012.

Peel, M. C., B. L. Finlayson, and T. A. McMahon, 2007: Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sci.*, **11 (5)**, 1433–1644, doi:10.5194/hess-11-1633-2007.

Peters, G., B. Fischer, and T. Andersson, 2002: Rain observations with a vertically looking Micro Rain Radar (MRR). *Boreal Environ. Res.*, **7 (4)**, 353–362.

Peters, G., B. Fischer, and M. Clemens, 2010: Rain attenuation of radar echoes considering finite-range resolution and using drop size distributions. *Journal of Atmospheric and Oceanic Technology*, **27 (5)**, 829–842, doi:10.1175/2009JTECHA1342.1.

Peters, G., B. Fischer, H. Münster, M. Clemens, and A. Wagner, 2005: Profiles of raindrop size distributions as retrieved by microrain radars. *J. Appl. Meteor.*, **44 (12)**, 1930–1949, doi:10.1175/JAM2316.1.

Petersen, W. A., A. Tokay, P. N. Gatlin, and M. T. Wingo, 2014: GPM ground validation Autonomous Parsivel Unit (APU) IFloodS [APU quality controlled drop size distributions]. URL https://fcportal.nsstc.nasa.gov/pub/gpm_validation/ifloods/disdrometers_and_gauges/parsivel/, dataset available online from the NASA Global Hydrology Resource Center DAAC, Huntsville, Alabama, U.S.A., doi:10.5067/GPMGV/IFLOODS/APU/DATA301.

Pielke Jr, R. A., and M. W. Downton, 2000: Precipitation and damaging floods: Trends in the united states, 1932–97. *Journal of Climate*, **13 (20)**, 3625–3637, doi:10.1175/1520-0442(2000)013<3625:PADFTI>2.0.CO;2.

Pruppacher, H., and K. Beard, 1970: A wind tunnel investigation of the internal circulation and shape of water drops falling at terminal velocity in air. *Q. J. Roy. Meteor. Soc.*, **96 (408)**, 247–256, doi:10.1002/qj.49709640807.

Pruppacher, H. R., and J. D. Klett, 2000: *Microphysics of clouds and precipitation*. 2nd ed., Kluwer Academic Publishers, Dordrecht.

Raupach, T. H., and A. Berne, 2016: Small-scale variability of the raindrop size distribution and its effect on areal rainfall retrieval. *J. Hydrometeor.*, **17**, 2077–2104, doi:10.1175/JHM-D-15-0214.1.

Ricard, D., V. Ducrocq, and L. Auger, 2012: A climatology of the mesoscale environment associated with heavily precipitating events over a northwestern mediterranean area. *J. Appl. Meteor. Climate*, **51 (3)**, 468–488, doi:10.1175/JAMC-D-11-017.1.

Roe, G. H., 2005: Orographic precipitation. *Ann. Rev. Earth Planet Sci.*, **33**, 645–671, doi:10.1146/annurev.earth.33.092203.122541.

Rosenfeld, D., and C. W. Ulbrich, 2003: Cloud microphysical properties, processes, and rainfall estimation opportunities. *Radar and Atmospheric Science: A Collection of Essays in Honor of David Atlas, Meteor. Monogr*, R. M. Wakimoto, and R. Srivastava, Eds., American Meteorological Society, Boston, 237–258, doi:10.1007/978-1-878220-36-3_10.

Roulet, Y.-A., A. Reverdin, and F. van den Heuvel, 2014: Commissioning report of site of Weissfluhjoch. URL http://www.wmo.int/pages/prog/www/IMOP/intercomparisons/SPICE/Commissioning/Weissfluhjoch_V2.pdf, accessed 06.04.2016.

Royer, J.-F., A. Biaou, F. Chauvin, D. Schertzer, and S. Lovejoy, 2008: Multifractal analysis of the evolution of simulated precipitation over France in a climate scenario. *Comptes Rendus Geoscience*, **340 (7)**, 431–440, doi:10.1016/j.crte.2008.05.002.

