THESIS

ADVANCED PROCESSING OF DUAL POLARIZATION WEATHER RADAR SIGNALS

Submitted by

Shweta Haran

Department of Electrical and Computer Engineering

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2022

Master's Committee:

Advisor: V. Chandrasekar

Haonan Chen
Thomas Siller

ABSTRACT

ADVANCED PROCESSING OF DUAL POLARIZATION WEATHER RADAR SIGNALS

This research focuses on processing of radar data in spectral domain and analysis of micro-physical properties of hail and rain in severe convective and stratiform storms. This research also discusses the optimization of a parametric time domain method to separate cloud and drizzle data.

The microphysical and kinematic properties of hydrometeors present in a precipitation event can be studied using spectral domain processing and analysis of the radar moments. This study along with polarimetric information is called spectral polarimetry. For this study, the observations made by CSU-CHIVO (Colorado State University - C-band Hydrometeorological Instrument for Volumetric Observation) radar during the RELAMPAGO (Remote sensing of Electrification, Lightning, And Mesoscale/Microscale Processes with Adaptive Ground Observations) campaign is utilized. Features such as the slope in differential reflectivity, spectrum width, and spectral copolar correlation are studied which gives a better understanding of the storm microphysics. In this thesis, microphysical properties of different types of hydrometeors such as hail, rain, and large drops are studied using convective and stratiform storm observations.

A parametric time-domain method (PTDM) is utilized for the separation of cloud and drizzle data. To reduce the time latency present in processing the data, the processing code is optimized by deploying on a high-performance computer (HPC). The processing code is tested

on an HPC and automated to handle errors in processing. The run time is reduced by approximately 50%, hence increasing the data processing efficiency. This study shows that optimization of the run time using an HPC is an efficient method. Data processing using an HPC can be used to deploy similar time-consuming algorithms, hence increasing the efficiency and performance.

ACKNOWLEDGEMENTS

I would like to thank all the individuals whose constant encouragement and support have made the completion of this thesis possible. First and foremost, I would like to express my sincere gratitude to my advisor, Dr. V. Chandrasekar, who has guided me through the process of this study with his insightful and timely advice. I can say with utmost confidence that it was his constant support and timely advice which nurtured me during my research and guided me in the right direction. I would also like to thank my thesis committee members, Dr. Haonan Chen, and Dr. Thomas Siller for taking time out of their busy schedules and reviewing all my subject material. I also want to thank my lab partners Ivan Arias Hernandez, Sounak Biswas, and Shashank Joshil for their valuable input and constant support during my research. Last but not the least, I would like to thank my father Hariharan Sachidanandam, my mother Shoba Meenakshi Sundaram, and my brother Harsha Haran, for their constant words of encouragement and for supporting me all these years. Their unwavering confidence in me allowed me to pursue my dream. The RELAMPAGO program was supported by the National Science Foundation. This research was also supported by the DOE/ASR program.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 : INTRODUCTION

1.1. INTRODUCTION AND MOTIVATION

Dual-pol radars have been used in atmospheric science and remote sensing to study different hydrometeors and to predict weather effectively for decades. Dual-pol Doppler weather radars are effective because they transmit and receive electromagnetic pulses which are both horizontally and vertically polarized. Since these pulses are polarized in both the orientations, the information of the target in both orientations is known, which gives us a better estimation of the size, shape, distribution of the target, type of the target, etc.

A single-pol Doppler weather radar gives three basic radar moments known as reflectivity (Z), Doppler velocity (v), and spectral width (W). Doppler dual-pol weather radars have three additional radar moments which are differential reflectivity (Z_{dr}), differential phase (φ_{dp}), and co-polar correlation (ρ_{hv}) apart from the initial three radar moments. Useful information can be derived from these radar moments like hydrometeor classification, rainfall estimation in a region, melting layer detection, etc. The advancements in weather radar systems and enhanced information from weather radar are very useful for forecasters to predict weather and allows researchers to study the microphysical properties of the different precipitation types.

To study and classify the different precipitation types and their microphysical properties and characteristics of these precipitation, the radar moments obtained from the Doppler weather radar are processed in the spectral domain. Spectral processing and analyzing of the radar moments can give a better understanding of the properties and characteristics of targets present in a radar observation which can be clutter (noise or other non-meteorological targets like birds)

or precipitation. The received signal power gives an idea of the precipitation in the observed radar resolution volume, while the Doppler velocity can unveil information about the motion of the observed particles relative to the radar, the differential reflectivity can help identify the size of the particles, copolar correlation coefficient can help identify the homogeneity of the particles present and the differential phase can help in measuring the concentration of the particles along a radial angle. Spectral processing is calculating and estimating these radar moments from the dual-pol spectra in the Fourier domain. Some techniques used for spectral processing are discussed in Bringi and Chandrasekar [1], Wang et al [2] and Pfitzenmaier et al [3]. Spectral polarimetry is the study and characterization of the microphysical properties of the different precipitations present in a storm by spectral processing and analyzing the radar moments.

While spectral processing and analysis gives a better understanding of the microphysical properties and characteristics of the precipitation, noise removal and estimation of the radar moments in spectral domain is challenging when the noise concentration is higher than precipitation. If the radial velocities of the weather echoes are small, then noise removal can be difficult in the frequency domain because of spectral leakage of the noise into the radar moment estimates. This drawback can be overcome by estimating the radar moments in time domain as shown by Nguyen et al [4] by using a parametric time domain method (PTDM). In the time domain method both the noise and precipitation estimates are calculated simultaneously hence the error in accuracy is reduced significantly. This separation of noise from precipitation can also be applied to other applications like separation of cloud and drizzle estimates in an observed radar resolution, as discussed in Joshil et al [5]. These radar moment estimates are accurate even if the presence of cloud is higher than that of precipitation.

1.2. THESIS OVERVIEW

This section will give an overview of the different topics covered in each chapter. In this thesis, spectral processing, and analysis of different precipitation type in two different types of storms, characterizing the microphysical properties of the precipitation, error analysis by calculating the standard deviation of the spectral reflectivity and differential reflectivity, separation of cloud and drizzle moments by using a parametric time domain method and optimization of this technique's processing time using a high-performance computer is discussed.

In Chapter 1, the thesis gives a brief introduction on the various topics covered in the thesis. Doppler dual-pol weather radars and the dual-pol radar products obtained are explained. Spectral and time domain analysis and the application of processing and estimating the radar moments in each of these domains is explained. The advantage of processing radar moments in each domain is discussed.

In chapter 2, the basics of a dual-pol Doppler weather radar like the components that make up the radar, how the radar works, the different scanning techniques used, how the radar observes and records data, and how the dual-pol radar moments can be calculated from the recorded data is discussed.

In chapter 3, a brief introduction about the RELAMPAGO campaign and the CSU-CHIVO radar is given. We investigate stratiform and convective storms and study the microphysical properties of the particles present in different hydrometeors of these storms and characterize their microphysical properties. The radar products namely reflectivity, differential reflectivity and

copolar correlation ratio is spectral processed and analyzed and the error analysis is done by calculating the standard deviation of the analyzed range gate with respect to the nearby range gates.

In chapter 4, a PTDM model to separate cloud and drizzle data is discussed and this model is deployed on an HPC to optimize its runtime and reduce time latency in processing the data. The data used was collected during the ARM CAP MBL (Atmospheric Radiation Measurement Clouds, Aerosol and Precipitation in the Marine Boundary Layer) campaign. The effectiveness of deploying the PTDM model on the HPC and how a high-performance computer can be used to process computationally intensive algorithms and is discussed.

Chapter 5 summarizes the work discussed in this thesis and gives possible avenues for future work.

CHAPTER 2 : BACKGROUND

2.1 DOPPLER WEATHER RADAR BASICS

Initially Doppler weather radars transmitted electromagnetic waves which were only horizontally polarized, which restricted the radar output moments to the basic three radar moments reflectivity, Doppler velocity, and spectral width. With the advent of dual polarization techniques in Doppler weather radars, it allowed to view the horizontal and vertical properties of the targets, hence additional radar output moments were obtained like differential reflectivity, copolar correlation and specific differential phase. These radar moments give a better sense of the targets like size, shape, orientation, distribution, etc. In this chapter, the different components that make up a dual polarization radar, how the electromagnetic waves are transmitted and received, and how the raw I/Q radar data is processed to get the radar moments is discussed. A dual-pol Doppler weather radar systems block diagram is shown below in figure 2.1.

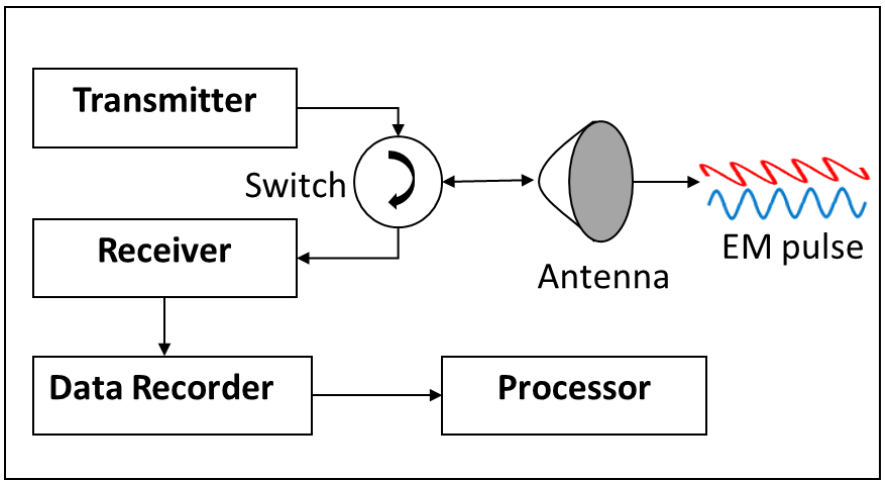

Figure 2.1 A block diagram of dual-pol Doppler radar system

A radar system can be classified based on the antenna as either a monostatic or bistatic. In a monostatic radar system, there is only one antenna which performs both the transmission and reception of the radar pulse signal, whereas is in a bistatic system there are two different antennas where each perform either the role of transmission or the reception of the radar pulse. Most weather radar systems are monostatic in nature where a single antenna performs both transmission and reception of the electromagnetic pulses. For spectral polarimetry, the dataset used was collected by the CSU-CHIVO radar which used a center-fed parabolic type reflector antenna. CSU-CHIVO is C-band radar which has a better sensitivity and attenuation comparatively with a gain of 45 dBi operating at a range of 5.4 GHz, having a beam width of less than 1° and a side lobe having less than -29 dB. The antenna also has a quasi-random beehive structure known as the radar dome. This is made of fiberglass which protects the antenna from the severe weather conditions that are present. For the separation of the cloud and precipitation using a PTDM the dataset was collected by a Scanning W-band ARM Cloud Radar (SWACR) during the ARM CAP-MBL campaign. The radar worked in a 95 GHz frequency range with a beam width of 0.32°. The antennae are usually put atop a pedestal and a rotatory joint which helps rotate and position the antenna to focus on the target. The radar pulses are generated by the transmitter which feeds the signals via the feedhorn to the antenna. The antenna then generates small beams by focusing the transmitted power and the antenna reflector is used to transmit the radar pulses and receive the signal which is bounced off the target.

The transmitter and receiver are an essential component of the radar systems. Transmitter is required to generate the radar pulse that is transmitted via the antenna. The transmitter consists of a waveform generator which generates the required radar signal pulse at a low power, which is then amplified and transmitted. Doppler weather radars generate a short duration high power

pulses of energy using a coaxial magnetron coupled to a solid-state modulator. The transmitter and the receiver both use the same antenna to transmit and receive the RF pulse signal respectively, hence, to separate the pulses a switch or a duplexer is used. This allows for isolation between the two components since the transmitter and receiver both work in different power levels. The receiver is fed the reciprocated radar pulse, which returns after hitting the target, collected by the antenna. The CSU-CHIVO uses a RVP900 Vaisala signal receiver and processor as shown in figure 2.2. The received RF pulse signal is digitized using an intermediate frequency digital receiver (IFDR) which gives us two outputs in-phase and quadrature phase (I and Q) samples. From the raw I/Q data the radar moments can be calculated. Here the radar moments are calculated by the signal processor RVP902 shown in figure 2.2.

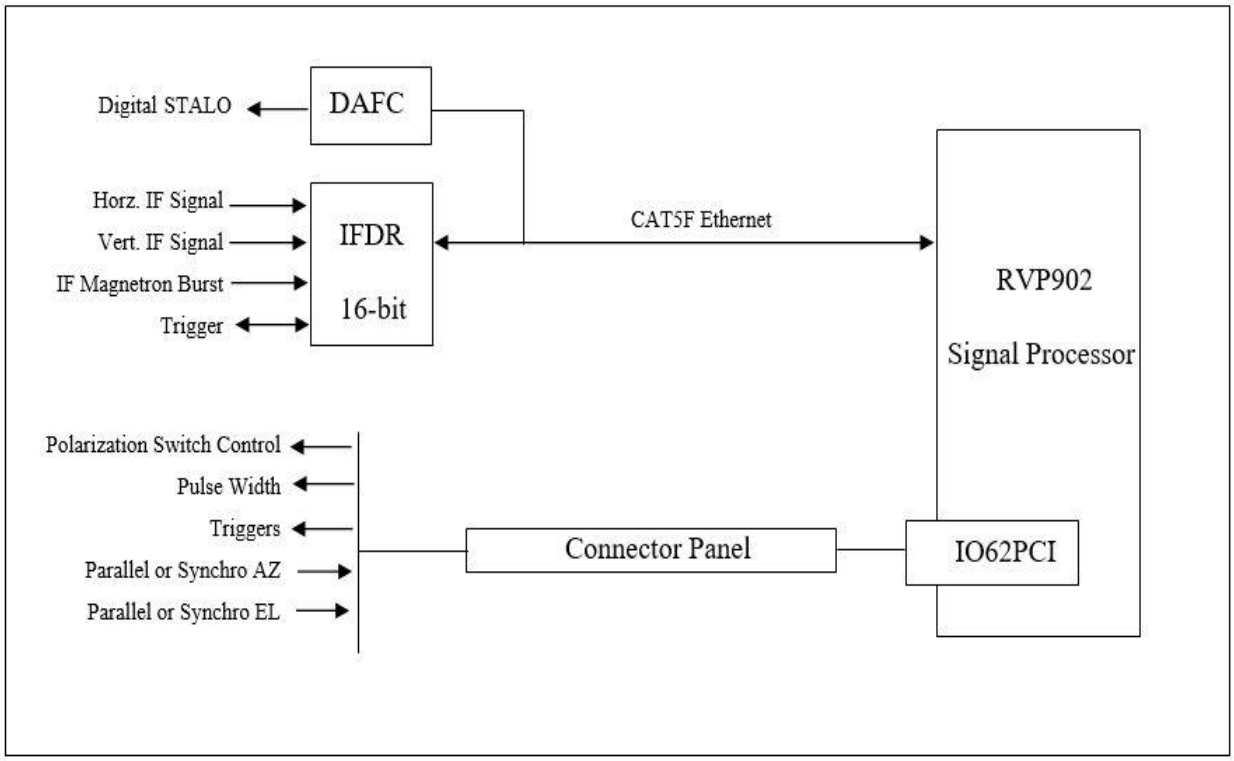

Figure 2.2 Block Diagram of Signal Receiver and Processor [6]

I/Q Processing is done to extract the radar moments like reflectivity, Doppler velocity, spectral width, differential reflectivity, specific differential phase and copolar correlation. The block diagram of the signal processor is shown in figure 2.3. The time series raw I and Q samples are passed into the clutter filter to remove noise. The output from the clutter filter is then processed to get autocorrelations R0, R1 and R2 which are derived through Pulse Pair, random phase or DFT/FFT modes. For the pulse pair model, the operations are performed in the time or frequency domain and the autocorrelations are calculated in the time domain. A Gaussian Model Adaptive Processing (GMAP) techniques is used for DFT/FFT mode, and the autocorrelation is calculated from inverse transform and for the random phase the filtering is done in frequency domain using adaptive algorithms and the autocorrelation is computed from the inverse transform. After this the autocorrelations of consecutive range bins are averaged after removing any range bin values with high clutter concentration so that the data isn't compromised. Next the radar moments are calculated using the autocorrelation functions which is shown in section 2.2. Thresholding can be done using signal to noise ratio (SNR) or the copolar correlation (ρ_{hv}) values which can be applied to the radar moments to remove noise from the signal.

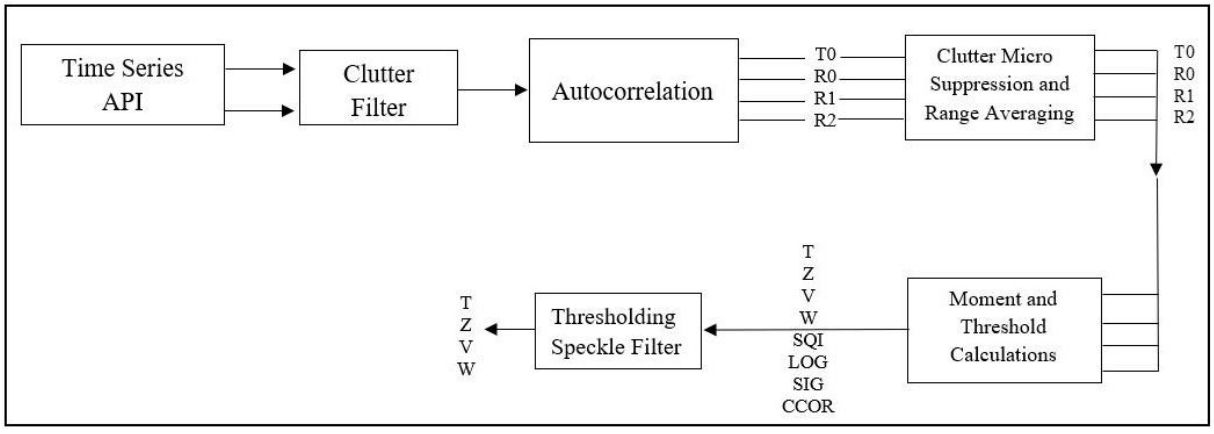

Figure 2.3 I/Q Processing for Weather Moment Extraction [6]

2.2 DOPPLER RADAR MOMENTS EXTRACTION AND SPECTRAL PROCESSING

The I and Q signals from the IFDR are processed by the signal processor to calculate the radar moments. For this pulse pair processing algorithm is used, where the I/Q samples are recorded and saved as a times-series data for every azimuth and elevation as a horizontal and vertical (H and V) polarization. Pulse-pair algorithm techniques is used because they give the least amount of bias and lowest variance in the calculated radar moments. Meteorologists and weather forecasters use these radar moments derived from processing the raw I/Q samples to perform further analysis to predict the weather accurately and study about the microphysical properties of the different precipitation. Spectral decomposition is the representation of the dual-pol radar moments in Fourier domain. In Bringi and Chandrasekar [1], the calculation of the Doppler radar moments is explained in detail. The spectral processing techniques as shown in Bringi and Chandrasekar [1] and Moisseev and Chandrasekar [7] is also discussed in this section. Spectral processing is used to calculate the power spectra and retrieve the radar moments and dual-pol parameters from the power spectra.

