

Crystal Clear or Frozen and Fuzzy: Ice Particle Classification Using Spectral Polarimetric Cloud Radar

Sophie Keemink

[This page is intentionally left blank]

Crystal Clear or Frozen and Fuzzy: Ice Particle Classification Using Spectral Polarimetric Cloud Radar

by

Sophie Keemink

in partial fulfilment of the requirement to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended and presented on Wednesday January 29, 2025 at 13:45.

Student number: 4663926
Thesis committee: Ir. Christine Unal, TU Delft, supervisor
Dr. Marc Schleiss, TU Delft
Dr. Isabelle. Steinke, TU Delft

Cover: Photograph of cirrus clouds, taken on 06-02-2024.

An electronic version of this thesis is available at <http://repository.tudelft.nl/>.



Preface

At last, I'm writing a preface. I didn't want to, but got peer-pressured into writing one anyway. After all, if there is one thing I have learned over the past years in Delft it is the importance of peers and, at times, the necessary pressure.

Therefore, I would like to start off by thanking my thesis committee. Christine, thank you for your patience with me and motivating words, and Marc and Isabelle for your valuable feedback. Ultimately, these meetings are what kept me going forward with my thesis.

In the process, my peers are what made me stop going insane. There seemed to be a direct correlation between the atmosphere in 'het hok' and my progress. Thank you for putting up with my endless side-tracking and for the countless coffees.

Finishing this thesis is not just handing in another assignment, but it means finishing my studies here in Delft. Again, it's my peers that made these years worth it. From the people that have been there since the very beginning: my friends at Aerospace Engineering, who showed me that you can be horribly smart, ambitious, and incredibly stupid at the same time; my teammates at Proteus, who showed you can be all of the above and athletic as well. Throughout the years, more people joined this list: my roommates at the AP, who were there for me during lockdown; the people at ELS, who brought in more work-life balance, and an appreciation for insane endurance feats; and the random groups of people who I climbed literal mountains with. While a thousand people walked in and out of my life, there were 4 that have always been there for me. Thank you, mam, pap, Esther, and Eline, for always being there, sometimes as peer and sometimes as pressure.

With that being said, whether you're just here for this preface, the pictures, you're a master student looking for inspiration, or you'll read every word: be like me, and take a look outside and up every now and then.

*Sophie Keemink
Delft, January 2025*

Abstract

Accurate classification of ice particles in clouds is essential for improving the understanding of cloud microphysics and improving weather and climate models. This thesis investigates the use of spectral polarimetry in millimetre-wavelengths, combined with a Discrete Dipole Approximation (DDA) and Gaussian Mixture Model (GMM) scattering database, to classify ice particles through fuzzy logic. Utilizing a dual-wavelength (94 and 35 GHz), dual-polarized cloud radar installed in Cabauw, this study analyses two non-precipitating ice cloud events. Spectral polarimetric variables, including differential reflectivity (Z_{DR}), Slanted Linear Depolarization Ratio (SLDR), backscattering phase (ϕ_{bs}), and Dual Spectral Ratio (DSR), were derived from radar measurements and compared with modelled values from the scattering database. Results indicated that different ice particle types exhibited distinct polarimetric characteristics, but a lot of overlap between particles remained.

A fuzzy logic classifier was developed, incorporating both 1D and 2D membership functions to improve differentiability between particle types. Adding temperature and liquid water path as variables was necessary to distinguish between branched planar, aggregates and graupel particles. The classification results were mostly consistent and as expected, though there was a high dependence on temperature, suggesting areas for further refinement. Through fuzzy logic outputs Q and Q -gap, the most probable type of ice particles is identified and a first assessment on the quality of this identification is given.

This study demonstrates that combining spectral polarimetric variables with an advanced scattering database has potential to improve the classification of ice particles. In particular, the proposed technique could allow the classification of possible different ice particle types for each radar observation volume. The method lays the basis for future developments in cloud microphysics and radar-based ice particle classification.

Contents

Preface	i
Abstract	ii
Nomenclature	vii
1 Introduction	1
1.1 State of the art	2
1.2 Research Aim	3
1.3 Plan of approach	4
2 Background knowledge	5
2.1 Ice crystals	5
2.2 Radar variables	6
2.2.1 Radar basics	6
2.2.2 Spectral polarimetric variables	7
2.3 Fuzzy logic	10
3 Data description	12
3.1 Particle Database	12
3.2 Cloud Radar Data	13
3.3 Radiometer Data	14
4 Method	15
4.1 Pre-processing database data	15
4.1.1 Incident angles	16
4.1.2 Omitting S_{hv}	16
4.2 Design of fuzzy logic classifier	17
4.2.1 1D fuzzy logic set-up	17
4.2.2 2D fuzzy logic set-up	19
4.2.3 Incorporating environmental variables	21
4.3 Radar data preparation	22
4.3.1 Events	22
4.3.2 Retrieval of polarimetric variables	24
5 Results	27
5.1 Comparing cloud radar variables against scattering database variables	27
5.1.1 Spread of the spectral polarimetric variables	27
5.1.2 Spectrogram of polarimetric variables 18:00	29
5.2 Fuzzy logic set-ups tested on database	30
5.2.1 1D vs 2D	30
5.2.2 Single frequency	31
5.3 Classification results on cloud radar data	31
5.3.1 Classification results Event A 18:00	32
5.3.2 Classification results Event A 18:30	32
5.3.3 Classification results on Event B	33
5.4 Sensitivity to different variables	34
5.4.1 Sensitivity to temperature	34
5.4.2 Classification with only Ka-band	35
5.4.3 Sensitivity to phase	35
5.5 Bulk results	36

6	Discussion	39
6.1	How do the data from the scattering database and the cloud radar compare?	39
6.2	Which choices were made when setting up the fuzzy logic classifier?	40
6.3	How is the quality of the classification results assessed?	40
6.4	Importance of different variables	42
6.5	Limitations	42
7	Conclusion	43
7.1	Conclusions on sub-questions	43
7.2	Overall conclusion	44
7.3	Recommendations	44
7.3.1	New events	44
7.3.2	Adjustments to the classifier	44
7.3.3	Analysis of results	45
7.3.4	Comparison with other classifiers	45
A	Unused polarimetric variables	51
B	Principal Component Analysis	53

List of Figures

1.1	Different ice crystal morphologies as a function of temperature and supersaturation, from Libbrecht [27]	1
1.2	Cloud radar reflectivity and hydrometeor classification result on January, 1st 2025 at Cabauw (The Netherlands). The classification algorithm, cloudnetpy, is used in Europe for all the cloud remote sensing national atmospheric sites. The inputs of cloudnetpy are cloud radar, microwave radiometer and lidar data. Note that there is no classification among the different ice particle types.	2
2.1	Columnar ice crystals photomicrographed by Wilson Bentley, used under CC BY 4.0 . .	6
2.2	Examples of conical graupel from the Multi-Angle Snowflake Camera (MASC). Source: Praz et al. [43] Fig. 3, cropped, used under CC BY 3.0	6
2.3	Planar ice crystals photomicrographed by Wilson Bentley, used under CC BY 4.0	6
2.4	Examples of dendrites from the Multi-Angle Snowflake Camera (MASC). Source: Praz et al. [43] Fig. 3, cropped, used under CC BY 3.0	6
2.5	Examples of aggregates from the Multi-Angle Snowflake Camera (MASC). Source: Praz et al. [43] Fig. 3, cropped, used under CC BY 3.0	6
2.6	Trapezoid membership function for one variable and one class. X_{min} and X_{max} are the minimum and maximum occurring values for that variables. $a, b, c,$ and d can be chosen. Often, a and d are the minimum and maximum values of this variable for this class, and b and c the 5 th and 95 th percentile.	10
3.1	Rotation angles of the particles. With the particle in xyz coordinates and $x''y''z''$ the scattering coordinate system, θ is the incident polar angle and ϕ the incident azimuth angle, from [31].	13
4.1	Schematic overview of approach to go from a particle database and cloud radar data to results with fuzzy logic.	15
4.2	Comparison between SLDR with and without taking into account S_{hv} . The relations are mostly linear, suggesting that the simplification of omitting S_{hv} has minimal impact. . .	16
4.3	Comparison between Z_{DR} vs SLDR with and without taking into account S_{hv}	17
4.4	Polarimetric variables for all particles in the database. Particles are coloured by mass, where the darkest shade for a given colour is the maximum mass for that particle type and the lightest colour is the lightest particle for that type.	18
4.5	1D trapezoidal membership functions.	19
4.6	2D membership functions based on the scattering database of Lu et al. [31] and the method of Al-Sakka et al. [1]. Bin size is $5dB$	20
4.7	Membership functions of temperature for aggregates, branched planar, columns and plates. The values correspond to Table 2.1.	21
4.8	Membership function of liquid water path for graupel.	21
4.9	The pre-processing procedure for the spectral cloud radar data to obtain spectral polarimetric variables.	22
4.10	Bulk variables, temperature, and relative humidity for Event A.	23
4.11	Liquid water path Event A	23
4.12	Bulk variables, temperature, and relative humidity for Event B.	24
4.13	Liquid water path for Event B.	24
4.14	Interpolation and bin size dv_d correction, for Event A $t = 18 : 00 : 36, h = 2762m$	25
4.15	SNR of 35GHz horizontal polarisation at event A, 18:00.	26

5.1	The 35GHz SNR for SLDR vs Z_{DR} for 35GHz cloud radar measurements of Event A with the 2D membership function for graupel (black lines).	28
5.2	The scatterplots of all cloud radar measurements of Event A after applying SNR=0[dB] (orange dots). The outlines of the 2D membership functions are given per particle type (black lines)	28
5.3	Phase membership functions (below) and all cloud radar measurements of Event A after applying SNR=0[dB](top).	29
5.4	Spectral polarimetric variables for Event A, 18:00.	29
5.5	Misclassification percentages for the different fuzzy logic algorithm set-ups, vertical is the input and horizontal is the output. A = Aggregates, G = Conical graupel, BP = Branched planar, C = columns, P = plates. E.g. for set-up 1, 22.7% of aggregates get misclassified as plates.	30
5.6	Misclassification percentages for a fuzzy logic algorithm based on the single frequency 94 GHz	31
5.7	Classification results for Event A for two different set-ups, 18:00.	32
5.8	Q at 18:00.	33
5.9	Q-gap at 18:00.	33
5.10	Classification outputs for Event A, 18:30	33
5.11	Classification outputs for Event B, 20:00	34
5.12	Classification results without the inclusion of temperature.	34
5.13	Classification outputs with only the 94GHz band.	35
5.14	Classification results without $\phi_{bs,35}$ and/or $\phi_{bs,94}$	36
5.15	Profiles of classification results of Event A.	37
5.16	Profiles of classification quality metrics of Event A.	37
5.17	Combined average classifications per spectrum for Event B.	37
5.18	Profiles of classification results of Event B. Note the different scales for columns and plates.	38
B.1	Absolute Loading for each variable and component in PCA.	53
B.2	Cumulative variance of each successive component in the PCA.	54

Nomenclature

Abbreviations

Abbreviation	Definition
BSA	Backward Scatter Alignment
DDA	Discrete Dipole Approximation
DSR	Dual Spectral Ratio
DWR	Dual Wavelength Ratio
FSA	Forward Scatter Alignment
GMM	Generalized Mie Method
HATPRO	Humidity And Temperature Profiler
HPBW	Half-Power Beam Width
LDR	Linear Depolarization Ratio
LWP	Liquid Water Path
PCA	Principal Component Analysis
SLDR	Slanted Linear Depolarization Ratio
SNR	Signal-to-Noise Ratio
STSR	Simultaneous Transmission Simultaneous Reception

Symbols

Symbol	Definition	Unit
$C_{hh,vv}$	Cross-spectrum	[mm ²]
D	Diameter	[m]
$ K ^2$	Dielectric factor	[-]
$N(D)$	Number of particles with diameter D	[-]
P	Mapped value in fuzzy logic	[-]
Q	Aggregated value in fuzzy logic	[-]
r	resolution volume	[-]
s	Prefix to indicate a spectral variable	
v_d	Doppler velocity	[m/s]
W	Variable weight	[-]
Z_{DR}	Differential reflectivity	[dB]
z_{hh}/Z_{hh}	Radar reflectivity in horizontal polarization	[mm ⁶ m ⁻³]/[dBZ]
z_{vv}/Z_{vv}	Radar reflectivity in vertical polarization	[mm ⁶ m ⁻³]/[dBZ]
θ	Incident polar angle	[°]
λ	Wavelength	[m]
ρ	Density	[kg/m ³]
σ_{co}	Co-polar scattering cross-section	[mm ²]
σ_{cx}	Cross-polar scattering cross-section	[mm ²]
σ_{hh}	Backscattering cross-section in horizontal polarization	[mm ²]
ϕ_{bs}	Differential backscatter phase	[°]
ϕ_{dp}	Differential propagation phase	[°]
ϕ_{sys}	System differential phase	[°]
Ψ	(Total) differential phase	[°]

Introduction

Clouds, and in particular ice clouds, are currently a major source of uncertainty in climate models. This is due to a lack of knowledge of their exact radiative properties and thus their role in the Earth's radiation budget [51]. This uncertainty is propagated into atmospheric circulation models and regional climate models [16, p. 1022]. The reason for the differences in scattering properties is that the ice particles in ice clouds can have many different shapes and sizes, each with different microphysical properties [4]. For example, Figure 1.1 shows typical examples of some of the different ice particles that can be found in a cloud.

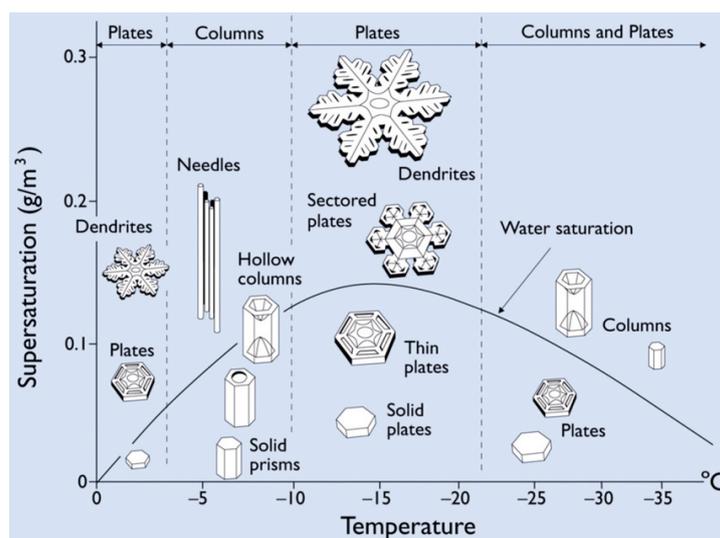


Figure 1.1: Different ice crystal morphologies as a function of temperature and supersaturation, from Libbrecht [27]

Within numerical weather and climate models, cloud microphysics are parameterised [4]. These parameterizations are based on classification profiles such as Figure 1.2 [10]. As can be seen, this classification currently does not distinguish between ice particles, instead grouping all pure ice particles under 'ice'. More detailed classification can decrease the uncertainty in numerical weather prediction [23] and global climate models [4], improve understanding of cloud microphysics in general, and numerous other applications such as severe weather surveillance, flight assistance, and now casting [35]. In short, it is worthwhile to improve the classification of ice particles within clouds.

This chapter starts with summarising the current state of the art in section 1.1. From this, the goal of the research and the corresponding research questions are formulated in section 1.2. Lastly, the structure of the report can be found in section 1.3.

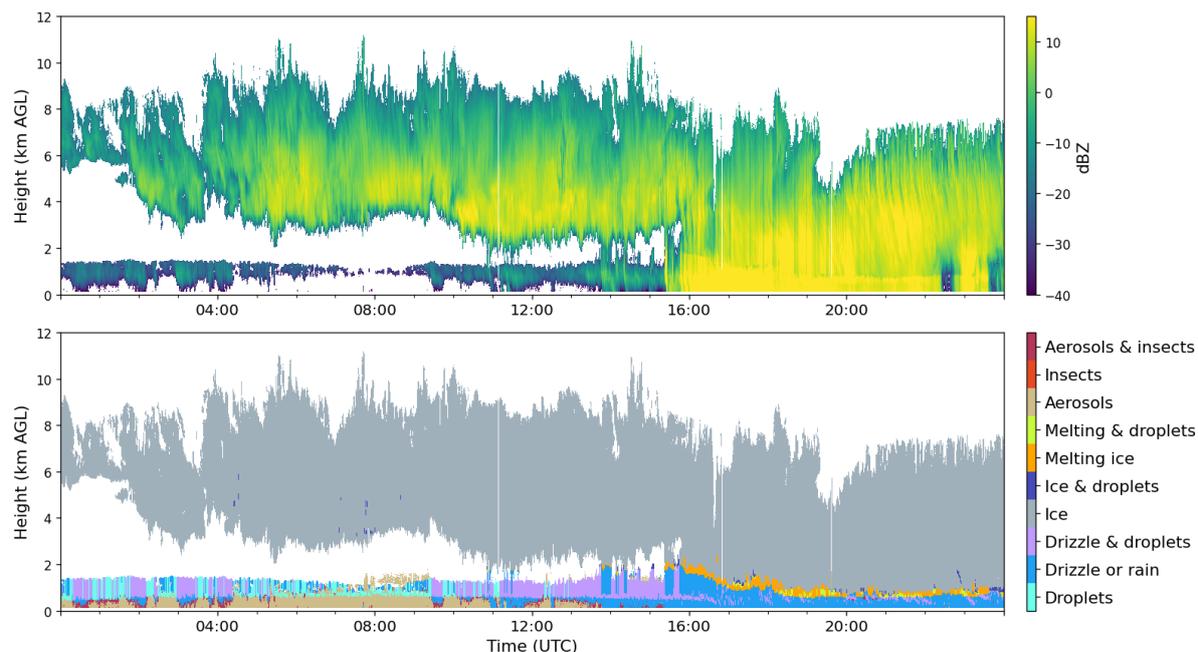


Figure 1.2: Cloud radar reflectivity and hydrometeor classification result on January, 1st 2025 at Cabauw (The Netherlands). The classification algorithm, cloudnetpy, is used in Europe for all the cloud remote sensing national atmospheric sites. The inputs of cloudnetpy are cloud radar, microwave radiometer and lidar data. Note that there is no classification among the different ice particle types.

1.1. State of the art

Different ways to observe ice particles in clouds are, among others, in situ (e.g. with airplanes), from space using satellites, and from the ground. From the ground, observing may be done with e.g. cm-wavelength radar or LiDAR, or cloud radar (mm-wavelength). Classification from the ground originated using S-band (long-range weather radar) to differentiate precipitation types, such as drizzle, rain, hail, and snow [19, 29, 42, 63]. Its success drove the use of the smaller C and X Bands as well [9]. When decreasing wavelength, the sensitivity to smaller particles increases. This allowed the rough distinction between different shapes [54]. With the introduction of cloud radar, sensitivity to smaller particles was increased, and thus more shapes could be discerned, in theory. Ka and W-Band were used in this case [3, 30].

Key in the progress has been the introduction of dual polarisation radar, introducing variables such as the differential reflectivity and linear depolarization ratio. Not every polarimetric variable contains information about each hydrometeor type. However, by combining as many as possible, more and more classes could be discerned [9]. Another way to add information to base classification on, is by combining different wavelengths, especially as the size of the particles and the wavelengths start overlapping. Not only does this double the amount of polarimetric variables, but the dual wavelength ratios can also be used [3, 25, 30, 13].

One characteristic of radar is that the measured reflectivities reflect an average of all particles in the resolution volume. Within these so-called bulk variables, the larger particles in the resolution volume have a big influence. A way to divide the particles within a resolution volume is by looking at the Doppler spectrum. This allows the separation of radar signatures in Doppler velocity bins. With that, so-called spectral polarimetric data can be generated. Dufournet and Russchenberg [13] proved that classification based on spectral polarimetry is possible, and can provide more detailed data to discern microphysical properties.

Regardless of the amount of variables, there is a lot of overlap between classes. This is partly because the ice crystals morph during their lifetime, so there really are no distinct boundaries. Because of this, Liu and Chandrasekar [29] promoted the use of fuzzy logic. This classifier deals with this lack of distinct classes through ‘fuzzification’. It is described in more detail in section 2.3. Its use has been proven

by numerous cases, such as in Thompson et al. [54], Hailong et al. [18], Park et al. [42] and Al-Sakka et al. [1]. For fuzzy logic, all the information that the classification is based on needs to be provided. Through this, an advantage is that it is not a 'black box', and weights and parameters can all be fine-tuned where necessary. Unlike machine-learning based techniques such as decision trees or vector support machines, fuzzy logic does not require any training data. Moreover, fuzzy logic classifiers are effective in situations where feature values are missing or noisy [34, 50].

Initially the properties of different precipitation types were based on physical knowledge of the particles such as shape and size. Romatschke and Vivekanandan [47] used in situ measurements to tune the classifier. Alternatively, classification can also come based on values from scattering models, notably T-matrix models [54, 12, 38, 1]. Because computing the scattering properties of ice crystals is computationally expensive, databases are used. Since it is impossible to encompass all possible wavelengths and ice particles, these databases are always limited. T-matrix, for example, is only applicable to rotationally symmetric particles. On top of that, it shows convergence problems as particles are larger or have high aspect ratios compared to the wavelength [25]. Sometimes authors choose to produce a database for their specific topic, omitting the need to have a general database, like Leinonen and Moisseev [26].

A foreseen limitation of this method is the lack of verification data. This problem has been addressed in several ways before, for example using airborne in-situ measurements as verification [56, 28], using a ceilometer, or comparing with other classification data.[47] Zrnic et al. [65] suggests using self-consistency checks and the fuzzy logic outputs to assess the quality of the results rather than using data from external sources.

1.2. Research Aim

From section 1.1 it is clear that there has been research in using radar to classify ice particles in clouds for decades already. Yet, performance has been limited by the quality and quantity of both radar variables and scattering databases. However, combining dual-polarization, dual wavelength spectral data has, as far as known, not been attempted to build an ice particle classifier. At the end of 2020 a new cloud radar was installed in Cabauw, which provides Ka and W-band dual-polarized spectral data. The quality of the data is improved by increased spatial, time, and Doppler resolution.

On top of that, a relatively new database of the scattering properties of ice particles by Lu et al. [31] became available. This database contains the scattering data of about 1600 different modelled particles. Because it uses the discrete dipole approximation (DDA) or generalised Mie method (GMM) to compute the scattering properties, it is more accurate than the T-matrix methods that are often used. Importantly, the scattering database provides the full scattering matrices of the modelled ice particles, thus making it possible to derive polarimetric variables.

Combining the wealth of data the Cabauw cloud radar provides, and the scattering database could improve the classification of ice particles. To achieve this, several steps need to be taken. Polarimetric variables need to be extracted from both the cloud radar and the scattering database, and then compared. From that, a fuzzy logic classifier needs to be built. Through analysis of the results, the most important variables can be identified to increase understanding. Finally, the quality of the results needs to be assessed.

As such, the main research question is:

How can ice particles be classified with fuzzy logic using the combination of spectral polarimetry in mm-wavelengths and a DDA/GMM scattering database?

To compile a complete answer, the following subquestions are formulated:

- How can cloud radar variables be related to those modelled by a state-of-the-art scattering database?
- How can fuzzy logic be used to build an ice particle classifier?
- What are the most important radar variables for differentiating between different types of particles?
- How can the quality of the classification results be assessed?

Each of these questions describes a step in the complete process and help guide in forming a complete answer to the main research question.

1.3. Plan of approach

In the following chapters, the steps to reach the research questions are explained in detail. First, in chapter 2 the background information that forms the foundation of this thesis is explained. Here, some theory on ice particles can be found, as well as the derivation of all spectral polarimetric variables used for both the cloud radar and the scattering database. Additionally, the basics of fuzzy logic are explained. After that, the specifications of the scattering database and the cloud radar data are given in chapter 3.

In chapter 4 the steps taken to build the classifier from the scattering database are first explained. First, the polarimetric variables are extracted from the database. From the spread of the particles among these variables, the fuzzy logic classifier is made. The description of the events analysed and the cloud radar data preparation can be found here as well. The results are then presented in chapter 5, and interesting observations are noted here. Also in this chapter is the introduction of some metrics to assess the quality of the results. The results are then discussed in chapter 6. The method is evaluated here as well. Finally, the research questions are answered in chapter 7, with recommendations for future work.

2

Background knowledge

In this chapter, some key theoretical concepts necessary to understand the method used is introduced. First, the formation of the different classes of ice particles is explained in section 2.1. Then, spectral polarimetric variables are defined in section 2.2 from the cloud radar measurements and the scattering database outputs. Last, fuzzy logic is explained in section 2.3.

2.1. Ice crystals

In this thesis, ice particles are divided into the five distinct types Lu et al. [31] uses: plates, columns, branched planar, aggregates, and graupel. However, anyone might know that ‘no two snowflakes are alike’. The shapes ice crystals may take on are influenced by temperature and water vapour saturation levels, as well as available ice nucleation particles throughout the entire lifetime of the particle. Since the conditions for two ice particles are rarely exactly the same, neither are their shapes.