Sassi, M., H. Leijnse, and R. Uijlenhoet, 2014: Sensitivity of power functions to aggregation: Bias and uncertainty in radar rainfall retrieval. *Water Resour. Res.*, **50 (10)**, 8050–8065, doi:10.1002/2013WR015109.

Schertzer, D., and S. Lovejoy, 1987a: Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes. *J. Geophys. Res.*, **92 (D8)**, 9693–9714, doi:10.1029/JD092iD08p09693.

Schertzer, D., and S. Lovejoy, 1987b: Singularités anisotropes, divergences des moments en turbulence: invariance d'échelle généralisée et processus multiplicatifs. *Ann. Sci. Math. Que*, **11 (1)**, 139–181.

Schertzer, D., and S. Lovejoy, 1991: Nonlinear geodynamical variability: multiple singularities, universality and observables. *Non-Linear Variability in Geophysics*, Springer, 41–82.

Schertzer, D., and S. Lovejoy, 1997: Universal multifractals do exist!: Comments on "A statistical analysis of mesoscale rainfall as a random cascade". *JAM*, **36**, 1296–1303, doi:10.1175/1520-0450(1997)036<1296:UMDECO>2.0.CO;2.

Schertzer, D., and S. Lovejoy, 2011: Multifractals, generalized scale invariance and complexity in geophysics. *Int. J. Bifurcation and Chaos*, **21 (12)**, 3417–3456, doi:10.1142/S0218127411030647.

Schertzer, D., I. Tchiguirinskaia, S. Lovejoy, and P. Hubert, 2010: No monsters, no miracles: in nonlinear sciences hydrology is not an outlier! *Hydrol. Sci. J.*, **55 (6)**, 965–979, doi:10.1080/02626667.2010.505173.

Schleiss, M., and A. Berne, 2010: Identification of dry and rainy periods using telecommunication microwave links. *IEEE Geosci. Remote Sens. Lett.*, **7 (3)**, 611–615, doi:10.1109/LGRS.2010.2043052.

Schleiss, M., S. Chamoun, and A. Berne, 2014a: Non-stationarity in intermittent rainfall: the dry drift. *J. Hydrometeor.*, **15 (3)**, 1189–1204, doi:10.1175/JHM-D-13-095.1.

Schleiss, M., S. Chamoun, and A. Berne, 2014b: Stochastic simulation of intermittent rainfall using the concept of dry drift. *Water Resour. Res.*, **50 (3)**, 2329–2349, doi:10.1002/2013WR014641.

Schleiss, M., J. Jaffrain, and A. Berne, 2011: Statistical analysis of rainfall intermittency at small spatial and temporal scales. *Geophys. Res. Lett.*, **38 (18)**, doi:10.1029/2011GL049000.

Schleiss, M., J. Jaffrain, and A. Berne, 2012: Stochastic simulation of intermittent DSD fields in time. *J. Hydrometeor.*, **13 (2)**, 621–637, doi:10.1175/JHM-D-11-018.1.

Schleiss, M., and J. Smith, 2015: A method to estimate the 3D–time structure of the raindrop size distribution using radar and disdrometer data. *Journal of Hydrometeorology*, **16 (3)**, 1222–1242, doi:10.1175/JHM-D-14-0182.1.

Schneebeli, M., N. Dawes, M. Lehning, and A. Berne, 2013: High-resolution vertical profiles of polarimetric X-band weather radar observables during snowfall in the Swiss Alps. *J. Appl. Meteor. Climate*, **52 (2)**, 378–394, doi:10.1175/JAMC-D-12-015.1.

Schneebeli, M., J. Grazioli, and A. Berne, 2014: Improved estimation of the specific differential phase shift using a compilation of Kalman filter ensembles. *IEEE T. Geosci. Remote Sens.*, **52 (8)**, 5137–5149, doi:10.1109/TGRS.2013.2287017.

