UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

APPLICATION OF MACHINE LEARNING TO MULTIPLE RADAR MISSIONS AND OPERATIONS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

YUNISH SHRESTHA
Norman, Oklahoma
2022

APPLICATION OF MACHINE LEARNING
TO
MULTIPLE RADAR MISSIONS AND OPERATIONS

A DISSERTATION APPROVED FOR THE
SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

BY THE COMMITTEE CONSISTING OF

Dr. Yan Zhang, Chair

Dr. Amy McGovern

Dr. Justin Metcalf

Dr. Alexander Ryzhkov

Dr. Tian-You Yu

Acknowledgments

I want to thank my dissertation committee for their guidance, support, and feedback throughout the process. I would especially like to thank Dr. Zhang for the immense help and guidance in my academic journey and for guiding through all stages of my dissertation. Thank you, Dr. Zhang , for being my mentor and providing countless opportunities to learn the skills that were important to me. I want you to know how much I appreciate it all.

Similarly, I would like to thank my committe members Dr. McGovern, Dr. Metcalf, Dr. Ryzhkov, Dr. Yu for their continuous feedback throughout my academic journey. I am also thankful to all the members of ECE and ARRC for all the help and support I received in the last 5 years.

I am grateful of the grant supports from HKO and NASA Langley Research Center. I would also like to thank Mr.PW Chan (HKO), Dr. William Blake (Garmin), Dr. Greg Mcfarquhar (CIWRO) and Mr. Steven Harrah (NASA) for their support and feedback.

Contents

List Of Tables

List Of Figures

Abstract

This dissertation investigated the application of Machine Learning (ML) in multiple radar missions. With the increasing computational power and data availability, machine learning is becoming a convenient tool in developing radar algorithms. The overall goal of the dissertation was to improve the transportation safety. Three specific applications were studied: improving safety in the airport operations, safer air travel and safer road travel. First, in the operations around airports, lightning prediction is necessary to enhance safety of the ground handling workers. Information about the future lightning can help the workers take necessary actions to avoid lightning related injuries. The mission was to investigate the use of ML algorithms with measurements produced by an S-band weather radar to predict the lightning flash rate. This study used radar variables, single pol and dual-pol, measured throughout a year to train the machine learning algorithm. The effectiveness of dual-pol radar variables for lighting flash rate prediction was validated, and Pearson's coefficient of about 0.88 was achieved in the selected ML scheme. Second, the detection of High Ice Water Content (HIWC),which impact the jet engine operations at high altitudes, is necessary to improve the safety of air transportation. The detection information help aircraft pilots avoid hazardous HIWC condition. The mission was to detect HIWC using ML and the X-band airborne weather radar. Due to the insufficiency of measured data, radar data was synthesized using an end-to-end airborne weather system simulator. The simulation employed the information about ice crystals' particle size distribution (PSDs), axial ratios, and orientation to generate the polarimetric radar

variables. The simulated radar variables were used to train the machine learning to detect HIWC and estimate the IWC values. Pearson's coefficient of about 0.99 was achieved for this mission. The third mission included the improvement of angular resolution and explored the machine learning based target classification using an automotive radar. In an autonomous vehicle system, the classification of targets enhances the safety of ground transportation. The angular resolution was improved using Multiple Input Multiple Output (MIMO) techniques. The mission also involved classifying the targets (pedestrian vs. vehicle) using micro-Doppler features. The classification accuracy of about 94% was achieved.