Roughly said, there are two main ways an ice particle can form. The first is homogeneous nucleation, which happens at temperatures lower than -36°C . Heterogeneous nucleation requires ice nucleation particles, which are aerosol particles, with surface particles that allow water molecules to form ice structures on them. This means that once water reaches a temperature below 0°C it does not immediately turn into an ice crystal; rather, it needs to come into contact with a particle with the right specific surface properties, or it needs to cool down further until about -36°C . Water droplets that are below 0°C are called supercooled and play an important role in the formation of graupel. [7, 44]

Once an ice particle is formed, it can grow in various ways. The three most important to consider in our case are vapour deposition, riming, and aggregation. Vapour deposition happens when water vapour molecules attach to an existing ice particle. Interestingly, the preferred side on which a water molecule will attach depends on temperature and supersaturation levels: either on the faces (leading to columns, see Figure 2.1) or on the edges or corners (leading to plates, Figure 2.3). Experiments show that as an ice particle moves through different temperatures and humidities, so does the preferred growth direction, leading to almost infinite different shapes. The transition temperatures are near -4 , -9 and -22°C [27]. When supersaturation is high and the temperature levels reflect preferred growth on corners and edges, branched planar crystals (also called dendrites) form. Some examples are shown in Figure 2.4. The ideal temperatures for branched planars to form is between -20 and -12°C [52]

Riming happens when an ice particle comes into contact with a supercooled water droplet, which will freeze onto it upon contact. Once the original shape of the ice crystal is no longer distinguishable, the particle is called graupel. Exact definitions do not exist here. Graupel usually has a density less than $0.8\text{g}/\text{cm}^3$ and can take on a rounded, conical, or more irregular shape. [44] Some typical examples of graupel can be seen in Figure 2.2. For graupel, sufficient supercooled liquid water needs to be present. Usually a liquid water path of $100\text{g}/\text{m}^2$ is maintained as a requirement for graupel to form, though Fitch and Garrett [15] have also shown graupel at concentrations as low as $50\text{g}/\text{m}^2$.

Aggregates form through the collision of multiple ice crystals, as seen in Figure 2.5. As such, they do



Figure 2.1: Columnar ice crystals photomicrographed by Wilson Bentley, used under CC BY 4.0

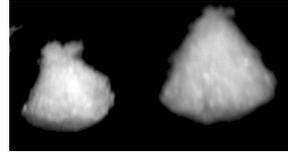


Figure 2.2: Examples of conical graupel from the Multi-Angle Snowflake Camera (MASC). Source: Praz et al. [43] Fig. 3, cropped, used under CC BY 3.0

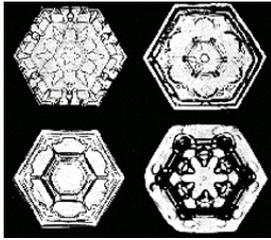


Figure 2.3: Planar ice crystals photomicrographed by Wilson Bentley, used under CC BY 4.0

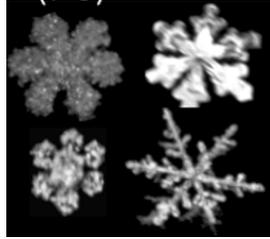


Figure 2.4: Examples of dendrites from the Multi-Angle Snowflake Camera (MASC). Source: Praz et al. [43] Fig. 3, cropped, used under CC BY 3.0

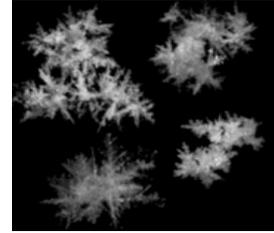


Figure 2.5: Examples of aggregates from the Multi-Angle Snowflake Camera (MASC). Source: Praz et al. [43] Fig. 3, cropped, used under CC BY 3.0

not have a distinct shape and can grow to a couple of millimetres. The chance of aggregates forming increases with higher temperatures, as crystals are more likely to stick to each other. [44]

The temperature ranges where the above-mentioned particles are most likely to grow are summarised in Table 2.1.

Type	Temperature [$^{\circ}C$]
Plates	<0
Columns	Growing from -32 to -22 and -10 to -3
Branched Planars	Growing from -40 to 0, most likely between -20 to -10
Aggregates	Growing from -20 and 6, most likely between -10 and 5
Graupel	Any

Table 2.1: Growth regions for different ice particles [12, 44, 62, 2, 27]

2.2. Radar variables

The radar measurements are mostly influenced by the size distribution of the particles in the resolution volume and therefore do not reflect the properties of the particles themselves. One way to reduce the influence of the particle size distribution is by looking at the Doppler spectrum. On top of that, by looking at ratios, the influence of the number concentration is dropped. In this section, first some radar basics are given and the Doppler spectrum is explained, and then the spectral polarimetric variables are derived for both the database and the cloud radar measurements.

2.2.1. Radar basics

Polarimetric radar works by transmitting electromagnetic waves in both horizontal and vertical polarization. As the wave encounters a particle, part of the wave is scattered back. The radar then receives the horizontal and vertical components of the backscattered wave. For each particle, the following equation holds:

$$\begin{bmatrix} E_h^s \\ E_v^s \end{bmatrix} = \begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix} \begin{bmatrix} E_h^i \\ E_v^i \end{bmatrix}, \quad (2.1)$$

where E is the electric field with i for incident and s for scattered. v and h denote the vertical and horizontal polarizations, and S is the scattering matrix.

The radar reflectivity measured with horizontal polarization in transmission and reception, z_{hh} , is proportional to the sum of all backscattering radar cross sections $\sigma_{hh} = 4\pi|S_{hh}|^2$. Here, $|S|^2 = S \cdot S^*$, with S^* the complex conjugate of S . z_{hh} is then an integral of the diameter:

$$z_{hh} = \frac{\lambda^4}{\pi^5|K|^2} \int N(D)\sigma_{hh}(D)dD = \frac{4\lambda^4}{\pi^4|K|^2} \int N(D)|S_{hh}|^2 dD \quad [mm^6 m^{-3}] \quad (2.2)$$

where $|K|^2$ is the dielectric factor and $N(D)$ the number of particles with max diameter D . The dielectric factor is a radar property, and for the two frequencies it is $|K_{35}|^2 = 0.90$ and $|K_{94}|^2 = 0.74$. Often, reflectivities are given in decibels, which can be distinguished by using capital letters:

$$Z_{hh} = 10 \log_{10}(z_{hh}) \quad [dBZ] \quad (2.3)$$

Equation 2.2 only takes single scattering into account. However, the electromagnetic wave may scatter multiple times before reaching the receiver again, see [64, p. 96]. This 'multiple scattering' effect increases for smaller wavelengths, but after correcting for attenuation effects, it is negligible when using a small beam [5], except in unusual conditions like strong hail storms [66].

Doppler spectrum

z_{hh} thus contains the sum of all particles within the resolution volume. These particles generally have different velocities, e.g. due to mass and size [21]. A Doppler radar can measure this difference in velocity by measuring the Doppler shift, which describes the phase shift a wave goes through as it encounters a moving object. Each particle within the radar volume has its particular backscattering properties as well as a Doppler shift. A Doppler radar divides the backscattered signal among Doppler velocity bins, allowing the calculation of polarimetric variables per Doppler bin. In general, heavier particles will have a larger fall velocity [37, 20]. Thus, the size and density of particles will spread them along the Doppler spectrum, facilitating the separation of particles within a resolution volume. Integrating over the Doppler spectrum will give the total reflectivity of the resolution volume:

$$z_{hh} = \int sz_{hh}(v_d)dv_d = \frac{\lambda^4}{\pi^5|K|^2} \int N(D(v_d))\sigma_{hh}(D(v_d))\left|\frac{dD}{dv_d}\right|dv_d. \quad (2.4)$$

Then,

$$sz_{hh}(v_d)dv_d = \frac{\lambda^4}{\pi^5|K|^2} N(D(v_d))\sigma_{hh}(D(v_d))\left|\frac{dD}{dv_d}\right|dv_d \quad (2.5)$$

The s indicates that sz_{hh} is spectral measurement.

For the measured cross-spectrum, $sC_{hh,vv}dv_d$, is

$$sC_{hh,vv}dv_d = \frac{4\lambda^4}{\pi^4|K|^2} N(D(v_d))S_{hh}^*(D(v_d))S_{vv}(D(v_d))\left|\frac{dD}{dv_d}\right|dv_d. \quad (2.6)$$

2.2.2. Spectral polarimetric variables

In this subsection, the used spectral polarimetric variables are derived or defined. Because the scattering database provides scattering matrices while the cloud radar measurements provide spectral reflectivities, each variable is defined for both cases. Some other polarimetric variables have been considered but not used. They can be found in Appendix A.

Differential Reflectivity

The differential reflectivity in decibels (Z_{DR}) is defined as [8, p. 60]:

$$Z_{DR} = 10 \log_{10}\left(\frac{\sigma_{hh}}{\sigma_{vv}}\right), \quad [dB] \quad (2.7)$$

for the scattering database. This relates to the Z_{DR} for the spectral radar measurement:

$$Z_{DR} = 10 \log_{10} \left(\frac{sz_{hh} dv_d}{sz_{vv} dv_d} \right). \quad [dB] \quad (2.8)$$

Note that in Equation 2.7, $\frac{\lambda^4}{\pi^5 |K|^2} N \left| \frac{dD}{dv_d} \right| dv_d$ drops from Equation 2.5 because it is the same in the numerator and denominator. When an oblate particle is more horizontally aligned, $Z_{DR} > 0$, and vice versa. With the particles above, that mostly means that graupel will produce negative Z_{DR} .

Slanted Linear Depolarization Ratio

The slanted linear depolarization ratio (SLDR) is calculated with the co-polar scattering cross-section (σ_{co}) and the cross-polar scattering cross-section (σ_{cx}). These cross sections are obtained in the case of the polarization basis ($45^\circ, 135^\circ$), as opposed to LDR (linear depolarization ratio) which is acquired using the polarization basis ($0^\circ, 90^\circ$) [8, p. 482–483].

$$\sigma_{cx} = \pi |S_{vv} - S_{hh}|^2 \quad [mm^2] \quad (2.9)$$

$$\sigma_{co} = \pi |S_{hh} + 2 \cdot S_{hv} + S_{vv}|^2 \quad [mm^2] \quad (2.10)$$

$$SLDR = 10 \log_{10} \left(\frac{\sigma_{cx}}{\sigma_{co}} \right) \quad [dB] \quad (2.11)$$

For the scattering database, the above equation can be used. However, as the cloud radar does not provide S_{hv} , the SLDR cannot be used in this way to compare database and radar results. As justified in subsection 4.1.2, because of this S_{hv} will be set to 0 in the database calculations. Then the equation can be rewritten to radar variables¹:

$$\begin{aligned} \frac{\sigma_{cx}}{\sigma_{co}} &= \frac{|S_{hh} - S_{vv}|^2}{|S_{hh} + S_{vv}|^2} \\ &= \frac{(S_{hh} - S_{vv})(S_{hh} - S_{vv})^*}{(S_{hh} + S_{vv})(S_{hh} + S_{vv})^*} \\ &= \frac{(S_{hh} - S_{vv})(S_{hh}^* - S_{vv}^*)}{(S_{hh} + S_{vv})(S_{hh}^* + S_{vv}^*)} \\ &= \frac{S_{hh}S_{hh}^* + S_{vv}S_{vv}^* - S_{hh}S_{vv}^* - S_{hh}^*S_{vv}}{S_{hh}S_{hh}^* + S_{vv}S_{vv}^* + S_{hh}S_{vv}^* + S_{hh}^*S_{vv}} \\ &= \frac{|S_{hh}|^2 + |S_{vv}|^2 - S_{hh}S_{vv}^* - (S_{hh}S_{vv}^*)^*}{|S_{hh}|^2 + |S_{vv}|^2 + S_{hh}S_{vv}^* + (S_{hh}S_{vv}^*)^*} \\ &= \frac{|S_{hh}|^2 + |S_{vv}|^2 - \text{Re}(S_{hh}S_{vv}^*) - \text{Im}(S_{hh}S_{vv}^*)j - \text{Re}(S_{hh}S_{vv}^*) + \text{Im}(S_{hh}S_{vv}^*)j}{|S_{hh}|^2 + |S_{vv}|^2 + \text{Re}(S_{hh}S_{vv}^*) + \text{Im}(S_{hh}S_{vv}^*)j + \text{Re}(S_{hh}S_{vv}^*) - \text{Im}(S_{hh}S_{vv}^*)j} \\ &= \frac{|S_{hh}|^2 + |S_{vv}|^2 - 2\text{Re}(S_{hh}S_{vv}^*)}{|S_{hh}|^2 + |S_{vv}|^2 + 2\text{Re}(S_{hh}S_{vv}^*)} \\ &= \frac{sz_{hh} + sz_{vv} - 2\text{Re}(sz_{hh,vv})}{sz_{hh} + sz_{vv} + 2\text{Re}(sz_{hh,vv})} \end{aligned} \quad (2.12)$$

Similarly to the Z_{DR} , in the last step π , K , λ , N and dv_d drop from the equation. Matrosov et al. [38] showed that the SLDR contains information on ice particles, particularly the structure.

Dual Spectral Ratio

The Dual Wavelength Ratio (DWR) is proportional to the ratio of the copolar reflectivity (hh or vv) at two different wavelengths. For example, for horizontal polarization and the full radar resolution volume:

$$DWR_{hh} = 10 \log_{10} \left(\frac{z_{hh,35}}{z_{hh,94}} \right) \quad [dB] \quad (2.13)$$

¹ $(S_{hh}S_{vv}^*) = \text{Re}(S_{hh}S_{vv}^*) + \text{Im}(S_{hh}S_{vv}^*)j$, so $(S_{hh}S_{vv}^*)^* = \text{Re}(S_{hh}S_{vv}^*) - \text{Im}(S_{hh}S_{vv}^*)j$

Inserting Equation 2.5 for the spectral variable leads to

$$DWR_{hh} = 10 \log_{10} \left(\frac{sz_{hh,35} dv_{35}}{sz_{hh,94} dv_{94}} \right) \quad (2.14)$$

$$= 10 \log_{10} \left(\frac{\frac{\lambda_{35}^4}{\pi^5 |K_{35}|^2} N(D(v_{35})) \sigma_{hh,35}(D(v_{35})) \left| \frac{dD}{dv_{35}} \right| dv_{35}}{\frac{\lambda_{94}^4}{\pi^5 |K_{94}|^2} N(D(v_{94})) \sigma_{hh,94}(D(v_{94})) \left| \frac{dD}{dv_{94}} \right| dv_{94}} \right) \quad (2.15)$$

$$= 10 \log_{10} \left(\frac{\lambda_{35}^4 |K_{94}|^2 \sigma_{hh,35}}{\lambda_{94}^4 |K_{35}|^2 \sigma_{hh,94}} \right) \quad (2.16)$$

The simplification occurring from Equation 2.15 to Equation 2.16 holds when the same Doppler velocity resolution is used for the frequencies 35 and 94 GHz, namely $dv_{35} = dv_{94}$.

Like Montopoli et al. [39], the Dual Spectral Ratio (DSR) can be defined, which is slightly different from the DWR:

$$DSR_{hh} = 10 \log_{10} \left(\frac{\lambda_{35}^4 \sigma_{hh,35}}{\lambda_{94}^4 \sigma_{hh,94}} \right) \quad (\text{database}) \quad (2.17)$$

$$= 10 \log_{10} \left(\frac{|K_{35}|^2 sz_{hh,35}}{|K_{94}|^2 sz_{hh,94}} \right) \quad (\text{radar data}) \quad (2.18)$$

To use these simplified expressions, the difference in Doppler velocity resolution dv between the two wavelengths is addressed in subsection 4.3.2.

In the case of the DSR in vertical polarization, hh is substituted for vv . The DSR relates to the size of the particles, especially as they are in the range of the wavelengths used [53, 17].

Differential Phase

The differential phase is the difference between the phase of the horizontal and vertical signals. Each particle has a differential backscatter phase ϕ_{bs} , which is defined as:

$$\phi_{bs} = \arctan \left(\frac{\text{Im}(S_{hh} S_{vv}^*)}{\text{Re}(S_{hh} S_{vv}^*)} \right) \quad [^\circ] \quad (2.19)$$

From the cloud radar spectral data, the differential phase is retrieved as:

$$\Psi = \arctan \left(\frac{-\text{Im}(sC_{hh,vv} dv_d)}{\text{Re}(sC_{hh,vv} dv_d)} \right) \quad [^\circ] \quad (2.20)$$

However, next to the differential backscatter phase, this data contains two additional differential phases: the differential propagation phase and the differential system phase, such that

$$\Psi = \phi_{bs} + \phi_{dp} + \phi_{sys} \quad [^\circ] \quad (2.21)$$

The differential propagation phase ϕ_{dp} consists of the cumulative phase difference from the scatterers along the propagation path. ϕ_{sys} is a radar hardware property and can be dealt with through calibration.

The differential phase will increase for non-spherical particles [24, p. 36-37], as well as for larger particles [55].

2.3. Fuzzy logic

To go from the scattering database to classified particles, a classifier needs to be build. As mentioned in chapter 1, a fuzzy logic classifier was chosen. In this section, the basics of the fuzzy logic as used in this thesis are explained, such that it can be applied in chapter 4.

A fuzzy logic algorithm consists of three parts: fuzzification, aggregation, and defuzzification.

Fuzzification of input variables happens by mapping them to a membership function. These functions have values ranging from 0 (not a member) to 1 (member). For each class, membership functions have to be constructed for each input variable, such as the trapezoidal membership function in Figure 2.6. Here, X_{min} and X_{max} are the minimum and maximum values for this variable. Per class, a, b, c and d can be determined. A sensible example is setting a as the minimum value for that class, b the 5th, c the 95th percentile, and d the maximum value.

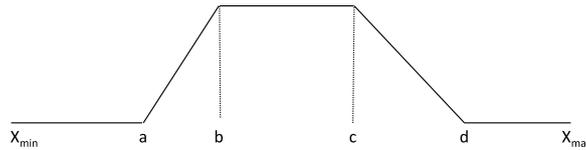


Figure 2.6: Trapezoid membership function for one variable and one class. X_{min} and X_{max} are the minimum and maximum occurring values for that variables. $a, b, c,$ and d can be chosen. Often, a and d are the minimum and maximum values of this variable for this class, and b and c the 5th and 95th percentile.

By 'mapping' the membership function to the measurement value is meant: taking the value of the measurement and finding the corresponding value between 0 and 1 per measurement and per class. The mapped value is P_i^j , with i the feature input and j the class.

The aggregation happens by assigning a weight to each variable, W_i . Then, for the number of classes n , the aggregated value Q^j can be calculated:

$$Q^j = \sum_{i=0}^n W_i P_i^j \quad (2.22)$$

Defuzzification happens by taking the class with the highest value for Q as output. As Zrnec et al. [65] points out, Q contains information about the confidence of the classification. The classification is more likely to be correct when Q is higher and when the gap to the second-highest value is large.

Due to the nature of the way the classifier is set up, the information in the correlation between two variables is lost. To include this information, 2D membership functions can be used [1, 2, 8, p. 522]. Al-Sakka et al. [1] has defined a simple way to set up a 2D membership function between variable A and B :

1. Divide the A range in bins.
2. For each range bin, determine the highest and lowest value of B for each class.
3. For a data point p , the bin corresponding to the value of A_p gets assigned to its corresponding bin.
4. If then the measurement of this data point B_p falls in between the minimum and maximum for that bin, a 1 is returned. Else, a 0 is returned.

Apart from the weights in Equation 2.22 there is another way to combine variables and include importance. When including a variable instead of adding it in the summation, the value can be multiplied. Take variable x with mapped values P_x , P_x can be included within the summation, using

$$Q^j = \left(\sum_{i=0}^n W_i P_i^j \right) \cdot P_x^j \quad (2.23)$$

Since P_x^j can take up a value between 0 and 1, adding variable x like this makes sense if it concerns a condition that makes it (im)possible for an outcome to happen. To illustrate, Al-Sakka et al. [1] used temperature as a multiplicative term, e.g. the variable 'temperature' for the class 'ice' $P_T^{ice} = 0$ if $T > 5^\circ C$. Using Equation 2.23, the aggregated value for ice Q^{ice} is 0 for temperatures too high.

3

Data description

This chapter contains a short description of the data used. First, the particle database is introduced in section 3.1. Then, the radar data is described in section 3.2. Temperature and other profiles from a radiometer are also used, and described in section 3.3.

3.1. Particle Database

As mentioned in chapter 1, the database used in this thesis is the one of Lu et al. [31]. The particles represented are divided in 5 types: plates, columns, branched planar crystals, aggregates, and conical graupel. An overview is given in Table 3.1. For the plates and columns the shapes are relatively straight-forward, which is why their numbers are limited. The branched planar crystals and aggregates have many different shapes they can take on, as explained in section 2.1. Notably, the models of aggregates also differ in the way the crystals are attached together. The graupel particles span across densities of 0.05 to 0.9[g/cm³] and cone angles. The exact formation and definition of the different crystals can be found in the corresponding paper of the database. [31]

Type	Number	Dimension range [mm]	Thickness ratios
Plates	44	0.1–2.52	0.5, 1.0, 2.0
Columns	30	0.18–4.31	1.0, 2.0
Branched Planar Crystals	405	0.5–5.63	0.5, 1.0
Aggregates	660	0.38–62.58	NA
Conical Graupel	640	0.2–2.5	NA

Table 3.1: Overview of particle number and dimensions in the database by Lu et al. [31].

The database contains smaller files with the most important information per particle. For each particle, the scattering matrix

$$S = \begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix} \quad (3.1)$$

is given. Then, for example S_{hh} , is given in $\text{Re}(S_{hh}^f)$, $\text{Im}(S_{hh}^f)$, $\text{Re}(S_{hh}^b)$ and $\text{Im}(S_{hh}^b)$, with f for forward and b for backward. This scattering matrix is given in the FSA (Forward Scatter Alignment) convention. Since most polarimetric variables are used as defined in Bringi and Chandrasekar [8], which uses the BSA (Backward Scatter Alignment) convention, the scattering matrix can be transformed to BSA as follows:

$$S_{BSA} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} S_{FSA} = \begin{bmatrix} -S_{hh} & -S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix}_{FSA} \quad (3.2)$$

For each particle in the database, scattering matrices are given depending on wavelength, incident polar angle, and incident azimuth angle. The wavelengths are 3.19, 8.4, 22.4, and 31.86 mm, corresponding

to 94, 35, 13 and 9 GHz, respectively. The incident angles are given in increments of 10° starting at 0° . In Figure 3.1 note how $\theta = 0$ corresponds to a 90° elevation angle.

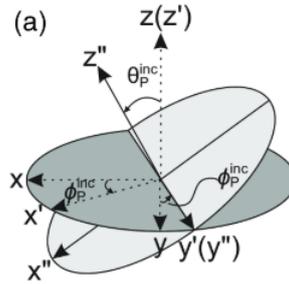


Figure 3.1: Rotation angles of the particles. With the particle in xyz coordinates and $x''y''z''$ the scattering coordinate system, θ is the incident polar angle and ϕ the incident azimuth angle, from [31].

There are some advantages to using this database specifically. First, the database consists of several types of particles, and has different thickness ratios and maximum dimensions for each type, leading to quite an extensive amount of modelled particles. Second, the database contains the scattering properties depending on incident angle, as opposed to averaging over orientations or choosing a random orientation. This allows for the computation of polarimetric variables [57]. Furthermore, the scattering data is computed using either DDA or GMM. These methods are more accurate at mm-wavelength for non-homogeneous and complex shaped particles than the often used T-matrix method [25]. Moreover, the database contains the full scattering matrix for each particle, as opposed to just providing backscattering radar cross sections. This too is necessary to compute some of the polarimetric variables. Finally, the database is described in detail such that limitations and assumptions are clear, and it is freely available online.

Of course, there are some limitations as well, and Lu et al. identifies three. First, the database can never encompass all possible particle types. Notably, it also does not include rimed or melting particles. Second, because the calculations are done with incident polar angle increments of 10° , interpolation will be necessary when taking other incident angles. Last, the absorption cross sections are based on older dielectric constants [22]. However, this mostly poses a problem when working with passive observation [14].

3.2. Cloud Radar Data

The radar used is a dual-frequency polarimetric scanning cloud radar, specifically the RPG-FMCW-DP-KW type, located in Cabauw. It is manufactured by Radiometer Physics GmbH. The radar operates in Simultaneous Transmission Simultaneous Reception (STSR) mode at 94 and 35 GHz. The beam widths are $0.85^\circ \pm 0.05^\circ$ and $0.56^\circ \pm 0.03^\circ$ (HPBW), respectively. Detailed specifications can be found in Myagkov and Rose [40]. During the chosen events, the radar was operating in slanted position, with an elevation angle of 45° . The data is divided in 3 chirps, with slightly different configurations, see Table 3.2.

	Chirp 1	Chirp 2	Chirp 3
Range FFTs	256	256	128
Height [m]	84–843	864–3458	3503–10587
Range resolution [m]	29.8	29.8	55.0
35 GHz Doppler velocity resolution [m/s]	0.1548	0.1261	0.1678
94 GHz Doppler velocity resolution [m/s]	0.0576	0.0470	0.0625

Table 3.2: Chirp table for the cloud radar in Cabauw.