Schönhuber, M., G. Lammer, and W. Randeu, 2007: One decade of imaging precipitation measurement by 2D-video-distrometer. *Adv. Geosci.*, **10 (10)**, 85–90, doi:10.5194/adgeo-10-85-2007.

Schönhuber, M., G. Lammer, and W. L. Randeu, 2008: The 2D-video-distrometer. *Precipitation: Advances in measurement, estimation and prediction*, Springer, 3–31, doi:10.1007/978-3-540-77655-0_1.

Schuurmans, J. M., M. F. P. Bierkens, E. J. Pebesma, and R. Uijlenhoet, 2007: Automatic prediction of high-resolution daily rainfall fields for multiple extents: The potential of operational radar. *J. Hydrometeor.*, **8 (6)**, 1204–1224, doi:10.1175/2007JHM792.1.

Seifert, A., and D. K. Beheng, 2006: A two-moment cloud microphysics parameterization for mixed-phase clouds. Part 1: Model description. *Meteorol. Atmos. Phys.*, **92 (1)**, 45–66, doi:10.1007/s00703-005-0112-4.

Seifert, A., U. Blahak, K. Stephan, M. Baldauf, and J.-P. Schulz, 2011: Documentation of the changes in the COSMO-model, version 4.21. URL http://www.cosmo-model.org/content/model/releases/histories/cosmo_4.21.htm, accessed 04.08.2015.

Sekhon, R. S., and R. Srivastava, 1971: Doppler radar observations of drop-size distributions in a thunderstorm. *J. Atmos. Sci.*, **28 (6)**, 983–994, doi:10.1175/1520-0469(1971)028<0983:DROODS>2.0.CO;20.

Seliga, T., and V. Bringi, 1976: Potential use of radar differential reflectivity measurements at orthogonal polarizations for measuring precipitation. *J. Appl. Meteor.*, **15 (1)**, 69–76, doi:10.1175/1520-0450(1976)015<0069:PUORDR>2.0.CO;2.

Sempere-Torres, D., J. M. Porrà, and J.-D. Creutin, 1994: A general formulation for raindrop size distribution. *J. Appl. Meteor.*, **33 (12)**, 1494–1502, doi:10.1175/1520-0450(1994)033<1494:AGFFRS>2.0.CO;2.

Sempere-Torres, D., J. M. Porrà, and J.-D. Creutin, 1998: Experimental evidence of a general description for raindrop size distribution properties. *Journal of Geophysical Research: Atmospheres*, **103 (D2)**, 1785–1797, doi:10.1029/97JD02065.

Sempere-Torres, D., R. Sanchez-Diezma, I. Zawadzki, and J. Creutin, 2000: Identification of stratiform and convective areas using radar data with application to the improvement of DSD analysis and ZR relations. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, **25 (10)**, 985–990, doi:10.1016/S1464-1909(00)00138-6.

Seto, S., T. Iguchi, and T. Oki, 2013: The basic performance of a precipitation retrieval algorithm for the Global Precipitation Measurement mission's single/dual-frequency radar measurements. *IEEE T. Geosci. Remote Sens.*, **51 (12)**, 5239–5251, doi:10.1109/TGRS.2012.2231686.

Shook, K., and D. M. Gray, 1996: Small-scale spatial structure of shallow snowcovers. *Hydrological Processes*, **10 (10)**, 1283–1292, doi:10.1002/(SICI)1099-1085(199610)10:10<1283::AID-HYP460>3.0.CO;2-M.

Shook, K., and D. M. Gray, 1997: Synthesizing shallow seasonal snow covers. *Water Resour. Res.*, **33 (3)**, 419–426, doi:10.1029/96WR03532.

Shook, K., and J. Pomeroy, 2010: Hydrological effects of the temporal variability of the multi-scaling of snowfall on the Canadian prairies. *Hydrology and Earth System Sciences*, **14 (7)**, 1195–1203, doi:10.5194/hess-14-1195-2010.

Stocker, T., and Coauthors, 2013: Technical summary. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley, Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Strauch, R. G., 1976: Theory and application of the FW-CW doppler radar. *Ph.D. Thesis Colorado Univ., Boulder*.