Chapter 1

Introduction

1.1 Introduction

Radar has proven to be a key component in meteorological studies and weather forecasting. Weather radar can provide precipitation measurements and accurate information on the type of clouds and their microphysics based on the analysis and processing of the reflected electromagnetic signals. Several groundbreaking developments in the field of weather radar, like the development of Pulse-Doppler radar, dual-polarization weather radar, and the development of algorithms, have aided significantly in meteorological investigations. As of this time, accurate prediction of weather type and its hazardous levels, determination of form, shape, and orientation of hydrometeors have been possible with the aid of existing weather radar technology (Branch 2012; Doviak et al. 2006; Ryzhkov and Zrnic 2019). However, most of the current operational weather prediction model is based on analytical or fuzzy-logic models. The increasing availability of weather data for researchers and the increasing processing ability of modern computers have facilitated the use of Artificial intelligence (AI)/Machine Learning (MI) in the field of meteorology and hence in the field of weather radars (McGovern et al. 2017). In addition to the stored data, the modern-day computer allows the synthesis of simulated data for machine-learning applications (Minhas et al. 2022). This research aims to identify and evaluate ML applications

in the field of weather radar to improve the existing weather forecasting models in terms of accuracy and efficiency. The study findings will be advantageous to academic researchers and industry partners in the development of newer weather-radar systems and their forecasting algorithms.

In recent years, the automotive industry has successfully demonstrated various safety functions such as parking assistance, collision avoidance, and adaptive cruise control using automotive radar. Complementary Metal Oxide Semiconductor (CMOS) in radar technology makes automotive radar an ideal sensor in terms of cost, size, weight, and power requirements (Ragonese et al. 2022). In addition, the radar can operate in darkness and extreme weather conditions, in which other sensors may not work. However, for all its advantages, the angular and range resolution of automotive radar is inferior. The inferior resolution can be explained based on the functioning of the antenna and the bandwidth available for an automotive radar. Therefore, the classification of targets using radar and ML algorithms or other perception algorithms is imperfect. There is, therefore, a need to improve the resolution of an automotive radar for classification purposes. This dissertation also explores various ways to improve the resolution in automotive radar, thereby improving the classification results.

This chapter will discuss the background and context first, followed by the research topic, the research aims, objectives, questions, importance, and finally, the limits.

1.2 Backgrounds

1.2.1 Weather Radar

The first type of radar used in the dissertation is Weather Radar. Various ground-based (both fixed and mobile) and airborne weather radars have been developed for weather observations. The basic working principle of weather radars is similar, although they vary in the frequency band and correspondingly size and platform. Most

weather radar systems are mono-static radars that employ a pulsed waveform. In the case of a weather radar, a point target is replaced by multiple hydrometeor scatterers within a resolution volume. These individual scatters contribute to the signals being reflected. Based on the signal statistics of the reflected signal, information about the weather can be inferred. The reflected signals are processed to estimate the Reflectivity, Velocity, and Spectrum Width. In the case of dual-pol radar, in which the electromagnetic waves are transmitted in the horizontal and vertical directions, additional variables like differential reflectivity (Z_{dr}), correlation coefficient (ρ_{HV}), differential phase (Φ_{dp}), and specific differential phase (K_{dp}) are computed. These variables, alone or in combination, can provide information about the weather and cloud microphysics.

1.2.2 Millimeter-Wave, Collision Avoidance Radar

Another type of radar used in this dissertation is the millimeter-wave (mmWave) radar designed for collision avoidance, most commonly for ground-road-traffic classification, used in automotive, self-driving cars (Engels et al. 2021; Waldschmidt et al. 2021). The features provided by this type of radar are simpler than polarimetric weather radar. However, the angular and range resolutions significantly impact target classification and estimations accuracy. In contrast to pulsed radar, such as weather radar, automotive radar uses a Frequency Modulated Continuous Waveform (FMCW) in which the frequency of the waveform varies linearly over time. To identify the range of the target, the reflected signal is mixed with the copy of the transmitted signal to produce the intermediate signal, the frequency of which indicates the range to the target. For collision avoidance applications, the range, speed, and angle of arrival are measured. The automobile radars use short-wave electromagnetic or mmWave, with frequencies ranging from 76 to 81 GHz. The advantages of using higher frequency is the higher accuracy (better resolutions), compared to the lower frequency system.

For example, larger bandwidth, which improves the range resolution, is available for higher frequencies. Higher frequency also means smaller components' sizes, thereby reducing the overall size of the radar system.