The radar provides data in two levels: LV0 and LV1. Since LV0 contains the spectral data, this one is focused on. The LV0 contain the Doppler spectrum at horizontal and vertical polarization, and the

real and imaginary part of the cross-spectrum.

$$\begin{aligned} \text{CHSpec} &= sz_{hh}(r, v_d)dv_d \\ \text{CVSpec} &= sz_{vv}(r, v_d)dv_d \\ \text{CReVHSpec} &= \text{Re}(sC_{hh,vv}(r, v_d)dv_d) \\ \text{CImVHSpec} &= \text{Im}(sC_{hh,vv}(r, v_d)dv_d) \end{aligned}$$

Here, r is the resolution volume, and v_d the Doppler velocity.

The LV0 data is internally calibrated every 1000 seconds. Additional calibration is applied as described in subsection 4.3.2.

3.3. Radiometer Data

Next to the cloud radar is a microwave radiometer, a HATPRO (Humidity And Temperature Profiler) [48]. It operates at two frequency bands, 22-31 GHz and 51-58 GHz and provides humidity and temperature profiles, as well as integrated water path, liquid water path. The temperature and humidity profiles are used to calculate the attenuation affecting the reflectivity values, and the temperature and liquid water path are used as variables in the classifier. The profiles are given in 93 altitude layers, from 0 to 10000 m, with the height resolution increasing with altitude, from 10 to 200m.

4

Method

This chapter describes which steps are taken from the data as described in chapter 3, to the results in chapter 5. The method is built around three main components: the scattering database, the cloud radar data, and the fuzzy logic classifier tying them together. An overview of the necessary steps is given in Figure 4.1.

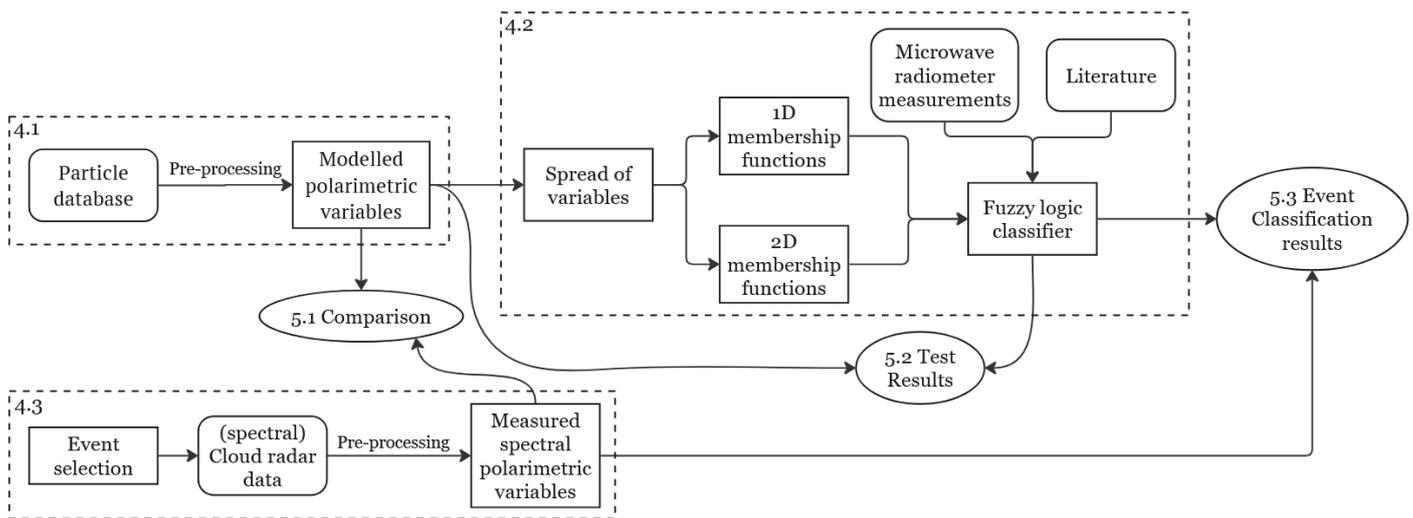


Figure 4.1: Schematic overview of approach to go from a particle database and cloud radar data to results with fuzzy logic.

As can be seen, both the particle database data and the cloud radar measurements need to be pre-processed before obtaining the polarimetric variables. The respective steps taken are explained in section 4.1 and section 4.3. From the variables of the scattering database, the fuzzy logic classifier is made. The approach taken here is explained in section 4.2. Because there is no validation data, extra care is taken to introduce intermediate results to check the method. The results are found in chapter 5.

4.1. Pre-processing database data

In order to use the database data, two assumptions are made. First, that the elevation angle of 45° can be achieved by averaging over 40° and 50° . Second, that the S_{hv} part of the scattering matrix can be set to 0. These assumptions are justified in this section.

4.1.1. Incident angles

When retrieving the scattering properties of a particle from the scattering database, the incident azimuth and incident polar angles need to be specified. The values from the scattering matrix are taken as the average over all incident azimuth angles (see Figure 3.1). This does not give notable changes compared to using just one incident azimuth angle. As for the incident polar angle, the average of the values for 40° and 50° is taken. To check if this is a valid approach, the backward scattering magnitudes have been plotted against the incident polar angle. The curve was near-linear around $40\text{-}50^\circ$ for all particles, which confirms that interpolation is a valid method.

Going forward, unless specified otherwise, scattering properties from the database contain the imaginary and the real part ($S_{hh} = S_{hh_real} + S_{hh_imag} * j$), averaged over all incident azimuth angles and averaged between 40° and 50° incident polar angles.

4.1.2. Omitting S_{hv}

With the LV0 variables, there is no access to S_{hv} itself. Therefore, the SLDR cannot be computed using σ_{co} as defined in Equation 2.10. However, as long as the SLDR is computed the same from the database as from the radar measurements, and SLDR still contains the same amount of information on the particles, S_{hv} could be omitted. For that purpose, the magnitudes of σ_{hv} compared to σ_{hh} and σ_{vv} from the database are shown in Table 4.1. The values for σ_{hv} are significantly lower than σ_{hh} and σ_{vv} , suggesting

	mean $\sigma_{hv}^b [mm^2]$	mean $\sigma_{hh}^b [mm^2]$	mean $\sigma_{vv}^b [mm^2]$	$\sigma_{hv}^b / \sigma_{hh}^b$	$\sigma_{hv}^b / \sigma_{vv}^b$
aggregates	0.0009	0.139	0.06	0.006	0.015
conical graupel	7e-13	3.16	6.27	2e-13	1e-13
dendrites	1e-08	0.012	0.008	1e-06	2e-06
plates	0.0002	0.02	0.005	0.009	0.04
columns	3e-09	0.01	0.007	4e-07	5e-07

Table 4.1: Comparison of magnitudes of the backscattering cross-sections of σ_{hh} and σ_{vv} with σ_{hv} for different particle types from the database.

it might not have a big influence on the overall scattering properties of the particles in the scattering database. As can be seen in the last two columns of the table, the value for σ_{hv} is largest compared to σ_{hh} and σ_{vv} in plates and aggregates. Therefore, leaving out S_{hv} will probably affect those particle types most.

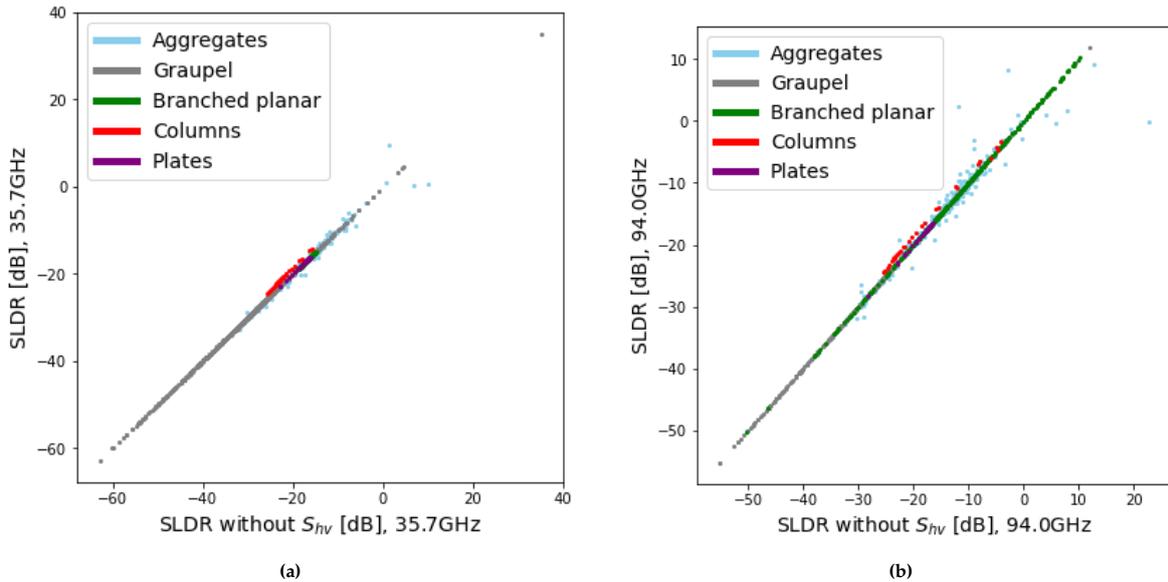


Figure 4.2: Comparison between SLDR with and without taking into account S_{hv} . The relations are mostly linear, suggesting that the simplification of omitting S_{hv} has minimal impact.

When comparing the SLDR with and without S_{hv} in Figure 4.2, it can be seen that overall, leaving out S_{hv} has little to no influence on the SLDR values. Again, an influence does not matter, as long as it is calculated the same for the cloud radar data and the scattering database data, and it can be used to differentiate between particle types. To that end, Figure 4.3 is plotted. Indeed, comparing the scatterplots of Z_{DR} against SLDR with or without S_{hv} (in Figure 4.3) shows little difference except for some aggregates and the columns as a whole. In fact, the difference between columns and plates becomes a bit bigger, making it easier to differentiate between these two types.

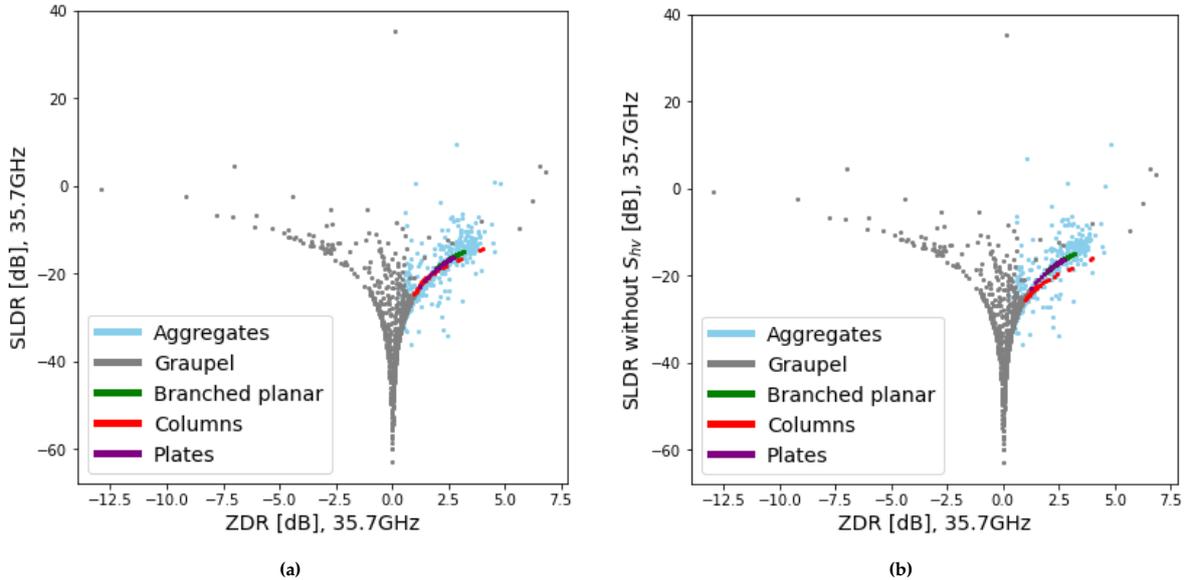


Figure 4.3: Comparison between Z_{DR} vs SLDR with and without taking into account S_{hv} .

So, when working with the database, $S_{hv} = 0$ in order to compare scattering model results and radar measurements.

With the above modifications, the polarimetric variables are calculated as they are defined for the scattering database in subsection 2.2.2.

4.2. Design of fuzzy logic classifier

With the polarimetric variables for every particle in the scattering database, the difference between ice particle types can be investigated. These differences are then used to build a fuzzy logic classifier. To illustrate the spread of the particles across the different variables, they are mapped against each other in a 2D space in Figure 4.4.

From these plots it can be seen that, as expected, the different types have a lot of overlap. However, each type also shows distinct features, like graupel being the only type to have negative Z_{DR} in Ka-Band, and branched planar having ϕ at 0 for Ka-Band but not for W-Band. The difference between plates and columns is subtle, but their features are distinct nonetheless. With the spread per type per variable known, the fuzzy logic classifier can be made. First, a 1D version is made. Then, the option of using 2D fuzzy logic is investigated and applied in subsection 4.2.2. At the end, three different set-ups with the membership functions are defined. After, the way environmental variables are included is explained in subsection 4.2.3.

4.2.1. 1D fuzzy logic set-up

The first step in making a fuzzy logic classifier is constructing the membership functions from the polarimetric variables. As explained in section 2.3, trapezoidal membership functions are made from the extent of the values the variables take on per type. When taking the lowest and highest value a variable can take on for one type and setting the membership function to 1 between these two values, this first

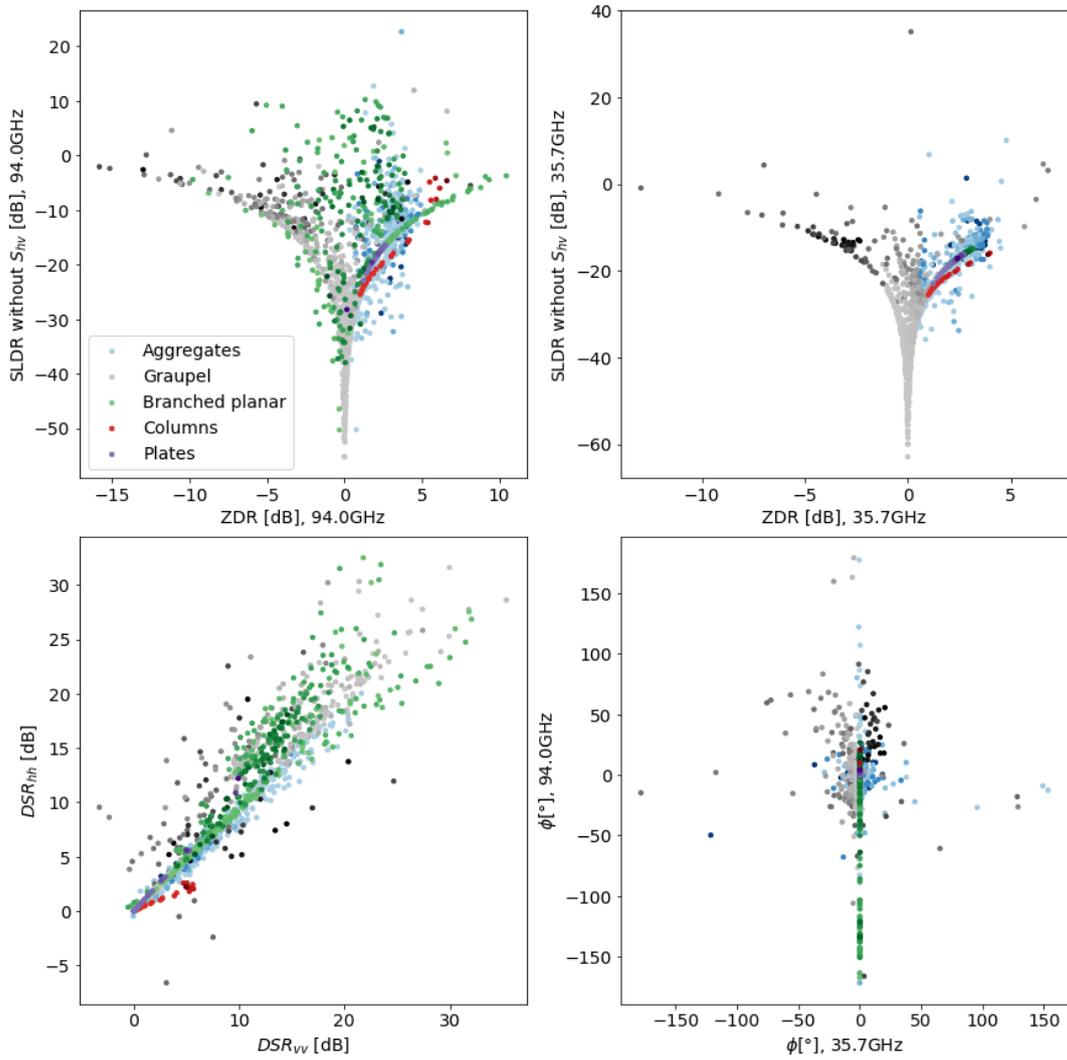


Figure 4.4: Polarimetric variables for all particles in the database. Particles are coloured by mass, where the darkest shade for a given colour is the maximum mass for that particle type and the lightest colour is the lightest particle for that type.

step is completed in its simplest form. Doing this, two problems are encountered: the membership functions become either too wide or too narrow.

To illustrate this first problem, consider the Z_{DR} at both frequencies for graupel. While most conical graupel particles show negative Z_{DR} , from the database we also get some positive values. So, when setting the membership function at 1 from highest to lowest value, it will simply be 1 for every single value of Z_{DR} . However, arguably when the Z_{DR} is positive, it is most likely not a conical graupel particle, and thus the membership function should be at least lower than 1. A way to deal with this was already proposed in section 2.3: by using trapezoidal membership functions, where the function linearly decreases towards the most extreme values on the left and right. Referring back to Figure 2.6, as suggested there, the membership function decreases linearly outside 5-95% of the values for that particle type.

The second problem concerns the complete opposite: the range of values is so narrow that the membership function barely exists. For this, consider the differential phase of plates, columns, and branched planars at Ka-Band. From the database, all particles have a differential phase of 0° . When making a membership function of that, a measurement of 0.001° will already fall outside the function. For that, a rough fix has been implemented: if the range of values for a variable for a type encompasses less than 5% of all possible values, the membership function is widened by 1% on both sides. So, for Φ_{35}

of plates, the lowest value is 0° and the highest 0.21° , which is less than 5% of the possible values. The range of values from the database particles goes from -178° to 153° for this particular variable, so the membership function of plates for Φ_{35} is set to range from -3.3° to 3.5° . A second argument for enlarging this range of values is the presence of some light noise in the radar data.

When constructing the membership functions as described above, one more detail arises as a possible problem, namely the size of the particle groups. Again regarding the Z_{DR} of graupel, the particles that account for the values around, say, $5[dB]$ can be regarded as outliers. The ability to represent them with a lower value of membership is a strength of the fuzzy logic classification. However, the column and plate particles show values that are all close together and are in line with expectations. Excluding the most extreme values for these types just narrows the membership function needlessly. For now, this problem is not addressed in the making of the membership functions but is regarded again in the discussion.

With those two additions, the membership functions can be constructed. For each variable, the functions are visualized in Figure 4.5. The weights are set at $W = 1$.

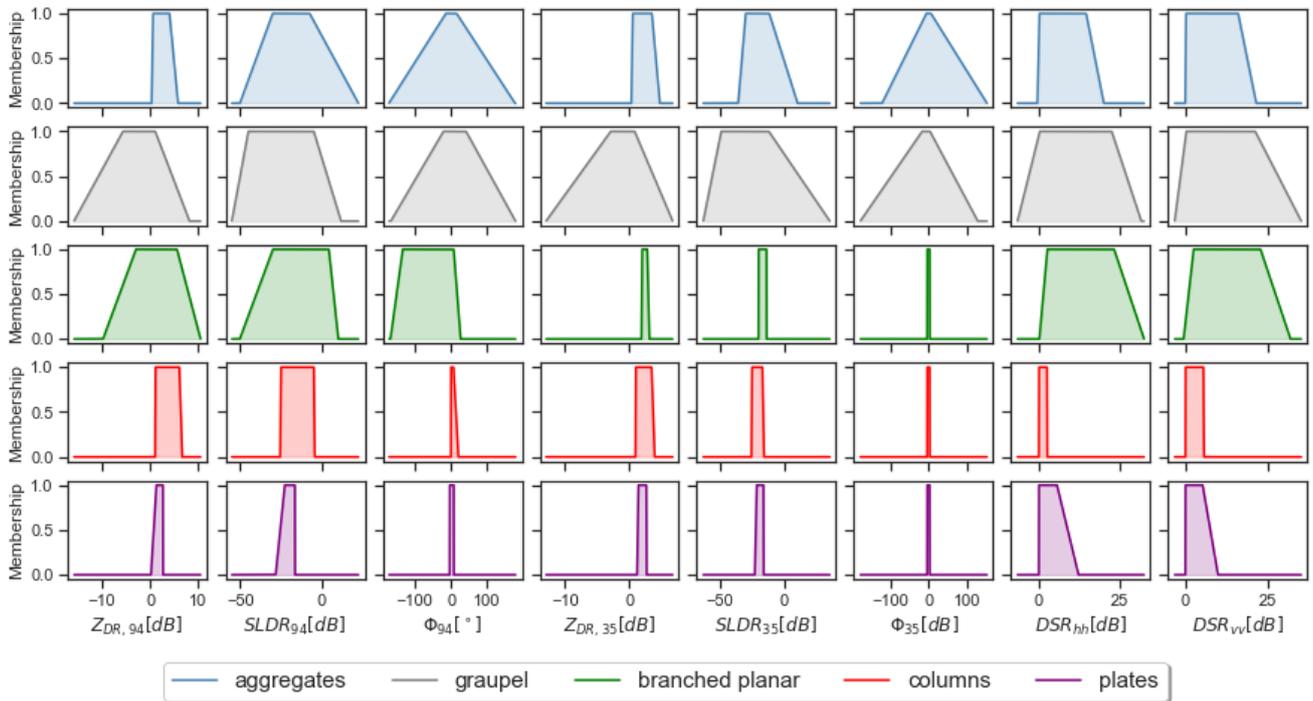


Figure 4.5: 1D trapezoidal membership functions.

4.2.2. 2D fuzzy logic set-up

Looking at Figure 4.5, the overlapping problem that is inherent to this ice particle classification becomes clear. When these graphs are all the classification is based on, a measurement of a column can easily be classified as plate, for example. However, in the scatterplots from Figure 4.4, these two types are distinct when looking at the DSR plot. Though subtle, together with the differences that can also be seen in the Z_{DR} -SLDR plots the two particle types can be differentiated. However, this 2D information is lost when moving to the 1D membership functions in Figure 4.5. As introduced in section 2.3, Al-Sakka et al. [1], among others, have constructed a way to include this 2D information by using 2D membership functions. Based on that, a simple way to set up a 2D membership function is as follows:

While taking SLDR and Z_{DR} as an example:

1. The range of SLDR is divided into bins.
2. For each bin, the highest and lowest value of Z_{DR} for each particle type are determined.
3. For a data point, the SLDR is assigned to its corresponding bin.

4. When the Z_{DR} measurement is in between the minimum and maximum of that bin, a 1 is returned for that data point. Else, a 0 is returned.

Applying the method above leads to the 2D membership functions as in Figure 4.6. The bin size is set at $5[dB]$. This value was based on visual inspection: balancing keeping enough detail in the shape of the function, while not having too many holes in the range.