Syed, K. H., D. C. Goodrich, D. E. Myers, and S. Sorooshian, 2003: Spatial characteristics of thunderstorm rainfall fields and their relation to runoff. *J. Hydrol.*, **271 (1-4)**, 1–21, doi:10.1016/S0022-1694(02)00311-6.

Tapiador, F., R. Checa, and M. De Castro, 2010: An experiment to measure the spatial variability of rain drop size distribution using sixteen laser disdrometers. *Geophys. Res. Lett.*, **37 (16)**, doi:10.1029/2010GL044120.

Tessier, Y., S. Lovejoy, P. Hubert, D. Schertzer, and S. Pecknold, 1996: Multifractal analysis and modeling of rainfall and river flows and scaling, causal transfer functions. *J. Geophys. Res: Atmos.*, **101 (D21)**, 26 427–26 440, doi:10.1029/96JD01799.

Tessier, Y., S. Lovejoy, and D. Schertzer, 1993: Universal multifractals: theory and observations for rain and clouds. *Journal of Applied Meteorology*, **32 (2)**, 223–250, doi:10.1175/1520-0450(1993)032<0223:UMTAOF>2.0.CO;2.

Testud, J., E. L. Bouar, E. Obligis, and M. Ali-Mehenni, 2000: The rain profiling algorithm applied to polarimetric weather radar. *J. Atmos. Oceanic Technol.*, **17 (3)**, 332–356, doi:10.1175/1520-0426(2000)017<0332:TRPAAT>2.0.CO;2.

Testud, J., S. Oury, R. A. Black, P. Amayenc, and X. Dou, 2001: The concept of "normalized" distribution to describe raindrop spectra: a tool for cloud physics and cloud remote sensing. *J. Appl. Meteor.*, **40 (6)**, 1118–1140, doi:10.1175/1520-0450(2001)040<1118:TCONDT>2.0.CO;2.

Thurai, M., V. Bringi, L. Carey, P. Gatlin, E. Schultz, and W. Petersen, 2012: Estimating the accuracy of polarimetric radar-based retrievals of drop-size distribution parameters and rain

rate: An application of error variance separation using radar-derived spatial correlations. *Journal of Hydrometeorology*, **13 (3)**, 1066–1079, doi:10.1175/JHM-D-11-070.1.

Thurai, M., and V. N. Bringi, 2005: Drop axis ratios from a 2D video disdrometer. *J. Atmos. Oceanic Technol.*, **22 (7)**, 966–978, doi:10.1175/JTECH1767.1.

Thurai, M., G. Huang, V. Bringi, W. Randeu, and M. Schönhuber, 2007: Drop shapes, model comparisons, and calculations of polarimetric radar parameters in rain. *J. Atmos. Oceanic Technol.*, **24 (6)**, 1019–1032, doi:10.1175/JTECH2051.1.

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

Tobin, C., L. Nicotina, M. B. Parlange, A. Berne, and A. Rinaldo, 2011: Improved interpolation of meteorological forcings for hydrologic applications in a Swiss alpine region. *J. Hydrol.*, **401 (1-2)**, 77 – 89, doi:10.1016/j.jhydrol.2011.02.010.

Tokay, A., and P. G. Bashor, 2010: An experimental study of small-scale variability of raindrop size distribution. *J. Appl. Meteor. Climate*, **49 (11)**, 2348–2365, doi:10.1175/2010JAMC2269.1.

Tokay, A., A. Kruger, and W. Krajewski, 2001: Comparison of drop size distribution measurements by impact and optical disdrometers. *J. Appl. Meteor.*, **40 (11)**, 2083–2097, doi:10.1175/1520-0450(2001)040<2083:CODSDM>2.0.CO;2.

Tokay, A., W. A. Petersen, P. Gatlin, and M. Wingo, 2013: Comparison of raindrop size distribution measurements by collocated disdrometers. *J. Atmos. Oceanic Technol.*, **30 (8)**, 1672–1690, doi:10.1175/JTECH-D-12-00163.1.