1.2.3 Application of Machine-Learning

ML is a subset of AI that has been greatly admired in numerous academic domains. There are three major categories of ML: Reinforcement Learning, Supervised Learning, and Unsupervised Learning. In Reinforcement learning, an agent is trained to make a sequence of decisions. For each decision, the agent gets either rewards or penalties. The goal is to maximize the reward. In Supervised Learning, the machine is trained using the data that has both inputs and outputs. The machine then detects the underlying patterns and relationships between the input and output data. In Unsupervised Learning, the machine is trained using the data that has only inputs, i.e., unlabeled datasets. The machine then discovers the patterns or clusters in the dataset. Supervised ML is used in this dissertation. For example, weather radar measurements are part of the "input feature vector" for ML, while some atmospheric quantity like the lightning flash rate is the output. The deep learning tools provided by MATLAB are used in this dissertation, including primary Artificial Neural Network (ANN), Convolutional Neural Network (CNN), and other variations as the key algorithms.

1.3 Research Objectives

The scope of the research is the application of a generic ML solution to observe multiple environmental "features". These environmental features can be both natural and man-made. The three objectives listed below are considered specific applications of the ML framework. The first study uses dual-pol radar variables as ML features to

predict the lightning flash rate. In the second study, the meteorological quantity of interest is High Ice Water Content (HIWC) which is one of the natural aviation hazards to be monitored. Extensive development of physical models of this phenomenon was developed and validated first, and then simulated polarimetric airborne weather radar data were used to verify the ML-based algorithms. In the final study, the goal was to investigate the classification of various traffic using mmWave radars and ML processing.

1.3.1 Prediction of Lightning Flash Rate Using Ground-based Weather Radar Observations

Although overall lightning-related fatalities may be decreasing, the number of people vulnerable to lightning in their workplaces is on the rise (Holle 2016). For example, the airport has been documented as one of the vulnerable workplaces. Personnel working outdoors responsible for maintaining airports or servicing aircraft are vulnerable to lightning-related injuries or deaths. Therefore, it is essential to provide timely alerts about possible Cloud to Ground (CG) lightning strikes for the safety (Steiner et al. 2012). The current lightning alert system at airports is based on visual warning cues and a course of actions related to ramp activities. This type of system requires the initiation of lightning strikes in a near vicinity which is detected by a lightning detection network and uses color-coded signal lights to indicate the distance of the strikes from the airport (Heitkemper et al. 2008). Because it is not always easy to detect lightning initiations, it is plausible to have a lightning alert system that can provide future lightning information.

One of the possible ways to predict lightning is to use ML algorithms. Previous researches have investigated some radar variables associated with lightnings (Carey and Rutledge 2000; Chase et al. 2022; Hondl and Eilts 1994; Woodard et al. 2012). The

data collected from the Lightning Location Information System (LLIS) can be combined with data from radars to study the nature of the storms that create lightning. Based on these studied data, a ML algorithm may be implemented to predict lightning activities. This study introduces a ML algorithm applied to combined radar and LLIS datasets and predicts the flash rates up to a 30-minute time window. Previous works have also looked into these variables (Kohn et al. 2011; Lakshmanan and Smith 2009). However, they were limited to using ML algorithms to predict the occurrence of lightning and not for nowcasting flash rates. Using the proposed algorithm from this work, the flash rate can be predicted in addition to the detection of lightning. This is important as a greater flash rate is usually associated with higher injuries, and damage (Mazzetti and Fuelberg 2017). Flash rates can also predict severe weather such as wind, hail, and tornadoes (Mazzetti and Fuelberg 2017; Schultz et al. 2009; Williams et al. 1989).

In this work, ANN is used to predict the flash rate for different types of storm cells. Storm cells-based lightning nowcasting has been made possible with storm tracking algorithms. Using storm identification and tracking algorithms, it is possible to extract the physical parameters of storm cells like their area, volume, and height, which can be used as the input parameters for the supervised ML algorithms. Storm physical parameters are related to different atmospheric parameters, which can be favorable for a lightning generation. Instead of using a single ANN to make the predictions, the Ensemble Learning Approach (ELA) is used. Here, a cluster of neural networks are trained, and their respective outputs are averaged out to give a final output to improve the accuracy of prediction (Smolyakov 2017).