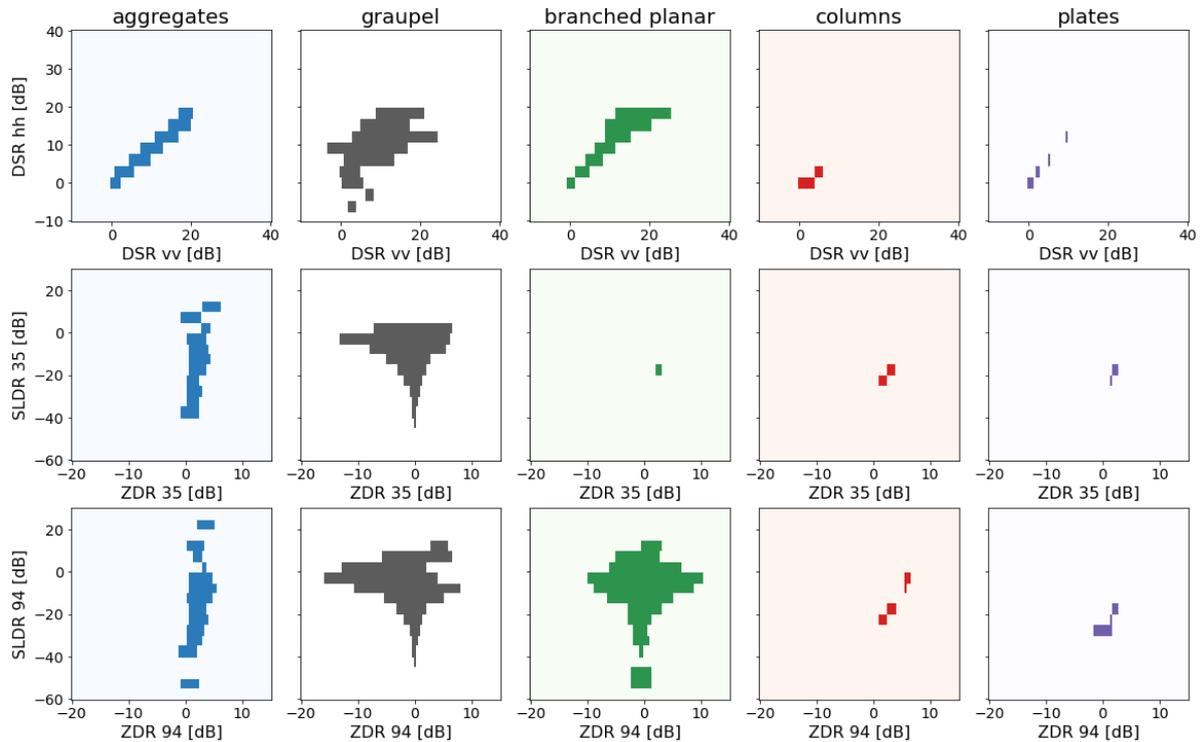


Figure 4.6: 2D membership functions based on the scattering database of Lu et al. [31] and the method of Al-Sakka et al. [1]. Bin size is $5dB$.

A consideration of 2D membership functions is their shape. On one hand, the 2D membership function is added to decrease the problem of overlap. As such, the areas should not be too big, or the problem remains. On the other hand, the particles in the database only span a range of possibilities, so leaving out the area between particles may leave out a lot of important information. For now, for example, the membership function Z_{DR} of the conical graupel now still encompasses all positive values. At the same time, one could imagine that a columnar ice particle can take on a Z_{DR} and $SLDR$ value at $94GHz$ that falls exactly in between the two ‘islands’ that can now be seen in Figure 4.6. Looking at Figure 4.4 $Z_{DR,94} = 5dB$ and $SLDR = -10dB$ would fit exactly in the row of columnar particles, but would return as a 0 from the 2D membership function.

Another consideration in 2D membership function is the question of which variables to combine. Keeping a robust classifier in mind that might also perform when leaving out one frequency, the combination $SLDR-Z_{DR}$ for both 94 and 35GHz and $DSR_{94}-DSR_{35}$ is chosen. As such, when one frequency is not available or is very noisy, at least one 2D function can be used. However, this scenario does highlight a weakness of using 2D functions in general: it is more sensitive to noisy or bad data. Looking at the differential backscatter phase in Figure 5.2, the 2D overview does not add above the 1D functions. Therefore, the phase is not considered as an input for the 2D functions. The option to use principal component analysis to form new variables that can discriminate particle types better has been investigated in Appendix B.

With the 1D and 2D membership functions, three different set-ups are considered:

1. All variables have 1D trapezoidal membership functions.

2. SLDR- Z_{DR} are 2D for both frequencies; the rest have trapezoidal membership functions.
3. Both SLDR- Z_{DR} and DSR have 2D membership functions; only ϕ still has 1D membership functions.

These set-ups are compared against each other in section 5.2.

4.2.3. Incorporating environmental variables

Apart from polarimetric variables, there are other factors that can help classify ice particles. From section 2.1 it is clear that temperature and water vapour saturation influence the formation and growing of ice particles. Since the temperature and liquid water path are available from the microwave radiometer, and because the polarimetric variables alone might not be enough to classify reliably, these environmental variables can be added in the fuzzy logic classifier.

The addition of these variables has been done before. Aydin and Singh [2] use temperature as a 1D input to the fuzzy logic algorithm. They only differentiate columnar crystals from other ice particles and give graupel a value of 1 for all temperatures. Al-Sakka et al. [1] on the other hand, uses temperature as a multiplicative term. This means that, if the temperature falls outside the range of a particle, it is immediately deemed impossible. In this way, temperature gets a more important role. As explained in section 2.1 there are many factors influencing the exact growing of ice particles, and temperature alone cannot distinguish particles. Al-Sakka et al. [1] does not differentiate between many particles, but instead between 'ice' and 'rain'. Indeed, in that case it makes more sense to use temperature as a definitive factor.

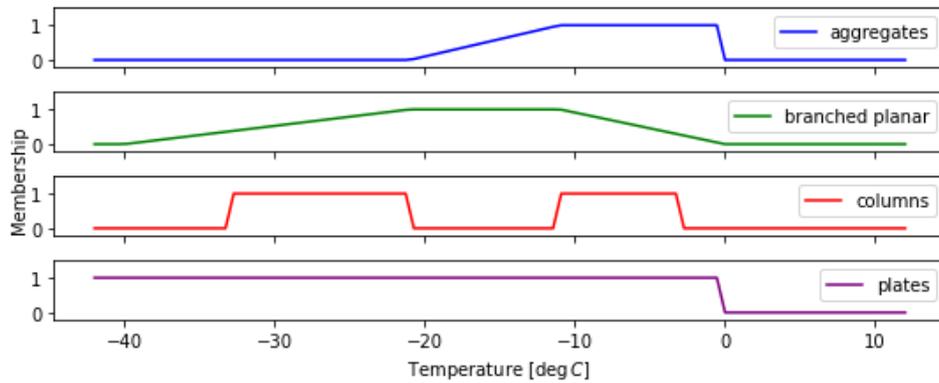


Figure 4.7: Membership functions of temperature for aggregates, branched planar, columns and plates. The values correspond to Table 2.1.

The situation is different when incorporating the LWP. The formation of graupel is dependent on a minimum value for the LWP. As mentioned in section 2.1 graupel can exist with liquid water paths from about $50g/m^2$, but is typically seen from $100g/m^2$. Therefore, the membership function of the LWP is as in Figure 4.8. Because here the LWP is a prerequisite for graupel, it is incorporated as a multiplicative term, as in Equation 2.23. In other words, when the LWP is below $50g/m^2$, the Q value for graupel becomes 0. For all other particle types, the membership function is simply 1, meaning it has no influence.

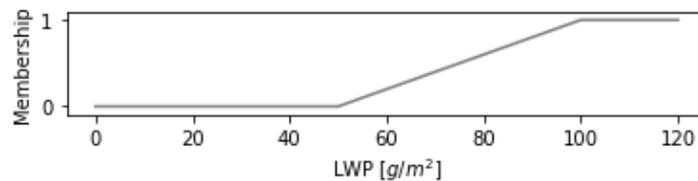


Figure 4.8: Membership function of liquid water path for graupel.

Thus, the temperature is only added as a 1D fuzzy logic membership function in the fuzzy logic al-

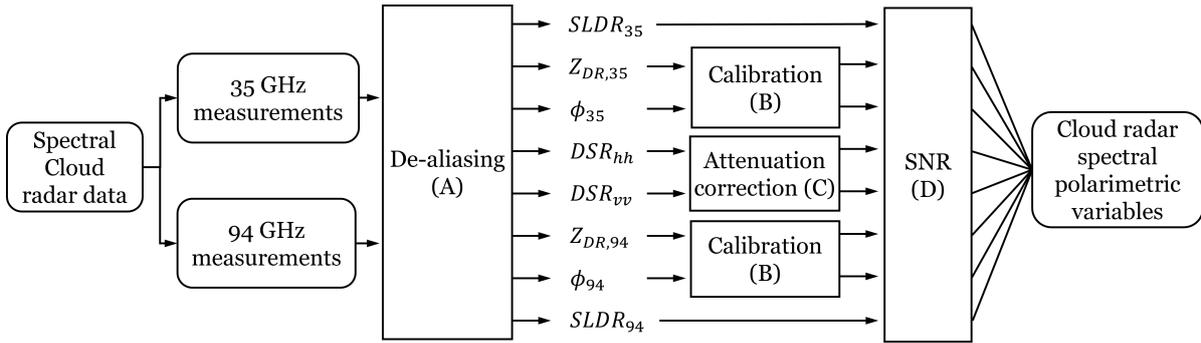


Figure 4.9: The pre-processing procedure for the spectral cloud radar data to obtain spectral polarimetric variables.

gorithm. The membership function of the LWP is applied to graupel as a multiplicative term. All weights are set at $W = 1$. With the fuzzy logic algorithm, it is possible for the value of Q to be the same. When two or more particle types have the highest value of Q , the classifier returns the classification as 'none'.

4.3. Radar data preparation

When obtaining the radar data, the measurements cannot be used as an input for the classifier immediately. The steps that are needed are visualised in Figure 4.9. The first step is to de-alias the Doppler spectra and interpolate them. Only then the polarimetric variables are calculated as described in section 2.2. Further, for Z_{DR} and ϕ additional calibration is available and thus applied. Since attenuation depends on the wavelength, the SLDR needs to be corrected for this, as it contains two different wavelengths. These steps are described in subsection 4.3.2. But first, the events need to be chosen.

4.3.1. Events

Throughout the definition of the spectral polarimetric variables, some necessary assumptions are mentioned. The first is that a non-precipitating event needs to be chosen. This is because precipitation increases attenuation, which is then different for polarizations and frequency. This would influence the accuracy of all variables using two polarizations or two wavelengths. Moreover, precipitation decreases the range of the radar. Since ice clouds are often quite high in the Netherlands, maximizing range is preferable. By choosing a non-precipitating event, the differential propagation phase is kept to a minimum as well. Another requirement is, of course, that the cloud case has to be an ice cloud. In order to use spectral polarimetric variables, the cloud radar needs to have an intermediate elevation angle in the range of 30° - 45° . The elevation angle 45° is chosen.

As introduced in section 2.2, the phase as measured by the cloud radar consists partly of the differential propagation phase ϕ_{dp} . Because in comparison with the database only the differential backscatter phase ϕ_{bs} is used, this ϕ_{dp} needs to be minimized as well. Again, this is done by omitting precipitating events.

Two study cases are considered. The first case is a single-layer ice cloud and the second one consists of a two-layer ice cloud, which may increase the complexity of the media to be classified. They both occurred on the 26th of January 2021 and were documented in Wang [61]. Documented ice cloud cases are first deemed for the analysis of the proposed new classification method.

Event A

The first event occurred in the time interval 18:00-19:00 UTC. Here we have a single cloud layer. In Figure 4.10 some of the bulk variables of this case are depicted. The differential phase is left out on purpose, because all values were practically 0° . In this plot, Z_{DR} is calibrated. Note that the SLDR is depicted instead of SLDR. This is because we are considering bulk variables and thus SLDR instead of the spectral SLDR. As a first check, indeed this case does not include precipitation. Looking at the temperature, the cloud seems to be from $-20^\circ C$ to $-30^\circ C$, so most likely an ice cloud.

Throughout this event, the liquid water path heavily fluctuates between 10 and $70 g/m^2$, see Figure 4.11. Wang [61] explains that convection at the top of the cloud leads to ice particles, which then fall down

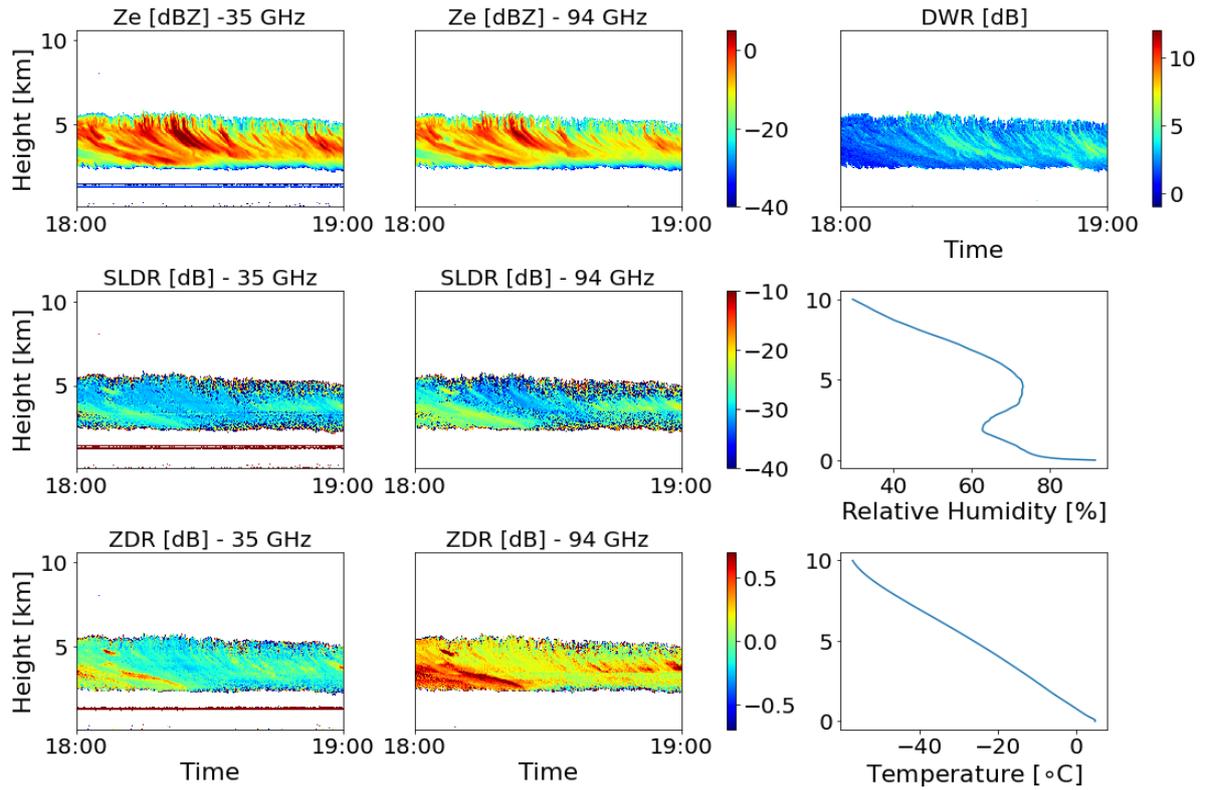


Figure 4.10: Bulk variables, temperature, and relative humidity for Event A.

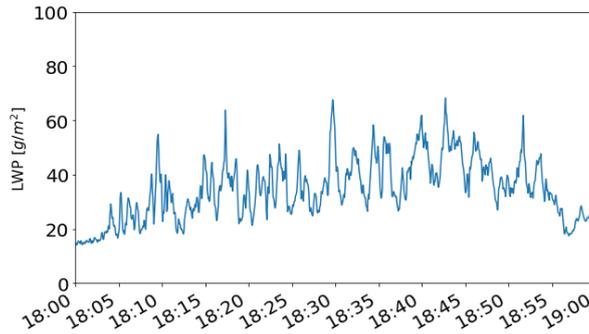


Figure 4.11: Liquid water path Event A

the fall streaks clearly visible in the Z_e profiles. Near 18:00, the Z_{DR} is slightly higher and the SLDR quite low, which might indicate small, non-spherical particles. Later, SLDR increases and Z_{DR} nears 0, indicating larger and rounder particles.

Event B

Event B occurred later on the same day, between 20:00-21:00. Because these graphs will be compared in the final bulk plot, they are included here as well in Figure 4.12. The cloud can be split into two parts, divided at a height of about 4000m. Especially the high DWR stands out, as well as streaks of high and low SLDR in the lower part. There, there is not only a big difference between the two wavelengths, but also between the two polarizations, as can be seen in the Z_{DR} . Some fall streaks are seen as well, albeit weaker than in event A. The negative temperature indicates a full ice cloud, and the LWP is low enough for analysis; see Figure 4.13.

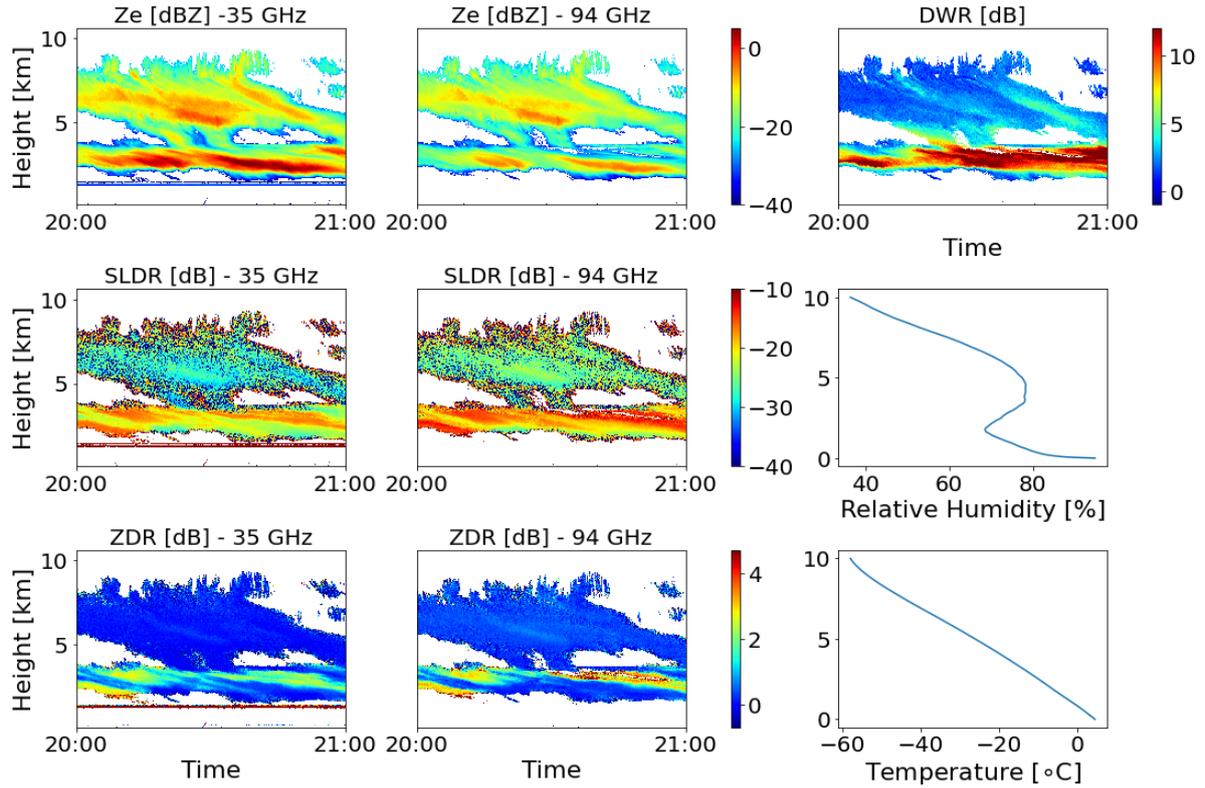


Figure 4.12: Bulk variables, temperature, and relative humidity for Event B.

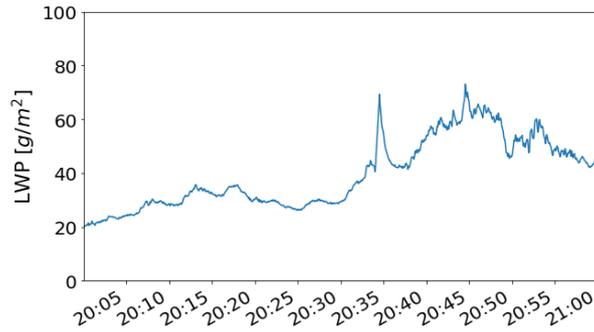


Figure 4.13: Liquid water path for Event B.

4.3.2. Retrieval of polarimetric variables

To obtain the spectral polarimetric variables from the cloud radar data, the two spectra of the two bands need to be matched, for which the dealiasing code of Wang [60] is used. The dealiased spectra are used to calculate the spectral polarimetric variables. These are then calibrated, and the SLDR is corrected for attenuation.

A: De-aliasing

When combining the Doppler spectra of the two wavelengths, a problem appears: the Doppler velocity resolution dv_d is not the same for the two spectra, and neither are the Nyquist velocities. In Wang [60] a code to dealias the Doppler spectra has been made. The output of this code consists of a spectrum per height and time index. Each spectrum is defined by a minimum Doppler velocity and step (bin) size, for both frequencies.

As the step size in $35GHz$ is bigger than that of $94GHz$, the $94GHz$ spectra are linearly interpolated to those of $35GHz$. This is shown in Figure 4.14. The magnitude of the $94GHz$ spectrum is smaller

because the dv is smaller, and thus, as per Equation 2.5, sZ_{hh} is smaller as well.

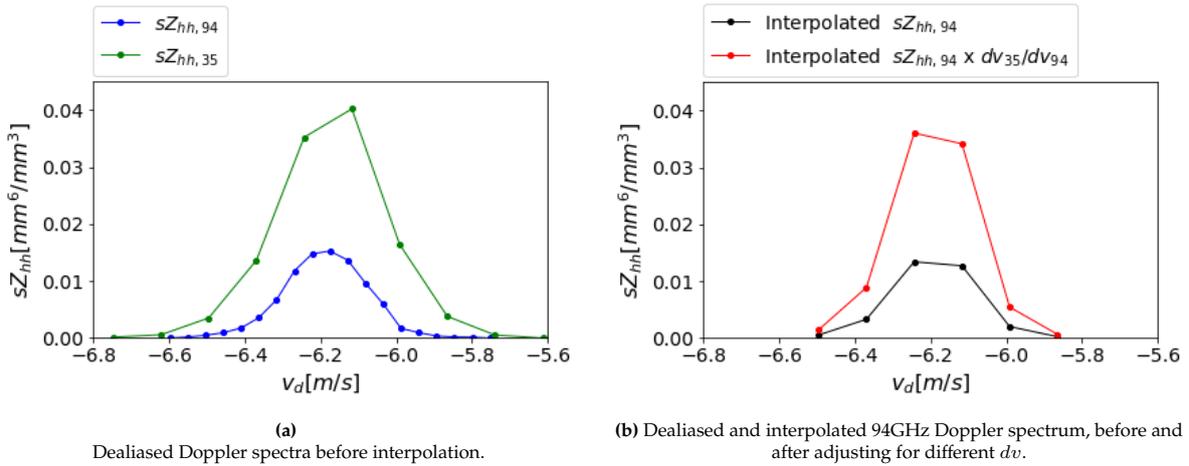


Figure 4.14: Interpolation and bin size dv_d correction, for Event A $t = 18 : 00 : 36, h = 2762m$.

A disadvantage of interpolating like this is that information is lost on the $35GHz$ spectrum by not using the tails, as well as on the $94GHz$ spectrum by interpolating.

After dealiasing, the polarimetric variables are calculated. This is done as described in chapter 3. The fact that $dv_{94} \neq dv_{35}$ in Equation 2.16 is corrected for by multiplying $sZ_{hh,94}$ with dv_{35}/dv_{94} as shown in Figure 4.14b.

B: Calibrating

Calibration was applied to both Z_{DR} and ϕ for both frequencies. This was done using the method of Mak [33]. She provided a calibration profile, which the dealiasing variables were multiplied by (Z_{DR}) or added to (ϕ). This profile is based on zenith-pointing radar data from a light stratiform rain day as close in time to the events in her thesis, namely 19-05-2021.

C: Correcting for attenuation

While a radar beam travels through the atmosphere, some of its energy is lost through interaction with gasses in the atmosphere and particles on its path. This is known as attenuation [45]. Attenuation is larger in smaller wavelengths [6], and thus needs to be taken into account when computing the DSR. Considering the attenuation due to the presence of gasses and/or spherical particles, the polarimetric variables are all ratios where the numerator and denominator are expressed at the same frequency and thus experience the same attenuation, which is then cancelled out.

Calculating attenuation is done using recommendations from ITU R676-10 [11]. For that, temperature, pressure, and water vapour density are needed. Temperature and water vapour density (absolute humidity) profiles are taken from the microwave radiometer. Since the surface pressure (p_{surf}) is known, the pressure at any height can be calculated with surface temperature T_{surf} , the gas constant R , gravitational acceleration $g = 9.81m/s^2$, the lapse rate $a = -0.006K/m$ and the temperature T [59]:

$$p = p_{surf} \left(\frac{T}{T_{surf}} \right)^{-\frac{g}{aR}} \quad (4.1)$$

The attenuation of ice particles is ignored, since it is several orders of magnitude smaller than other components in the atmosphere [58]. The liquid water attenuation is also ignored, since non-precipitating events are chosen. (Non-precipitating) water particles suspended in the air are generally small and round, and thus influence the horizontal and vertical polarizations equally for one wavelength. An alternative to calculating pressure profiles is obtaining them from ECMWF models [41]. This shows at most a 4% difference in pressure from the method above, and is not used due to the extra steps needed.

The attenuation in $[dB/m]$ is calculated for each measurement level of the microwave radiometer, and added together to form a total attenuation in $[dB]$ for each height of the cloud radar. This is done using a Python package [11].