The autocorrelation function of zeroth lag can be calculated from the horizontally polarized (H-pol) raw time series samples as shown below in equation 1, where n is the sample size and N is the length of the sequence. $R0_h$ is the zeroth lag calculated from the H-pol raw I/Q data. We can calculate the power of the signal, $R0_h$, along any given k^{th} elevation (if RHI scan) or azimuth (if PPI scan) angle and at any given range gate *r* kilometers from the radar.

$$R0_h = \frac{1}{N} \sum_{n=1}^{N} \mathrm{H(n)H^*(n)} \tag{1}$$

The autocorrelation function at first lag, which is $R1_h$ in the figure 2.3, can also be calculated from the H-pol time series as shown is equation 2,

$$R1_h = \frac{1}{N+1} \sum_{n=1}^{N} \mathrm{H(n)H^*(n+1)} \tag{2}$$

The autocorrelation function at second lag, which is $R2_h$ in the figure 2.3, can also be calculated from the H-pol time series as shown is equation 3,

$$R2_h = \frac{1}{N+2} \sum_{n=1}^{N} \mathrm{H(n)H^*(n+2)} \tag{3}$$

The auto correlation function R_h at m^{th} lag can be calculated from H-pol time series of a length N and with a sampling index of n as shown in equation 4,

$$R_h(m) = \frac{1}{N} \sum_{n=0}^{N-1} \mathrm{H(n+m)H^*(n)} \tag{4}$$

The Doppler power spectrum can be computed using discrete Fourier transform (DFT) of the above derived autocorrelation function at a particular range gate as shown in equation 5,

$$S_h(r,k) = \sum_{n=-L}^{L} R_h(n) e^{-\frac{j2\pi nk}{N}} \tag{5}$$

Where k=0,1,2….N-1 and is the DFT sample indices and the total number of samples are L≤N. S_{hh} is the auto-spectra since it is the DFT of the auto-correlation function $R_{hh}(n)$.

The cross-correlation function at zeroth lag, which is T0 in the figure 2.3, is calculated from the H-pol and V-pol time series raw data as shown is equation 6,

$$T0 = \frac{1}{N}\sum_{n=1}^{N} \mathrm{H(n)V^{*}(n)} \tag{6}$$

We can also calculate the cross-correlation functions R_{hv} and S_{hv} in a similar manner. They are the cross correlation between the horizontal and vertical data samples. The cross correlation is calculated as shown in equation 7 and the corresponding cross spectra is derived from equation 7 as shown in equation 8.

$$R_{hv}(m) = \frac{1}{N}\sum_{n=0}^{N-1} \mathrm{H(n+m)V^{*}(n)} \tag{7}$$

$$S_{hv}(r,k) = \sum_{n=-L}^{L} R_{hv}(n)e^{-\frac{j2\pi nk}{N}} \tag{8}$$

Using autocorrelation and cross correlation functions calculated above the radar moments, reflectivity, Doppler velocity and spectral width can be calculated, which are the zeroth, first and second order moments of the power spectrum respectively. The H-pol reflectivity, Z_h, of the radar is calculated as given by equation 9, where C is the radar constant and r is the range gate in km from the radar.

$$Z_h = 10\log_{10}(R0_h) + 20\log_{10}(\mathrm{r}) + \mathrm{C} \tag{9}$$

From equations 4, 5, 7 and 8, we can obtain the spectral decomposition of the radar moments and dual-pol parameters like reflectivity, differential reflectivity, co-polar correlation, etc. Spectral Polarimetry is the study of dual-pol spectral properties and characterizing the micro-physics of the different meteorological events. The spectrum of the reflectivity can be derived

from the power spectrum as shown in equation 10 and the first radar moment, reflectivity, Z_h, can also be derived from the power spectrum for a range r km away from the radar as shown in equation 11.

$$S(Z_h(r,k)) = 10\log_{10}(S_h(r,k)) + 20\log_{10}(\mathrm{r}) + \mathrm{C} \tag{10}$$

$$Z_h(r) = 10 \times \log_{10}(\sum_{k=0}^{N-1} S_h(r,k)) + 20 \times \log_{10}(\mathrm{r}) + \mathrm{C} \tag{11}$$

The Doppler velocity, v, the second radar moment, is calculated as shown in equation 12 below using equation 2. The Doppler velocity for a range, r, from the radar can also be calculated from the power spectral densities as shown in equation 13

$$v_h = \frac{\lambda}{4\pi T_s} \arg\,[R1_h] \tag{12}$$

$$v(r) = \frac{\sum_{k=0}^{N-1} v_k S_h(r,k)}{\sum_{k=0}^{N-1} S_h(r,k)} \tag{13}$$

where λ is the wavelength of the radar and T_s is the sampling time period, also known as the pulse repetition time (PRT). The third radar moment, spectral width, which is the measure of the sheer and turbulence within a radar resolution volume is calculated using equation 14 below and equation 15 shows how to retrieve the width of the spectrum from the power spectral densities along a range r km from the radar.

$$W = \sqrt{\frac{2\ln\frac{R0}{|R1|}}{\pi}} \tag{14}$$

$$W(r) = \frac{\sum_{k=0}^{N-1}(v_k - v(r))^2 S_h(r,k)}{\sum_{k=0}^{N-1} S_h(r,k)} \quad (15)$$

The dual-pol radar products like differential reflectivity, co-polar correlation, and differential phase are calculated using the parameters calculated above. Differential reflectivity, Z_{DR}, is defined as the difference between the horizontal and vertical reflectivity. It can be calculated using equation 16, shown below and the spectral differential reflectivity is calculated as shown in equation 17. Similarly, the differential reflectivity can also be retrieved from the power spectrum as shown in equation 18.

$$Z_{dr} = 10\log_{10}\left(\frac{Z_h}{Z_v}\right) \quad (16)$$

$$S\left(Z_{dr}(r,k)\right) = \frac{S_h(r,k)}{S_v(r,k)} \quad (17)$$

$$Z_{dr}(r) = 10\log_{10}\left\{\frac{\sum_{k=0}^{N-1} S_h(r,k)}{\sum_{k=0}^{N-1} S_v(r,k)}\right\} \quad (18)$$

The copolar correlation coefficient is used to get an idea of the meteorological targets. It is the similarity of the returned power from the H-pol and V-pol parameters. If the value is higher than 0.75 than the target is a meteorological target, if not it can be dismissed as noise from the surroundings. This is because meteorological targets change both the phase and power of H-pol and V-pol uniformly, and hence the copolar correlation value is closer to 1, compared to non-meteorological targets. The copolar correlation value is calculated as shown in equation 19, and the spectral decomposition of the copolar correlation coefficient is calculated as shown in

equation 20. The copolar correlation value can be calculated from the power spectrum also as shown in equation 21.

$$\rho_{hv} = \frac{|T0|}{\sqrt{R0_h \; R0_v}} \tag{19}$$

$$S(\rho_{hv}(r,k)) = \frac{|S_{hv}(r,k)|}{\sqrt{S_h(r,k) S_v(r,k)}} \tag{20}$$

$$\rho_{hv}(r) = \frac{\sum_{k=0}^{N-1} |S_{hv}(r,k)|}{\sqrt{\sum_{k=0}^{N-1} S_h(r,k) \sum_{k=0}^{N-1} S_v(r,k)}} \tag{21}$$

Using the copolar correlation calculated from equation 19, we can calculate the differential phase as shown in equation 22, the spectrum of the differential phase is calculated as shown in equation 23 and the

$$\varphi_{dp} = \arg[\rho_{hv}] \tag{22}$$

$$S(\varphi_{dp}(r,k)) = \arg[\rho_{hv}(r,k)) \tag{23}$$

$$\varphi_{dp}(r) = \arg[\sum_{k=0}^{N-1} S_{hv}(r,k)] \tag{24}$$

Signal Quality Index (SQI) helps in thresholding the signal echoes. If the values are closer to 1 than the signal is good quality signal. The SQI is calculated using equation 25,

$$SQI = \frac{|R1_h|}{R0_h} \tag{25}$$

Spectral decomposition plays an important role in noise suppression in the Fourier domain and differentiating precipitation from noise. It also helps us to get a more detailed insight into a ray in an elevation or range height indicator (RHI) scan or an azimuth or a plan position indicator scan (PPI). Spectral analysis plays an important role in understanding the microphysical properties and characteristics of the different hydrometeors detected in a storm better.

CHAPTER 3 : SPECTRAL POLARIMETRY OF CSU-CHIVO WEATHER RADAR DATA TO STUDY PRECIPITATION

3.1. INTRODUCTION

Spectral processing plays an important role in hydrometer classification, quality control of the data by effectively removing or suppressing noise, weather forecasting and quantitative precipitation estimation as discussed in Chandrasekar et al [8], Park et al [9], and Ryzhkov et al [10]. Polarimetric variables like reflectivity, differential reflectivity, copolar correlation, differential phase, and specific differential phase help in differentiating the various observed targets in a radar resolution volume as hydrometers (precipitation) or other non-meteorological particles like organic targets (bird, insects, etc) or ground clutter (example: buildings). Spectral processing and analysis has allowed researchers to classify different precipitations in weather echoes with respect to their size, shape and orientation and gives an insight into the mechanisms occurring in higher altitudes where snow or hail melt to form rain and other phenomenon like convective updraft in storms which increase the size of the hail. A combination of both the Doppler parameters and polarimetric properties can be studied for this purpose and this is known as spectral polarimetry.

In this chapter, we'll be looking at how spectral polarimetry can be used to study different hydrometeors present in storms and classify the microphysical properties. Representation in Fourier domain of the dual-pol parameters like reflectivity, differential reflectivity, copolar correlation, etc. is called spectral decomposition of dual-pol parameters. The slope of the differential reflectivity in the spectral domain can give us many insights about the precipitation characteristics, especially of the hail core in storms, this is discussed in [2], where they also show

that size sorting of the hail occurs due to the presence of vertical shear in a turbulent environment like a convective storm.

Precipitation is not always constant, that is, when there is rain, it doesn't have to necessarily imply that the rain droplets were formed in the beginning. Rain and drizzle drops sometimes are formed by a combination of different processes like auto conversion process in collision-coalescence process or because of Bergeron process, as discussed in Fabry and Zawadzki [11]. This suggests that when there is a collision between warm rain droplets and cold snow the result is a supercooled drizzle phenomenon. Hail can also exist during extreme condition of updraft where the water molecules are constantly lifted above the melting point layer where they freeze and while they travel down they melt and collect other water molecules. This constant motion is carried out repetitively until the hail stone is large enough that it can't be carried up by the updraft anymore and hence falls to the ground.

In this chapter we'll be studying the different precipitations that existed during three storms observed by the CSU-CHIVO radar during the RELAMPAGO campaign. Two of these storms were convective storms, one which occurred on 25th January 2019 and was one of the tallest storms that was ever observed by a dual-pol radar and another storm that occurred on 12th December 2018. The third storm which occurred on 18th November 2018 and was a stratiform storm. During the time that the CSU-CHIVO radar was deployed many severe weather conditions and storms were observed and the time series data was collected. The radars used for this campaign are C-band radars, even though in the United States S-band radars are used extensively. C-band radars are much smaller and more inexpensive to install and also the S-Band radar signals are not extensively attenuated by heavy rains like C-band. To study the microphysical properties of the different precipitation particles, especially if they are larger in

size like hail, then C-band is better suited because a lot of the larger particles are not affected by the resonance effects at this wavelength. The next section gives an in-depth discussion about the RELAMPAGO campaign and how the data was collected.

3.2. OVERVIEW OF RELAMPAGO CAMPAIGN

Remote sensing of **E**lectrification **L**ightning **A**nd **M**esoscale/microscale **P**rocesses with **A**daptive **G**round **O**bservations, RELAMPAGO for short, was a joint venture between various organizations, and also funded by them like the National Oceanographic and Atmospheric Administration (NOAA), US National Science Foundation (NSF), National Aeronautics and Space Administration (NASA), Ministry of Education, Science and Technology of Argentina (MinCyT), Servicio Meteorologico Nacional (SMN), Province of Cordoba, Brazil, and INVAP. The main objective of the campaign was to record and observe the convective storms and severe weather conditions, which are one of the causes of high impact weather in the Andes Mountain region in Argentina. Figure 3.1 below shows the study region during the RELAMPAGO campaign. This region has severe weather conditions where there are unpredictable storms, heavy rain and large hail. As discussed in Nesbitt et al [12], the campaign lasted for 11 months from 1st June 2018 to 30th April 2019, which was called the extended observing period, while the intensive observing period lasted for a month and a half from 1st November 2018 to 18th December 2018. The RELAMPAGO campaign deployed a thick network of radars and sensors on the ground which allowed researchers to collect and study severe weather phenomenon like tall storms and strong updrafts.

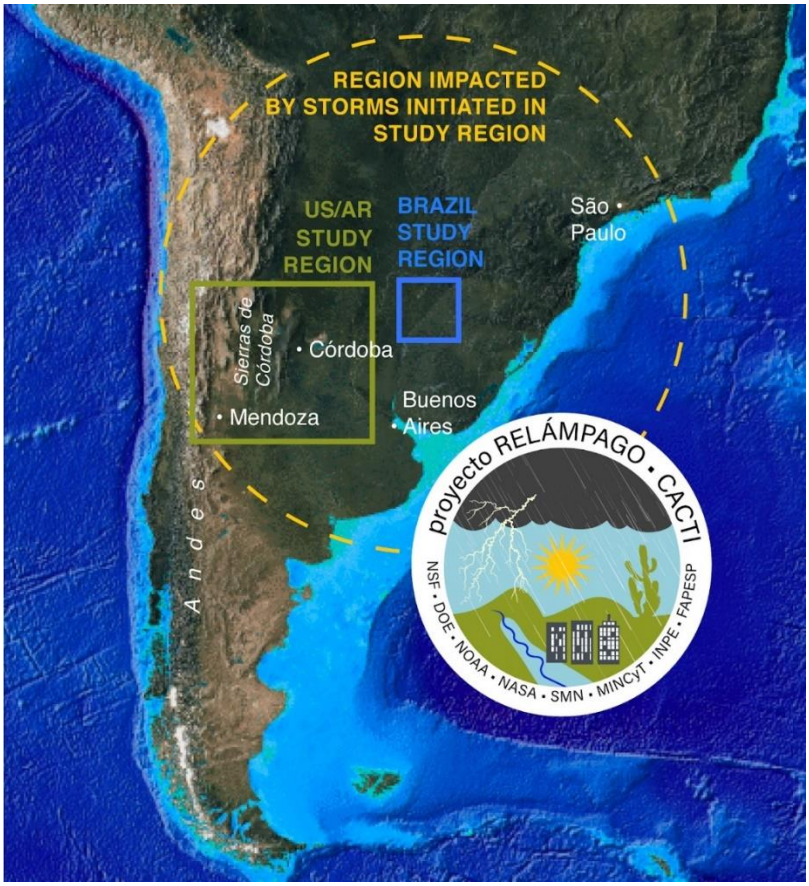

Figure 3.1 RELAMPAGO campaign study site. Figure above is the courtesy of Steve, Nesbitt, University of Illinois [13]

During the RELAMPAGO campaign, a network of radars was deployed to monitor and observe the severe weather condition. One of the radars deployed was the **C**-band **H**ydrometeorological **I**nstrument for **V**olumetric **O**bservation (CHIVO), which is a research radar from CSU, and it was deployed in the south of the Cordoba city. Apart from the three main radar moments, the CSU-CHIVO radar also gave the dual-pol variables differential reflectivity, copolar correlation and differential phase. It was operational for almost 3 months and started observing and recording data from 10th November 2018 to 31st January 2019. The CHIVO radar is a C-band rand commissioned by CSU at was located at 31.63°S latitude and 64.17°W longitude at an altitude of 421 m above the mean sea level. Since the one of the goals

of the researchers were to study the storm from beginning to the end and since most of the tallest storms were in the regions along the Andes and Sierra de Cordoba hills, most of the radars deployed were along this region including CSU-CHIVO.

Figure 3.2 The location of the CSAPR2 radar and CSU-CHIVO radar

Another radar that is discussed in this chapter is the C-band Scanning ARM Precipitation Radar 2 (CSAPR2). This is also a dual-pol Doppler C-band radar like CSU-CHIVO, but works in a STAR mode, which is simultaneous transmit and receive and it can transmit in either single polarization or dual polarization. The radar was located at 32.7°S latitude and 64.43°W longitude at an altitude of 1141 m above the sea level. The radar was operations from 23rd

September 2018 to 1st May 2019. The CSAPR2 also recorded the radar moments and the dual-pol variables like CSU-CHIVO radar. This makes rainfall estimation and classification of the precipitations easier. Figure 3.2 above shows the location of both the radars that will be discussed in this chapter and are approximately 78 km away from each other. The CSAPR2 and CHIVO radar both collected radar resolutions in RHI (Range Height Indicator), PPI (Plan Position Indicator) and Bird Bath (vertical) scans. The CSAPR2 was located more closer to the foothills of the Andes Mountain ranges.

TABLE 3.1 Calibration specifications of CSU-CHIVO radar

Parameter	**Specifications**
Frequency	5.4 GHz
Pulse repetition frequency	1 kHz
Gate spacing	150 m
Footprint	994 m at 60 km
Antenna beamwidth	< 1°

The CSU-CHIVO radar operates at a frequency range of 5.4 GHz with a peak power of 250 kW, it follows a pulse repetition frequency (PRF) method with the pulses transmitted at a frequency of 50 to 2400 Hz. The CSU-CHIVO radar has a Vaisala's RVP900 signal processor, as explained in the previous chapter with azimuth averaging from 2 to 1024 pulsers and uses a Gaussian Model Adaptive Processing (GMAP) filter for clutter and noise removal, this provides more than 50 dB clutter rejection. The CSU-CHIVO radar used a Sigmet processor but some of the hard-wired functionalities like the attenuation correction and ground clutter removal were

turned off to receive as much as pure unprocessed observation data as possible, through which some unique data can be captured. The CSU-CHIVO radar specifications are given above in table 3.1.

The figure 3.3 shows the CSU-CHOVO radars. The CSU-CHIVO performed all 3 scan types as mentioned previously, RHI, PPI and Bird bath scans. The RHI scan is performed along the height or elevation, while maintaining the azimuth constant, which is from the ground to the sky along the vertical plane. An PPI scan is performed along the horizontal plane, which is keeping the elevation constant for a scan and varying the azimuth. Hence it is also known as a surveillance scan.

Figure 3.3 CSU-CHIVO radar [14]

3.3. SPECTRAL ANALYSIS OF DUAL-POL RADAR MOMENTS:

Spectral analysis of weather echoes helps in detection of noise and clutter which can be suppressed using filters. In Bachmann et al [15], the spectral densities of dual-pol parameters are estimated to differentiate the scatter of noise from precipitation signal in a radar resolution volume. Spectral polarimetry helps in understanding the microphysics of different type of precipitation in a radar resolution volume as discussed in Dutta et al [16]. The convective storms that were recorded during this campaign has some of the tallest storms ever recorded, one of them which occurred on 25th January 2019 and was as tall as 20km. One of the properties of the convective storm region are they have a high reflectivity core and maximum echo top, which is a significant indication of hail in the storms. In the next section we'll look at the two different storm types, convective and stratiform storms. For the convective storm we'll investigate the 25th of January 2019 and 12th of December 2018 storms and for the stratiform storm we'll be looking into the 30th of November 2018 storm.

3.3.1 Convective Storm Cases

Convective storms are turbulent in nature with presence of intense heavy rain, hail, and thunderstorms. The convective storms from this region have hail of very large sizes. In this section, the mixture of rain and hail, large drops and heavy rain regions from two convective storms that occurred on 25th of January 2019 and 12th of December 2018 are studied.