1.3.2 Airborne Weather Radar Detection of High Ice Water Content

HIWC represent a significant aeronautical hazard and a threat to jet engine operations at high altitudes (Dye et al. 2004; Mason et al. 2006; Mason and Grzych 2011; Ratvasky et al. 2015) including commercial vehicles, rockets, super-sonic vehicles, and other space exploration mission platforms. The ice crystals get lodged into the jet engines causing engine power loss and engine rollbacks. International studies, flight campaigns, and modeling studies have been ongoing to address the HIWC challenge (Strapp et al. 2016a). Currently, scientific teams at National Aeronautics and Space Administration (NASA), The University of Oklahoma (OU), and elsewhere are working to better understand the microphysical processes causing HIWC at high altitudes. Despite the research progress, there are still significant gaps between current research and the operational needs of NASA and the industry. The active Federal Aviation Administration (FAA)-Radio Technical Commission for Aeronautics (RTCA) working group (RTCA 2019) on HIWC detection has been seeking to develop the minimum operational performance standards for HIWC detection for commercial airborne weather radars. Critical and fundamental questions need to be answered within an urgent schedule, such as the criteria for declaring the hazards, the detection ranges, and the validation of radar capabilities. In terms of testing and evaluation, past flight campaigns were mainly limited to collecting in-situ measurements and some collocated radar measurements. There is also a strong need to incorporate dual-polarization capability into the airborne radar sensor, since polarimetric radar observations are critical for better characterizing microphysical properties of clouds (Bringi and Chandrasekar 2001; Ryzhkov and Zrnic 2019; Zhang 2016). Radar polarimetry offers one of the most promising solutions for HIWC identification, and quantification (Li et al. 2019; Ryzhkov and Zrnic 2019; Zhang 2016). In this study, the truth IWC data has been collected from different flight campaigns. This truth IWC data, along with the

information about the microphysical properties of ice particles, measured during the flight campaigns and knowledge based on the literature, have been used to simulate radar features. The simulation depends on the scattering properties of ice particles within a resolution volume and the parameters like the waveform and antenna pattern of the radar system. These radar features, along with some atmospheric properties like the temperature, have been used as input features for the supervised learning-based ML approach. ML is used for the classification purpose (whether the IWC condition is sufficient to declare a hazard or not) and the estimation of IWC values.

1.3.3 Environment Objects Classification Based on Vehicle-Mounted Radar Observations

The driving safety and driving comfort offered by autonomous driving have attracted many researchers. Autonomous driving requires information from sensors like cameras and radars to make the autonomous vehicle perceive the environment. Various sensors have their pros and cons. For example, a camera is a superior sensor in terms of resolution when compared to radar. However, it falls behind radar regarding its operational ability during the night or in poor light conditions. mmWave FMCW radar is very common for automotive applications. The range resolution depends upon the bandwidth of the chirp signal, and angular resolution depends upon the number of receiving antennas. The higher the number of receiving antennas, the better will be the resolution. This can be achieved by increasing receiving antenna elements or implementing the Multiple Input Multiple Output (MIMO) techniques. By increasing the number of transmitting elements with proper spacing, it is possible to create an increased number of virtual receiving antennas (Rao 2018). The goals of this study are to demonstrate the techniques to improve the resolutions of the radar sensors and use radar features for classifying pedestrians from vehicles. A complete simulation of the automotive radar case is discussed. A W-band radar has

been simulated with two transmitting antennas and four receiving antennas in the MATLAB environment. In addition to this, results from the initial measurement using commercial automotive radar are presented.