With the attenuation $a[dB]$ of both frequencies from the ground to a given height, the DSR is calculated with

$$DSR = DSR_{measured} - 2 \frac{a_{94} - a_{35}}{\sin(45^\circ)}. \quad (4.2)$$

Note that the attenuation is divided by $\sin(45^\circ)$ because the radar beam is slanted and thus experiences more attenuation than if it was pointing towards the zenith. The attenuation is multiplied by 2 because the path is travelled twice: forwards and backwards.

D: Signal-to-noise ratio

Applying a filter based on signal-to-noise ratio (SNR) can clean up the data before applying the classifier. The noise from the cloud radar data is given as $CHNoisePower$, which is the noise power per timestamp and height. To get to spectral noise power, it is divided by the amount of velocity bins:

$$sSNR(t, r, v_d) = CHNoisePower(t, r, v_d) / N_d(r), \quad (4.3)$$

with N_d the number of Doppler velocity bins. A negative SNR in the logarithmic scale means the noise

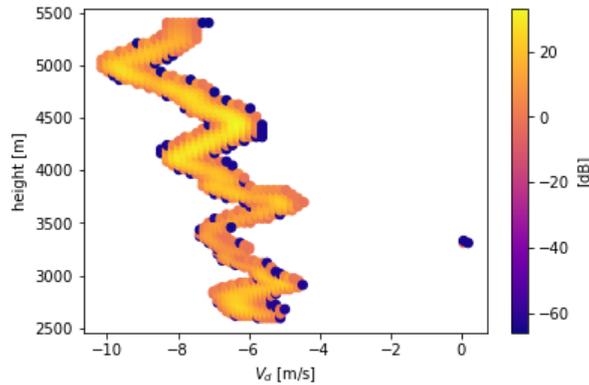


Figure 4.15: SNR of 35GHz horizontal polarisation at event A, 18:00.

is larger than the signal. In Figure 4.15 it can be seen that in the middle of the data the noise is the smallest. This is not strange considering that the noise is the same over every Doppler bin, but the signal strength follows a curve (like in Figure 4.14); the signal is weaker on the edges, and so the SNR is lower.

Applying a signal-to-noise ratio (SNR) can lead to cleaner data; however, it also leads to fewer data points. A SNR of 0dB removes 17% of all points, while a SNR of 5dB removes about 35% of data points. Looking at Figure 4.15, setting the SNR at a maximum of 0 already gets rid of the strangest values. When a value is significantly noisy, it will fall outside the membership functions, and thus this noisy measurement will be picked out by the classifier by having a low Q value (see Equation 2.22) for every particle type. The SNR is applied to all variables at $SNR = 0dB$ for both frequencies.

5

Results

Following the method from the previous chapter, the results can be obtained. This chapter contains these results and some observations and preliminary discussion points. The structure follows that of the research (sub)questions. and the method (see Figure 4.1). First, the spectral polarimetric cloud radar variables are compared with those as modelled by the scattering database in section 5.1. Then, theoretical performance of the different fuzzy logic set-ups is shown in section 5.2. The classifier is applied to specific timestamps of the events in, and the results are shown in section 5.3. The quality metrics of the results Q and Q -gap can be found here as well. Following, the sensitivity to some of the different variables is depicted in section 5.4. Finally, the results are summarized in some bulk plots, showing the entire events, in section 5.5.

5.1. Comparing cloud radar variables against scattering database variables

To answer the question: “How can cloud radar variables be related to those modelled by a state-of-the-art scattering database?” the polarimetric variables were derived in subsection 2.2.2. The polarimetric variables are obtained using the method in section 4.1 for the scattering database, and section 4.3 for cloud radar measurements. As introduced in section 1.2, an important assumption in the method used is that the ice particles in a resolution volume are spread out over the Doppler spectrum through different fall speeds, such that every velocity bin only contains one particle type. With this assumption, the scattering properties from the scattering database can be related to cloud radar data without the need for a particle distribution. To check the sensibility of the assumption and the pre-processing steps in the method, the polarimetric variables obtained from the scattering database and the cloud radar are compared in this section. First, all data points are compared to the spread of the modelled variables in subsection 5.1.1. Then, a spectrogram example is shown in subsection 5.1.2.

5.1.1. Spread of the spectral polarimetric variables

To see if spectral cloud radar measurements can be related to single scattering database particles, plots such as Figure 5.1 are made. The expected values in this case are those retrieved from the scattering database, which is why the outline of the 2D membership function is shown as well. For clarity’s sake only the outline of the 2D membership function of graupel is shown in this graph, as it is quite representative of the values that for SLDR and Z_{DR} that the particles from the scattering database take on. Clearly, there are some clusters of data points falling far outside the expected regions. Note that though it seems like a third of the cloud radar data points fall inside the expected region, there are 1.5 million points, and the point density near the 2D membership functions is higher. These plots are similar for the other spectral variables, with clusters of radar measurements far outside the regions as modelled by the scattering database. The reason for the clusters around SLDR values of -60 and $-120[dB]$ is explained by looking at the SNR. It becomes apparent that the somewhat strange clusters on the left correspond to very low signal-to-noise ratios.

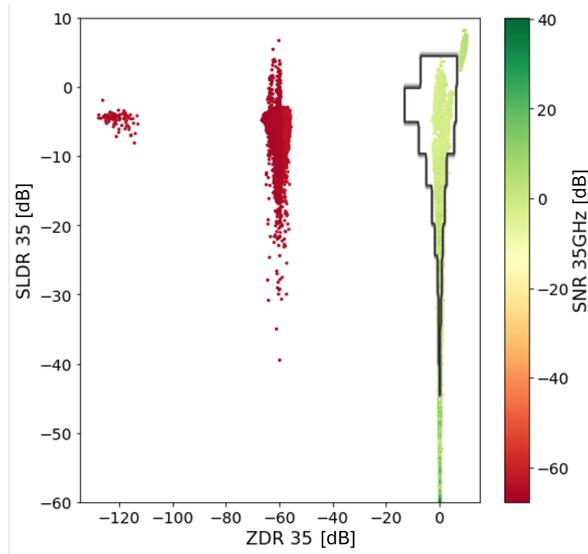


Figure 5.1: The 35GHz SNR for SLDR vs Z_{DR} for 35GHz cloud radar measurements of Event A with the 2D membership function for graupel (black lines).

After applying the SNR for both 35 and 94GHz, the 2D scatterplots of the variables are as in Figure 5.2. The comparison of the membership functions and the data points for the phases can be seen in Figure 5.3. The measurements are centred around 0, with a relatively small amount of points slightly positive or negative. This indicates that the particles are limited in size and most likely consist of dry ice [55]. In general, the measurements cover roughly the same region as those modelled by the scat-

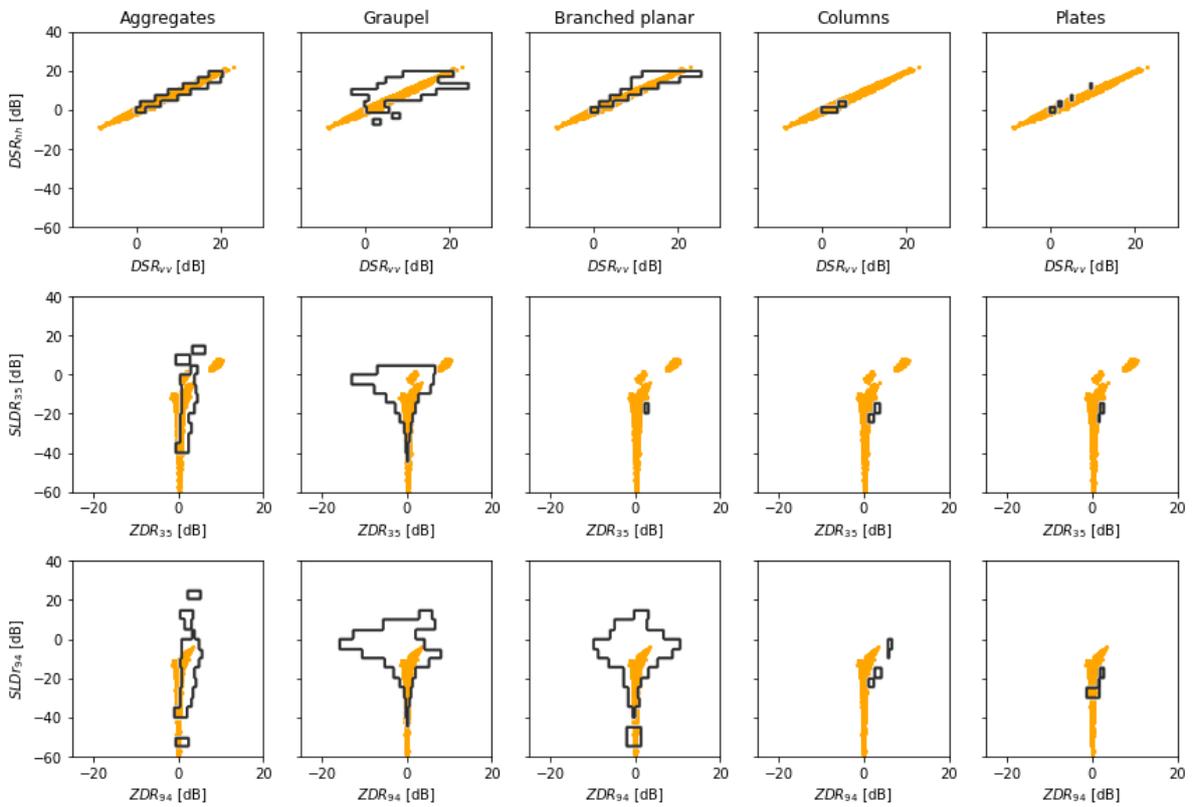


Figure 5.2: The scatterplots of all cloud radar measurements of Event A after applying SNR=0[dB] (orange dots). The outlines of the 2D membership functions are given per particle type (black lines)

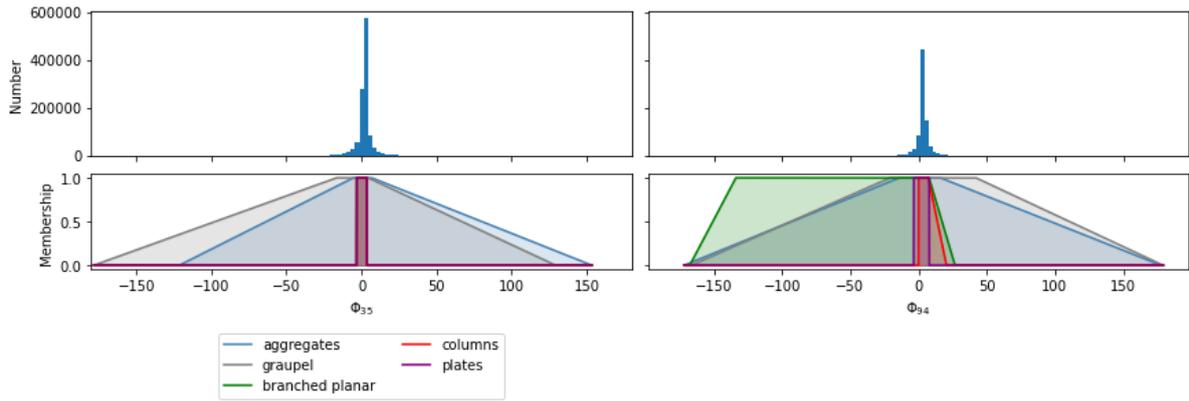


Figure 5.3: Phase membership functions (below) and all cloud radar measurements of Event A after applying $SNR=0$ [dB](top).

tering database. Figure 5.2 exposes one of the shortcomings of the way the 2D fuzzy logic classifier is set up right now. Since the membership functions are based on the (limited) values from the database, it was inevitable that some measurement data points would fall outside this region. Aside from the points with low SNR, there seem to be quite a lot of points with decent SNR that do not fit inside any of the membership functions. Actually, it seems the points showcasing particularly low SLDR values have the highest SNR . This shortcoming will show up in the classification results when the value of Q is 0 for all particle types.

5.1.2. Spectrogram of polarimetric variables 18:00

With the first check if the polarimetric variables from the scattering database and the cloud radar compare, the data can be analysed in more detail. For Event A, the spectral polarimetric variables for one timestamp look like Figure 5.4. With the Doppler velocity in the x-axis, the spectra look quite squiggly. This is because when the elevation angle is 45° the horizontal component of the velocity also plays a part, and thus horizontal winds do as well. This is also a reason why in this method the fall velocity

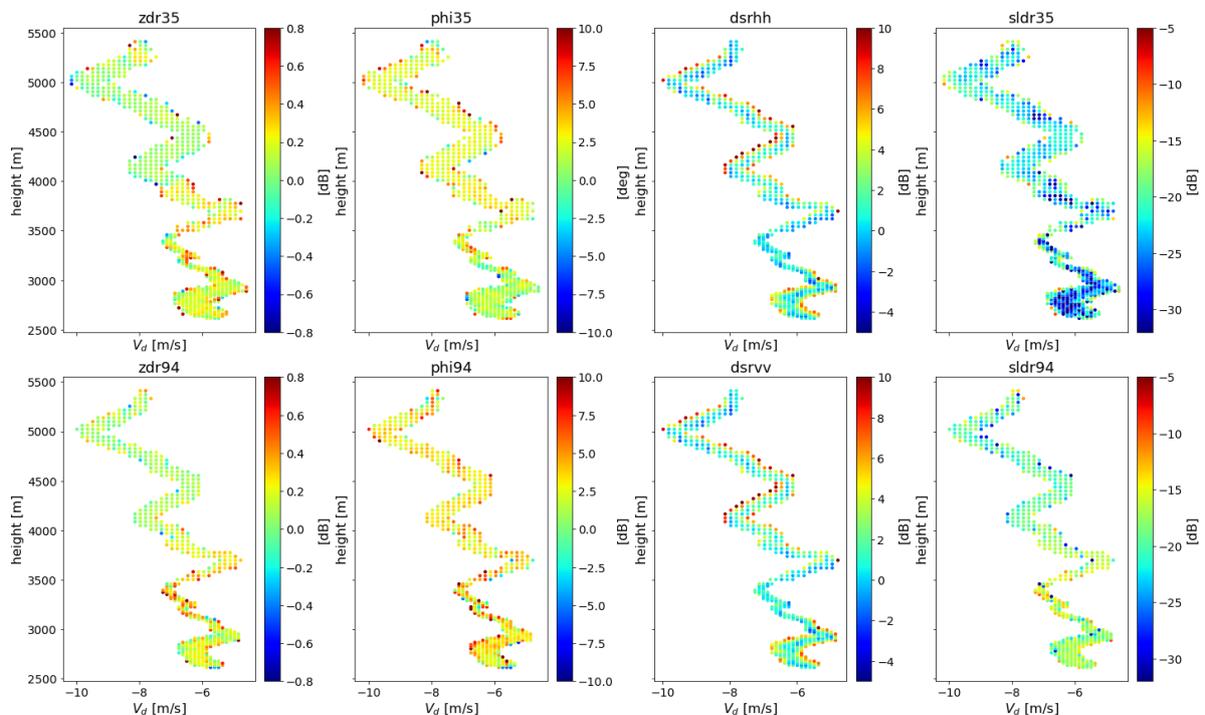


Figure 5.4: Spectral polarimetric variables for Event A, 18:00.

is not taken as an input to relate to particle mass and size; the influence of the wind velocity would have too big a part and is not measured accurately enough yet to be able to decompose the horizontal (wind-dependent) and vertical (particle-dependent) components of the Doppler velocity. Note that the data points up to a height of 3500[m] are closer together because the Doppler velocity bins are smaller below that height.

Looking at the variables, a couple of things can be noticed. Checking the measured values against the values that are expected from the scattering database, the magnitudes span about the same ranges. This affirms the methods used to calculate the variables both for the scattering database and the cloud radar data. Across the Doppler velocity axis, some different values can be seen. Especially the DSRs and SLDRs show some gradient. The DSR is influenced by the size of the particle, and since the Doppler velocity is dependent on the fall velocity and thus size of the particle, a gradient in DSR across the Doppler velocity is indeed expected. This because the DSR increases for particle size, and heavier particles are falling faster, thus spreading the particle sizes along the Doppler velocity spectrum. This difference in values across the spectrum confirms that ice particles are separated in different velocity bins according to shape. There is not a lot of difference in the kind of values at the bottom of the cloud versus at the top, suggesting that the cloud particles are quite consistent throughout the altitudes. With an SNR of 0 dB, some outliers may still be present at the edges of the spectra where the spectral SNR is the lowest. However, with the fuzzy logic classification, that may not be an issue, as explained in Equation 4.3.2.

5.2. Fuzzy logic set-ups tested on database

The fuzzy logic classifier is made with the variables from the scattering database. As in section 4.2, some choices were made concerning the classifier. In this section, different set-ups are tested against the particles of the database that the classifier was based on to compare them. First, the 1D is compared with the 2D set-up. Then, a set-up with only the 94GHz band is tested.

5.2.1. 1D vs 2D

In section 4.2, three different set-ups for the fuzzy logic classifier are defined. With these set-ups defined, there are two questions about the performance of the classifier. First, if the different particle types can be differentiated with the current approach, and which particles might have the largest overlap. And second, how the different set-ups compare.

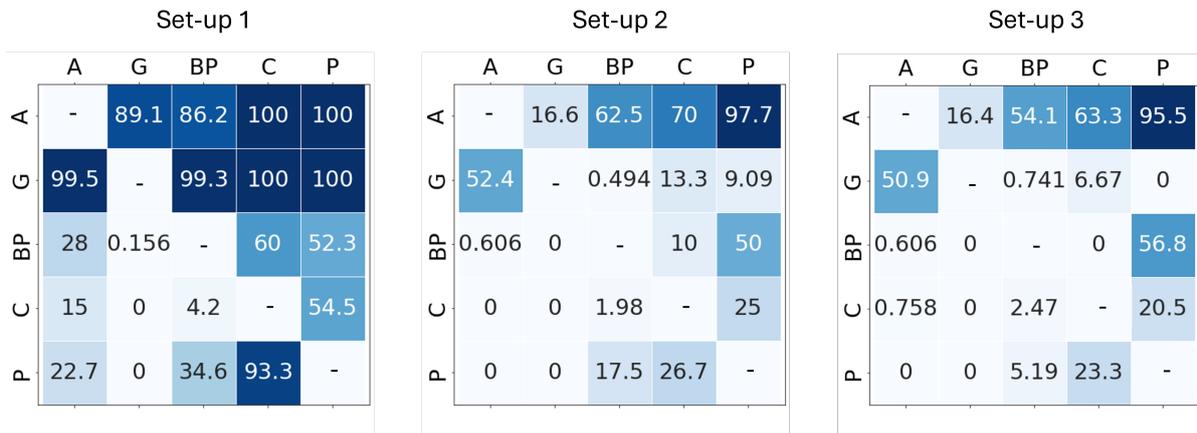


Figure 5.5: Misclassification percentages for the different fuzzy logic algorithm set-ups, vertical is the input and horizontal is the output. A = Aggregates, G = Conical graupel, BP = Branched planar, C = columns, P = plates. E.g. for set-up 1, 22.7% of aggregates get misclassified as plates.

To start answering these two questions, the particles from the database can be used as an input to the classifier. Of course, this is by no means a proper classifier test as the testing data are the exact data that the classifier is built with, but it serves merely as an indication if the classes are at all differentiable solely by the spectral polarimetric variables. In Figure 5.5 the three different set-ups have been tested. What

can be seen are the percentages of misclassifications. Since the input and the testing data is the exact same, each particle gets appropriately assigned to the correct class. However, it often also gets classified as another particle. That is, in the first set-up for any plate or column particle as input Q is maximum for aggregates and graupel as well. To illustrate: of all graupel particles in the scattering database, 89.1% also gets misclassified as an aggregate in Set-up 1. One takeaway from the first set-up is that graupel is not often misclassified, but other particles are easily classified as graupel. This is no surprise when looking at Figure 4.5. There we see that the 1D membership function of graupel encompasses almost every possible value. What Figure 5.5 clearly shows is that using at least the 2D $Z_{DR} - S_{LDR}$ matrices adds to the discrimination between types. Only plates still almost always get classified as aggregates. The third set-up seems to behave slightly better in most cases. Indeed, columns and plates are separate in the scatterplot of the $DSRs$ in Figure 4.4, and thus their 2D membership functions are different in Figure 4.6, where they are almost exactly the same in the 1D functions.

Keep in mind that these results do not include temperature and liquid water path as variables; they merely serve as a check for the classifier based on the polarimetric variables from the database. Actually, what they show is that with just the spectral polarimetric variables, there is still a big chance that particles get misclassified as aggregates. Further, without the inclusion of any other variable such as LWP, in an ideal case at least 50% of aggregates would also get classified as graupel. The expectation is thus that the temperature and LWP will play a big role in the performance of the classifier, specifically to differentiate between graupel, branched planars, and aggregates.

Since set-up 3 seems to perform best here, this is the one continued on with. Unless specified, the results shown are thus using this set-up.

5.2.2. Single frequency

Since most of the RPG type cloud radars in Europe operate at the single frequency of 94 GHz, it is interesting to know how the classifier performs only considering the variables available in that case. In this scenario only Z_{DR} , S_{LDR} , and ϕ_{bs} for 94GHz are available. This is less than half of the variables, as of course the $DSRs$ require a second wavelength. The results are visualized similarly to the dual wavelength results in Figure 5.6.

	A	G	BP	C	P
A	-	26.7	68.6	86.7	97.7
G	99.7	-	98.5	100	100
BP	97.6	88.6	-	100	100
C	31.5	1.25	52.1	-	95.5
P	22.9	0.938	21	60	-

(a) Set-up 1

	A	G	BP	C	P
A	-	28.4	64.4	70	97.7
G	75	-	57.5	26.7	27.3
BP	63	74.5	-	53.3	90.9
C	9.55	0.156	24.9	-	40.9
P	22.4	1.56	20.7	40	-

(b) Set-up 3

Figure 5.6: Misclassification percentages for a fuzzy logic algorithm based on the single frequency 94 GHz

Perhaps unsurprisingly, both of these set-ups perform significantly worse than those using the dual wavelength set-ups. Especially the 1D set-up is unable to differentiate between particle types.

5.3. Classification results on cloud radar data

With the fuzzy logic set-up chosen, the results of the classification of cloud radar data can be made. Some specific timestamp are highlighted to showcase some different observations. First, in subsection 5.3.1, the results for Event A at 18:00 are shown. Also visualised in this subsection are Q and Q -gap. According to the bulk polarimetric variables, perhaps some different classifications are expected at 18:30 in Event A. Therefore, the results on this timestamp are shown in subsection 5.3.2. Event B has

two cloud layers, and its results in the form of spectrograms is given in subsection 5.3.3.

5.3.1. Classification results Event A 18:00

Applying the fuzzy logic to the cloud radar for 18:00 leads to the classifications as Figure 5.7. Here, the results for set-up 1 and set-up 3 from subsection 5.2.1 are next to each other. The reason to still compare with the 1D set-up even after the results from the previous section is that the 1D method is simpler, and thus quicker to run. So if the results are similar, the 1D method is preferred.

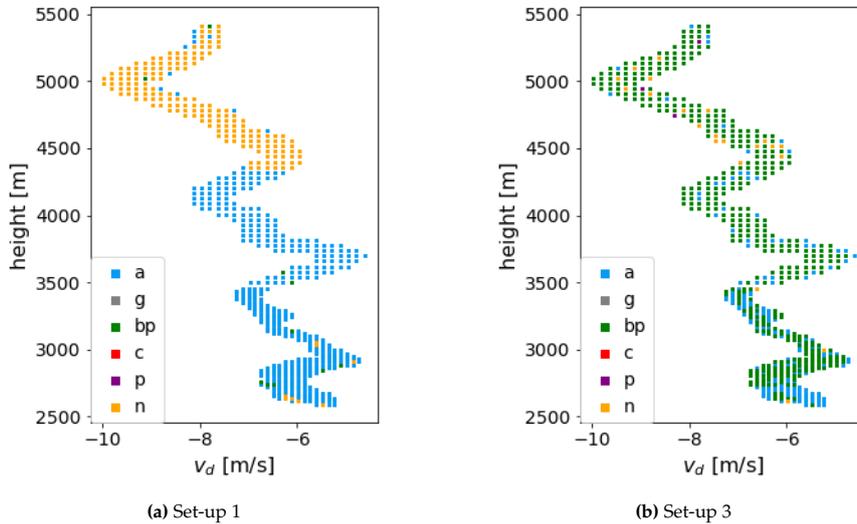


Figure 5.7: Classification results for Event A for two different set-ups, 18:00.

Remember that the class 'none' means the maximum value of Q occurs for two types. As expected, the 1D results do not seem good. That is, a lot of data points are not classified, and those that are, are uniformly the same. That is of course possible, but subsection 5.2.1 shows that the 2D method should behave better. So if the results are different, it is reasonable to say that the one using a 2D set-up is more accurate. In Figure 5.7b the results show two particle types: aggregates and branched planars. The absence of graupel can be explained by the low LWP, and as mentioned in subsection 4.3.1 and Wang [61] the main expected particle type in the cloud would be aggregates. However, the DSR values are quite low here at the first timestamp, so perhaps there could be some smaller particles such as plates or columns.