The figure below, figure 3.4, shows an RHI scan of the convective storm that occurred on 25th of January 2019 at 21:27:16 UTC along 277.4 Deg azimuth. In the figure the top left plot shows the reflectivity, top right is the differential reflectivity, bottom left is the unfolded velocity

plot and the bottom right is the hydrometeor identification index plot. The black line is plotted at an elevation of 6.7 Deg and the two areas of interest are shown in black crosses marked as A and B. The first area of interest is at a range of 13.4 km from the radar and is in the rain and hail region and the second area of interest is at a range of 7.8 km from the radar and in the heavy rain region.

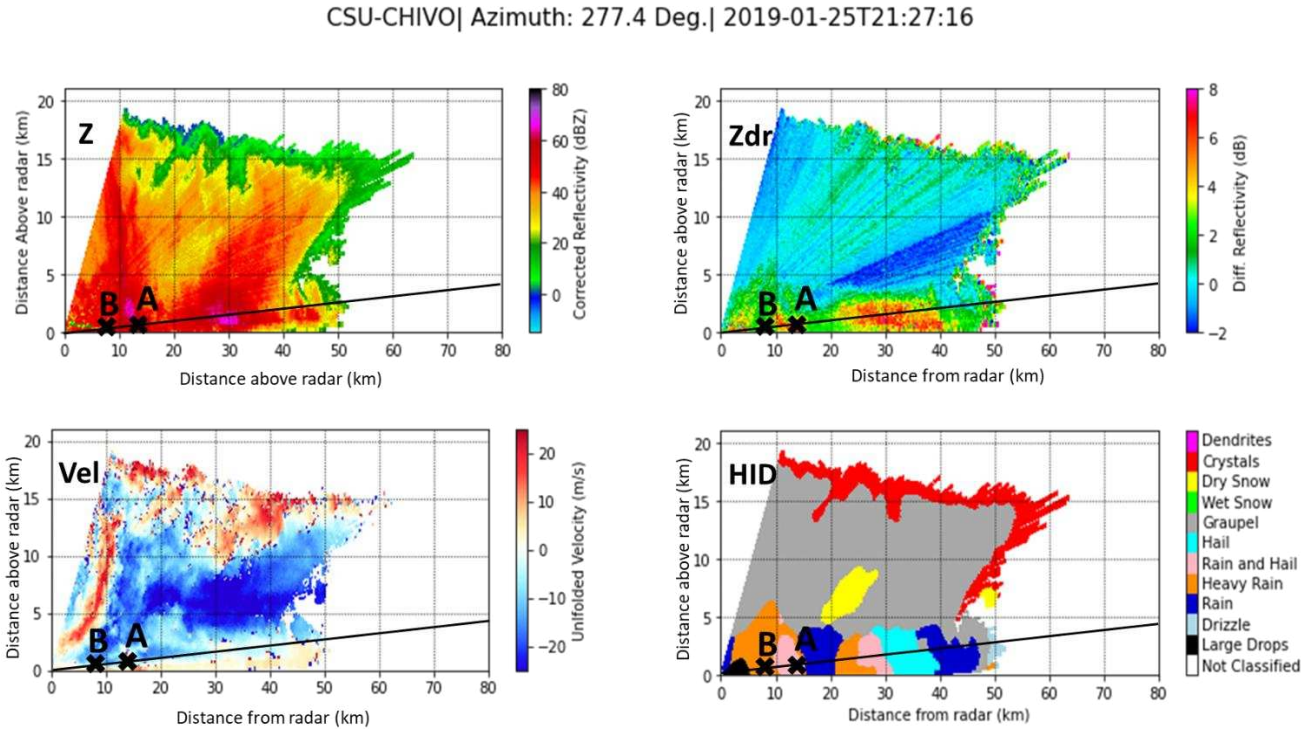

Figure 3.4 The RHI scan of the data observed on 25th January 2019 21:27:16 UTC along the 277.4 Deg. The black line is at plotted at an elevation of 6.7 Deg. The figure shows (from top left in clockwise direction) Reflectivity, Differential Reflectivity, Hydrometeor Identification Index and Velocity plots. The regions of interest A and B are at a range of 13.4 km and 7.8 km away from the radar respectively

- Rain and Hail Region (A):

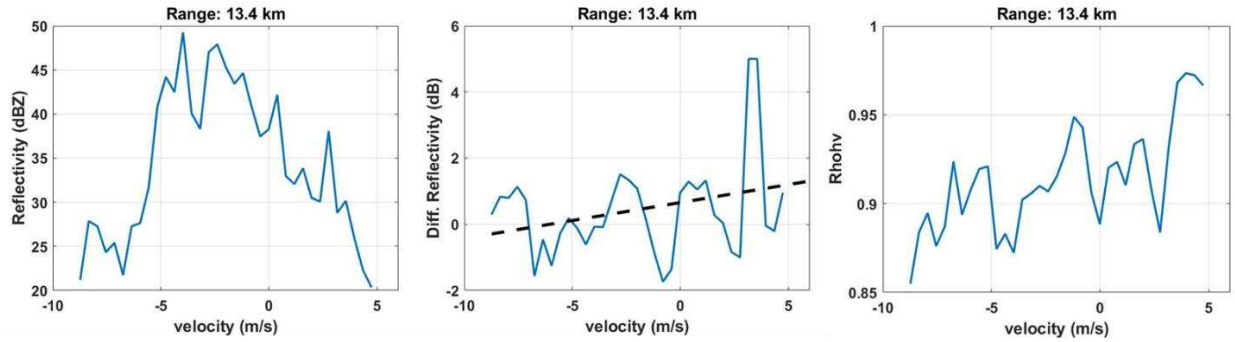

Figure 3.5 Dual-pol spectral analysis done at a range of 13.4 km away from the radar. The figure shows the spectral reflectivity (sZ), spectral differential reflectivity (sZdr) and spectral copolar correlation (sRhohv) from left to right in the region A shown in Figure 3.4

Dual-polarization spectra at the range of 13.4 km from the radar at an elevation of 6.7° along the 277.4° azimuth is done and the spectra of the reflectivity, differential reflectivity and the copolar correlation is shown above in image, figure 3.5. All the three spectra have been thresholded using spectral copolar correlation values where the lower copolar correlation values are masked to reduce the noise in the Doppler spectrum. From the spectra we can see that the reflectivity spectrum is broad and ranges from a value of 20 dBZ to 50 dBZ. Higher reflectivity values indicate hail and lower reflectivity values along the ranges of 20 dBZ to 30 dBZ indicate the presence of rain. Looking into the spectral differential reflectivity we see a positive slope of 0.14 $dB(ms^{-1})^{-1}$. Since the range we are looking at is in the rain and hail mixture as shown in the hydrometeor identification index in figure 3.5, we see a positive slope in the differential reflectivity spectra in figure 3.6 as discussed in [2]. The spectra of the copolar correlation shows values ranging from 0.85 to almost 1 in this region, which shows that the sizes of the particles present in this region is not uniform. This can be

due to the presence of rain and hail mixture in this region, where the raindrops are smaller is size compared to the hail.

- Heavy Rain Region (B):

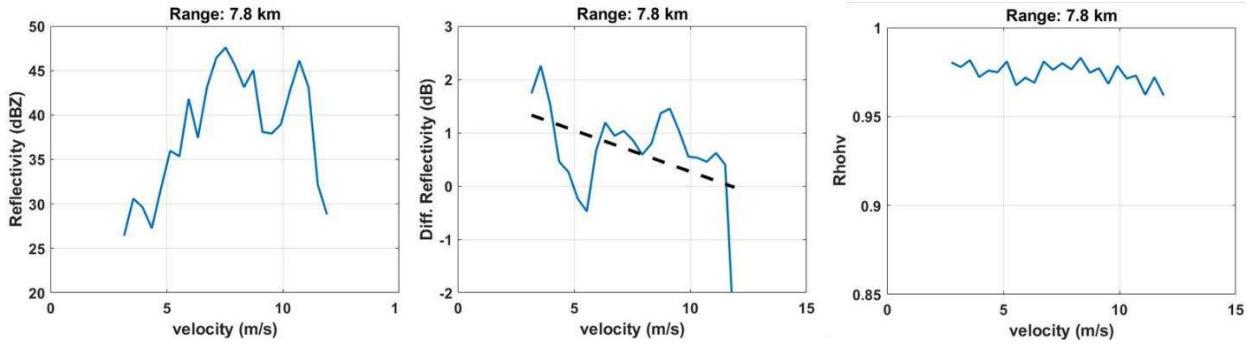

Figure 3.6 Dual-pol spectral analysis done at a range of 7.8 km away from the radar. The figure shows the spectral reflectivity (sZ), spectral differential reflectivity (sZdr) and spectral copolar correlation (sRhohv) from left to right in the region B shown in Figure 3.4

In the figure above, figure 3.6, the spectra of the dual polarized variables reflectivity, differential reflectivity and copolar correlation at a range of 7.8 km away from the radar at an elevation of 6.7° and along the 277.4° azimuth is shown. The spectral reflectivity is high in this region and the spectral differential reflectivity has low values in this region. The spectral differential reflectivity has a negative slope of -0.102 $dB(ms^{-1})^{-1}$.The spectral copolar correlation values are consistent with a very high with a value of almost 1. This high value shows that the particles in this region are of the same size and type. From the hydrometeor classification plot in figure 3.4 we can see that this is in the region of heavy rain. The high and consistent values in

the spectral copolar correlation also reflects in the spectral differential reflectivity where the fluctuation and the variance in the values are also minimal.

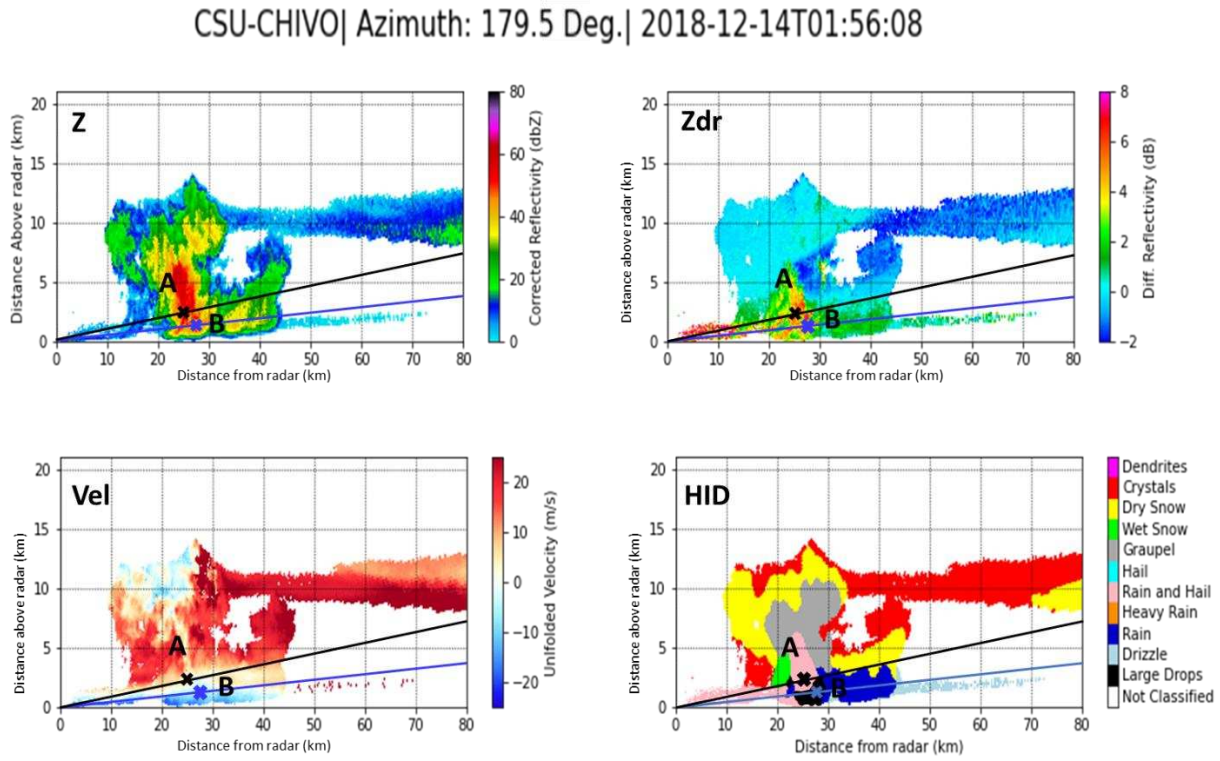

Figure 3.7 The RHI scan of the data recorded on 14th December 2018 01:56:08 UTC along the 179.5 Deg. The black line is at plotted at an elevation of 6.3 Deg and blue line at an elevation of 4°. The figure shows (from top left in clockwise direction) Reflectivity, Differential Reflectivity, Hydrometeor Identification Index and Velocity plots. The regions of interest A and B are in a range of 22 km and 28.4 km away from the radar respectively

The next convective storm that is used in this study was observed by the radar on 14th of December 2018 and the RHI plot for the time 1:56:08 UTC along the 179.5° azimuth is plotted and shown above in the figure 3.7. The figure shows the reflectivity, differential reflectivity, unfolded velocity and the hydrometeor classification as done by the DROPS2 code. The two regions of interest A and B are also shown the figure. The first region, denoted by a black cross

is in an elevation of 6.3° and at a range of 22 km from the radar. This is in the region of hail and rain mixture. The next region of interest, B, denoted by a blue cross, is in the region of large drops and is at an elevation of 4° at a range of 28.4 km from the radar.

- Rain and Hail Region (A):

The spectra are plotted for the region A shown in figure 3.7 which is in the region of hail and rain mixture. The spectra of the dual polarized parameters reflectivity, differential reflectivity and copolar correlation is shown in the figure 3.8 as sZ, sZdr and sRhohv respectively.

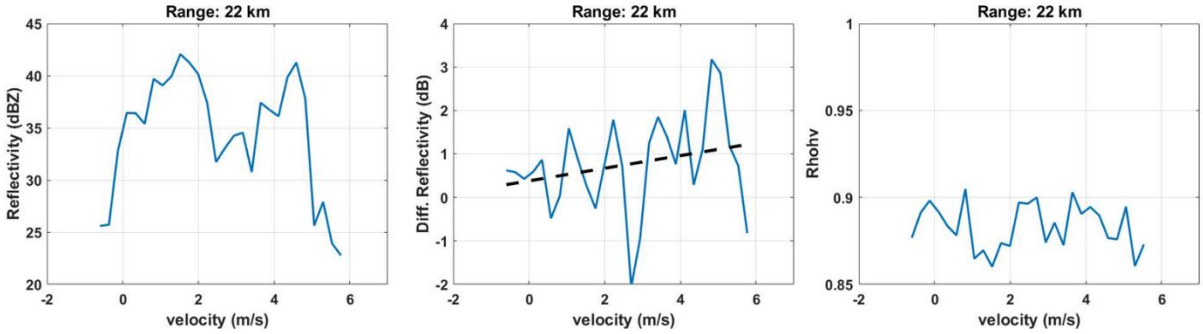

Figure 3.8 Dual-pol spectral analysis done at a range of 22 km away from the radar. The figure shows the spectral reflectivity (sZ), spectral differential reflectivity (sZdr) and spectral copolar correlation (sRhohv) from left to right of the region A shown in fig 3.7

Spectral thresholding of the dual polarized parameters is done using the copolar correlation values to eliminate noise present in the spectra. The differential reflectivity spectra has a lot of fluctuations and since this is in a region of mixture of hail and rain we see a positive slope of 0.15 $dB(ms^{-1})^{-1}$. The copolar correlation value is low around the range of 0.85 to 0.9, this can be because there is mixing of rain and hail, because the size of the rain drops and hail will differ, or it can also be present due to the melting of the hail into rain. The fluctuations in the differential

reflectivity might also be caused because of the low copolar correlation values. The reflectivity spectrum is broad and in the range of 20 dBZ to 40 dBZ, with traces of a bimodal spectrum. Since this region has a mixture of two hydrometeors, rain and hail, the bimodal spectrum can be representing both the hydrometeors. The spectra shown in figure 3.9 shows that the region has low copolar correlation value, which shows non-uniformity in the size of the particles present in this region and a positive slope in the differential reflectivity which corresponds to presence of hail as discussed in [2].

- Large Drops Region (B):

Next, we look into the other region of interest in figure 3.7, B, which is in the large drops region. The spectra of the dual polarized parameters that is reflectivity, differential reflectivity and copolar correlation are shown in figure 3.9 and is thresholded using the spectral copolar correlation value to reduce the noise present in the spectra. As we can see the spectra of the reflectivity is gaussian and high in the ranges of 40 to 50 dBZ which can be because the hail in the region above can be melting to form large drops. The spectral copolar correlation varies in value from 0.85 to 0.95 which shows some non-uniformity in the sizes of the rain drops in this region. The spectral differential reflectivity value is also high and a negative slope of 0.36 $dB(ms^{-1})^{-1}$ is seen. Size sorting of the drops could be happening here as discussed in [17] which could be a reason why a negative slope is seen in the differential reflectivity spectra. The low copolar correlation and the fluctuation in the value can be due to the presence of different sizes of drops of rain and this non-uniformity also causes fluctuations in the differential reflectivity spectra seen in the figure above.