Tokay, A., D. B. Wolff, and W. A. Petersen, 2014: Evaluation of the new version of the laser-optical disdrometer, OTT Parsivel^2. *J. Atmos. Oceanic Technol.*, **31**, 1276–1288, doi:10.1175/JTECH-D-13-00174.1.

Trenberth, K. E., A. Dai, R. M. Rasmussen, and D. B. Parsons, 2003: The changing character of precipitation. *Bull. Amer. Meteor. Soc.*, **84 (9)**, 1205–1217, doi:10.1175/BAMS-84-9-1205.

Tridon, F., J. Van Baelen, and Y. Pointin, 2011: Aliasing in Micro Rain Radar data due to strong vertical winds. *Geophys. Res. Lett.*, **38**, L02804, doi:10.1029/2010GL046018.

Uijlenhoet, R., 2001: Raindrop size distributions and radar reflectivity-rain rate relationships for radar hydrology. *Hydrology and Earth System Sciences*, **5 (4)**, 615–628, doi:10.5194/hess-5-615-2001.

Uijlenhoet, R., J. M. Porra, D. S. Torres, and J.-D. Creutin, 2006: Analytical solutions to sampling effects in drop size distribution measurements during stationary rainfall: Estimation of bulk rainfall variables. *J. Hydrol.*, **328 (1)**, 65–82, doi:10.1016/j.jhydrol.2005.11.043.

Uijlenhoet, R., and D. Sempere Torres, 2006: Measurement and parameterization of rainfall microstructure. *J. Hydrol.*, **328**, 1–7, doi:10.1016/j.jhydrol.2005.11.038.

Uijlenhoet, R., J. A. Smith, and M. Steiner, 2003a: The microphysical structure of extreme precipitation as inferred from ground-based raindrop spectra. *J. Atmos. Sci.*, **60 (10)**, 1220–1238.

Uijlenhoet, R., M. Steiner, and J. A. Smith, 2003b: Variability of raindrop size distributions in a squall line and implications for radar rainfall estimation. *J. Hydrometeor.*, **44 (4)**, 43–61, doi:10.1175/1525-7541(2003)004<0043:VORSDI>2.0.CO;2.

Ulbrich, C. W., 1983: Natural variations in the analytical form of the raindrop-size distribution. *J. Climate Appl. Meteor.*, **22 (10)**, 1764–1775, doi:10.1175/1520-0450(1983)022<1764:NVITAF>2.0.CO;2.

Ulbrich, C. W., 1985: The effects of drop size distribution truncation on rainfall integral parameters and empirical relations. *J. Climate Appl. Meteor.*, **24 (6)**, 580–590, doi:10.1175/1520-0450(1985)024<0580:TEODSD>2.0.CO;2.

van de Beek, C. Z., H. Leijnse, P. J. J. F. Torfs, and R. Uijlenhoet, 2011: Climatology of daily rainfall semi-variance in the netherlands. *Hydrol. Earth Syst. Sci.*, **15 (1)**, 171–183, doi:10.5194/hess-15-171-2011.

van Dijk, A., L. Bruijnzeel, and C. Rosewell, 2002: Rainfall intensity-kinetic energy relationships: a critical literature appraisal. *J. Hydrol.*, **261 (1-4)**, 1 – 23, doi:10.1016/S0022-1694(02)00020-3.

Van Weverberg, K., E. Goudenhoofdt, U. Blahak, E. Brisson, M. Demuzere, P. Marbaix, and J.-P. van Ypersele, 2014: Comparison of one-moment and two-moment bulk microphysics for high-resolution climate simulations of intense precipitation. *Atmos. Res.*, **147**, 145–161, doi:10.1016/j.atmosres.2014.05.012.

Velasco-Forero, C. A., D. Sempere-Torres, E. F. Cassiraga, and J. J. Gomez-Hernandez, 2008: A non-parametric automatic blending methodology to estimate rainfall fields from rain gauge and radar data. *Adv. Water Resour.*, **32 (7)**, 986–1002, doi:10.1016/j.advwatres.2008.10.004.