To further investigate the results, the confidence of the classification can be studied. This can be done in two ways: the value of Q for the classification, and the gap between the highest and the second-highest value of Q . See section 2.3 for the definition of Q . As a reminder, there are 6 membership functions: 2 for $SLDR - Z_{DR}$ and phase, plus $DSR_{hh} - DSR_{vv}$ and temperature, so the maximum value of Q in the 2D set-up is 6. Then, if the classification is made on $Q = 3$ the confidence is rather low. Similarly, if the gap between the highest and second-highest Q value is low (near 1), the confidence is also low. The value of Q might be 6, but if the second highest is 5.9, the classification is not convincing. These two indicators are shown in Figure 5.8 and Figure 5.9. Arguably, a value of $Q = 3$ is too low to classify. Namely, this could mean a classification is based solely on phase and temperature. There are some values showing a Q of 5, and a gap of around 1. These classifications are then quite plausible.

So, these results suggest that the 2D set-up performs better when applied to cloud radar data as well. The most ice particle types most occurring according to the classification are aggregates and branched planars. With a Q value of around 4 and a Q -gap value often below 1, the results are plausible, but not conclusive.

5.3.2. Classification results Event A 18:30

The change of DWR over time was briefly mentioned in subsection 4.3.1. In Figure 4.10 an increase in DWR can be seen over time, and a decrease in Z_{DR} . As such, it may be interesting to have a look at the

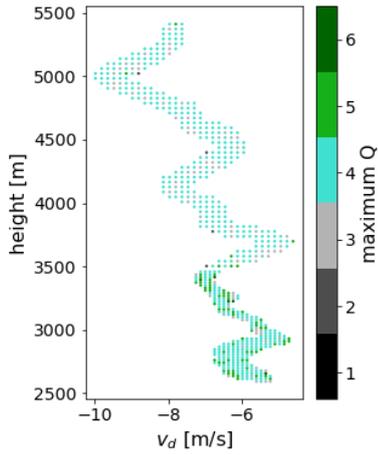


Figure 5.8: Q at 18:00.

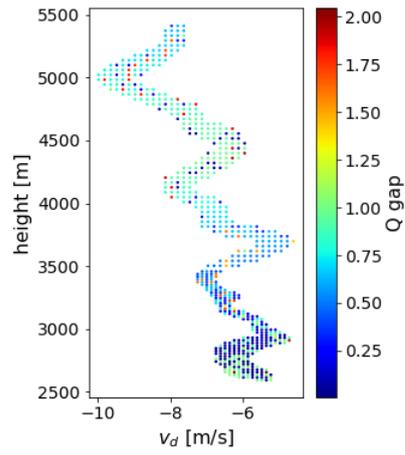


Figure 5.9: Q-gap at 18:00.

results for 18:30 as well. The classification results are shown in Figure 5.10a.

Unexpectedly, some more plates show up here, while larger particles were expected. There is not much different between the classification profile of 18:00 and 18:30, even though the bulk plots in Figure 4.10 show quite some different values. Figure Figure 5.10c shows that indeed, the DSR values at 18:30 tend to be a bit higher than at 18:00, with clear gradient across the Doppler velocity. However, this is not reflected in the classification results. This can be explained by referring back to the 2D membership functions in Figure 4.6. Smaller particles are associated with lower DSR, but these values are also part of the membership functions of aggregates and branched planars.

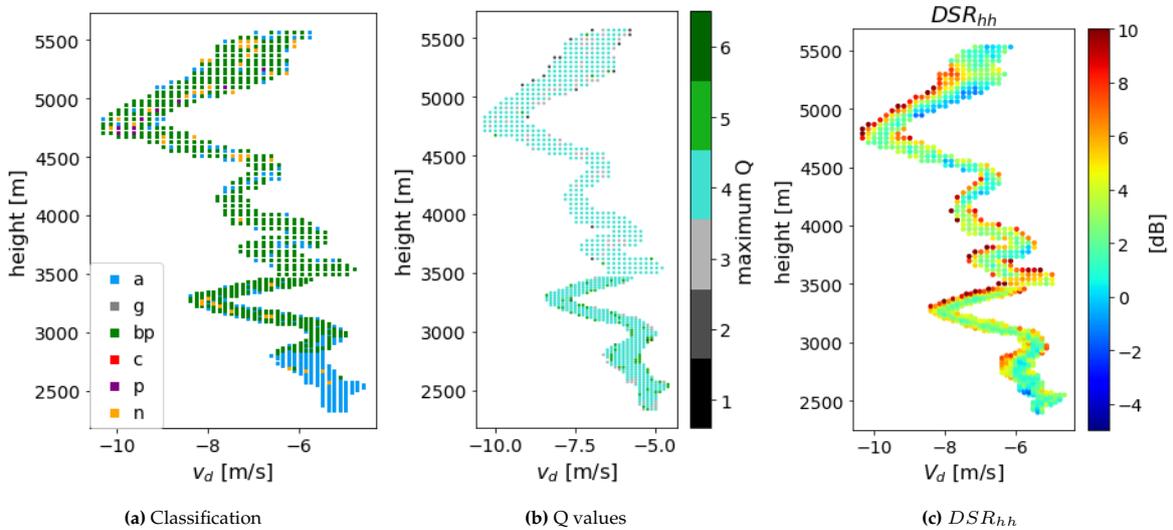


Figure 5.10: Classification outputs for Event A, 18:30

5.3.3. Classification results on Event B

Event B can also provide some insights into the performance of the classification. The most important difference in this event is that it is a higher cloud, see Figure 4.3.1. In Figure 5.11, note the confidence in which we can say that the bottom kilometre of the cloud consists of aggregates: Q is 6 and the gap is near 2. With increasing altitude, the Q value and Q -gap both decrease, thus decreasing the confidence. The SNR only decreases significantly after 7000[m], so it does not influence these results.

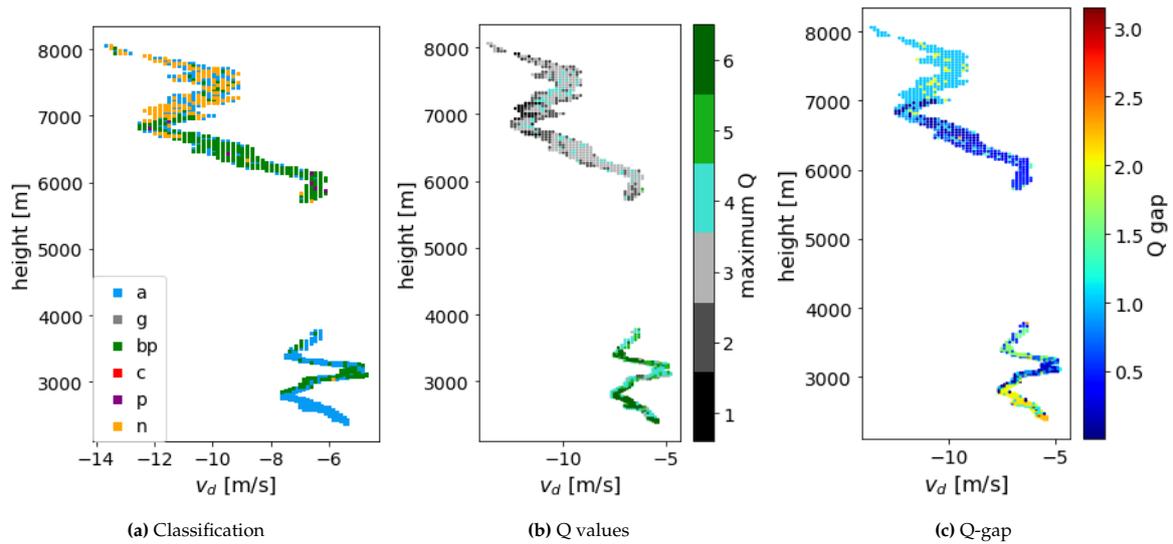


Figure 5.11: Classification outputs for Event B, 20:00

5.4. Sensitivity to different variables

The third subquestion in the research questions concerns the importance of different variables in the classification of ice particles. The hypothesis that temperature is important in the distinction between aggregates and branched planars from subsection 5.2.1 is tested. Then, the possibility of only using the 94GHz band is in subsection 5.4.2. Also the influence of differential backscattering phase is investigated in subsection 5.4.3. The findings are summarised in a table at the end of this section.

5.4.1. Sensitivity to temperature

The smooth transition of the Q -gap with height in Figure 5.9 is striking. It can be explained by looking at the membership function of temperature, Figure 4.7 and raises the question of how big the influence of temperature is in the classifier. To that end, the classification results without temperature are shown in Figure 5.12. This result shows that most classifications of branched planars are based on temperature.

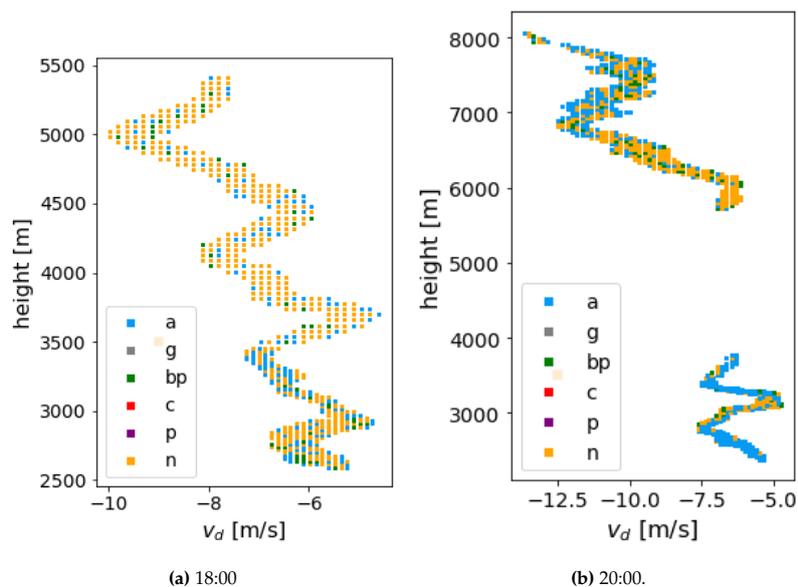


Figure 5.12: Classification results without the inclusion of temperature.

Recalling Figure 5.5 this is not really surprising. Already there, it was argued that aggregates and

branched planars had a large overlap even when including 2D membership functions, and temperature needs to be taken into account. For the classifier this means that reliable temperature measurements are imperative. On top of that, the inclusion of temperature in the classifier is based on the assumption that the cloud is vertically stable. If this is not the case, the inclusion of temperature can seriously skew the results. The effect of leaving out temperature are the same for the cases at 18:30 and 20:00, where aggregates are still mostly classified as aggregates, but what were previously branched planars are now either branched planars or aggregates. In short, the impression that temperature has a big influence on the classification results is correct.

5.4.2. Classification with only Ka-band

When leaving out the W-band in the classifier, thus being left with 94GHz, the results are as Figure 5.13. The point density of the 94GHz classifications is much higher than the other ones depicted in this chapter, as the 94GHz spectra do not need to be interpolated before using. Almost all data points score

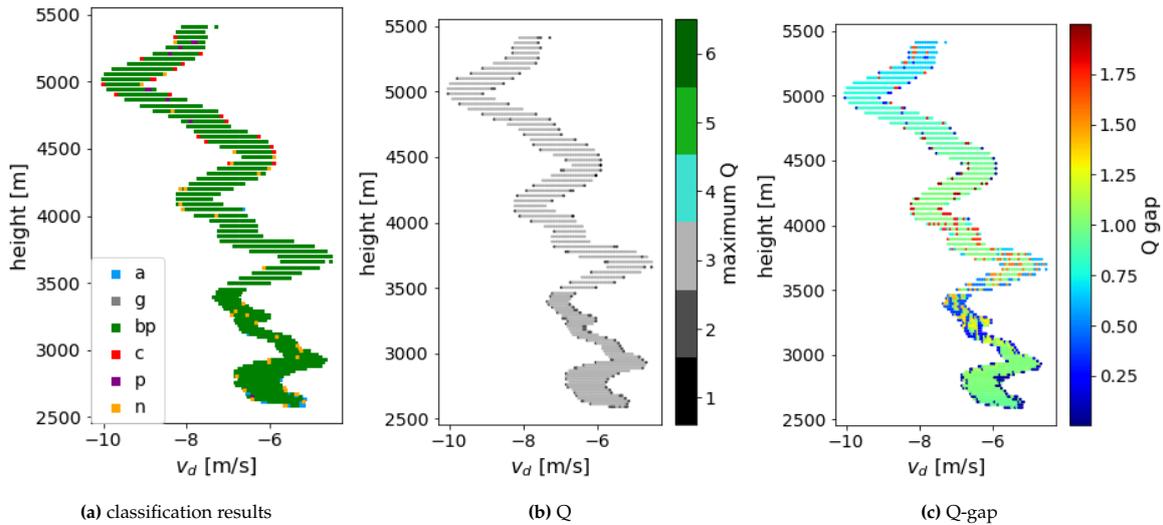


Figure 5.13: Classification outputs with only the 94GHz band.

full points on the classification (3 for $SLDR-Z_{DR}$ phase, and temperature). The Q -gap graph suggests that the classification is solely based upon temperature, and since the temperature is lower than -10°C , almost the entire spectrum gets classified as branched planar. An interesting observation is that some columns and plates show up at the edges. This would be as expected, with small particles at the edge of the spectrum, were it not that these points coincide with low Q -values. Indeed, when cross-checking with the output P matrix, they just fall outside some variable for all particle types, and match better with the temperature function of columns. Realistically, this classification is not useful.

5.4.3. Sensitivity to phase

One of the research questions to answer was which spectral polarimetric variables were key in distinguishing particle types. Of the polarimetric variables used, the phase is the only one applied in a 1D function. To that end, its influence might be minimal, and thus this might be worth investigating. Namely, leaving out variables where possible simplifies the classifier, and is a good indicator of where improvements in measurements are most useful.

Comparing Figure 5.14 with the Figure 5.7b results, one can see that without the phase, the points that get classified as aggregates are almost all swapped for branched planars and plates. Comparing Figure 5.14b and Figure 5.14c, mostly the 35GHz backscattering phase has this influence. This is explained by looking at the membership functions of ϕ_{bs} in Figure 4.5. In short, the backscattering phase does add to the classifier, and cannot be left out.

The differences made by leaving out variables are quantified in Table 5.1. These results confirm what was shown in the single timestamps above. Interestingly, the omission of the 94GHz backscattering phase seems to have minimal impact on the results. To understand this, Figure 5.3 can be consulted.

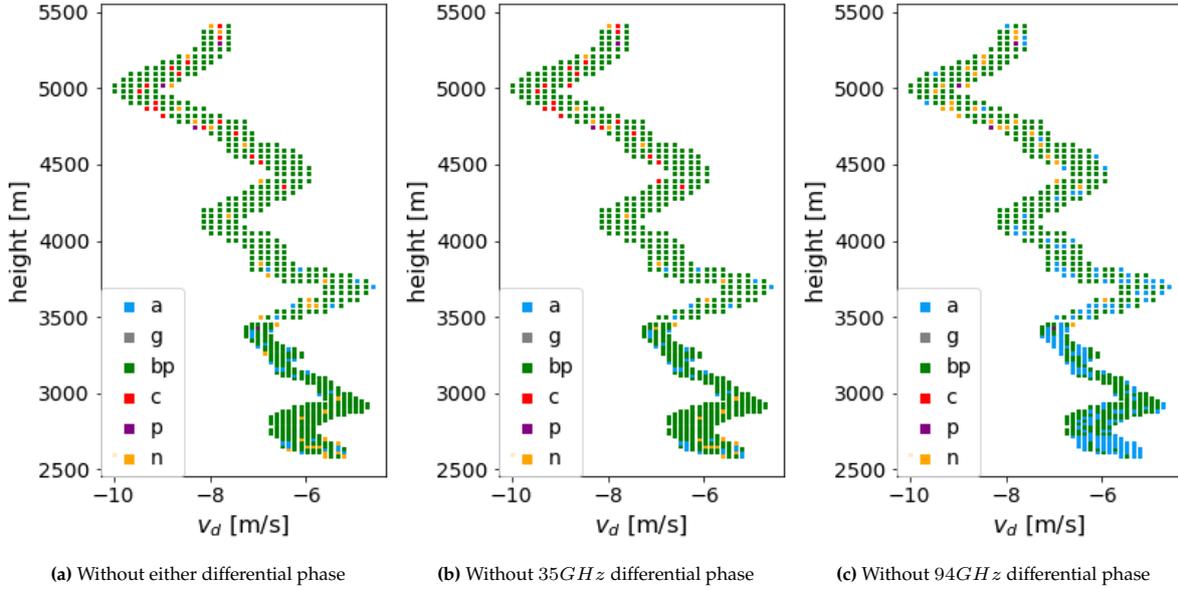


Figure 5.14: Classification results without $\phi_{bs,35}$ and/or $\phi_{bs,94}$.

The exact numbers of the classification with only 94GHz cannot be compared one-on-one, because there are many more data points since the spectrum is not interpolated to that of 35GHz.

	Set-up 3	No T	No $\phi_{bs,35}$	No $\phi_{bs,94}$	Only 94GHz
Branched planars [%]	64	5	87	64	72
Aggregates [%]	30	25	5	29	15
'None' [%]	5	70	5	6	10
Plates [%]	1	0	1	1	3
Columns [%]	0	0	3	0	1
Graupel [%]	0	0	0	0	0

Table 5.1: Percentage of types in output of different classification set-ups for Event A.

5.5. Bulk results

So far, the results are only analysed by looking at the spectrograms at specific timestamps. Looking at them for the entire event can provide more useful information. For Event A this is done for the occurrences of the classification of aggregates and branched planar. The brightness is determined as follows: for every height and time, there is a list of classifications on the Doppler velocity spectrum. In the figure, the percentage of the spectrum that is classified as that particle is depicted. For 18:00 to 19:00 (Event A) this looks like Figure 5.15.

Two things can be noted from this. First is the distinct line between aggregates and branched planar. This line perfectly coincides with the temperature line of $-10^{\circ}C$. Also, there is some structure corresponding to the polarimetric variables. The clusters of aggregates above 3000m coincide with the more negative Z_{DR} values. Only aggregates and branched planars are depicted because the other types are not occurring frequently enough to really show anything. However, it is interesting to note that most measurements classified as plates occur higher than the temperature line of $-22^{\circ}C$. In general, the confidence in the classification is on the low side, with a Q value of around 4 and a Q -gap value generally below 1.

The 20:00 (Event B) case is a bit more interesting to look at over time. All particle types are combined into the profile as in Figure 5.17. Again, there are two distinct lines: these are the temperature borders of $-10^{\circ}C$ at about 3000m and $-40^{\circ}C$ at about 7000m. Also, the structure of the classification corresponds to the bulk variables. Lighter patches do not necessarily mean bad classification results. Rather, it

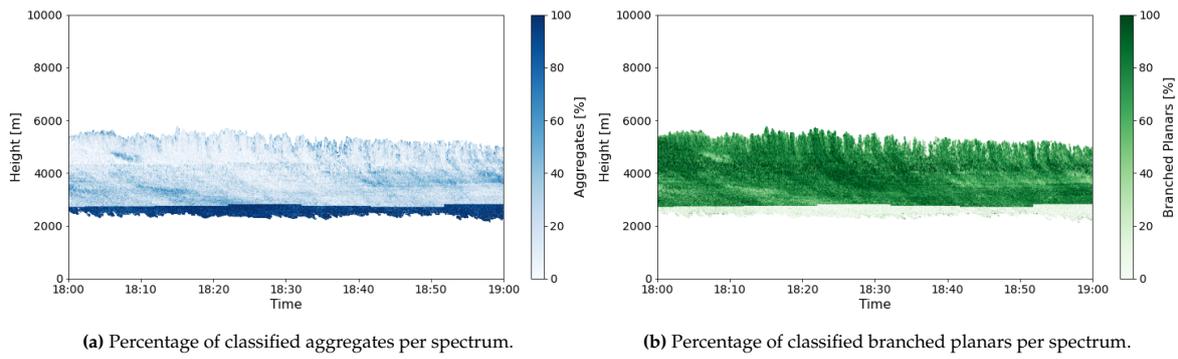


Figure 5.15: Profiles of classification results of Event A.

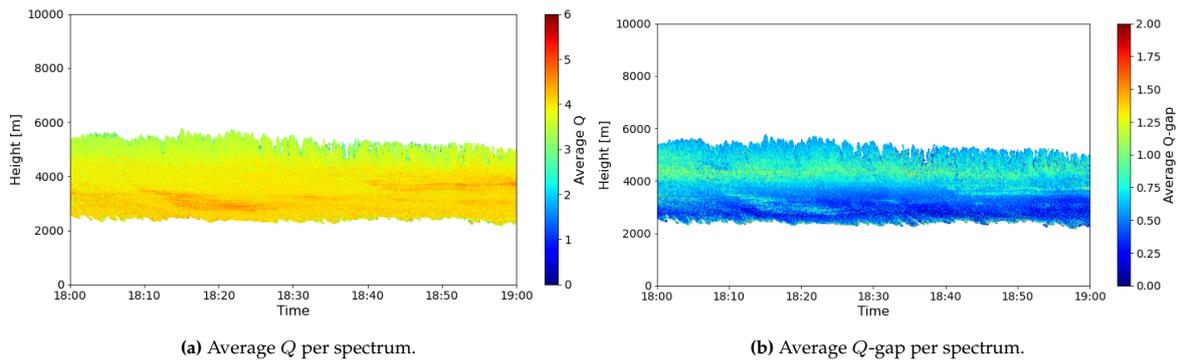


Figure 5.16: Profiles of classification quality metrics of Event A.

means that there is more than one type of particle within the Doppler spectrum. To that end, the average of the Q of the spectra is depicted in Figure 5.18e. As in the Figure 5.11b, the Q decreases with height. In Figure 5.18d and Figure 5.18c the occurrence of some columns and plates is observed as well. However, note the colourbar here, which goes up to 10 and 30 % respectively. In Figure 5.18f there is a strong line at around 7000[m]. This is not an indication of a strong classification. Rather, many points are classified as 'none' here, which would in fact mean the Q -gap = 0. However, in the code used, the

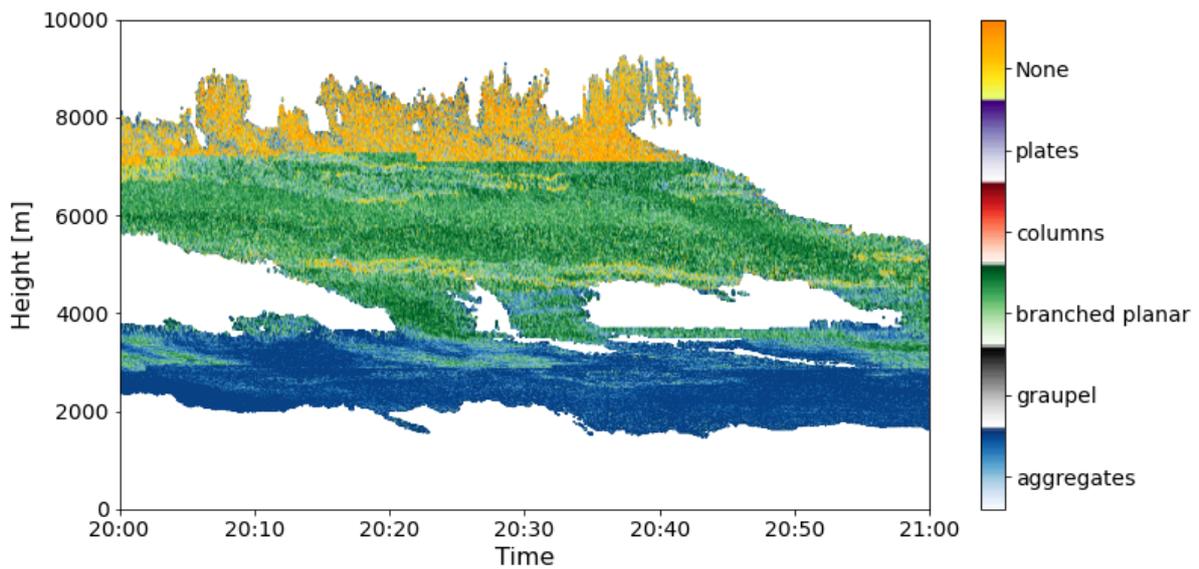


Figure 5.17: Combined average classifications per spectrum for Event B.