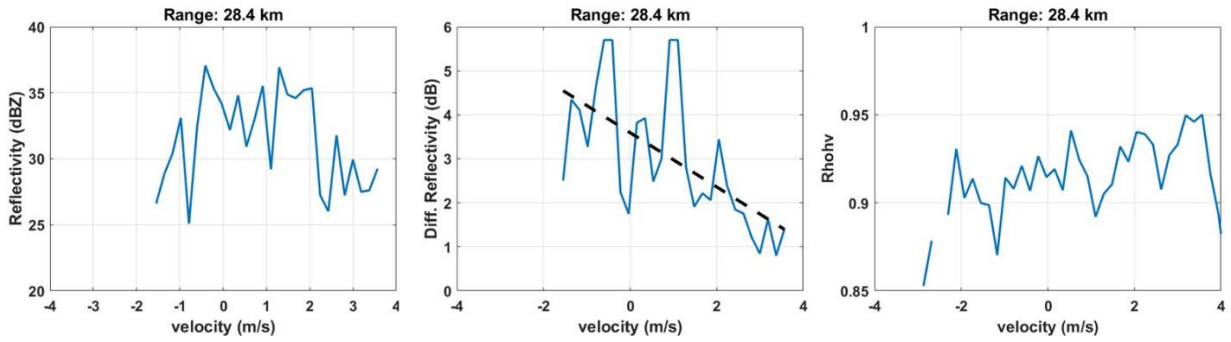

Figure 3.9 Dual-pol spectral analysis done at a range of 28.4 km away from the radar. The figure shows the spectral reflectivity (sZ), spectral differential reflectivity (sZdr) and spectral copolar correlation (sRhohv) from left to right of the region B shown in figure 3.7

3.3.2 Stratiform Storm Cases

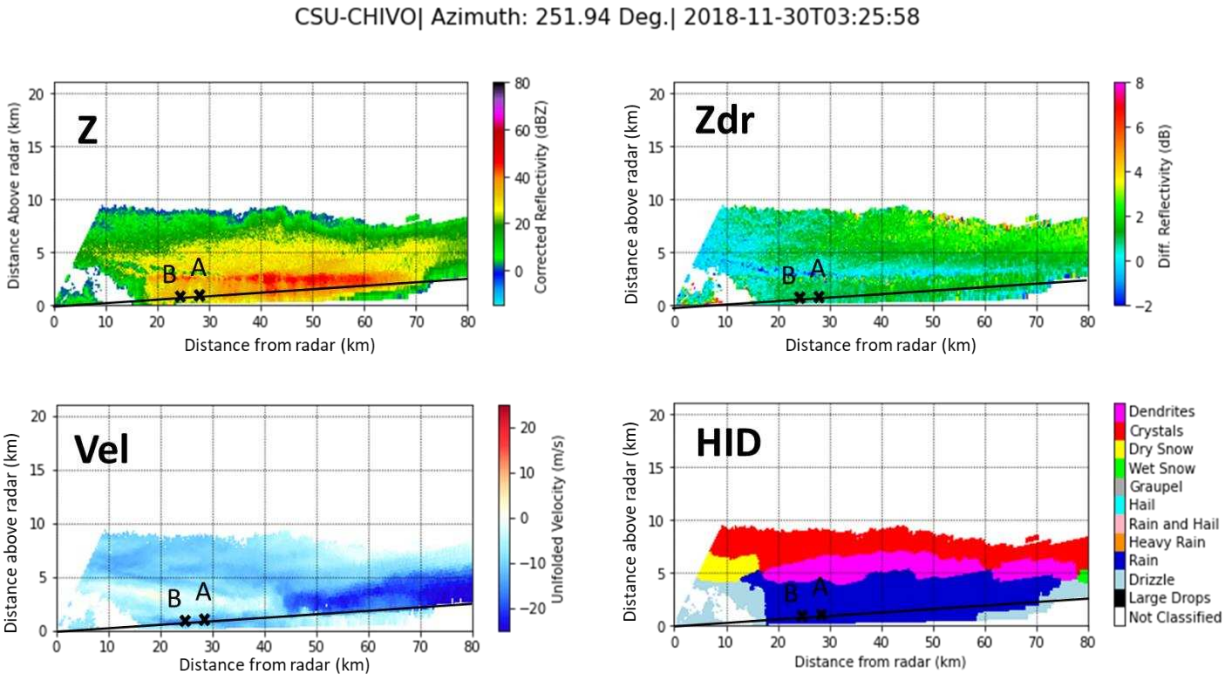

Figure 3.10 The RHI scan of the data observed on 30th November 2018 03:25:58 UTC along the 251.94 Deg. The black line is at plotted at an elevation of 5.3°, The figure shows (from top left in clockwise direction) Reflectivity, Differential Reflectivity, Hydrometeor Identification Index and Velocity plots. The regions of interest A and B are at a range of 27.4 km and 25 km away from the radar respectively

A stratiform rain case is categorized by the uniformity and steady rain in the region where the storm clouds are present, with no presence of vertical updrafts. They are simple rainstorms that cover a large area. In this section, we'll be looking into the stratiform rain case that occurred on 30th of November 2018 at 3:25:58 UTC. The figure above shows the RHI plot for the observed data on 30th November 2018 along the azimuth 251.94°. Since it is a stratiform storm we can see only the presence of rain below the melting layer in the hydrometeor classification plot above. The velocity plots also shows that there is no presence of vertical updraft in the storm. The two regions of interest, A and B, for which we will be spectral analyzing the radar moments is along the elevation of 5.3° at a range of 27.4 km and 25 km from the radar.

- Rain Region (A):

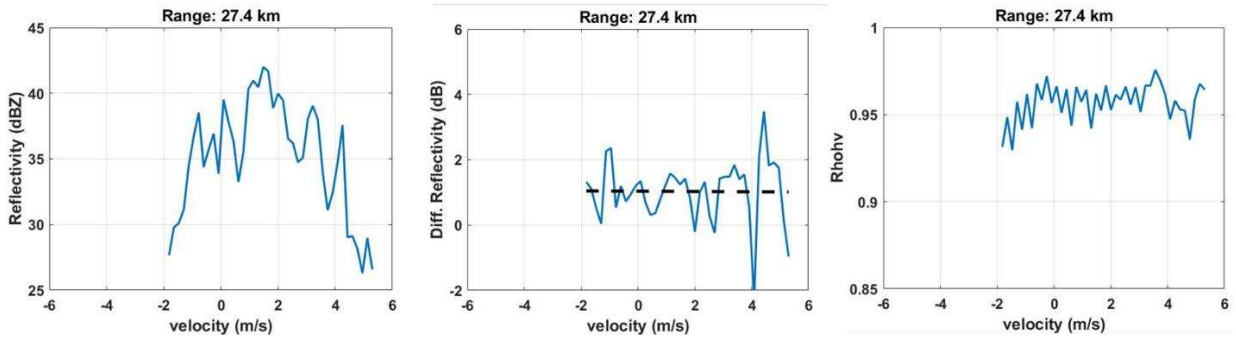

Figure 3.11 Dual-pol spectral analysis done in the range of 27.4 km away from the radar. The figure shows the spectral reflectivity (sZ), spectral differential reflectivity (sZdr) and spectral copolar correlation (sRhohv) from left to right of the region A shown in Fig 3.11

The figure above, figure 3.11, shows the spectra of the dual-pol radar moments reflectivity, differential reflectivity and copolar correlation of the first region of interest, A, at a range of 27.4 km from the radar and at an elevation of 5.3°. The reflectivity spectrum at this range is gaussian

and the differential reflectivity spectra is a consistent flat slope. The copolar correlation spectrum is high with values almost 1. This shows that the uniformity in particles present and the high value of the copolar correlation reflects on the differential reflectivity, leading to a flat slope that is seen. A slope in the differential reflectivity region is seen if there is a size sorting of the particles is occurring. The gaussian spectrum seen in the reflectivity is in accordance with the reflectivity spectrum seen in a typical rain region.

- Rain Region (B):

The next area of interest, B, is also in a rain region at a distance of 25 km from the radar at the same elevation of 5.3° as shown in figure 3.10. The spectra of the dual polarized parameters like reflectivity, differential reflectivity, and copolar correlation is shown below in the figure 3.12.

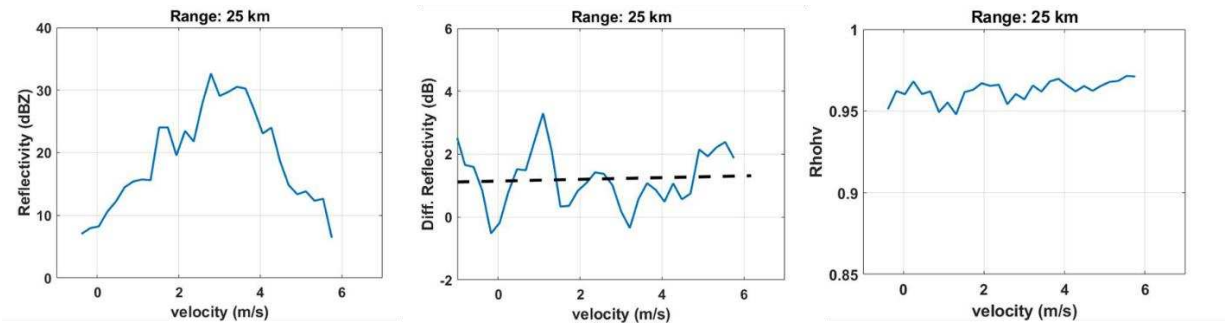

Figure 3.12 Dual-pol spectral analysis done in the range of 25 km away from the radar. The figure shows the spectral reflectivity (sZ), spectral differential reflectivity (sZdr) and spectral copolar correlation (sRhohv) from left to right of the region B shown in Fig 3.10

The copolar correlation is very high, which shows that the particles in this range are homogeneous and similar in size. The reflectivity spectrum is gaussian and which shows that

rain is present in this region as shown in the hydrometeor classification done by the DROPS2 code and plotted above in figure 3.10. A region with just rain has a narrow, gaussian, spectrum in the reflectivity with a high copolar correlation value and a flat slope in the differential reflectivity spectrum which are seen in the dual polarized spectra plotted for the two region of interest and shown in figure 3.12 and 3.13.

3.3.3 Error Analysis Using Standard Deviation

In this section we will investigate the error analysis done for the spectral reflectivity and spectral differential reflectivity for each of the cases discussed in the previous section. For this purpose, the standard deviation of the spectral reflectivity and spectral differential reflectivity is found with respect to the adjacent range gates nearby to the range where the spectral analysis has been performed in the previous section.

Six different range gates surrounding the range bin for which the error analysis is done is chosen to find the standard deviation of the spectral reflectivity and spectral differential reflectivity. The six points chosen are shown in an illustrated three-dimensional version in the image below, figure 3.13 and are as described as follows. The figure depicts the side view of the radar which will be similar to the RHI plots that are plotted in the previous section and the top view depicts the radar if viewed from an aerial shot. The black star symbol on the figure represents the radar's location and the black cross denotes the range bin for which error analysis is performed and will be called as the original point. The adjacent range gates which are chosen to calculate the standard deviation of the original point are represented in numbers ranging from one to six. The first point (marked in the figure as a red 1) is chosen as one range bin preceding the original point and is along the same elevation and azimuth as the original point. The second

point (marked in the figure as a red 2) is chosen as the subsequent range bin to the original point and is also along the same elevation and azimuth as the original point. The third point (marked in the figure as a blue 3) is chosen at the same range distance from the radar as the original point and along the same azimuth plane but at a 0.1° higher elevation than the original point. The fourth point (marked in the figure as a blue 4) is chosen at the same range distance from the radar and along the same azimuth plane as the original point but at a 0.1° lower elevation than the original point. The fifth point (marked in the figure as a green 5 in the right image) is chosen along the same range distance from the radar and at the same elevation angle as the original point but in 0.1° azimuth above the original point's azimuth angle. And lastly, the sixth point (marked in the figure as a green 6 in the left image) is chosen along the same range distance from the radar and at the same elevation angle as the original point but in 0.1° azimuth angle lower than the original point. These are the six points that are chosen to calculate the standard deviation of the spectral reflectivity and spectral differential reflectivity at the original point.

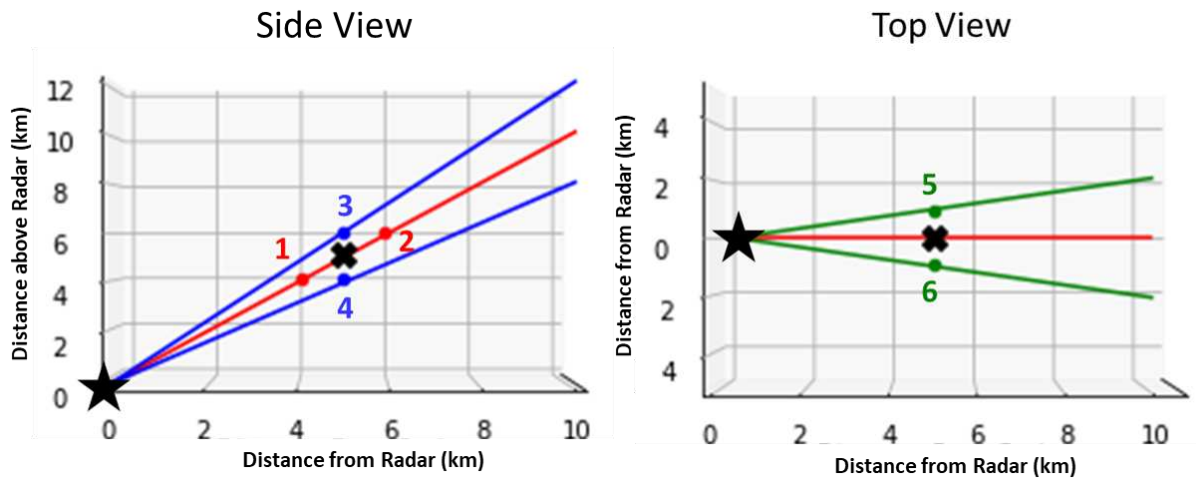

Figure 3.13 Range bins selected to calculate the standard deviation for error analysis. The black cross denotes the region for which the error analysis is performed, the numbered points are the range bins selected to calculate SD and the star denotes the location of the radar

- For Mixture of Hail and Rain Region Cases:

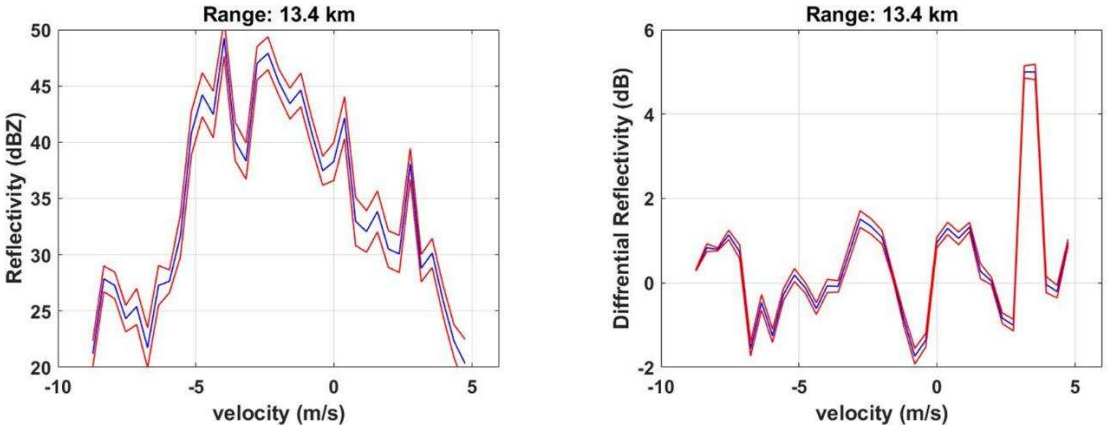

Figure 3.14 The spectral reflectivity (sZ) and spectral differential reflectivity (sZdr) with ±1 Standard Deviation (SD) for case A shown on figure 3.4

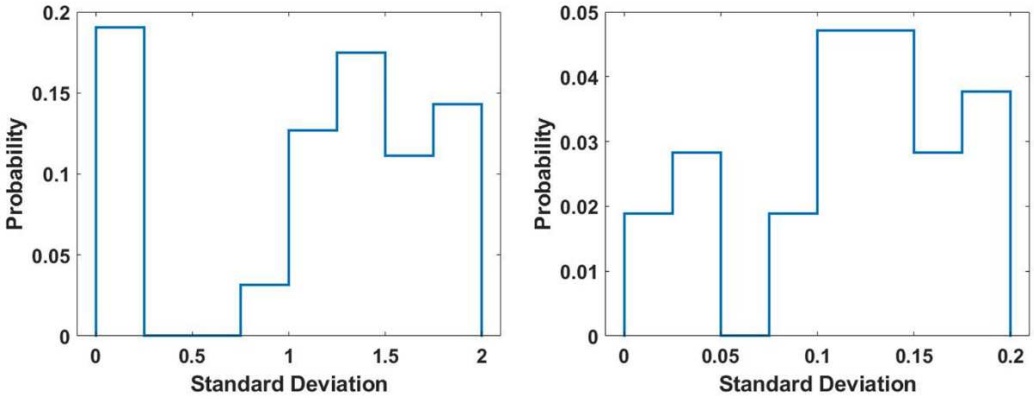

Figure 3.15 The Histogram for standard deviation of spectral reflectivity (sZ) and spectral differential reflectivity (sZdr) is shown above for case A shown on figure 3.4

The standard deviation (SD) is calculated for the spectral reflectivity and spectral differential reflectivity by considering the 6 range gates adjacent to the original range gate as mentioned previously for case A in figures 3.4 and 3.7. After the SD is calculated, the ±1 SD is plotted along with the sZ and sZdr of the original point. This is shown in figures 3.14 and 3.16.

In figure 3.15 and 3.17 the blue line is the actual spectral reflectivity or spectral differential reflectivity whereas the red line is the ±1 standard deviation calculated from the blue line. The histograms of the standard deviation calculated is shown in figure 3.15 and 3.17 for the cases marked as A in figures 3.4 and 3.7 respectively.

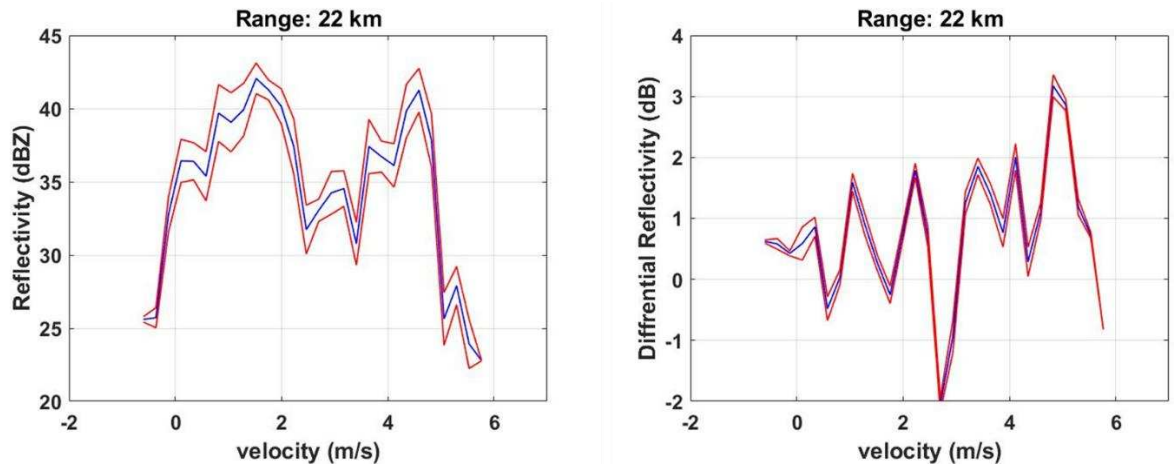

Figure 3.16 The spectral reflectivity (sZ) and spectral differential reflectivity (sZdr) with ±1 Standard Deviation (SD) for case A shown on figure 3.7

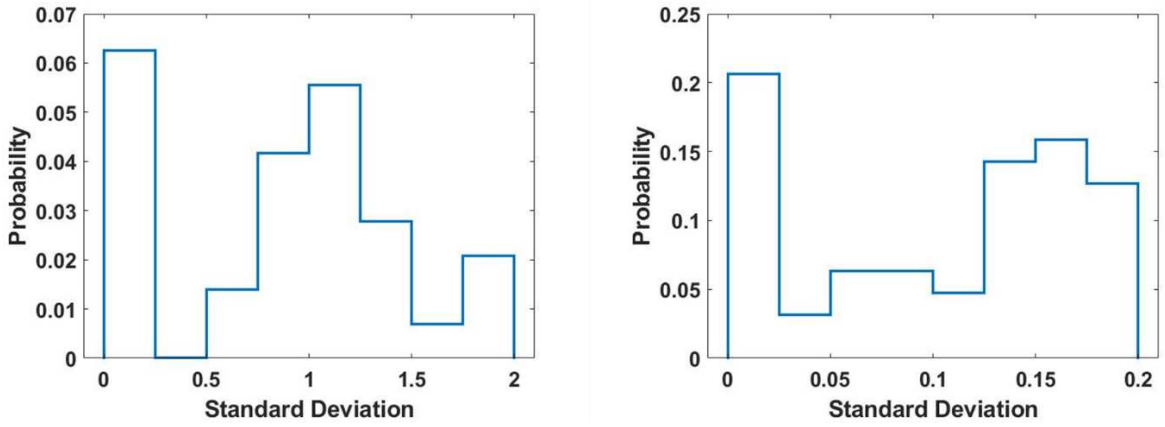

Figure 3.17 The Histogram for standard deviation of spectral reflectivity (sZ) and spectral differential reflectivity (sZdr) is shown above for case A shown on figure 3.7

- For Heavy Rain Region Case:

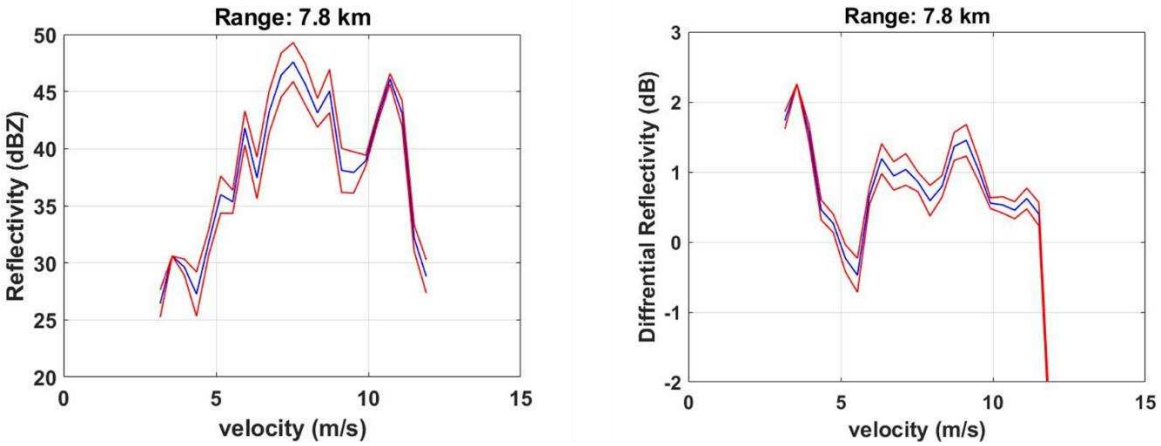

Figure 3.18 The spectral reflectivity (sZ) and spectral differential reflectivity (sZdr) with ±1 Standard Deviation (SD) for case B shown on figure 3.4

The standard deviation (SD) is calculated for the spectral reflectivity and spectral differential reflectivity for the region marked as B in figure 3.4. After the SD is calculated, the ±1 SD is plotted along with the sZ and sZdr of the original point. This is shown in figure 3.18. The histogram of the SD calculated for sZ and sZdr is shown in figure 3.19.