Previous research has explored the possibility of using MIMO techniques in an FMCW radar. Various techniques to implement orthogonality in an FMCW radar have been discussed in the previous article (De Wit et al. 2011; Zwanetski et al. 2013). A practical implementation of Time Division Multiplexing (TDM) and Binary Phased Modulation (BPM) has been demonstrated in the previous report (Rambach and Yang 2017; Rao 2018). The theoretical concepts of Frequncy Division Multiplexing (FDM) MIMO radar have been presented in articles (De Wit et al. 2011). In this study, an end-to-end automotive radar has been simulated to demonstrate the use of FDM based MIMO technique to improve angular resolution. Moreover, the choice of design parameters is discussed to consider the drawbacks of FDM-based MIMO identified in prior research.

In addition to improving the angular resolution, the micro-Doppler-based classification of targets is also explored in this study. Since a pedestrian and car have different micro-Doppler features, this can be used for the classification purpose in the absence of higher resolution offered by a camera or a lidar.

1.4 Organization of Dissertation

In Chapter 1, the background of the dissertation research has been introduced. Then, the scopes of research, objectives, and the significance of the work are briefly discussed.

In Chapter 2, the basic theory of radar systems will be reviewed. Architectural foundations of the radar and the signal processing techniques used in the study will be presented. In addition, the theory of ML will be presented.

In Chapter 3, the concept of lightning as a typical aviation hazard will be introduced. The methods and results of the application of the ML algorithm with polarimetric radar data for the lightning flash rate prediction will be presented.

In Chapter 4, the challenge of HIWC as an emerging aviation hazard will be further introduced. Much effort was spent on developing a reliable physics-based radar signature model of HIWC. The radar return model is then connected to the airborne weather radar system simulator, which provides an end-to-end tool for airborne radar developers and industrial partners. Finally, the application of ML solutions is demonstrated again using the simulation tool and the data it produces.

Chapter 5 introduces the mmWave-FMCW radar sensor and data collection for road traffic environment monitoring. Usage of MIMO technique to improve spatial resolution is also discussed. Similar ML algorithms are applied to classify different targets in an automotive environment. The features are selected specifically for the road traffic radar missions. The results from this Chapter demonstrate the impacts of spatial resolution on the sensor data quality.

In Chapter 6, a summary and conclusions will be presented, and the proposed future works will be outlined.

Chapter 2

Theoretical and Architecture Foundations

2.1 Introduction

2.1.1 Basics of pulsed radar

As its acronym suggests, radar (Radio Detection and Ranging) uses electromagnetic waves to detect and range targets. This is achieved by transmitting electromagnetic waves, which is reflected after hitting some target. The detection of the received signal is associated with detecting some targets, and the time it takes to return can be used to find the range to the target. If T and c represent the time delay and the speed of propagation of electromagnetic waves, then the range to the target is given as

$$r = \frac{c \cdot T}{2} \tag{2.1}$$

Most weather radars are pulsed-Doppler radars. A high-level block diagram of pulsed-Doppler radar is shown in Figure 2.1. Electromagnetic waves generated by a Stable Local Oscillator (STALO) are modulated to create the transmit pulse. The pulse is then amplified with a high-power amplifier (HPA). Ground-based weather radars such as Next-Generation Radar (NEXRAD) use klystron transmitters. In contrast, modern weather radar transmitters, such as airborne weather radar, have

been dominated by solid-state technology. The solid-state amplifiers have advantages such as low power consumption, high efficiency, small form factor, and increased precision (Hoang 2019; Schreurs and O'Droma 2009). The amplified signal then passes through the Tx/Rx switch and is finally radiated through an antenna. The Tx/Rx switch prevents amplified power from leaking into the receiver subsystems. The radar antenna may be a parabolic dish (as in NEXRAD), a slotted waveguide array (as in airborne weather radar), or a phased-array antenna. The received signal is mixed with the original signal and its 90° phase-shifted version, then finally low-pass filtered to obtain the in-phase and quad-phase (IQ) signals. The IQ signal can be represented with the equation 2.2, where r is the range to the target moving with the radial velocity of v_r.

$$V(t,r) = I + jQ = \frac{|A|}{2}U(t - \frac{2r}{c})e^{-j(\frac{4\pi r}{\lambda}+\frac{4\pi v_r t}{\lambda}-\psi_t-\psi_s)} \tag{2.2}$$

Here, $U(t-\frac{2r}{c})$ signifies the pulse of electromagnetic energy, ψ_t is the transmitter phase and ψ_s is the phase shift after scattering.