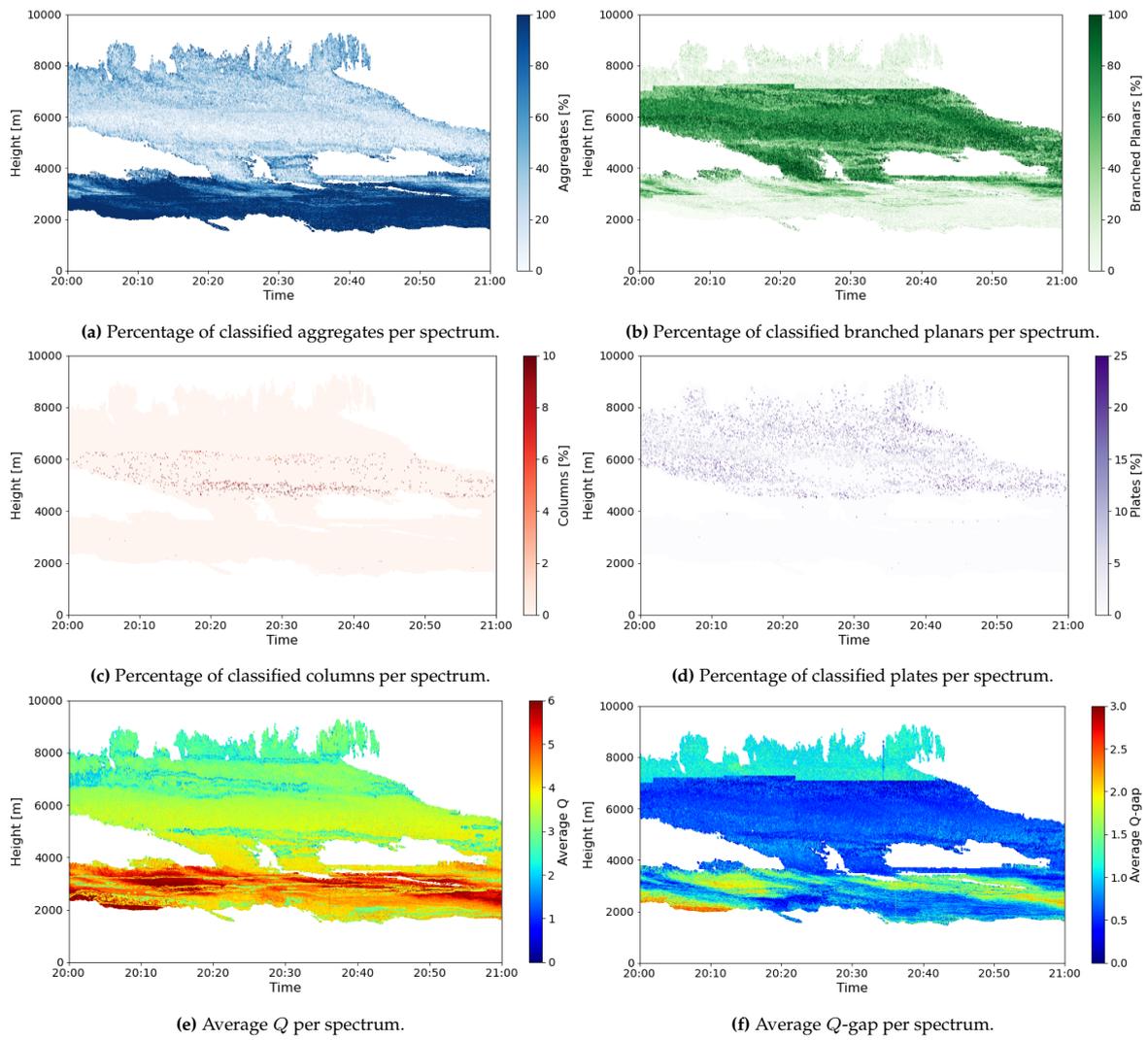


Figure 5.18: Profiles of classification results of Event B. Note the different scales for columns and plates.

difference between the highest and the second highest is shown. So, if the Q values are $[5, 5, 2, 1, 1.5]$, the classification is 'none' as two are the same, but Q -gap = 3 as the gap to the Q that is not maximum is 3.

These bulk plots serve as a check for the classifier. Similar to the spectral results, the expectation of a correct classifier is to show some structure in the cloud, as opposed to a completely white noise-like classification. It makes sense for the top of the cloud to contain smaller particles, and then get bigger as they fall and clump together to form aggregates, the growing processing described in Wang [61].

6

Discussion

In the previous chapters, the steps to answer the research questions have been taken. In this chapter, the degree to which they can now be answered will be discussed, as well as some of the choices made along the way and their implications. First, some observations of relating cloud radar data to the scattering database are discussed. Then, some of the choices made during the design of the fuzzy logic classifier are explained. Furthermore, the quality of the results is discussed after the quality metrics are explained in detail. Following is an analysis on the influence of the different variables during classification. Finally, some limitations of the methodology are mentioned.

6.1. How do the data from the scattering database and the cloud radar compare?

As introduced in chapter 4, the polarimetric variables retrieved from the scattering database and the cloud radar data are compared, to ensure some of the assumptions made in pre-processing are valid. The correspondence between the variables is key for the effectiveness of the classifier. As could be seen in Figure 5.2, generally the spectral polarimetric variables from both sources compare well. However, there are a lot of values that fall outside the values obtained from the scattering database. This can be due to three things: the method to retrieve the variables is wrong, the data is noisy, or the scattering database is not an inclusive enough representation of what is found in the cloud.

If there were an error in the retrieval method, the expectation is that there would be an offset in values between the database and the cloud radar. Because the centre of the values is approximately the same, this is not considered the most likely. In regard to the noise in the data, an SNR of 0 was applied, because when using fuzzy logic, noisy variables falling outside membership functions will not have influence on the classification. However, interestingly, cloud radar SLDR values that fall outside the scattering database ranges are associated with the highest SNR. To investigate the cause of this, comparison with literature can be done to see if this is caused by the methodology or because of a limitation of the scattering database.

Most likely, the cloud radar values falling outside those of the scattering database is caused by the limited amount of particles represented. This misrepresentation can be divided in three different kinds. First, not all variants of the available ice particle types are in the dataset. A method to include these can be to widen the membership functions. This needs careful consideration, as this will also increase overlap between membership functions. Literature or different databases can help here. Second, not all shapes of ice particles are represented in the database. Notable types of particles not in the scattering database are rimed particles or bullet rosettes. Last, the particles may be in transition between two different particles, like a plate growing into a branched planar. However, if this is the case, it is expected that its corresponding values of polarimetric variables would be in between the values of the two types, and not falling outside the overall range of the scattering database. In subsection 7.3.3 this is addressed again.

6.2. Which choices were made when setting up the fuzzy logic classifier?

Throughout the setting up of the fuzzy logic classifier, several choices have been made that had a big impact on the performance of the classification. Examples are the shape of the membership functions, the combination of variables in 2D functions, and the weights. In this section, these choices and their impact are discussed.

When constructing the membership functions, the particle type group sizes appear as a possible problem. This problem was first mentioned when setting up the 1D membership functions in subsection 4.2.1, where it seemed unnecessary to apply a trapezoidal function for the plates and columns, whereas this was absolutely necessary for the three other types. Since the 1D method was not finally chosen due to too much overlapping of the 1D membership functions, this problem was not further addressed in the making of 1D set-up. However, for the 2D membership functions the problem becomes apparent again. In the 2D membership functions Figure 4.6 there is some gap between the areas of, for example, columns around $SLDR_{94} = -15dB$, $Z_{DR,94} = 5dB$. Intuitively, this empty space would be part of the membership function as well. At the same time, there are parts of the membership functions that should perhaps not be part of it, such as the far positive Z_{DR} of graupel. One strength of using fuzzy logic is that these kinds of adjustments are fully within possibilities. A danger of this is that the adjustments are made too much on what might seem good based on literature or expected results.

The first problem, concerning the relatively small amount of plate and column particles, is not uncommon in machine learning. There are some common suggestions to improve such an imbalanced dataset. The first suggestions concern undersampling the larger datasets, which in this context means taking fewer particles as input for the larger datasets (so aggregates, branched planars, and graupel). However, in this case, this is not beneficial, as the spread of these larger datasets will still encompass the spread of the smaller ones. On top of that, an unfavourable effect is that the membership functions of the larger groups will be smaller and thus exclude a lot of valid measurements. Similarly, generating new data for the smaller sets will not add a lot to the classifier, as this will not decrease the overlap between the types. It might, however, close the gap in the 2D membership functions mentioned above.

One possible solution, other than filling in the gaps of the membership functions, is changing the way the shape of the 2D membership function is obtained. The current method was chosen for its simplicity, but slightly more complicated membership functions can be made that address some of the problems. For example, shaping the membership functions as a concave hull around the scattering points.

Right now, the 2D membership functions are not fully using the possibilities of fuzzy logic, because they do not contain values that are not 1 or 0. By adding some gradient as in the diagonal part of the trapezoidal membership functions, more nuance can be given to the border of the membership functions. A smooth transition between 0 and 1 can also decrease the total area under the membership functions of the ice particle types with a large number of points, while increasing the area of the membership functions of plates and columns.

One other aspect that has not been used is the changing of the weights. Again, this was done to not overcomplicate the classifier before obtaining results. Now, there are two instances where changing the weights could make sense. First, is with the inclusion of temperature. As later argued as well, the temperature has a huge impact on the result of the classification. Reducing the weight can reduce its impact, and give the polarimetric variables relatively more weight. Another consideration could be the weight of the phase functions. With the 3rd set-up, the maximum value of Q is 6. In this set-up, the phase was included as two functions, accounting to $1/3^{\text{rd}}$ of the maximum value of Q . However, the two variables of DSR_{hh} and DSR_{vv} can only amount to $1/6^{\text{th}}$. By decreasing the weight of the phase functions to $W = 0.5$, this imbalance would disappear.

6.3. How is the quality of the classification results assessed?

One known issue of hydrometeor classification in general is the lack of true data to validate the results. Because of this, the events chosen were ones that we already had some knowledge about, so it can be checked that the results were in line with the expectations. In a way, there is no real way to know that the results are correct. However, there are a few ways to know that the results are incorrect or

highly unlikely. These concern the consistency of the results, and the fuzziness in the output of the classification.

An expectation, for example, is that you find smaller particles higher up, where the temperature is lower as well as the DWR . A higher Z_{DR} should indicate more oblate particles. Lower in the cloud, there should be more aggregates as the temperature increases there, which encourages aggregation, and where particles have had the time to aggregate while falling. Consistent results are results that do not look arbitrary. A classification should be somewhat coherent with the points next to it and the input spectral polarimetric variables. An inconsistent classification would perhaps look random, like white noise, or show patterns different from those in the cloud radar data.

Fuzzy logic allows the quantification of confidence in the results in two ways: by using Q and Q -gap. Ideally, the Q is maximum at every classification, which is 6 for the normal 2D set-up. This number indicates to what degree the data point corresponds to the membership functions, and thus the scattering database, of the classified particle. When using a similar set-up as this in applications, there should be a minimum value of Q that is deemed good enough for classification. Because this study consists of a first analysis of the proposed methodology, it was not yet useful to set such a limit. The results as they are allow broad analysis, and should be used to build upon and improve, rather than take as true. Assuming that the method is correct, low Q values can mean a couple of things. First, that the measurement is noisy. If values with low Q -value get filtered out in the end result, this is not an issue. Second, it could mean that the measured particles are not part of the 5 types as defined in the scattering database. This is an expected problem when basing classification on a database: simply not every particle can be represented. Last, it could mean that the membership functions are not defined properly. This is discussed in section 6.2 as well, and can be avoided by carefully crafting the membership functions if it seems that this is the case.

The so-dubbed Q -gap value is also an indicator for confidence. As explained, when this value is high, the classification is quite certain for that type, and the chance that it is another type is low. A high value would be around 2 in the 2D case. When this value is very low, it is an indication that the classification is not certain. A small change in any variable could lead to a whole different classification.

Another difficulty in quantifying the confidence in the results is that an ambiguous classification could be correct. The way the classifier is set up, with the use of spectral polarimetry and a scattering database, relies on two assumptions. First, that all particles within one Doppler velocity bin are the same. Second, that the measured particles are similar to those in the database.

A low Q -gap value could mean that there are two or more types of particles within the Doppler velocity bin. Or, that the particles in question are somewhere between two particle types, like a plate that is growing to become a branched planar. In this last case, the Q value should be high for both. This is what is seen in the lower part of the cloud in Event B, where the Q -gap is near 0 and $Q > 5$.

An advantage of using Q values as opposed to expectation and consistency as a metric for quality is that it is more easily quantifiable and applicable to different events. Consistency in the results could be quantified, by for example taking into account the classifications of the data points around a point.

The three ways to indicate the confidence in the classification results have been applied to the events in section 5.3. As mentioned before, the expectation was that the cloud would consist mostly of branched planars and aggregates. For both events, this is indeed the case. For Event B, smaller particles were expected higher up. Indeed, the plates that were found, were found higher up in the cloud. In Event A, generally, the Q value is around 4, which is satisfactory, but is not convincing. Combined with the Q -gap which is below 1 nearly everywhere, the classification is not so confident. The low Q -gap is not strange given the large amount of overlap between branched planars and aggregates. As mentioned, Q -gap in Event A is mainly determined by temperature, and as such, so is the result of the classification. In Event B, the Q does reach up to 6, and there are streaks with a Q -gap of around 2, which is as good as possibly expected. However, all other areas show low Q -gap values of below 1.

For both these events, it is interesting to note which variables are most often within the membership functions. In a way, this is done by looking at the different variables, discussed in section 6.4. Especially in Event B the stark decrease in Q with height is quite interesting. Recommendations about further analysis of the results can be found in subsection 7.3.3.

6.4. Importance of different variables

The last subquestion formulated for this research concerns the importance of the different variables used in the classifier. This can help the future ice particle classifiers, and provide more insight into measuring ice particles with radar as a whole. To this end, three set-ups have been tested. Firstly, without phase because it is the only 1D variable in the 2D set-up. Secondly, without temperature, because it seemed the classifier depended a lot on it. Finally, with only W-Band because many of the cloud radars around Europe only have this frequency.

The results of these set-ups are quite clear. Leaving only the 94GHz variables leaves little to base the classification on. Especially the loss of the DSR relation makes it such that almost any particle also gets classified as branched planar. Plates and columns are now completely indiscernible by spectral polarimetric variables only. At this point in ice particle classification, it seems more variables are needed, and leaving out the valuable information of 35GHz is detrimental for the quality of the classification. Nevertheless, using variables that do not add any information only slows down the algorithm, costs time to set up, and possibly introduces errors. To that end, the results without ϕ_{dp} have been investigated as well. From Figure 5.14 it is shown that the phase also adds information. In this specific spectra, it can be seen that phase adds to differentiate between branched planar and aggregates. and fewer points get classified as plates. Especially the 35GHz contains information that helps drive the classification, which is no surprise looking at the 1D membership functions in Figure 4.5. In short, all the spectral polarimetric variables investigated in this research need to be considered for the ice particle classifier.

Concerning the use of spectral measurements, the hope was that the different particles within a resolution volume would be spread out along the Doppler spectrum. Indeed, this seems to be happening in the cases described by looking at the spectral DSR and SLDR. However, in the classification results this is not reflected. When data points are classified as columns or plates, this is more often rather random than on the side of the spectrum. Looking at the results at 18:00 and 18:30, a difference in DSR over the spectrum does not result in different classifications. Admittedly, whereas the DSR shows a stark difference over the spectrum, the other variables show much less variability, let alone as structured as the DSR does. Perhaps, the size changes over the spectrum, but the particle type does not. The results from the database show that also the 'larger' particles can have DSR values nearing 0dB.

From the results, it is clear that there is a distinct dependence on temperature in the classification. This dependence showcases itself in the Q -gap following the temperature gradient in Figure 5.9, the stark line in the bulk classification results in Figure 5.15a, and of course the results without the inclusion of temperature in Figure 5.12.

Testing the dependence of LWP in the classification has not been done. However, from Figure 5.5 we know that about 50% of the aggregate particles in the database get misclassified as graupel. On top of that, there is no temperature range for graupel. What is expected is that when the LWP is sufficiently high, a lot of the aggregates will get misclassified as graupel.

6.5. Limitations

Assumptions in the methodology mean that this classifier can only be applied to non-precipitating ice clouds, with little convection. On top of that, the radar used needs to be dual-frequency, dual-polarized cloud radar, limiting the locations where the classification can be applied. The addition of temperature in the classification means that accurate temperature measurements are crucial for accurate classification.

Another limitation in this study is that only two events have been analysed. From the way the classifier is set up, with low LWP it is expected that too many particles get classified as aggregates, as per Figure 5.5. And indeed, the results show mostly aggregates and branched planars. However, for the events chosen these were also the expected types, so just looking at the classification results one could say the classifier performs perfectly. Only by checking with an event where it is known that there are in fact smaller particles, it can be confirmed that there is a bias towards aggregates right now.

7

Conclusion

This chapter summarizes the findings presented in the previous chapters. First, some conclusions are given for each research question. This is followed by an overarching conclusion to answer the main research question. Finally, some recommendations are given for potential improvements.

7.1. Conclusions on sub-questions

To relate the cloud radar measurements to a scattering database, several spectral polarimetric variables were derived. The spectral polarimetric variables, Z_{DR} , $SLDR$, and ϕ_{bs} , for both 94GHz and 35GHz, and DSR_{hh} and DSR_{vv} , generally compare well. This means that the pre-processing steps and assumptions are likely to be valid. Comparison results show that the spread of the measurements was slightly larger than that of the scattering database over the different variables. This is most likely because the size of the scattering database is limited: not all types of particles are represented, and it is likely that the particles found in the cloud are more diverse than those in the scattering database. To account for this, changes need to be made to the functions used in the classifier. Spectral profiles of the cloud radar variables show a gradient in the DSR and the SLDR over the Doppler spectrum, suggesting a difference in particles within one resolution volume.

The polarimetric variables derived from the scattering database were used as a base for the fuzzy logic classifier. In this approach, the classifier was kept as simple as possible. At first, a 1D set-up was constructed with trapezoidal membership functions. However, the five ice particle types (aggregates, branched planars, graupel, columns, and plates). showed a significant overlap. Therefore, variables were combined in 2D membership to provide more information to base the classification on. The combinations chosen were $SLDR - Z_{DR}$ for 94 and 35 GHz, and $DSR_{hh} - DSR_{vv}$. Testing this on the database to check differentiability showed that the 2D method indeed performed better. However, in order to make sure particles do not get misclassified as aggregates, more information was needed. To that end, temperature and liquid water path are added as variables in the classifier.

The method adapted to obtain the membership functions meant there are some counterintuitive gaps in the 2D versions, and the functions do not extend past the values from the scattering database. Specifically, this could be a problem for the plates and columns, since are relatively underrepresented in the scattering database. Additionally, the 2D membership functions only consist of 0 and 1, which does not make full use of the possibilities of including fuzziness in a fuzzy logic classifier.

One of the main challenges in this study was the lack of a proper validation dataset to assess the performance of the classifier. The two case studies provided some useful qualitative insight. However, since the true particle types remain unknown, no quantitative performance could be performed.

Instead, four other metrics were used to assess the quality of the results. The first was comparing the results with prior expectations. Based on a previous study and prior knowledge of the influence of shape on the bulk variables, it was found that the ice clouds should mainly consist of aggregates and branched planars, with in Event B some smaller particles at higher altitudes. The second quality check concerned the consistency of the results. That is, completely random results, or artificial looking lines suggest that

the classification is improbable. Temperature appears to critically influence the classification. Indeed, the results showed a sharp line between branched planars and aggregates in the bulk profile results as a function of temperature. This is problematic and suggest that the classifier is too dependent on temperature compared to other variables. Besides these, the fuzzy logic classifier makes it possible to compute two additional metrics. One of these is Q , which indicates how much the measured variables match the modelled variables per particle type. A low Q indicates low confidence in the classification. A large difference between the highest Q and the second-highest Q values, the so-called Q -gap, means the classification is more certain. In this study, the Q -gap values highlighted a low confidence in the classifications. However, in Event B, there were some areas with particularly high Q and Q -gap values, implying a strong classification.

By analysing the different spectral polarimetric variables, the understanding of the polarimetric characteristics of ice particles can be better understood. Excluding variables showed significantly different results. For example, not including $\phi_{bs,35}$ decreases the amount of points classified as aggregates from 30 to 5%. Again, the dependence on temperature was highlighted, as leaving it out showed that the classification was dependent on temperature for 65% of all points.

While many cloud radars only have the 94GHz band, classifying with only this band resulted in too few variables to provide a meaningful classification. Results show that in its current set-up, the classifier is dependent on all variables in the classifier.

7.2. Overall conclusion

The main research question of this thesis is: *How can ice particles be classified using the combination of spectral polarimetry in mm-wavelengths and a DDA/GMM scattering database using fuzzy logic??*. The novelty of this work lies in the use of spectral measurements to potentially identify different particle types within one resolution volume. A key assumption was that each Doppler velocity bin contains only one particle type. Though the results showed differences in the polarimetric variables across the Doppler spectrum, the classification results did not reflect this. Thus, there is potential in using spectral data to classify ice particles, but the current classifier is insufficient. The scattering database provided a strong foundation for the set-up of the classifier, but its limitations in particle representation were evident. Overall, this thesis demonstrates that the combination of spectral polarimetric variables and a scattering database can form a basis of an ice particle classifier that can identify different particles within one resolution volume.

7.3. Recommendations

Though the findings in this thesis show that there is potential in classifying ice particles using the current methodology, there are many ways in which improvements can be made. In chapter 6 some possible changes have been mentioned, as well as limitations of the methodology. In this section, the changes that are deemed most important are mentioned, as well as some improvements that could prove especially valuable.

7.3.1. New events

One of the future work recommendations is to apply and analyse the classifier to new events. Additional events can give better understanding of the performance of the classifier. An event with higher LWP to see its influence on the classifier would be interesting. Also, ideally an event where it is known or suspected that there are smaller, pristine ice crystals would be beneficial. Besides giving more insight into the classifier, additional events also make sure that the classifier is not designed to the expectations of two events used now.

7.3.2. Adjustments to the classifier

As mentioned in section 6.2 the 2D membership functions should be adjusted to not contain gaps. Including a gradual decrease from 1 to 0 in the 2D membership functions can make the classifier more robust, and widen the functions to include some of the closer cloud radar data points that now fall outside these functions. A first suggestion would be to have the function decrease linearly within 5 [dB] in all directions. When

Looking at the 2D membership functions in Figure 4.6 and the results in general, perhaps the variables used are not enough to differentiate different particles. An obvious solution would be to add an extra variable. Some polarimetric variables are not considered, like the circular depolarization ratio, or the co-polar correlation coefficient. An explanation of them and reasons why they are not used are in Appendix A. Though it might not be as straightforward as the other spectral polarimetric variables, it can be investigated how these variables can add some information to the classifier nonetheless. Perhaps, similarly to temperature, relative humidity can be added as a variable, since it also has an influence on the shape of the particles formed. Just like temperature, this depends on values found in literature, and the cloud cannot have much vertical mixing. Briefly mentioned before, an additional way to add information is by using the Doppler decomposition to retrieve mass or density. For that, accurate wind measurements need to be available, with vertical wind components as well.

Though the use of principal component analysis was disregarded in Appendix B, there are some other ways the current variables could be adjusted. One way could be by transforming the variables such that the different classes are better separable. This could be especially beneficial in the 2D membership functions. On top of that, the combination of different variables in the 2D membership functions could be optimised, or variables could appear in multiple functions, like in Al-Sakka et al. [1].

7.3.3. Analysis of results

The current results should be analysed more thoroughly in order to gain more understanding of the way the classifier currently performs. Examples of this are to find out which percentage of measurements fall outside the membership functions, and if these are the same data points for all functions. Or, what the leading variable is that determines the decrease in Q with height. By retrieving the P -matrix in the classifier, this should not be complicated, but can highlight the influence of different variables. On top of that, this could show changes made to the membership functions with more detail.

When working towards a final product, a metric could be formed to show total confidence in the result using a combination of Q , Q -gap and consistency (vertically and in time). Perhaps, when a particle has a high Q value for two particle types, the result could give both as an output. Something like 'column-aggregate' could indicate that the particle found might be in between a column and an aggregate. However, this kind of results should be backed by some study to check if the scattering properties of, for example, a column slowly growing into an aggregate does indeed resemble the scattering properties of both.

7.3.4. Comparison with other classifiers

Lack of validation data makes it more complex for the design of the classification system to be iterative. However, there are other hydrometeor classifiers that are validated, and can be used for comparison. Though these classifiers might not contain spectral data, some aspects can be used. For example, the main particle type over a spectrum. Reference classifiers should be made for Ka or W-band radar with slanted elevation angle, and discern between at least 2 types of ice crystals, for example Maahn and Löhnert [32] and Matrosov et al. [38].