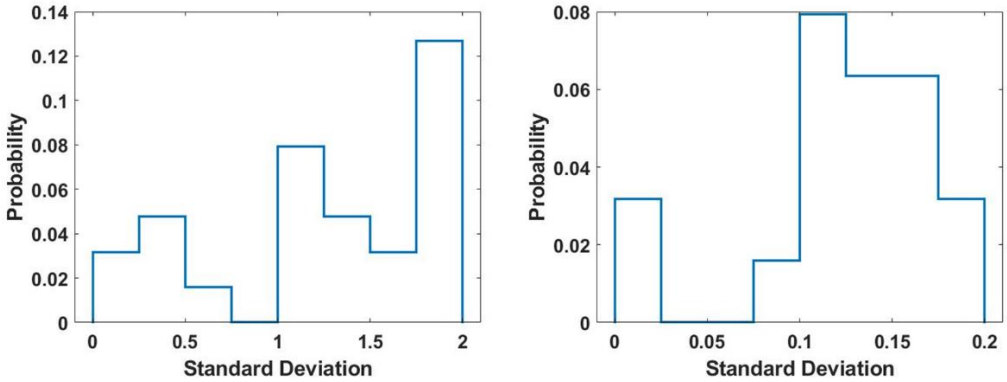

Figure 3.19 The Histogram for standard deviation of spectral reflectivity (sZ) and spectral differential reflectivity (sZdr) is shown above for case B shown on figure 3.4

- For Large Drops Region Case:

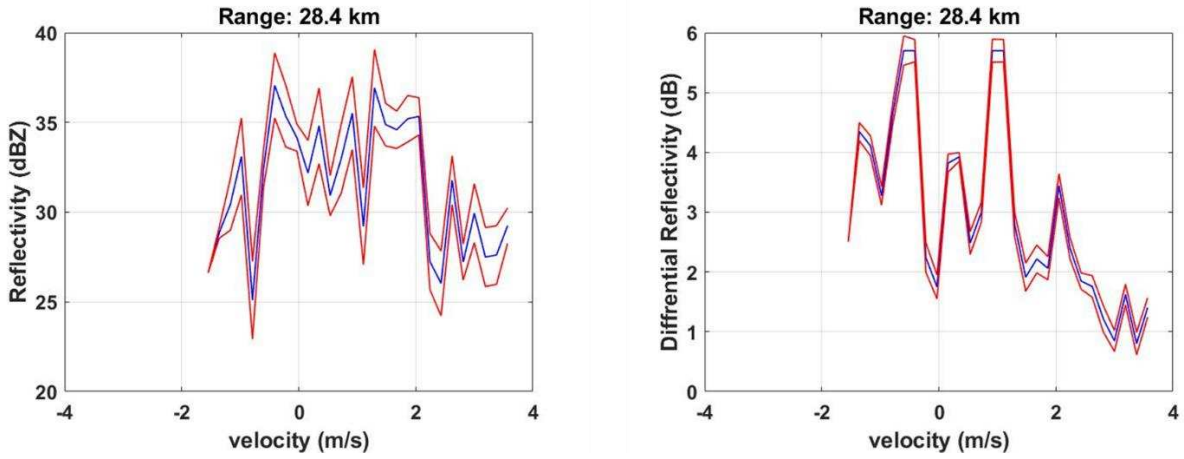

Figure 3.20 The spectral reflectivity (sZ) and spectral differential reflectivity (sZdr) with ±1 Standard Deviation (SD) for case B shown on figure 3.7

The standard deviation (SD) is calculated for sZ and sZdr for the region marked as B in figure 3.7. After the SD is calculated, the ±1 SD is plotted along with the sZ and sZdr of the original point, this is shown in figure 3.20. The histogram of the SD calculated for sZ and sZdr is shown in figure 3.21.

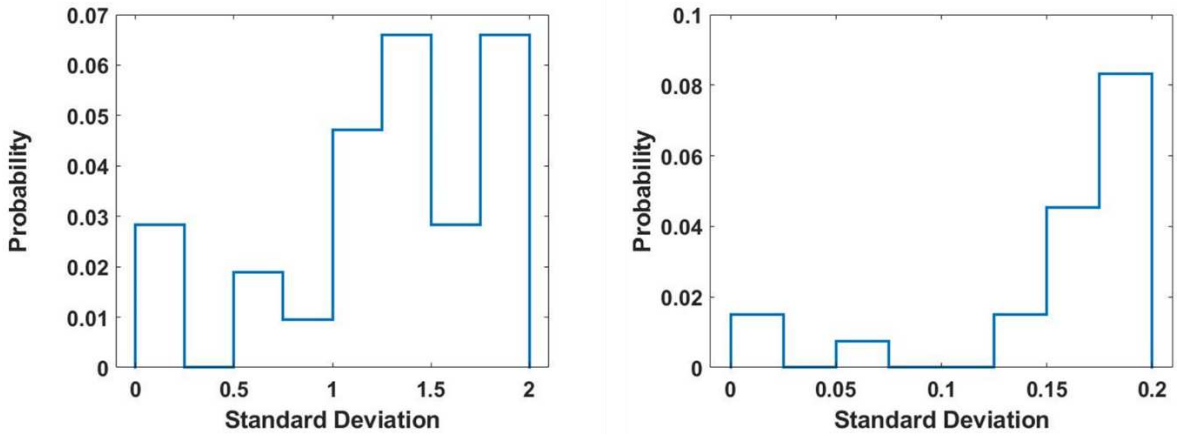

Figure 3.21 The Histogram for standard deviation of spectral reflectivity (sZ) and spectral differential reflectivity (sZdr) is shown above for case B shown on figure 3.7

- For Rain Region Cases:

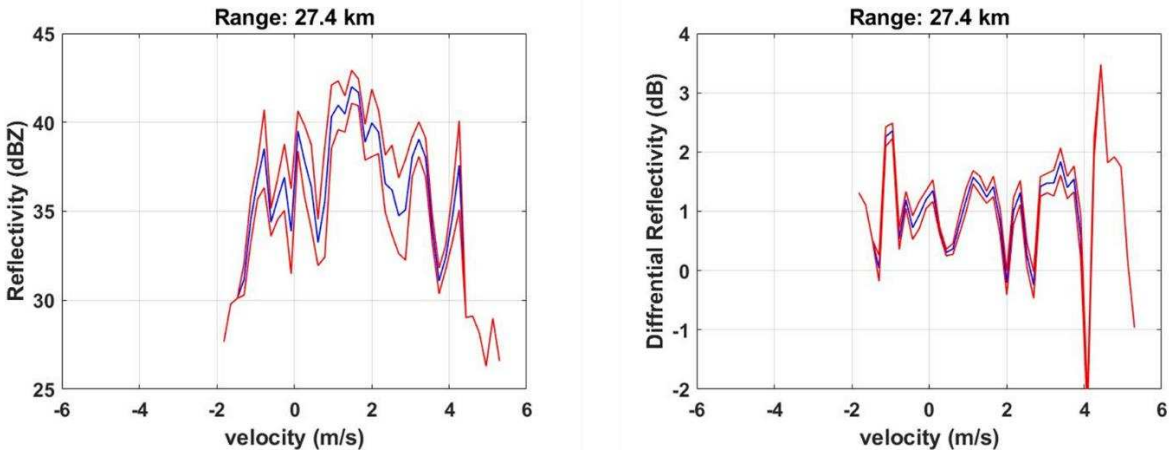

Figure 3.22 The spectral reflectivity (sZ) and spectral differential reflectivity (sZdr) with ±1 Standard Deviation (SD) for case A shown on figure 3.10

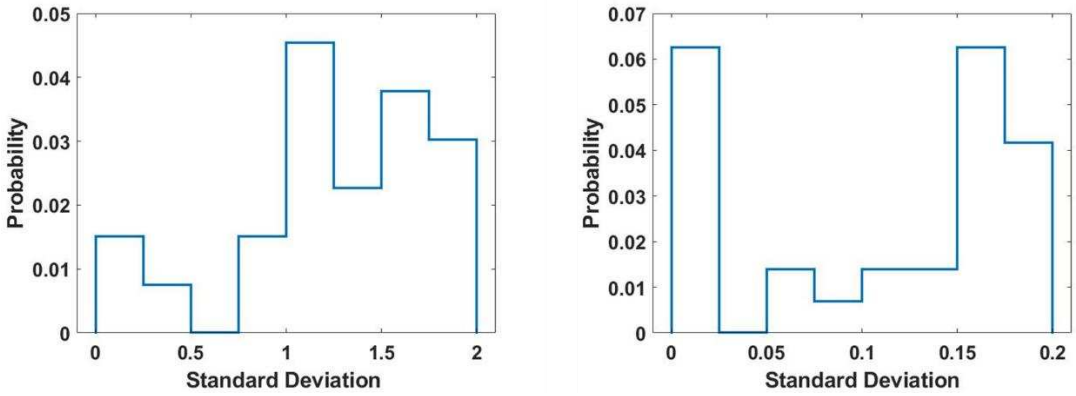

Figure 3.23 The Histogram for standard deviation of spectral reflectivity (sZ) and spectral differential reflectivity (sZdr) is shown above for case A shown on figure 3.10

The standard deviation (SD) is calculated for the spectral reflectivity and spectral differential reflectivity adjacent range gates of the original range gate for cases A and B in figure 3.10. After the SD is calculated, the ±1 SD is plotted along with the sZ and sZdr of the original point for both the cases and is shown in figures 3.22 and 3.24, respectively.

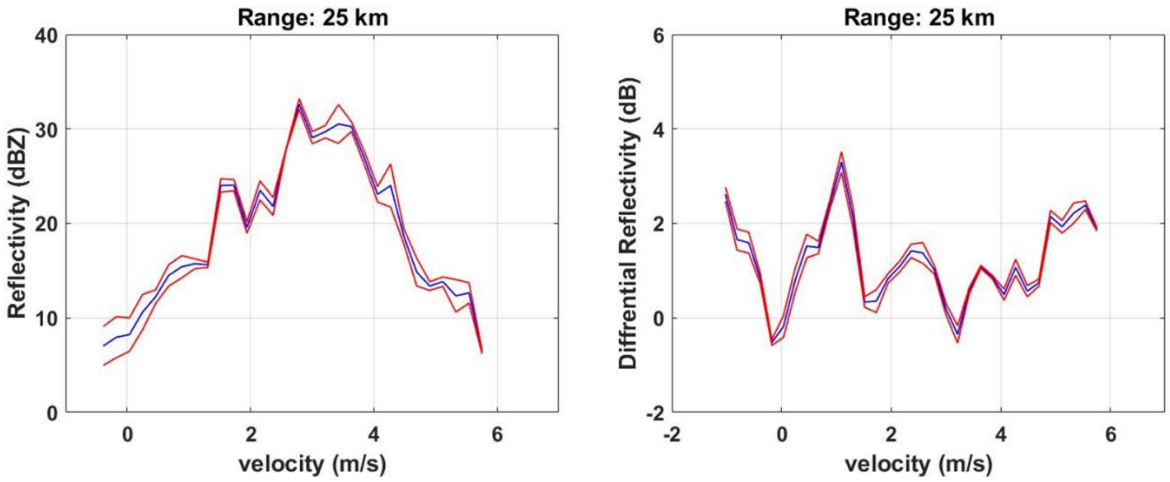

Figure 3.24 The spectral reflectivity (sZ) and spectral differential reflectivity (sZdr) with ±1 Standard Deviation (SD) for case B shown on figure 3.10

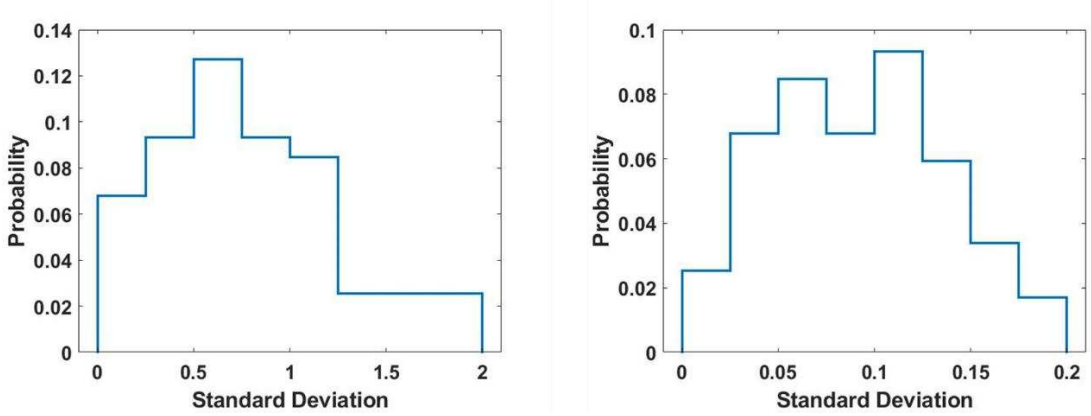

Figure 3.25 The Histogram for standard deviation of spectral reflectivity (sZ) and spectral differential reflectivity (sZdr) is shown above for case B shown on figure 3.10

The histograms of the calculated SD for the region of interest A and B represented in Fig 3.10 is shown in figures 3.23 and 3.25 respectively. From the standard deviations that were calculated for the various cases we can see that the maximum standard deviation of the spectral reflectivity is ±2 dBZ and maximum standard deviation of the spectral differential reflectivity is ±0.2 dBZ. This is shown in the table below.

TABLE 3.2 Maximum standard deviation values for spectral reflectivity and differential reflectivity

Metrics	Reflectivity (dBZ)	Differential Reflectivity (dB)
Maximum Standard Deviation	±2	±0.2

3.3.4 Discussion

The different hydrometeor regions studied in section 3.3.1 and 3.3.2 helps in identifying the microphysical properties of the particles present in the region. From these sections it can be interpreted that in a region of mixture of hail and rain, the reflectivity spectrum can be bimodal, as seen in figure 3.8, because of the presence of two different type of hydrometeors. The differential reflectivity spectra has a positive slope which shows that size sorting of the different hydrometeor is taking place, as seen in figure 3.5 and 3.8. The spectral copolar correlation will be relatively low because of the non-uniformity of the particles present. In the region of large drops, shown in figure 3.9, the reflectivity spectrum is gaussian, the differential reflectivity spectrum has a negative slope indicating size sorting of the different sizes of large raindrops are occurring. In the heavy rain region, shown in images 3.6, the spectral copolar correlation values are high which can indicate the presence of uniform drops of rain and the spectral differential reflectivity values have low values with a negative slope. Low differential reflectivity indicated the presence of rain. In the rain regions that were analyzed in the stratiform storms, from figures 3.11 and 3.12 the reflectivity spectrum is gaussian and the differential reflectivity has a low

value with no slope which indicates the presence of rain. The copolar correlation coefficient is very high indicating uniformity in the sizes of the particles present in this ranges.

In the next section 3.3.3, error analysis of the range gates which are spectral analyzed is done. The purpose of this analysis is to calculate the deviation of the spectrum and look for errors which may not give an accurate spectrum of the radar moments hence tampering with the conclusion we arrive at from the spectral decomposition of the radar moments. From the standard deviation that was calculated we can see that the average standard deviation for the reflectivity is $\pm$1.5 dBZ and the average standard deviation for the differential reflectivity is $\pm$0.15 dB. The maximum deviations occur in the mixture of hail and rain cases, seen in figures 3.14 and 3.16 and the standard deviation is the least in the stratiform rain cases as seen in figures 3.22 and 3.24. This difference in the standard deviations can be due to the presence of different hydrometeors or particles of different sizes present in the region. If the standard deviation values are higher than $\pm$2.5 dBZ for the reflectivity spectrum and $\pm$0.25 dB for the differential reflectivity spectrum then the region the region is not suitable for spectral analysis, since the error is high. From the error analysis, we can conclude that the range gates considered for the spectral decomposition was a region with minimal error.

3.4. CONCLUSION

In this chapter, the spectral analysis of the dual-pol radar moments like reflectivity, differential reflectivity and copolar correlation was performed on different hydrometeors present in convective and stratiform storms. Spectral polarimetry helps in identifying and classifying the microphysical properties of the hydrometeors present in these precipitation events to gain a better understanding of them. The error analysis is done by calculating the standard deviation of

the reflectivity and differential reflectivity from the adjacent range gates to determine the accuracy of the spectra and it is seen that the standard deviation is moderate which shows that the region of study is not an anomaly. Microphysical properties of the storm can be characterized from spectral decomposition and analysis which is essential to learn more about the storms.

CHAPTER 4 : OPTIMIZATION OF PTDM TECHNIQUE TO SEPARATE CLOUD AND DRIZZLE DATA USING HIGH PERFORMANCE COMPUTER

In the previous chapter, we have seen how the microphysical properties of the different types of precipitation can be studied using spectral polarimetry. In this chapter, the microphysical properties of the cloud and drizzle is studied using weather radar data. The data collected from weather radar observing cloud and drizzle may have resolution volumes in which both cloud and drizzle are present. So, we need to first separate the cloud and drizzle moments from the radar data to interpret the data accurately. Previously, many methods have been used to separate these moments using the Doppler spectra skewness and properties which have been analyzed to separate the cloud and drizzle layer, but one of the drawbacks of this method is that when drizzle is stronger it is ineffective. In [5] they have discussed a PTDM (**P**arametric **T**ime **D**omain **M**ethod) approach to effectively separate the cloud and drizzle data by analyzing the Doppler spectra and does not assume any predetermined properties of the cloud or drizzle data. This method processes the data in the time domain to obtain the estimates of the cloud and drizzle data. The only drawback of this method arises in the time taken to process the data. The PTDM technique processes the data twice to retrieve all the properties and estimates of the observed radar data signal which makes it effective but also time consuming. The time taken to process one hour worth of data is almost two to four hours depending on the data. To overcome this drawback, we can deploy the PTDM codes on a **H**igh-Performance **C**omputer (HPC) and process the data on the HPC which reduces the time significantly, almost by 60%, hence increasing the overall time latency of this technique. In this chapter we'll be looking into some of the details of PTDM technique and how it is deployed on the HPC.