The radar received power (P_r) for a point target is given by

$$P_r = \frac{P_t g^2 \lambda^2 \sigma_b f^4(\theta,\phi)}{(4\pi)^3 r^4} \tag{2.3}$$

Here, P_t is the transmit power, g is the antenna gain, σ_b is the backscattering cross area.

Range resolution, maximum unambiguous range, and aliasing velocity of a general radar are given by: .

$$\Delta r = \frac{c \cdot \tau}{2} \tag{2.4}$$

$$r_a = \frac{c \cdot T_s}{2} \tag{2.5}$$

$$v_a = \frac{\lambda}{4T_s} \tag{2.6}$$

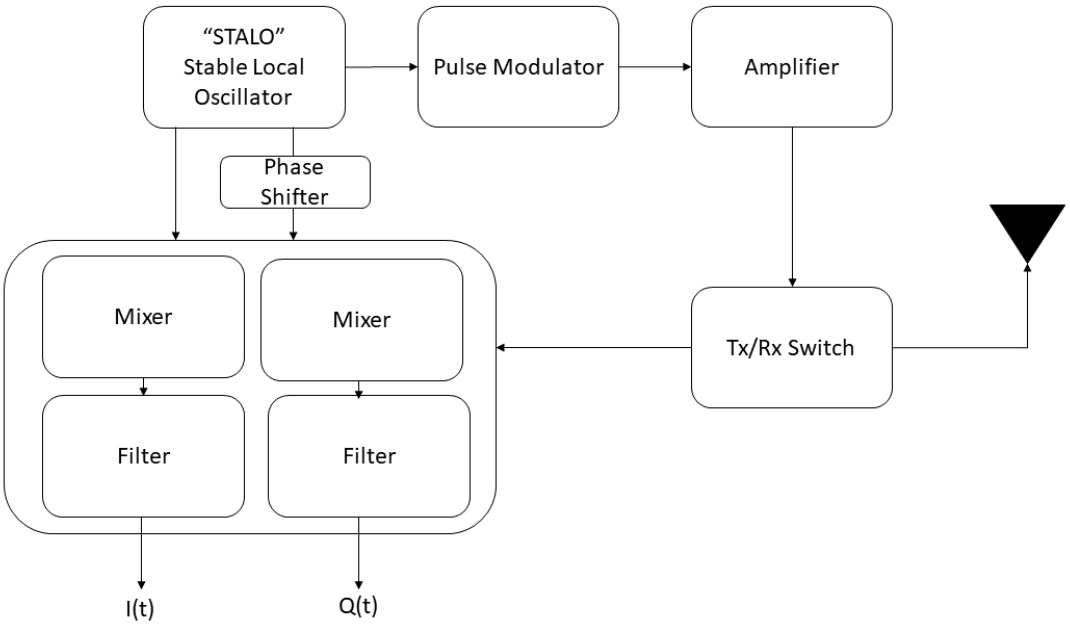


Figure 2.1: Basic block diagram of Weather Radar

2.1.2 Introduction to Weather Radars

The weather radar is a particular application of the Pulse-Doppler radar in which, instead of a single target "hard target" being the target of interest, it scans through the storm, which consists of many individual scatterers within the resolution volume. The resolution volume is defined by the range resolution and the beamwidth of the antenna. An illustration of the resolution volume is shown in Figure 2.2.