Verrier, S., L. Barthès, and C. Mallet, 2013: Theoretical and empirical scale dependency of Z-R relationships: Evidence, impacts, and correction. *J. Geophys. Res: Atmos.*, **118 (14)**, 7435–7449, doi:10.1002/jgrd.50557.

Verrier, S., L. De Montera, L. Barthès, and C. Mallet, 2010: Multifractal analysis of African monsoon rain fields, taking into account the zero rain-rate problem. *J. Hydrol.*, **389 (1)**, 111–120, doi:10.1016/j.jhydrol.2010.05.035.

Vivekanandan, J., G. Zhang, and E. Brandes, 2004: Polarimetric radar estimators based on a constrained gamma drop size distribution model. *J. Appl. Meteor.*, **43 (2)**, 217–230, doi:10.1175/1520-0450(2004)043<0217:PREBOA>2.0.CO;2.

Vulpiani, G., F. S. Marzano, V. Chandrasekar, A. Berne, and R. Uijlenhoet, 2006: Polarimetric weather radar retrieval of raindrop size distribution by means of a regularized artificial neural network. *IEEE T. Geosci. Remote Sens.*, **44 (11)**, 3262–3275, doi:10.1109/TGRS.2006.878438.

Wallace, J. M., and P. V. Hobbs, 2006: *Atmospheric science: an introductory survey*, Vol. 92. 2nd ed., Academic press.

Willis, P. T., 1984: Functional fits to some observed drop size distributions and parameterization of rain. *J. Atmos. Sci.*, **41 (9)**, 1648–1661, doi:10.1175/1520-0469(1984)041<1648:FFTSOD>2.0.CO;2.

Wolfensberger, D., D. Scipion, and A. Berne, 2015: Detection and characterization of the melting layer based on polarimetric radar scans. *Quarterly Journal of the Royal Meteorological Society*, doi:10.1002/qj.2672.

Yoshikawa, E., V. Chandrasekar, and T. Ushio, 2014: Raindrop size distribution (DSD) retrieval for X-band dual-polarization radar. *J. Atmos. Oceanic Technol.*, **31 (2)**, 387–403, doi:10.1175/JTECH-D-12-00248.1.

Yuter, S. E., D. E. Kingsmill, L. B. Nance, and M. Löffler-Mang, 2006: Observations of precipitation size and fall speed characteristics within coexisting rain and wet snow. *J. Appl. Meteor. Climate*, **45 (10)**, 1450–1464, doi:10.1175/JAM2406.1.

Zawadzki, I. I., 1975: On radar-raingage comparison. *J. Appl. Meteor.*, **14 (8)**, 1430–1436, doi:10.1175/1520-0450(1975)014<1430:ORRC>2.0.CO;2.

Zhang, G., J. Vivekanandan, and E. Brandes, 2001: A method for estimating rain rate and drop size distribution from polarimetric radar measurements. *IEEE T. Geosci. Remote Sens.*, **39 (4)**, 830–841, doi:10.1109/36.917906.

Zhang, G., J. Vivekanandan, E. A. Brandes, R. Meneghini, and T. Kozu, 2003: The shape-slope relation in observed gamma raindrop size distributions: Statistical error or useful information? *Journal of Atmospheric and Oceanic Technology*, **20 (8)**, 1106–1119, doi:10.1175/1520-0426(2003)020<1106:TSRIOG>2.0.CO;2.

T. H. Raupach, curriculum vitae

Details

Name Timothy Hugh Raupach
DOB 29.05.1982
Nationality Australian
Address Av. de France 47, 1004 Lausanne, Switzerland

Education

2012–2016 PhD in hydrometeorology, École Polytechnique Fédéral de Lausanne, Switzerland.
2005–2009 MPhil (Computer Science, part time), Australian National University, Australia.
2001–2004 Bachelor of Software Engineering, Australian National University, Australia.