Bibliography

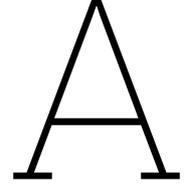
- [1] H. Al-Sakka, A. A. Boumahmoud, B. Fradon, S. J. Frasier, and P. Tabary. A new fuzzy logic hydrometeor classification scheme applied to the french x-, c-, and s-band polarimetric radars. *Journal of Applied Meteorology and Climatology*, 52:2328–2344, 10 2013. ISSN 1558-8424. doi: 10.1175/JAMC-D-12-0236.1.
- [2] K. Aydin and J. Singh. Cloud ice crystal classification using a 95-ghz polarimetric radar. *Journal of Atmospheric and Oceanic Technology*, 21:1679–1688, 11 2004. ISSN 0739-0572. doi: 10.1175/JTECH1671.1.
- [3] K. Aydin and T. M. Walsh. Separation of millimeter-wave radar reflectivities of aggregates and pristine ice crystals in a cloud. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 1: 440–442, 1998. doi: 10.1109/IGARSS.1998.702932.
- [4] A. J. Baran. From the single-scattering properties of ice crystals to climate prediction: A way forward. *Atmospheric Research*, pages 45–69, 2012. doi: 10.1016/j.atmosres.2012.04.010.
- [5] A. Battaglia, S. Tanelli, S. Kobayashi, D. Zrnica, R. J. Hogan, and C. Simmer. Multiple-scattering in radar systems: A review. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 111(6):917–947, 2010. ISSN 0022-4073. doi: <https://doi.org/10.1016/j.jqsrt.2009.11.024>.
- [6] A. Battaglia, S. Tanelli, G. M. Heymsfield, and L. Tian. The dual wavelength ratio knee: A signature of multiple scattering in airborne ku-ka observations. *Journal of Applied Meteorology and Climatology*, 53:1790–1808, 7 2014. ISSN 1558-8424. doi: 10.1175/JAMC-D-13-0341.1.
- [7] A.-C. Billault-Roux. *Snowfall microphysics: a dual-frequency and Doppler spectral radar perspective*. PhD thesis, EPFL, 2023.
- [8] V. N. Bringi and V. Chandrasekar. *Polarimetric Doppler Weather Radar: Principles and Applications*. Cambridge University Press, 2001. doi: 10.1017/cbo9780511541094.
- [9] V. Chandrasekar, R. Keränen, S. Lim, and D. Moisseev. Recent advances in classification of observations from dual polarization weather radars. *Atmospheric Research*, 119:97–111, 2013. ISSN 0169-8095. doi: <https://doi.org/10.1016/j.atmosres.2011.08.014>. ADVANCES IN PRECIPITATION SCIENCE.
- [10] M. de Haij, C. Unal, E. O’Connor, and A. Apituley. Classification data from cabauw on 3 february 2022, 2024. Data is volatile and may be updated in the future.
- [11] I. del Portillo. *Itu-rpy: A python implementation of the itu-r p. recommendations to compute atmospheric attenuation in slant and horizontal paths.*, 2017. URL <https://github.com/inigodelportillo/ITU-Rpy/>.
- [12] B. Dolan and S. A. Rutledge. A theory-based hydrometeor identification algorithm for x-band polarimetric radars. *Journal of Atmospheric and Oceanic Technology*, 26(10):2071 – 2088, 2009. doi: 10.1175/2009JTECHA1208.1.
- [13] Y. Dufournet and H. W. Russchenberg. Towards the improvement of cloud microphysical retrievals using simultaneous doppler and polarimetric radar measurements. *Atmospheric Measurement Techniques*, 4:2163–2178, 2011. ISSN 18671381. doi: 10.5194/AMT-4-2163-2011.
- [14] P. Eriksson, R. Ekelund, J. Mendrok, M. Brath, O. Lemke, and S. A. Buehler. A general database of hydrometeor single scattering properties at microwave and sub-millimetre wavelengths. *Earth Syst. Sci. Data*, 10:1301–1326, 2018. doi: 10.5194/essd-10-1301-2018. een andere database?!

- [15] K. E. Fitch and T. J. Garrett. Graupel precipitating from thin arctic clouds with liquid water paths less than 50 g m^{-2} . *Geophysical Research Letters*, 49, 1 2022. ISSN 19448007. doi: 10.1029/2021GL094075.
- [16] P. Forster, T. Storelvmo, K. Armour, W. Collins, J.-L. Dufresne, D. Frame, D. Lunt, T. Mauritsen, M. Palmer, M. Watanabe, M. Wild, and H. Zhang. The earth’s energy budget, climate feedbacks, and climate sensitivity. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou, editors, *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, book section 7, pages 923–1054. Cambridge University Press, Cambridge, UK and New York, NY, USA, 2021. doi: 10.1017/9781009157896.009. URL https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_Chapter07.pdf.
- [17] C. Grasmick, B. Geerts, J. R. French, S. Haimov, and R. M. Rauber. Estimating microphysics properties in ice-dominated clouds from airborne ka-w-band dual-wavelength ratio reflectivity factor in close proximity to in situ probes. *Journal of Atmospheric and Oceanic Technology*, 39(11):1815 – 1833, 2022. doi: 10.1175/JTECH-D-21-0147.1.
- [18] W. Hailong, S. Zhao, and M. Jianli. The application of hydrometeor classification algorithm for china new generation doppler radar. *Proceedings of 2019 International Conference on Meteorology Observations, ICMO 2019*, 12 2019. doi: 10.1109/ICMO49322.2019.9026070. this one describes the classification algorithm.
- [19] M. P. Hall, J. W. Goddard, and S. M. Cherry. Identification of hydrometeors and other targets by dual-polarization radar. *Radio Science*, 19:132–140, 1984. ISSN 1944799X. doi: 10.1029/RS019I001P00132.
- [20] A. J. Heymsfield. Properties of tropical and midlatitude ice cloud particle ensembles. part i: Median mass diameters and terminal velocities. *Journal of the Atmospheric Sciences*, 60(21):2573 – 2591, 2003. doi: 10.1175/1520-0469(2003)060<2573:POTAMI>2.0.CO;2. URL https://journals.ametsoc.org/view/journals/atsc/60/21/1520-0469_2003_060_2573_potami_2.0.co_2.xml.
- [21] A. J. Heymsfield and C. D. Westbrook. Advances in the estimation of ice particle fall speeds using laboratory and field measurements. *Journal of the Atmospheric Sciences*, 67(8):2469 – 2482, 2010. doi: 10.1175/2010JAS3379.1. URL <https://journals.ametsoc.org/view/journals/atsc/67/8/2010jas3379.1.xml>.
- [22] S. Kneifel, J. Leinonen, J. Tyynelä, D. Ori, and A. Battaglia. Scattering of hydrometeors. In V. Levizzani, C. Kidd, D. B. Kirschbaum, C. D. Kummerow, K. Nakamura, and F. J. Turk, editors, *Satellite Precipitation Measurement: Volume 1*, pages 249–276. Springer International Publishing, Cham, 2020. ISBN 978-3-030-24568-9. doi: 10.1007/978-3-030-24568-9_15. URL https://doi.org/10.1007/978-3-030-24568-9_15.
- [23] G. Köcher, T. Zinner, C. Knote, E. Tetoni, F. Ewald, and M. Hagen. Evaluation of convective cloud microphysics in numerical weather prediction models with dual-wavelength polarimetric radar observations: methods and examples. *Atmospheric Measurement Techniques*, 15(4):1033–1054, 2022. doi: 10.5194/amt-15-1033-2022.
- [24] M. R. Kumjian. Weather radars. In C. Andronache, editor, *Remote Sensing of Clouds and Precipitation*, volume 1, pages 15–65. Springer International Publishing, 2018. ISBN 978-3-319-72582-6. doi: 10.1007/978-3-319-72583-3.
- [25] J. Leinonen. High-level interface to t-matrix scattering calculations: architecture, capabilities and limitations. *Optics Express*, 22:1655, 1 2014. ISSN 10944087. doi: 10.1364/oe.22.001655.
- [26] J. Leinonen and D. Moisseev. What do triple-frequency radar signatures reveal about aggregate snowflakes? *Journal of Geophysical Research: Atmospheres*, 120:229–239, 1 2015. ISSN 2169-8996. doi: 10.1002/2014JD022072.

- [27] K. G. Libbrecht. The physics of snow crystals. *Reports on Progress in Physics*, 68:855–895, 3 2005. doi: 10.1088/0034-4885/68/4/R03.
- [28] S. Lim, V. Chandrasekar, and V. N. Bringi. Hydrometeor classification system using dual-polarization radar measurements: Model improvements and in situ verification. *IEEE Transactions on Geoscience and Remote Sensing*, 43:792–801, 4 2005. ISSN 01962892. doi: 10.1109/TGRS.2004.843077.
- [29] H. Liu and V. Chandrasekar. Classification of hydrometeors based on polarimetric radar measurements: Development of fuzzy logic and neuro-fuzzy systems, and in situ verification, 2000.
- [30] S. P. Lohmeier, S. M. Sekelsky, J. M. Firda, G. A. Sadowy, and R. E. Mcintosh. Classification of particles in stratiform clouds using the 33 and 95 ghz polarimetric cloud profiling radar system (cprs). *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, 35:256–270, 3 1997.
- [31] Y. Lu, Z. Jiang, K. Aydin, J. Verlinde, E. Clothiaux, and G. Botta. A polarimetric scattering database for non-spherical ice particles at microwave wavelengths. *Atmospheric Measurement Techniques*, 9: 5119–5134, 10 2016. ISSN 18678548. doi: 10.5194/AMT-9-5119-2016.
- [32] M. Maahn and U. Löhnert. Potential of higher-order moments and slopes of the radar doppler spectrum for retrieving microphysical and kinematic properties of arctic ice clouds. *Journal of Applied Meteorology and Climatology*, 56:263–282, 2 2017. ISSN 1558-8424. doi: 10.1175/JAMC-D-16-0020.1.
- [33] L. Mak. Peering into the heart of thunderstorm clouds. Master’s thesis, Delft University of Technology, 2023. URL <https://resolver.tudelft.nl/uuid:f1a41622-b59f-42a6-bc7a-f7371038278f>.
- [34] F. S. Marzano, D. Scaranari, M. Celano, P. P. Alberoni, G. Vulpiani, and M. Montopoli. Hydrometeor classification from dual-polarized weather radar: extending fuzzy logic from s-band to c-band data. *Advances in Geosciences*, 7:109–114, 2006. doi: 10.5194/adgeo-7-109-2006.
- [35] F. S. Marzano, D. Scaranari, M. Montopoli, and G. Vulpiani. Supervised classification and estimation of hydrometeors from c-band dual-polarized radars: A bayesian approach. *IEEE Transactions on Geoscience and Remote Sensing*, 46(1):85–98, 2008. doi: 10.1109/TGRS.2007.906476.
- [36] S. Y. Matrosov. Depolarization estimates from linear h and v measurements with weather radars operating in simultaneous transmission–simultaneous receiving mode. *Journal of Atmospheric and Oceanic Technology*, 21:574–583, 2004.
- [37] S. Y. Matrosov and A. J. Heymsfield. Use of doppler radar to assess ice cloud particle fall velocity-size relations for remote sensing and climate studies. *Journal of Geophysical Research: Atmospheres*, 105(D17):22427–22436, 2000. doi: <https://doi.org/10.1029/2000JD900353>.
- [38] S. Y. Matrosov, G. G. Mace, R. Marchand, M. D. Shupe, A. G. Hallar, and I. B. Mccubbin. Observations of ice crystal habits with a scanning polarimetric w-band radar at slant linear depolarization ratio mode. *Journal of Atmospheric and Oceanic Technology*, 29:989–1008, 8 2012. ISSN 0739-0572. doi: 10.1175/JTECH-D-11-00131.1.
- [39] M. Montopoli, A. Bracci, E. Adirosi, M. Iarlori, S. Di Fabio, R. Lidori, A. Balotti, L. Baldini, and V. Rizi. Cloud and precipitation profiling radars: The first combined w- and k-band radar profiler measurements in italy. *Sensors*, 23(12), 2023. ISSN 1424-8220. doi: 10.3390/s23125524.
- [40] A. Myagkov and T. Rose. Rpg dual polarization dual frequency scanning cloud radar systems: Configurations and applications, 2022. URL https://www.radiometer-physics.de/download/PDF/Cloud%20Radar/RPG_FMCW_Brochure%202021.pdf.
- [41] E. O’Connor. Model data from cabauw on 26 january 2021, 2022. URL <https://hdl.handle.net/21.12132/1.64e2c7e77dd84130>.

- [42] H. S. Park, A. V. Ryzhkov, D. S. Zrnić, and K. E. Kim. The hydrometeor classification algorithm for the polarimetric wsr-88d: Description and application to an mcs. *Weather and Forecasting*, 24: 730–748, 6 2009. ISSN 1520-0434. doi: 10.1175/2008WAF2222205.1.
- [43] C. Praz, Y. A. Roulet, and A. Berne. Solid hydrometeor classification and riming degree estimation from pictures collected with a multi-angle snowflake camera. *Atmospheric Measurement Techniques*, 10:1335–1357, 4 2017. ISSN 18678548. doi: 10.5194/AMT-10-1335-2017.
- [44] H. Pruppacher and J. Klett. *Microphysics of Clouds and Precipitation*. Springer Dordrecht, 2 edition, 2010. ISBN 978-0-306-48100-0. doi: 10.1007/978-0-306-48100-0.
- [45] W. G. Rees. *Physical Principles of Remote Sensing*. University Press, 3 edition, 2012. ISBN 978-1-107-00473-3.
- [46] A. V. Ryzhkov and D. S. Zrnic. *Radar Polarimetry for Weather Observations*. Springer Cham, 1 edition, 3 2019. ISBN 978-3-030-05093-1.
- [47] U. Romatschke and J. Vivekanandan. Cloud and precipitation particle identification using cloud radar and lidar measurements: Retrieval technique and validation. *Earth and Space Science*, 9, 5 2022. ISSN 23335084. doi: 10.1029/2022EA002299.
- [48] T. Rose, S. Crewell, U. Löhnert, and C. Simmer. A network suitable microwave radiometer for operational monitoring of the cloudy atmosphere. *Atmospheric Research*, 75:183–200, 5 2005. ISSN 0169-8095. doi: 10.1016/J.ATMOSRES.2004.12.005.
- [49] A. Ryzhkov, S. Y. Matrosov, V. Melnikov, D. Zrnic, P. Zhang, Q. Cao, M. Knight, C. Simmer, and S. Troemel. Estimation of depolarization ratio using weather radars with simultaneous transmission/reception. *Journal of Applied Meteorology and Climatology*, 56:1797–1816, 7 2017. ISSN 1558-8424. doi: 10.1175/JAMC-D-16-0098.1.
- [50] J. Saez, J. Luengo, and F. Herrera. On the suitability of fuzzy rule-based classification systems with noisy data. *IEEE Transactions on Fuzzy Systems*, pages 1–1, 2012. doi: 10.1109/TFUZZ.2012.2182774.
- [51] C. G. Schmitt, A. J. Heymsfield, P. Connolly, E. Järvinen, and M. Schnaiter. A global view of atmospheric ice particle complexity. *Geophysical Research Letters*, 43(22):11,913–11,920, 2016. doi: <https://doi.org/10.1002/2016GL071267>.
- [52] T. Takahashi. Influence of liquid water content and temperature on the form and growth of branched planar snow crystals in a cloud. *Journal of the Atmospheric Sciences*, 71:4127–4142, 11 2014. ISSN 0022-4928. doi: 10.1175/JAS-D-14-0043.1.
- [53] E. Tetoni, F. Ewald, M. Hagen, G. Köcher, T. Zinner, and S. Groß. Retrievals of ice microphysical properties using dual-wavelength polarimetric radar observations during stratiform precipitation events. *Atmospheric Measurement Techniques*, 15:3969–3999, 7 2022. ISSN 18678548. doi: 10.5194/AMT-15-3969-2022.
- [54] E. J. Thompson, S. A. Rutledge, B. Dolan, V. Chandrasekar, and B. L. Cheong. A dual-polarization radar hydrometeor classification algorithm for winter precipitation. *Journal of Atmospheric and Oceanic Technology*, 31:1457–1481, 7 2014. ISSN 0739-0572. doi: 10.1175/JTECH-D-13-00119.1.
- [55] S. Trömel, M. R. Kumjian, A. V. Ryzhkov, C. Simmer, and M. Diederich. Backscatter differential phase—estimation and variability. *Journal of Applied Meteorology and Climatology*, 52:2529–2548, 11 2013. ISSN 1558-8424. doi: 10.1175/JAMC-D-13-0124.1.
- [56] J. Tyynelä and V. Chandrasekar. Characterizing falling snow using multifrequency dual-polarization measurements. *Journal of Geophysical Research*, 119:8268–8283, 7 2014. ISSN 21562202. doi: 10.1002/2013JD021369.
- [57] J. Tyynelä and A. von Lerber. Validation of microphysical snow models using in situ, multifrequency, and dual-polarization radar measurements in finland. *Journal of Geophysical Research: Atmospheres*, 124:13273–13290, 12 2019. ISSN 21698996. doi: 10.1029/2019JD030721.

- [58] F. Ulaby, D. Long, W. Blackwell, C. Elachi, A. Fung, C. Ruf, K. Sarabandi, J. Zyl, and H. Zebker. *Microwave Radar and Radiometric Remote Sensing*. University of Michigan Press, 01 2014. ISBN 978-0-472-11935-6.
- [59] J. M. Wallace and P. V. Hobbs. Atmospheric science: An introductory survey: Second edition. *Atmospheric Science: An Introductory Survey: Second Edition*, pages 1–488, 2006. doi: 10.1016/C2009-0-00034-8.
- [60] P. Wang. Dealiasing of doppler spectra for cloud radar at 94 ghz, student report, 2023. URL <https://resolver.tudelft.nl/uuid:2326e5f4-f0ef-4f37-af4c-64cc67b1c9d3>.
- [61] P. Wang. Microphysical retrievals in mixed-phase clouds with low lwp using cloud radar. Master's thesis, Delft Univeristy of Technology, 2024. URL <https://resolver.tudelft.nl/uuid:72573b61-7cad-4707-adb2-9115a438e31e>.
- [62] G. Wen, A. Protat, P. T. May, X. Wang, and W. Moran. A cluster-based method for hydrometeor classification using polarimetric variables. part i: Interpretation and analysis. *Journal of Atmospheric and Oceanic Technology*, 32:1320–1340, 7 2015. ISSN 0739-0572. doi: 10.1175/JTECH-D-13-00178.1.
- [63] G. Wen, A. Protat, P. T. May, W. Moran, and M. Dixon. A cluster-based method for hydrometeor classification using polarimetric variables. part ii: Classification. *Journal of Atmospheric and Oceanic Technology*, 33:45–60, 1 2016. ISSN 0739-0572. doi: 10.1175/JTECH-D-14-00084.1.
- [64] G. Zhang. *Weather Radar Polarimetry*. CRC Press, Inc., USA, 1st edition, 2016. ISBN 1439869588.
- [65] D. S. Zrnica, A. Ryzhkov, J. Straka, Y. Liu, and J. Vivekanandan. Testing a procedure for automatic classification of hydrometeor types. *Journal of Atmospheric and Oceanic Technology*, 18, 2000. doi: 10.1175/1520-0426(2001)018<0892:TAPFAC>2.0.CO;2.
- [66] D. S. Zrnica, G. Zhang, V. Melnikov, and J. Andric. Three-body scattering and hail size. *Journal of Applied Meteorology and Climatology*, 49(4):687 – 700, 2010. doi: 10.1175/2009JAMC2300.1. URL <https://journals.ametsoc.org/view/journals/apme/49/4/2009jamc2300.1.xml>.



Unused polarimetric variables

Circular Depolarization Ratio

The circular depolarization ratio (CDR) is similar to the SLDR. It depends less on the orientation of the scatterer compared to LDR and SLDR, and more on the shape and phase [46, p. 126-127]. It is obtained from radars that operate in circular polarization bases (opposite of those in linear polarization basis, thus horizontal and vertical polarizations).[36]. The CDR is given as[56]:

$$CDR = \frac{Z_{e,rr}}{Z_{e,rl}} \quad (A.1)$$

$$= \frac{\sigma_{rr}}{\sigma_{rl}} \quad (A.2)$$

$$= \frac{|S_{rr}|^2}{|S_{rl}|^2} \quad (A.3)$$

$$= \frac{|\frac{1}{2}(S_{vv} + S_{hv} + i(S_{hh} + S_{vh}))|^2}{|\frac{1}{2}(S_{vv} + S_{hv} - i(S_{hh} + S_{vh}))|^2} \quad (A.4)$$

Converting to the BSA alignment gives

$$CDR = \frac{|\frac{1}{2}(S_{vv_backward} + -S_{hv_backward} + i(-S_{hh_backward} + S_{vh_backward}))|^2}{|\frac{1}{2}(S_{vv_backward} + -S_{hv_backward} - i(-S_{hh_backward} + S_{vh_backward}))|^2}. \quad (A.5)$$

However, after omitting the S_{hv} the CDR now contains the exact same information as ϕ_{bs} , brought to attention by a Spearman correlation coefficient of 1. Matrosov [36] and Ryzhkov et al. [49] deal with this issue by using a special signal processing method on top of a high-power phase shifter.

Co-polar Correlation Coefficient

The co-polar correlation coefficient ρ_{hv} is calculated by [8, p. 404]

$$\rho_{hv} = \frac{|-S_{hh}^* S_{vv}|}{\sqrt{|-S_{hh}|^2 |S_{vv}|^2}} \quad (A.6)$$

$$\rho_{hv} = \frac{|-Shh_backward^* Svv_backward|}{\sqrt{|-Shh_backward|^2 |Svv_backward|^2}}. \quad (A.7)$$

Here, the * indicates the complex conjugate. In the case that there is only one scatterer in the resolution volume, $\rho_{hv} = 1$. [46, p. 139] Therefore, it is not possible to use these in combination with the database, for the polarimetric variables from the database will always be just one particle. However, it could be used to test the assumption that the Doppler velocity bins indeed do contain only one particle type.

Cross Polar Correlation Coefficient

The cross polar correlation is given by,

$$\rho_{cx} = \frac{|S_{vv}^* \cdot -S_{hv}|}{\sqrt{|S_{vv}|^2 - S_{hv}^2}} \quad (\text{A.8})$$

$$\rho_{cx} = \frac{|S_{vv_backward}^* \cdot -S_{hv_backward}|}{\sqrt{|S_{vv_backward}|^2 \cdot | - S_{hv_backward}|^2}}. \quad (\text{A.9})$$

in a vertical polarized wave [64, p. 125]. Note the minus sign in front of S_{hh} and S_{hv} to make use of the BSA convention.

Just like the co-polar correlation coefficient, this variable is not used because it will always be 1 when computing for a single particle, as is done when using the scattering database.

B

Principal Component Analysis

The option to use principal component analysis (PCA) to create new variables. The idea is that by combining variables into new variables, so-called principal components, the variance per component is maximized, the amount of dimensions in the classifier is reduced, and variables that are highly correlated are identified. The principal components are linear combinations of the original variables, and they are uncorrelated to each other. Doing this can negate the issue of losing information when working with 1D membership functions.

Using the polarimetric variables, the absolute loading for each principal component is shown in Figure B.1. Their cumulative variance can be seen in Figure B.2.

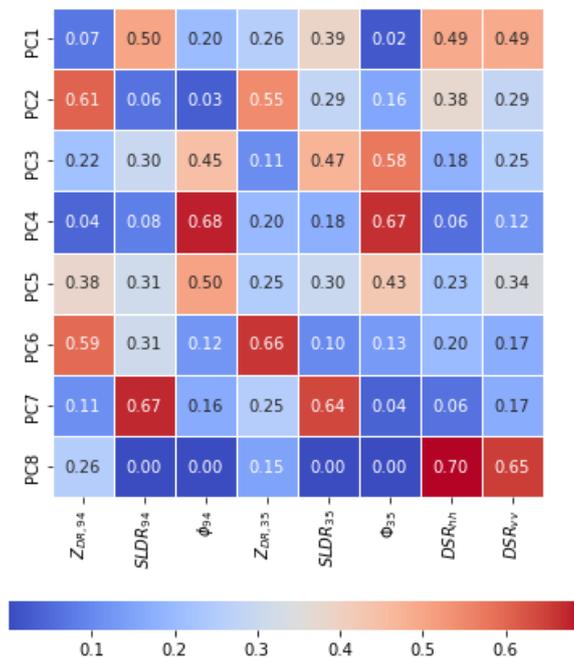


Figure B.1: Absolute Loading for each variable and component in PCA.

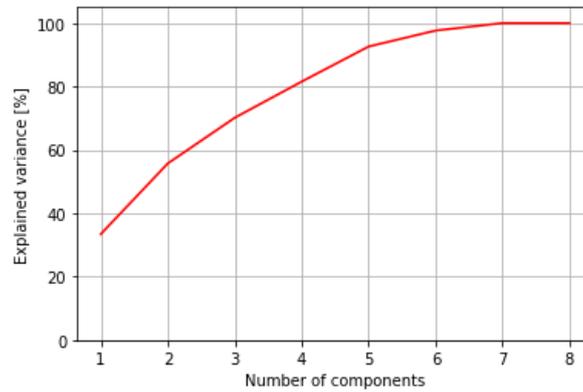


Figure B.2: Cumulative variance of each successive component in the PCA.

With these results in mind, there are several reasons not to use PCA variables in the classifier:

1. Variables such as Z_{dr} have been widely used, and are physically interpretable. Because of that, classification results can be related to measurements, which would be lost when using linear combinations of these variables.
2. Noisy measurements or calibration issues will not stand out as clearly when using principal components, because each measurement on its own is given less weight within the component, or it is harder to derive where odd values are coming from.
3. A strength of fuzzy logic was the ability to leave out measurements when they are in fact too noisy or simply not available. With using principal components, whole components could be lost if this variable has a large weight.