4.1. INTRODUCTION

The need to separate the cloud and drizzle data and study them arises because clouds are important for earth's radiation control, especially boundary layer clouds which have a higher albedo and hence a larger cloud layer is required to manage the radiation reaching the ground, which is also discussed in Hartmann et al [18] and Hanh and Warren [19]. Drizzle mostly occurs from the lower cloud layers, and frequent drizzle can reform the boundary layer cloud structures which will in turn impact the level of earth's radiation, and these are discussed in Wood [20], Yamaguchi et al [21] and Zhou et al [22]. Given the fact that recent climate change and global warming has increased the overall temperature of the earth, this can also impact the process of cloud formation and drizzle which will affect the earth's radiation levels. Since all these steps are interdependent on each other we need to study them to understand the nuances and the impact it will have on the future climate of earth.

The issue arises because we don't have an estimate of the how the increase in temperature of the earth due to global warming and radiation will impact the cloud and hence precipitation processes. For this reason, we need to have a better understanding about the cloud formation processes which leads to drizzle. The life cycle of the cloud can be explained in five steps. First activation, then condensation followed by evaporation of vapor, leading to collision-coalescence and in the end sedimentation. The most important of all these steps is the collision-coalescence because this is where the cloud turns into precipitation, and this is driven by two mechanisms: auto conversion and accretion. These mechanisms can also affect the amount of rainfall in regions where the cloud is present as discussed in Weber and Quaas [23], Takahashi et al [24]. In the wake of global warming, the presence of aerosols in the atmosphere has also increased and

also there is a steady rise in the temperature the earth's surface, these can also influence the cloud formation process as discussed in Jing et al [25] and Golaz et al [26] respectively.

The importance of getting a better understanding of the collision-coalescence process to study the cloud and drizzle properties and the difficulty in separating the two data from the observed radar data is explained in [5]. In that paper, they have also discussed some of the previous techniques that were implemented and why they fail in effectively separating the two data, cloud and drizzle is also explained. So, they propose the PTDM methodology which had been implemented to separate the precipitation from the ground clutter by Nguyen et al [4], to separate the cloud and drizzle data. The drawback of the previously mentioned Doppler spectra technique to separate cloud and drizzle data arises when these two data overlap with each other. The PTDM methodology can handle data where two different data overlap with each other, and so it can overcome the drawbacks of the Doppler spectra techniques, hence the PTDM technique only strengthens the Doppler spectra analysis method. The PTDM technique does not assume any prior properties of cloud or drizzle from the observed radar data during the processing and estimates the properties during processing hence effectively separating the cloud and drizzle data.

The PTDM method, proposed by [5] is discussed in detail in section 4.2, one of the major setbacks of this method is the time taken to process the data. Though this method is successful it takes a long time to process the data. The processed data should be further analyzed to study the impacts it has on cloud formation process. To overcome this time latency, we can optimize the run time by deploying it on an HPC system. HPCs have a higher computational power which can be used to process and analyze either large amounts of data or to run programs that are computationally intensive. This is the case here for the PTDM technique where we need to

process and analyze a large amount of data collected and also run a computationally intensive process PTDM which process the dataset twice to effectively separate cloud and drizzle data. The process of adapting the PTDM codes to run successfully on the HPC, testing and deployment of the codes on the HPC server is discussed further in detail in section 4.3 of this chapter.

4.2. PTDM MODEL

The PTDM technique proposed by Nguyen et al [4] is utilized for separation of cloud and drizzle data by [5]. If the variable V represent the received voltage samples of the observed radar data, then,

$$R_V = VV^H \tag{26}$$

$$R = E(VV^H) \tag{27}$$

where equation 26 represents the sample covariance matrix and equation 27 gives the covariance matrix. The H in V^H is used to represent the transpose conjugate. The scattering recorded in the radar resolution volume of the individual signals have similar statistical properties and so the joint probability density function (pdf) of this signal can be assumed to be a normal distribution with a zero mean, as discussed in [1].

In [5], the Doppler spectra of the received cloud and drizzle data is assumed to have a Gaussian distribution, so they characterize the radar moments as explained in chapter 2, with P representing the received signal power, v representing the mean velocity and W representing the spectral width. As explained previously, if there is only cloud data present and no drizzle then the distribution can be approximately represented by a single gaussian distribution, but if both

cloud and drizzle data is present we can represent it as a summation of both the gaussian distributions and this is expressed in equation 28, where the spectra (S) of the observed radar data containing both the cloud and drizzle content is shown as a function of the radar moments.

$$S(v) = \frac{P_1}{\sqrt{2\pi W_1}} \exp\left[\frac{1}{2}\left(\frac{v - v'_1}{W_1}\right)^2\right] + \frac{P_2}{\sqrt{2\pi \sigma_2}} \exp\left[\frac{1}{2}\left(\frac{v - v'_2}{W_2}\right)^2\right] + \frac{2T_s}{\lambda} W_N^2 \qquad (28)$$

where the subscripts 1 and 2 denote cloud and drizzle respectively. Equation (28) can be seen as where the first part is the gaussian function of the cloud, the second part is the gaussian function of the drizzle and the constants, T_s is the sampling time period of the radar, λ is the wavelength used and W_N^2 is the power of the noise present. Using this equation, we can now arrive at the measured radar signal's covariance matrix, $R_{k,l}$ shown below in equation 29, as discussed in Nguyen et al [26],

$$R_{k,l} = P_1 \exp\left[-\frac{8\pi^2 W_1^2 (k-l)^2 T_S^2}{\lambda^2}\right] \exp\left[-j\frac{4\pi v'_1 (k-l) T_s}{\lambda}\right] + P_2 \exp\left[-\frac{8\pi^2 W_2^2 (k-l)^2 T_S^2}{\lambda^2}\right] \exp\left[-j\frac{4\pi v'_2 (k-l) T_s}{\lambda}\right] + W_N^2 \delta(k-l) \qquad (29)$$

where the limits on k and l is given as $k, l = 1, 2, 3, \ldots . N$, and denote the matrix elements' index and N is the sample size. Since both in equations 28 and 29 we assume both cloud and drizzle data are present the function of the spectra and the covariance matrix is shown as a combination of two Gaussian distribution, this is known as a two-echo PTDM model. These above functions can be altered to accommodate the presence of either cloud or drizzle data present, where they will no longer be a two-echo but a one-echo PTDM model.

Now, let us assume a variable, α, which contains all the radar spectral moments of cloud, drizzle and noise. We can define that variable as shown in equation 30, we can estimate this variable by using a standard maximum-likelihood technique.

$$\alpha = [P_2, v'_2, W_{2,} P_1, v'_1, W_1, \sigma_N^2] \tag{30}$$

The PTDM technique is carried out in steps, as proposed in [5], and the steps are shown in the flowchart below in figure 4.1. Primarily all the received radar signals are compared with the one-echo PTDM model to determine whether the radar resolution volume has more than one-echo model, i.e., either cloud or drizzle. The cloud base height which is provided by ceilometers, limits the scene's determination, which depends on the how well a Gaussian distribution fits the signal. The signal is classified as a cloud signal if the radar gates are above the cloud base height and as drizzle if the radar gates are below this cloud base height if the distribution is unimodal. The two-echo PTDM model is applied only if a non-Gaussian spectrum is detected above the cloud base height because this warrants for a separation of the cloud and drizzle layer. In the end we arrive at the estimated spectral moment value for the cloud and/or drizzle at each individual radar gate. This approach effectively separates the cloud and drizzle moments from each other. The developed PTDM technique is tested using the data collected from the ARM CAP-MBL campaign. The results of the testing of the PTDM codes after deploying the codes on the high-performance computer is shown in the next section.

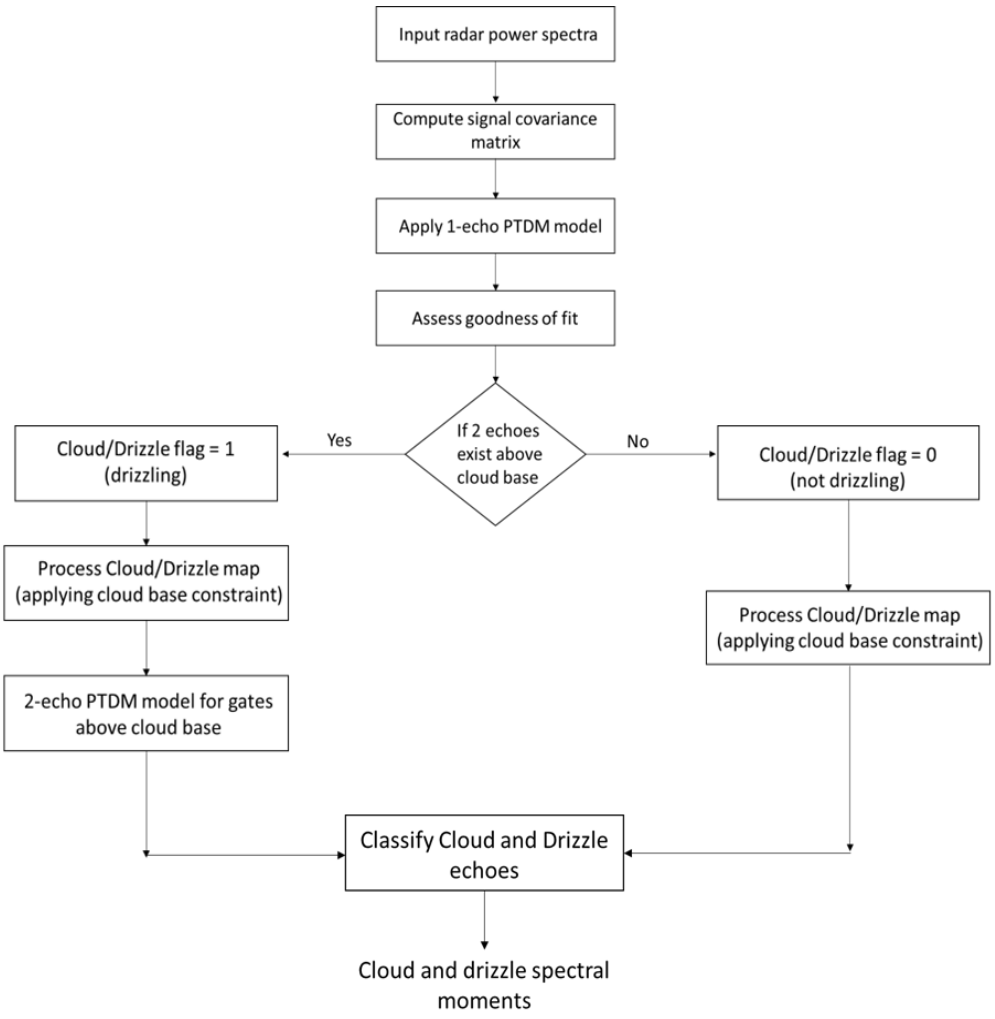

Figure 4.1 Flowchart depicting the various steps involved in the retrieval of the Cloud and Drizzle spectral moments using the PTDM technique [5]

This ARM CAP-MBL campaign dataset is utilized in this chapter for testing the PTDM codes. The SWACR is located in the ARM Mobile Facility (AMF) deployed on the Graciosa Island in the Portugal archipelago called as Azores. The main goal of this campaign was to collect cloud, aerosol, and precipitation data present in the Azores region. The AMF is in an ideal region to sample the change from overcast stratocumulus regime during the spring to the broken trade cumulus regime during the summer. This dataset has proven to be valuable when studying the processes that control the radiative properties of the boundary layer clouds and their

microphysics. The operating parameters of the ARM W-Band cloud radar which collected the dataset which is used here to test the PTDM codes is given below in Table 4.1.

TABLE 4.1 Calibration of ARM W-band Cloud Doppler Radar used in PTDM testing [5]

Parameter	**Specifications**
Frequency	95 GHz
Pulse repetition frequency	10 kHz
Gate spacing	42.857 m
Unambiguous velocity	7.885 ms^{-1}
Velocity resolution	0.0616 ms^{-1}
Integration Time	2.048 s
FFT Length	256
No. of spectral averages	160
Peak transmit power	1500 W
Antenna beamwidth	0.32°

4.3. OPTIMIZATION AND IMPLEMENTATION ON HPC

4.3.1. Introduction

High-Performance Computing (HPC) plays a significant role when we need to process large datasets and/or complex computing algorithms. An HPC can perform computing much faster and in parallel because of the large number of resources available on the system which are much advanced than the resources available on a personal computer or laptop. HPCs are usually a supercomputer or clusters of computers with high processing power. A personal system like a

laptop or PC usually has a 8-10 GB RAM with quad-core, 2 GHz processor frequency with a memory space of 512GB or 1TB, with these specifications we can do approximately 2 billion calculations per second. Even if this does look like a vast amount and fast enough than what a human mind may be able to perform, these calculations aren't fast enough for a complex algorithm processing a large dataset. We need computers that are faster for these types of complex calculations, but with the need for higher processing power also comes the cost for these systems. More processing power implies that the cost of these resources is also going to be high and so is the maintenance of these systems. The need for a cost-efficient, fast processing unit is where HPC comes into play.

HPC initially started as a supercomputer, where the HPC was just one computer system with a very high processing speed, where the clock speed of the computing nodes were much higher than the normal PC or laptops, hence giving us an edge over the personal computer systems. Apart from having a higher frequency per core per node they also had a large memory and other high-end technologies, like larger storage memory, higher RAM, etc. The only problem was that the supercomputer was one single system. In recent times the HPC is a computer cluster, these refer to a group of high-performing computing nodes which communicate with each other through a high-speed network and have individual memory but share a file system. Since each node can be considered a computer in its own sense, they can vary among them by the clock speed or the individual memory allocation or by whether they allow parallel processing, we can request the type of node we deem fit to process our computation and the number of nodes we require to complete our job in the most efficient way. Job or task scheduling on an HPC is done through a third-party software for cluster workload management. With the more modern advancement in recent times the term supercomputers and

computer clusters are used interchangeably for an HPC. A node in a typical computer cluster in a HPC will have a higher processing capability than an average personal system and so HPCs can perform 1000 trillion computations a second, because of these advantages they are being used widely now a days for both scientific and commercial applications as discussed in He et al [27].

The PTDM model described in section 4.2 took an average of three hours to process one hour of data, depending on the size of the data and the number of profiles that need to be processed. This puts a huge time crunch especially when the need arises to process data that has been recorded over an extended period of time. This is where the HPC systems come in helpful, because of their large computational prowess and speed. In this section we will be discussing about the HPC cluster, Summit, used for deploying the PTDM model and the results of this work.

4.3.2. SUMMIT Supercomputer

The Summit supercomputer was designed and deployed by the funding provided by the National Science Foundation (NSF). The Summit system was first introduced in 2016 as the second generation of supercomputers. The HPC system specification is shown below in table 4.2. The 3 main components of a good HPC system are computing, network, and storage. The Summit supercomputing cluster consists of high-performance CPUs, GPUs, and a large storage system on a first-generation Intel Omni-Path Architecture, which ensures a high throughput of 100 GB/s/port while consuming low power and communication latency. The Intel Omni-Path architecture has a 16/edge core links and 32/edge ports and provides links to IBM General Parallel File System (GPFS) scratch filing system. The Summit system gives each a user a

memory allocation of 10.252 TB to store their data and apart from this additional memory is available at a cost.

TABLE 4.2 Summit Supercomputer Specifications [28]

PARAMETER	SPECIFICATION
Architecture	1st gen 2:1 Intel Omni-Path
	(8 core switches, 16/edge Core links)
Computational nodes	Intel Xeon "Haswell"
	Nvidia Tesla K80 (2/node) for GPU
	Intel Xeon Phi
Storage	DDN SFA 14k GRIDScaler storage appliance
	IBM GPFS File system
	4.722 TB Metadata storage
	1.5 PB Scratch storage
Operating system	Red Hat Enterprise Linux 7
Workload management system	SLURM

There are more than 500+ nodes in total available on the Summit cluster. A detailed specification of the nodes are given below in Table 4.3. The general nodes are available to perform just high computational processes which do not require parallel processing or a high memory requirement to run jobs, so they are used to run jobs which are computationally intensive. An additional 20 general nodes were added after deployment of Summit with advanced processing power. The GPU nodes are best used for parallel processing with multiple

threads in the nodes performing a specific independent task. The high memory nodes are best used for computations that are memory intensive, that is require a large memory allocation for the data that is to be processed and while running the job.

TABLE 4.3 Summit Compute Node Specifications [28]

TYPES OF NODES	NO. OF NODES	CPU	MEMORY	LOCAL STORAGE
GENERAL	452	Intel Xeon E5-2680 v3 @2.50GHz (2 CPUs/node, 24 cores/node)	2133 MT/s, Dual Rank, x4 Data Width RDIMM, (4.84 GB/core)	200 GB SSD (1/node)
GENERAL	20	Intel(R) Xeon(R) Gold 6126 CPU @ 2.60GHz (2 CPUs/node, 24 cores/node)	2666 MT/s, Dual Rank, x4 Data Width RDIMM (192 GiB/node)	200 GB SSD (1/node)
GPU	10	Intel Xeon E5-2680 v3 @2.50GHz (2 CPUs/node, 24 cores/node)	2133 MT/s, Dual Rank, x4 Data Width RDIMM, (4.84 GB/core)	200 GB SSD (1/node)
HIGH MEMORY	5	Intel Xeon COU E704830 v3@2.10GHz (4 CPUs/node, 48 cores/node)	2133 MT/s, Dual Rank, x4 Data Width RDIMM, (42.7 GB/Core)	1TB 7.2K RPM, 6Gbps Near Line SAS 2.5" Hard Drive (12/node)
PHI	20	Intel Xeon Phi "Knights Landing" processor (1/node)	112 GiB/node, 6*16 GiB DIMMs and local Phi memory	200 GB SSD

The PHI nodes are Intel's Xeon Phi "Knight's Landing", which has a higher on-chip memory with higher memory bandwidth. These resources are available for all researchers of CUB, CSU,

and RMACC. These resources are allocated among all the contributors, where 10% of the total resources are reserved for RMACC users, 22.5% of the total resources are allocated for CSU users and the remaining 67.5% of the total resources are allocated for CUB users.

To compute on the HPC, the users can submit their jobs by batch scripting using the SLURM commands. SLURM is a workload manager and job scheduler for Linux-based OS systems. When a user types in any one of the SLURM commands they are passed on to the SLURM controller daemons, which then process the command. If the client command is regarding a job scheduling, then the requested compute node is accessed to submit the job, if the node is available the job starts running immediately, if not the job is added to a queueing list and the job starts running once the node becomes available. If the client command is an information collection, then the database is accessed to retrieve the information requested by the user. The job scheduler also prioritizes each user based on the amount of previous jobs that have been submitted by the user to maintain a fair share priority usage of the resources, so that no one user over utilizes the system unfairly. As discussed in J.H. Abawajy [29], an efficient method should be adapted to schedule jobs on the HPC. Users can also install any software they would on require performing their computation on the HPC to use on the Summit systems, there are some basic modules that are installed on Summit and can be used by any user. The jobs can also be an interactive session with the GUI interface provided.