PhD thesis

Title Variability of the raindrop size distribution across scales in Mediterranean rainfall: characterisation and stochastic simulation.
Supervisor Prof. A. Berne.

Experience

2012–2015 **Teaching assistant**, École Polytechnique Fédéral de Lausanne, Switzerland, for M.Sc. course "Spatial statistics and analysis", and undergraduate courses "Analyse I", "Analyse II" and "Remote sensing".
2010–2012 **Project support officer**, CSIRO Land and Water, Canberra, Australia. Scientific programming and spatial data analysis for hydrological remote sensing research.
2005–2010 **Scientific programmer**, National ICT Australia, Canberra, Australia. Programming for computer vision research.
2003–2004 **IT support staff member**, CSIRO Land and Water, Canberra, Australia.
2001–2003 **Programmer** Digital (Digest) Data Design (4D), Canberra, Australia.

Publications

2016 Raupach, T. H. and A. Berne. Small-scale variability of the raindrop size distribution and its effect on areal rainfall retrieval. *J. Hydrometeor.*, 17:2077–2104, July 2016. doi: 10.1175/JHM-D-15-0214.1 .

2016 Raupach, T. H. and A. Berne. Spatial interpolation of experimental raindrop size distribution spectra. *Q. J. Roy. Meteor. Soc.*, May 2016. doi: 10.1002/qj.2801.

2015 Raupach, T. H. and A. Berne. Correction of raindrop size distributions measured by parsivel disdrometers, using a two-dimensional video disdrometer as a reference. *Atmos. Meas. Tech.*, 8(1):343–365, January 2015. doi: 10.5194/amt-8-343-2015.

2014 Huza, J., A. J. Teuling, I. Braud, J. Grazioli, L. A. Melsen, G. Nord, T. H. Raupach, and R. Uijlenhoet. Precipitation, soil moisture and runoff variability in a small river catchment (Ardéche, France) during HyMeX Special Observation Period 1, *J. Hydrol.*, 516, 330–342, August 2014. doi: 10.1016/j.jhydrol.2014.01.041.

2014 Grazioli, J., D. Tuia, S. Monhart, M. Schneebeli, T. H. Raupach, and A. Berne. Hydrometeor classification from two-dimensional video disdrometer data, *Atmos. Meas. Tech.*, 7, 2869–2882.

2013 Brawata, R, T. H. Raupach, and T. Neeman. Techniques for monitoring carnivore behavior using automatic thermal video, *Wildlife Society Bulletin*, 37(4), 862–871, July 2013, doi: 10.1002/wsb.315.

2013 Chappell, A., L. J. Renzullo, T. H. Raupach, and M. Haylock. Evaluating geostatistical methods of blending satellite and gauge data to estimate near real-time daily rainfall for Australia, *Journal of Hydrology*, 493, 105–114, June 2013, doi: 10.1016/j.jhydrol.2013.04.024.

2006 Prakash, S., P.Y. Lee, T. Caelli, and T. Raupach. Robust thermal camera calibration and 3D mapping of object surface temperatures, *Proc. of SPIE*, 6205, 2006.

Awards

2012 Commonwealth Scientific and Industrial Research Organisation (CSIRO) Land and Water Young Scientist Award

Languages

English (native), **French** (advanced).

Other academic activities

2015–present Reviewer for *Journal of Hydrology*, *Journal of Hydrometeorology*, *Hydrology and Earth System Sciences*, and *Journal of Applied Meteorology and Climatology*.

Computing skills

Languages R, C, C++, C#, Java, Python, Perl, PHP, bash, Matlab.

Computing High-performance computing (parallel/distributed processing, GPU programming), algorithm design, web development, Linux, LaTeX.

Other skills

Remote sensing Processing and interpretation of disdrometer, rain gauge, weather station, polarimetric radar and satellite data.

Mathematics Geostatistics, applied statistics, spatial analysis, statistical programming in R, stochastic simulation, time series analysis.

Communication English writing and oral presentation.