4.3.3. Deployment of PTDM on Summit

In this section, the deployment of the PTDM model on the Summit system and the performance on the Summit system compared to a normal personal system is discussed. The comparisons are made for the two similar cases observed by [5] discussed in the previous section

4.2. The steps involved in testing the codes on the HPC and adapting them to deploy on the HPC is shown in the flowchart, figure 4.2. To process the data using the PTDM codes on an HPC, the codes were first tested on the HPC using an interactive node. The interactive node uses a GUI interface which makes editing the codes and viewing the outputs more easier and user friendly. During the testing process it was noticed that the PTDM algorithm does not converge for certain data and so the processing is stopped in between. To optimize the code even further we need to automate the process, for this step instead of inputting the data stream for every hour that we need to process, we run the model in a loop for the list of hours that we require the processed data for. Another advantage of the HPC system over the personal system is that they have multiple cores and nodes available and as long as we don't overutilize our allocation on the system and stay well within our fair share, we can run multiple simulations of the same code simultaneously. This scales back the time required to process the data significantly.

Since the PTDM code processes data for every hour, each hour data should be inputted to process via the code. This is automated through a bash code to input the data automatically for every hour and repeat the processing for each hour in a loop. While implementing the automation it was noticed that the process ends abruptly, and the loop exits if the PTDM code doesn't converge for the data for that hour and must be manually repeated from the start by expelling the data that the code didn't converge for. To avoid this unwanted exiting of the loop, the code check whether the data is compatible and can be processed by the PTDM code. If there is an error and the data cannot be processed by the PTDM code, then the code ignores the data for that hour and processes for the next hour automatically instead of exiting the loop.

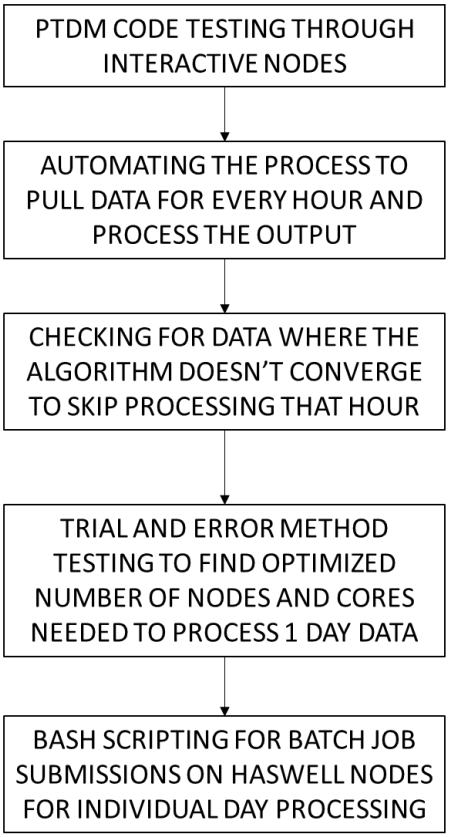

Figure 4.2 Flowchart depicting steps involved to adapt PTDM codes for HPC and deployment on HPC

The experiments are performed as for two cases 29th Nov 2009 for the time 1500 to 1600 UTC and 27th July 2010 for the time 1000 to 1100 UTC also discussed by Mann et al [30] and Luke and Kollias [31] respectively. The cases are processed by the PTDM models on the HPC and the retrieved reflectivity and velocity for the cloud and drizzle data are plotted and this is seen in figure 4.3 and figure 4.4 respectively. The evaluation is done with respect to the findings discussed in [5]. To submit this job on the HPC and to optimize the use of the resources available we must find the most efficient values for the nodes and cores on the system. After trial-and-error method testing, we found that one node with fifteen cores is the most optimized usage of the resources for fast processing. Comparing the outputs, we receive from the codes after

processing on the HPC we arrive at the conclusion that the reflectivity and velocity values retrieved from the PTDM model processed on the HPC are similar to the values retrieved when processed not on the HPC.

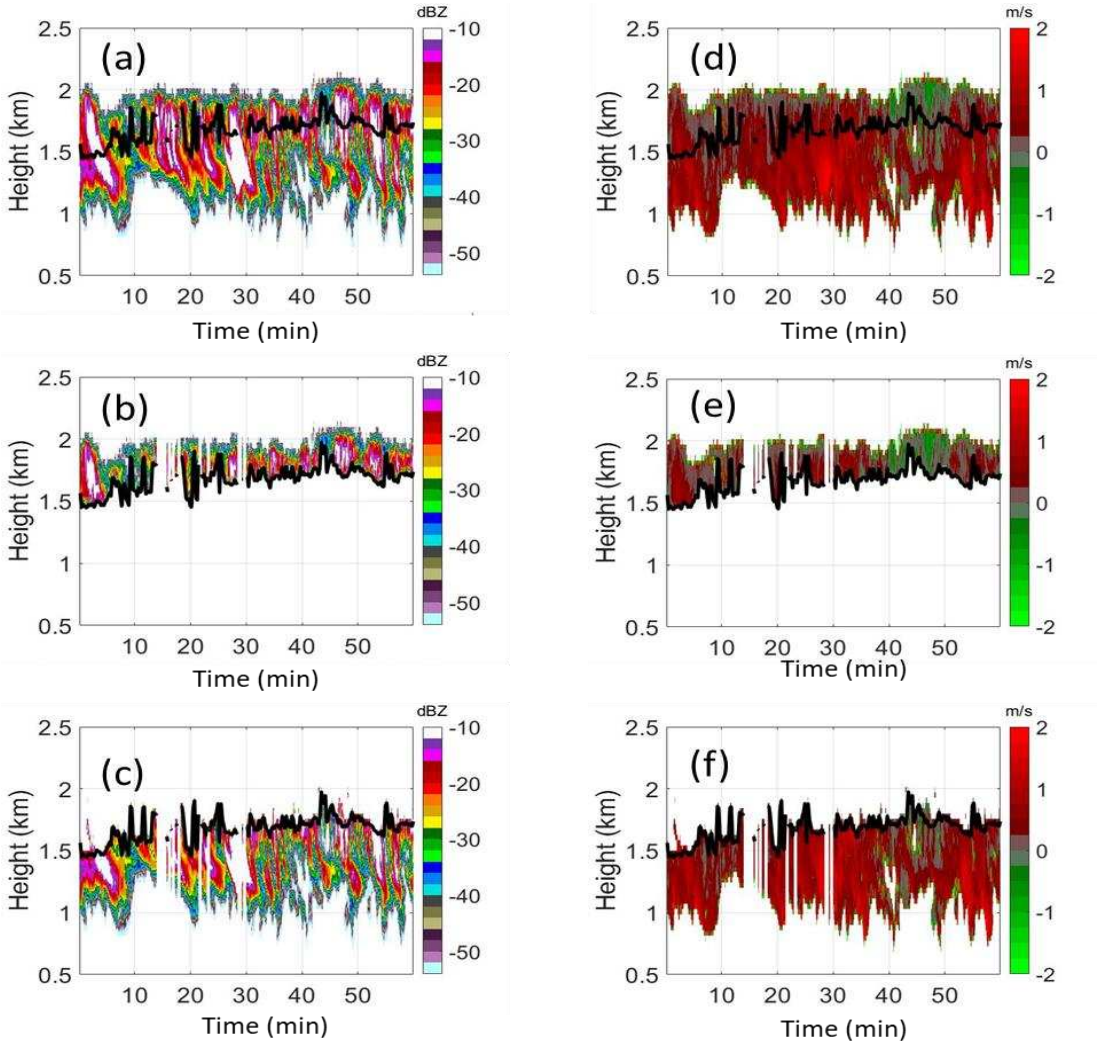

Figure 4.3 Time Height plots of (a) observed radar reflectivity, (b) retrieved reflectivity of cloud droplets, (c) retrieved reflectivity of drizzle droplets, (d) observed radar velocity, (e) retrieved velocity of cloud droplets, (f) retrieved velocity of drizzle droplets
For 29th Nov 2009 1500-1600 UTC processed on the HPC

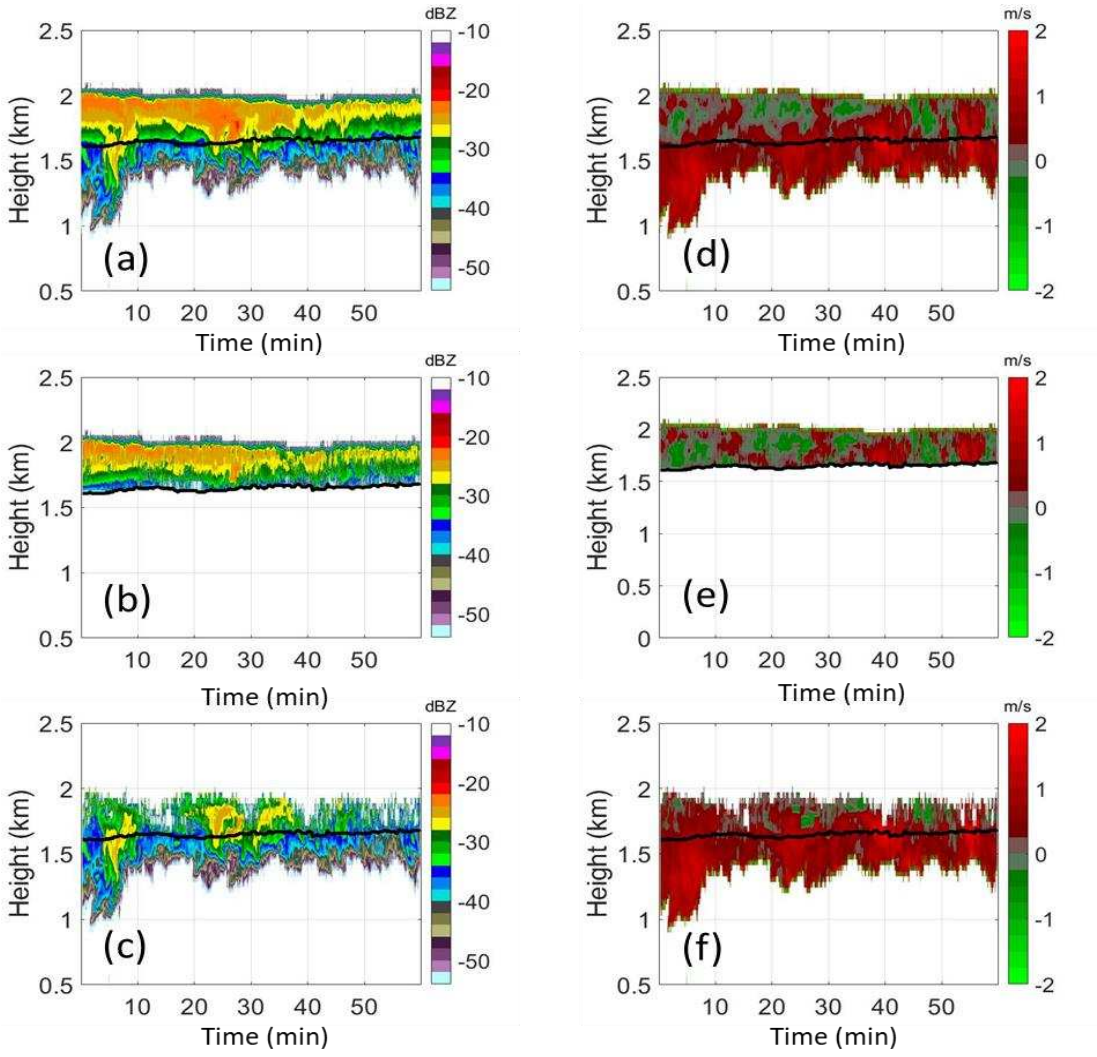

Figure 4.4 Time Height plots of (a) observed radar reflectivity, (b) retrieved reflectivity of cloud droplets, (c) retrieved reflectivity of drizzle droplets, (d) observed radar velocity, (e) retrieved velocity of cloud droplets, (f) retrieved velocity of drizzle droplets
For 27th July 2010 1000-1100 UTC processed on the HPC

As we can see from the comparison below in Table 4.4, the comparison is done between running the codes on a personal computer (PC), a Dell G3 with Intel core i7 processor, with 8GB DDR4 RAM with a clock speed of 2.66MHz using a Qualcomm DW1810 1x1 for networking communications and running on the HPC on a general node with the specifications as explained in Table 4.3. As we can see in both the cases the processing time is increased when the number

of profiles required to process is increased. Also, by comparing the last 2 columns we can see that the run time is decreased significantly, in first case the run time has been optimized by 55% and for the second case by 57.5%. The HPC has a pronounced effect on the run time and given the fact that we can process multiple instances of the model for different cases at a given time simultaneously, we can increase the overall runtime efficiency of the PTDM model while processing datasets from a longer time frame.

TABLE 4.4 Comprehensive comparison between the two cases analyzed

Case	No. of Profiles	Runtime on PC (Above average processing power)	Runtime on HPC
29th Nov 2009	841	2 hours 10 mins	59 mins
27th July 2010	1642	4 hours 45 mins	2 hours 1 min

From the above comprehensive table, we can conclude that the number of profiles that are required to process increases the time latency. The number of profiles is increased in size, doubling in its quantity, in the observed data from 1st December 2009. This can be attributed to the fact that the time between profiles observed is reduced by half which increases the number of profiles by a factor of two. This change is only evident in data collected after 30th November 2009, that is from 1st December 2009. The time interval between the profiles is reduced, where previously the data is observed and collected for every 4.28 s initially, for the second case it is reduced by half and the radar data is observed is in intervals of 2.14 s. This increases the number of spectra that is to be processed by data collected in case 2, till November 31st, 2009, the number of spectra is 26684 but after that, it is increased by a factor of two and comes to 58743,

which increases the number of profiles. Since the number of profiles that are to be processed is increased by two, the time taken to process the data is also increased, which is been shown below in table 4.5.

TABLE 4.5 Metrics affecting the time latency

Cases	**No. of Profiles**	**Time between profiles (s)**
29th Nov 2009	841	4.28
27th July 2010	1642	2.14

The PTDM model uses the dataset from ARM CAP-MBL campaign to separate cloud and drizzle moments. By separating the cloud and drizzle moments, we can study the boundary the layer clouds and how the aerosol contents present in the atmosphere affect them. For this we need to process data over several months to study the changes in the cloud formation. Since we need to analyze the data after it is processed by the PTDM code to study the aerosol effect on the cloud the data processing should be quick. The HPC speeds up this process effectively hence shortening the wait time during the processing. The table 4.6 shows the time taken to process the data for different time periods such as one hour worth of data, one day worth of data and so on. The time taken to process that data is an approximation of the time that was taken to process the data on the HPC.

TABLE 4.6 Time taken to process the data

Data period	**Time taken to process (Excluding the wait time for resource allocation)**
1 hour	0.75 hours
1 day	16.45 hours
1 week	4.5 days
1 month	18 days
6 months (Aug 2009 - Jan 2010)	119 days

After the initial testing on the HPC, the data has been processed in batches for every month. The data is processed for the months where drizzle was present. The data is available from April 2009 to January 2011. The drizzle moments are more prominent for the months of June 2009 and till August 2009 and after that the winter months starting from November 2009 till March 2010 have drizzle activity. This processed data is further analyzed to see how the increase in aerosol content in the atmosphere impacts boundary layer clouds and their formation. This will be very useful to study global warming and its effects on boundary layer clouds.

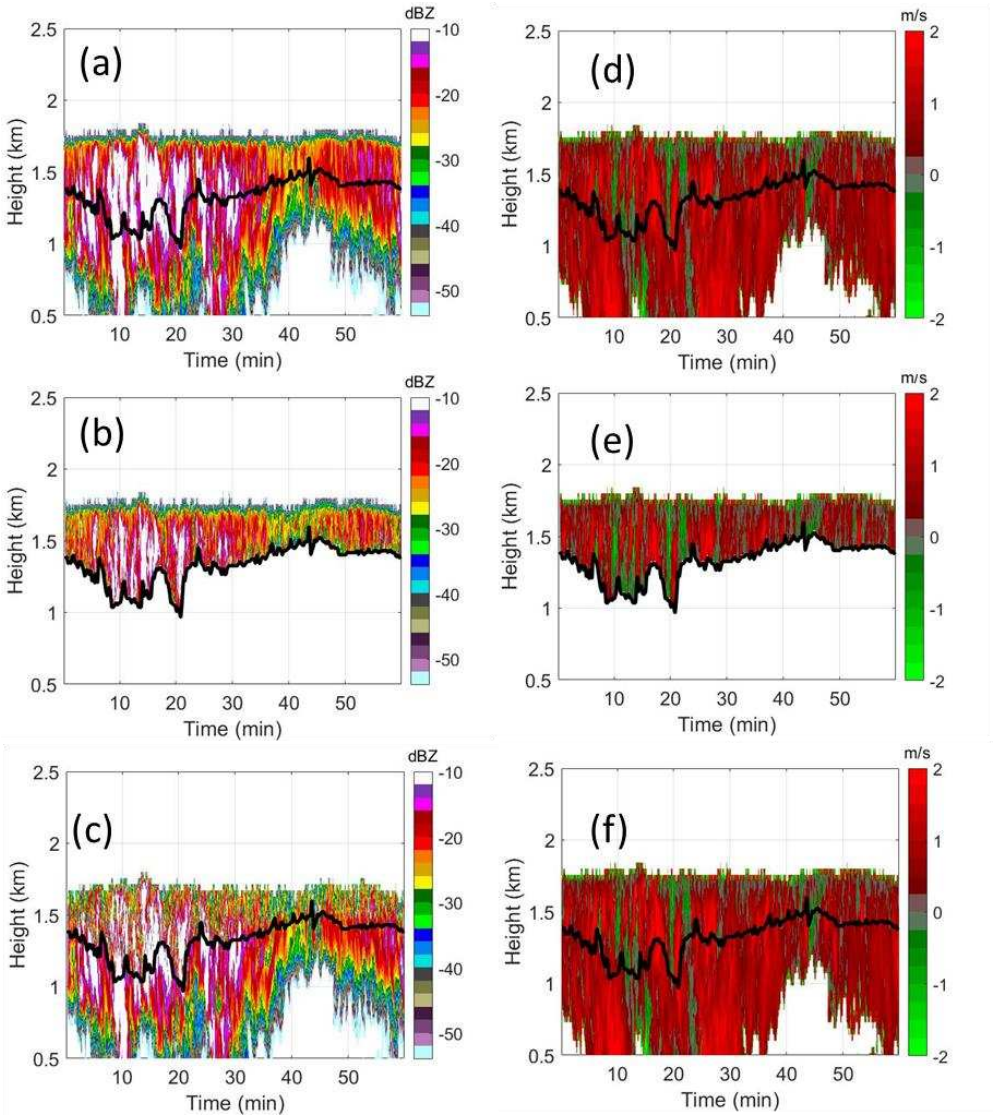

Figure 4.5 Time Height plots of (a) observed radar reflectivity, (b) retrieved reflectivity of cloud droplets, (c) retrieved reflectivity of drizzle droplets, (d) observed radar velocity, (e) retrieved velocity of cloud droplets, (f) retrieved velocity of drizzle droplets For 22nd November 2009 0600-0700 UTC processed on the HPC

The figure shown above, figure 4.5 shows the observed and retrieved reflectivity and velocity plots for the cloud and drizzle moments for 22nd November 2009 0600-0700 UTC. The black line on the plots shows the cloud base line. The image below, figure 4.6, shows the data processed for 22nd November 2009 1900-2000 UTC. The case shown in figure 4.6 has a

moderate drizzle activity while the case shown in figure 4.5 has a moderate to high drizzle activity

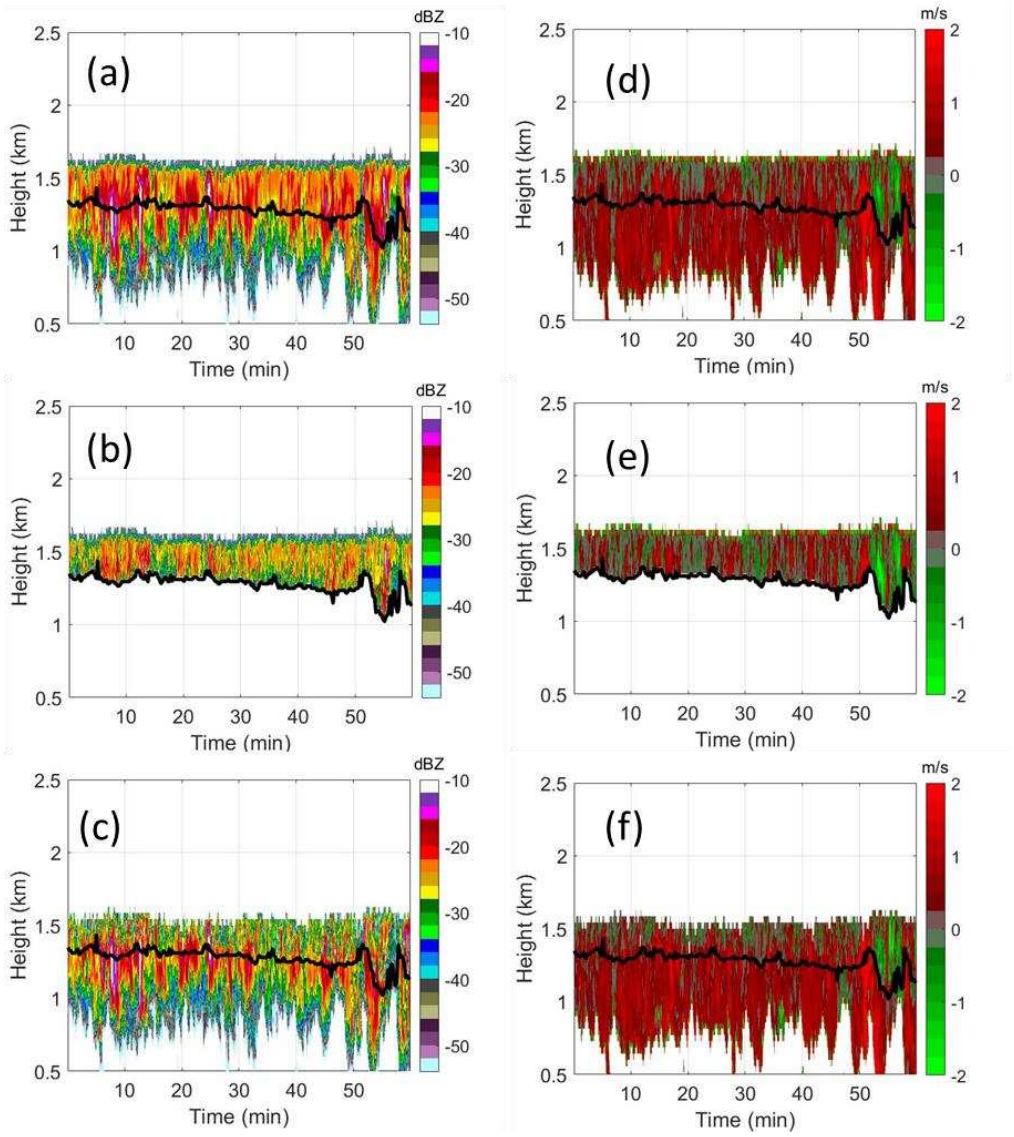

Figure 4.6 Time Height plots of (a) observed radar reflectivity, (b) retrieved reflectivity of cloud droplets, (c) retrieved reflectivity of drizzle droplets, (d) observed radar velocity, (e) retrieved velocity of cloud droplets, (f) retrieved velocity of drizzle droplets For 22nd November 2009 1900-2000 UTC processed on the HPC

4.4. CONCLUSION

In this chapter, we have discussed an efficient technique for separating cloud and drizzle data using Doppler spectra and a PTDM model, introduced in [5], which is a meticulous mathematical framework assuming that the Doppler spectra of both the cloud data and drizzle data is a Gaussian distribution. The setback of this PTDM model, being the run time taken to process the data was overcome by deploying on the HPC Summit and the results were cross verified for different cases having light and moderate drizzle, the November 29th 2009, and July 27th 2010.

The analysis found that the runtime depends on two parameters, the processor speed on the computing platform where the PTDM code is processed and the number of profiles that are to be processed. Since the number of profiles depends on the dataset, that is not a controllable parameter. The processor on which the PTDM model is deployed can be changed and hence we can opt for an HPC which has a higher processing speed and allows parallel computing, hence optimizing the runtime by more than 50%. In the future, this optimized PTDM model method can be applied to W-band cloud radar operated in a scan mode but can be applied to vertically pointing Ka-band radars.

CHAPTER 5 : CONCLUSION AND FUTURE WORK

5.1. RESEARCH SUMMARY

This thesis studied the microphysical properties of particles in different hydrometeors such as rain, heavy rain, large drops, and a mixture of hail and rain from convective and stratiform storms. The PTDM technique developed by [5] to separate cloud and drizzle data in the time domain is optimized by reducing the processing time by deploying it on a high-performance computer. The CSU-CHIVO radar was deployed in Argentina during the RELAMPAGO campaign to study the severe weather condition and intensive storms that are present in the foothills of the Andes mountains. Spectral processing of radar observations and analyzing the radar moments that have been collected from such an unpredictable and severe weather condition can give us an insight into storms and storm microphysics. The next part discussed in this thesis was the optimization of the PTDM technique to reduce the time latency by processing the codes on a high-performance computer. This allows the data to be processed significantly faster because of the higher computational power present in an HPC.

In Chapter 2, the basics of weather radar and how the radar observes, and records data are discussed in brief. The main components that make up a radar such as an antenna, transmitter, receiver, and the signal processor is discussed and how they work is explained. An introduction to the three different radar moments that are important for spectral processing of the data to retrieve valuable information is discussed. In this chapter, an introduction to spectral analysis and polarimetry and how the radar moments are calculated from the raw I/Q data is also seen.

The mathematical formulas required to calculate the spectral radar moments of the radar observations and how the radar moments can be retrieved after spectral processing is also reviewed. The importance of spectral domain analysis and its need are discussed in brief.

In chapter 3, spectral processing and analysis of the data collected during the RELAMPAGO campaign by the CSU-CHIVO radar are discussed. By studying the data collected during these intensive storm periods we can get a better understanding of the microphysics of the hydrometeors present in different type of storms. First, the study of two different convective storm cases with different hydrometeors such as heavy rain, large drops, mixture of hail and rain is done. The I/Q data collected during these convective storms is used to perform spectral processing to derive the radar moments in different regions of the hydrometeors. The analysis of different features such as the slope in the differential reflectivity, the copolar correlation values, and the modality of the reflectivity spectra is done to study the microphysics of the hydrometeors and classify them. An error analysis of the reflectivity and the differential reflectivity for the range gates which were spectral processed and analyzed is done by calculating the standard deviation with respect to the adjacent range gates. The standard deviation is found to be a maximum of ± 2 dBZ for the spectral reflectivity and ± 0.2 dB for the differential reflectivity spectra. The error analysis is done to verify whether the region of study for the spectral analysis is and since the maximum standard deviation value is low it shows that the data has less scatter.

In chapter 4, the optimization of the PTDM technique to separate cloud and drizzle moments from the observed radar data is performed by deploying it on an HPC to reduce the time taken by the algorithm to process the data. The PTDM techniques proposed by Nguyen et al [4] is modified and adapted to better suit the separation of cloud and drizzle data which may at

times overlap with one another. The PTDM technique to separate cloud and drizzle data is an efficient approach to separate the two data in the time domain, but the issue arises when a large amount of data needs to be processed. To study the effect of global warming and the increase in aerosol content in the atmospheric boundary layer clouds and drizzle, a long period of observed radar data must be processed. To speed up the processing time of this technique we use a high computational powered computer. The PTDM codes were tested using the data collected during the ARM CAP-MBL campaign. The run-time performance of the algorithm was evaluated by comparing the time taken to process the same data on an above-average computational resource and on the HPC. It was seen that the HPC reduced the run time by approximately 50%. HPC can be used to process data with different software and environments which can be easily replicated on an HPC.

5.2. FUTURE WORK

In chapter 3, the spectral analysis of the different hydrometeors present in a storm was performed and the classification of the hydrometeors was done. The importance of spectral analysis and study of spectral polarimetry is to get a better understanding of the hydrometeors present in severe weather conditions. A detailed and thorough analysis of the various hydrometeors in such severe storms must be conducted to get a more detailed understanding of the storm's patterns and how it evolves.

The processed data retrieved from the PTDM technique to separate cloud and drizzle data is used to analyze further how boundary layer clouds may change in the future due to global warming and the impact the increase in aerosol content in the atmosphere has on these boundary

layer clouds and their formations. For this, a large amount of data needs to be processed to analyze the impact accurately. Most of the observed radar data have a large size and to analyze and study features of the radar moments, large quantities of data need to be processed with complex algorithms. These techniques and algorithms can also be deployed on HPCs to reduce their processing time and overcome storage problems that may arise while processing a large dataset. HPCs have a large memory space allocation and hence can also store the large data that is required to be processed and their outputs. Another advantage of HPC is the vast computational resources available like high-performance CPU and GPU. By utilizing GPUs we can parallelize the workflow and efficiently process these data simultaneously by multithreading.

BIBLIOGRAPHY

[1] V. N. Bringi and V. Chandrasekar, Polarimetric Doppler Weather Radar: Principles and Applications, Cambridge, New York: Cambridge University Press, 2001.

[2] Y. Wang, T.-Y. Yu, A. V. Ryzhkov and M. R. Kumjian, "Application of Spectral Polarimetry to a Hailstorm at Low Elevation Angle," *Journal of Atmospheric and Oceanic Technology,* vol. 36, no. 4, pp. 567-583, 2019.

[3] L. Pfitzenmaier, C. M. H. Unal, Y. Dufournet and H. W. J. Russchenberg, "Observing ice particle growth along fall streaks in mixed-phase clouds using spectral polarimetric radar data," *Atmospheric Chemistry and Physics,* vol. 18, no. 1, pp. 7843 - 7862, 2018.

[4] C. M. Nguyen, D. N. Moisseev and V. Chandrasekar, "A parametric time domain method for spectral moment estimation and clutter mitigation for weather radars," *Journal of Atmospheric and Oceanic Technology,* vol. 25, no. 1, pp. 83-92, 2008.

[5] S. S. Joshil, C. M. Nguyen, V. Chandrasekar, J. C. Chiu and Y. Blanchard, "Separating Cloud and Drizzle Signals in Radar Doppler Spectra Using a Parametric Time Domain Method," *Journal of Atmospheric and Oceanic Technology,* vol. 37, no. 9, pp. 1669-1680, 2020.

[6] Vaisala, "User's Manual RVP900 Digital Receiver and Signal Processor," 2017. [Online]. Available: https://www.manualslib.com/manual/906126/Vaisala-Rvp900.html?#manual.

[7] D. N. Moisseev and V. Chandrasekar, "Polarimetric spectral filter for adaptive clutter and noise suppression," *Journal of Atmospheric and Oceanic Technology,* vol. 26, no. 2, pp. 215-228, 2009.

[8] V. Chandrasekar, R. Keranen, S. Lim and D. N. Moiseev, "Recent Advances in classification observation from dual polarization weather radars," *Atmospheric Research,* vol. 119, pp. 97-111, 2013.

[9] H. 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 Forecasting,* vol. 24, no. 3, pp. 730-748, 2009.

[10] A. V. Ryzhkov, T. J. Schuur, D. W. Burgess, S. Giangrande and D. S. Zrnić, "The Joint Polarization Experiment: Polarimetric rainfall measurements and hydrometeor classification," *Bulletin of the American Meteorological Society,* vol. 86, no. 6, pp. 809-824, 2005.

[11] F. Fabry and I. Zawadzki, "Long-Term Radar Observations of the Melting Layer of precipitation and their interpretation," *Journal of Atmospheric Sciences,* vol. 52, no. 7, pp. 838-851, 1995.

[12] S. W. Nesbitt, P. V. Salio, E. Ávila, P. Bitzer, L. Carey, V. Chandrasekar, W. Deierling and et al, "A storm safari in subtropical South America: Proyecto RELAMPAGO," *Bulletin of the American Meteorological Society,* vol. 102, no. 8, pp. E1621-E1644, 2021.

[13] A. Varble, "RELAMPAGO-CACTI WORKSHOP ALL ABOUT THE DATA," 20 2020 January. [Online]. Available: https://www.arm.gov/news/blog/post/59580.

[14] I. Arias and V. Chandrasekar, "Cross Validation of the Network of Ground-Based Radar with GPM during the Remote Sensing of Electrification, Lightning, And Mesoscale/Microscale Processes with Adaptive Ground Observations (RELAMPAGO) Field Campaign," *Journal of the Meteorological Society of Japan,* vol. 99, no. 6, pp. 1423-1438, 2021.

[15] S. Bachmann, V. DeBrunner, D. S. Zrnić and M. Yeary, "Spectral Analysis of Polarimetric Weather Radar Data with Multiple Processes in a Resolution Volume," in *IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07*, Honolulu, HI, USA, 2007.

[16] A. Dutta, V. Chandrasekar and S. K. Biswas, "Spectral polarimetry for microphysical studies of rain and hail during RELAMPAGO campaign-initial results," *American Geophysical Union Fall Meeting Abstracts,* vol. 2019, pp. A53U-3046, 2019.

[17] M. R. Kumjian and A. V. Ryzhkov, "The impact of size sorting on the polarimetric radar variables," *Journal of the Atmospheric Sciences,* vol. 69, no. 6, pp. .2042-2060, 2012.

[18] D. L. Hartmann, M. E. Ockert-Bell and M. L. Michelsen, "The Effect of Cloud Type on Earth's Energy Balance: Global Analysis," *Journal of Climate,* vol. 5, no. 11, pp. 1281-1304, 1992.

[19] C. J. Hanh and S. G. Warren, "A Gridded Climatology of Clouds over Land (1971-1996) and Ocean (1954-2008) from Surface Observations Worldwide (NDP-026E)*," Carbon Dioxide Information Analysis Center (CDIAC), Oak

Ridge National Laboratory (ORNL), Oak Ridge, TN (USA), 2007.

[20] R. Wood, "Stratocumulus Clouds," vol. 140, no. 8, pp. 2373-2423, August 2012.

[21] T. Yamaguchi, G. Feingold and J. Kazil, "Stratocumulus to cumulus transition by drizzle," *Journal of Advanced Model Earth System,* vol. 9, no. 6, pp. 2333-2349, 2017.

[22] X. Zhou, T. Heus and P. Kollias, "Influences of drizzle on stratocumulus cloudiness and organization," *Journal of Geophysical Research: Atmospheres,* vol. 122, no. 13, pp. 6989-7003, 2017.

[23] T. Weber and J. Quass, "Incorporating the subgrid-scale variability of clouds in the autoconversion parameterization using a PDF-scheme," *Journal of Advanced Model Earth Systems,* vol. 4, no. 4, p. M11003, 2012.

[24] H. Takahashi, M. Lebsock, K. Suzuki, G. Stephens and M. Wang , "An investigation of microphysics and subgrid-scale variability in warm-rain clouds using the A-Train observations and a multiscale modeling framework," *Journal of Geophysical Research: Atmospheres,* vol. 122, no. 14, p. 7493–7504, 2017.

[25] X. Jing, K. Suzuki and T. Michibata, "The key role of warm rain parameterization in determining the aerosol indirect effect in a global climate model," *Journal of Climate,* vol. 32, no. 14, p. 4409–4430, 2019.

[26] J.-C. Golaz, L. W. Horowitz and H. Levy II, "Cloud tuning in a coupled climate model: Impact on 20th century warming," *Geophysical Research Letters,* vol. 40, no. 10, p. 2246–2251, 2013.

[27] L. He, S. A. Jarvis, D. P. Spooner, H. Jiang, D. N. Dillenberger and G. R. Nudd,

"Allocating non-real-time and soft real-time jobs in multiclusters," *IEEE Transactions on Parallel and Distributed Systems,* vol. 17, no. 2, pp. 99-112, 2006.

[28] O. o. I. T. Research Computing, "What is RMACC Summit?," 2017. [Online]. Available: https://www.colorado.edu/rc/articles/rmaccsummit.

[29] J. H. Abawajy, "An efficient adaptive scheduling policy for high-performance computing," *Future Generation Computer Systems,* vol. 25, no. 3, pp. 364-370, 2009.

[30] J. A. Mann, J. C. Chiu, R. J. Hogan, E. J. O'Connor, T. S. L'Ecuyer, T. H. Stein and A. Jefferson, "Aerosol impacts on drizzle properties in warm clouds from ARM Mobile Facility maritime and continental deployments," *Journal of Geophysical Research: Atmospheres,* vol. 119, no. 7, p. 4136–4148, 2014.

[31] E. P. Luke and P. Kollias, "Separating cloud and drizzle radar moments during precipitation onset using Doppler spectra," *Journal of Atmospheric and Oceanic Technology,* vol. 30, no. 8, pp. 1656-1671, 2013.

LIST OF ABBREVIATIONS

Z: Reflectivity; Unit - dBZ (decibel relative to Z)

Z_{dr}: Differential Reflectivity; Unit- dB (decibel)

ρ_{hv} : Copolar Correlation

φ_{dp} : Differential Phase; Unit- ° (degree)

IFDR : : Intermediate Frequency Digital Receiver

PPI: Plan Position Indicator

RHI: Range Height Indicator

PTDM: Parametric Time-Domain Method

SQI: Signal Quality Index

GMAP: Gaussian Model Adaptive Processing

I/Q: In-phase, Quadrature-phase

DFT: Discrete Fourier Transform

H-pol: Horizontal polarization

V-pol: Vertical polarization

PSD: Power Spectral Density

CSU: Colorado State University

SNR: Signal to Noise Ratio

MLE: Minimum Likelihood Estimate

IOP: Intensive Observing Period

CUB: University of Colorado, Boulder

DROPS2: Dual-Polarization Radar Operational Processing System 2

RELAMPAGO: Remote sensing of Electrification, Lightning, And Mesoscale/Microscale Processes with Adaptive Ground Observations

CHIVO: C-band Hydrometeorological Instrument for Volumetric Observation

CSAPR2: C-band Scanning ARM Precipitation Radar 2

ARM: Atmospheric Radiation Measurement

CAP-MBL: Clouds, Aerosol and Precipitation in the Marine Boundary Layer

RMACC: Rocky Mountain Advanced Computing Consortium

SWACR: Scanning W-band ARM Cloud Radar

HPC: High Performance Computing

SD: Standard Deviation