The echo voltage is given by

$$V(T_S) = \frac{1}{\sqrt{2}} \sum_{i=1}^{Ns} A_i W_i e^{-j\frac{4\pi}{\lambda} r_i} \tag{2.7}$$

,

where N_S is the number of scatterers. A_i is dependent on the antenna pattern, and W_i is the range weighting factor. Here, $r_i = \frac{4\pi r_i}{\lambda} + \frac{4\pi v_{r_i} T_s}{\lambda} - \Psi_{s_i} - B_i$.

The expected power is given by

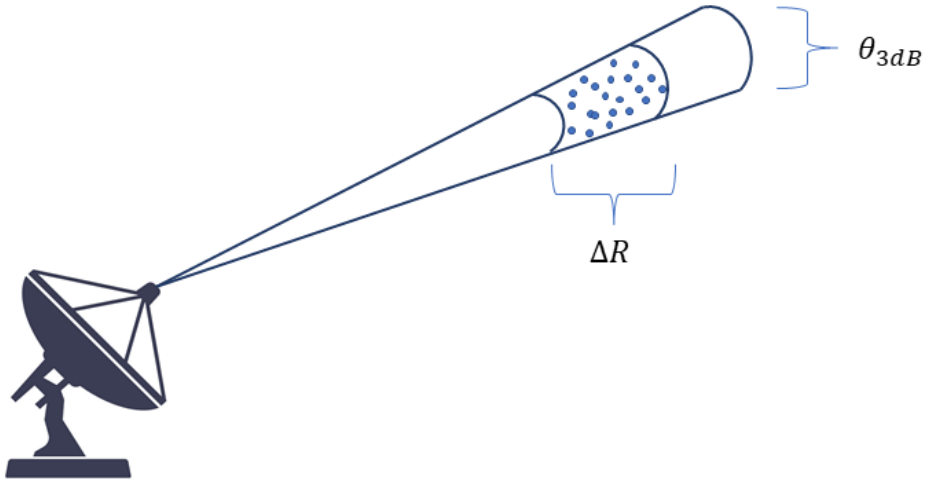


Figure 2.2: Operational Scenario of the Weather Radar.

$$E[P(r_\circ)] = \frac{P_t g^2 \lambda \eta}{(4\pi)^3 r_\circ^2 l^2} \left[\frac{c}{2}\right] \left[\frac{\pi \theta_i^2}{8 \ln 2}\right] \tag{2.8}$$

Note that the dependency of the received power is squared of the range ($\frac{1}{r_\circ^2}$), whereas for the point target, it is $\frac{1}{r_\circ^4}$

The most important variables that are measured by weather radars are reflectivity, velocity, and spectrum width. If the weather radar is dual-polarized, additional variables such as Z_{dr}, Φ_{dp}, K_{dp}, and ρ_{HV} play significant roles in meteorological studies, particularly for hydrometeor classification. Reflectivity is derived from the radar return power and is the estimate of the reflectivity factor(η). Reflectivity factor is backscattering cross-section per unit volume. Mathematically it is written as

$$\begin{aligned} \eta &= \frac{\pi^5}{\lambda^4} |K_w|^2 Z \\ Z = \frac{1}{\Delta V} \sum_i D_i^6 &= \int_0^\infty N(D, r) D^6 dD \end{aligned} \tag{2.9}$$

Here, Z is the Reflectivity. Reflectivity is often interpreted in logarithmic units as $10log_{10}(Z)$. The value of reflectivity ranges from 0 dBZ in light precipitation to more than 60 dBZ in heavy rainfall.

Velocity is the mean radial velocities of all the particles inside the resolution volume, and the spectrum width represents the standard deviation of these velocities. Spectrum width represents the relative radial motions of the particles inside the resolution volume. More turbulent the weather phenomenon, a wider spectrum width is expected. Velocity and spectrum width can be better explained by the Doppler spectrum or the power spectral density. Doppler spectrum is defined as the power-weighted distribution of the radial velocities within the resolution volume. It is the velocity distribution function of the particles inside the resolution volume. The power spectral density can be modeled as a Gaussian Distribution with the mean of the radial velocity and the standard deviation of the spectrum width.

Z_{dr} is the difference between horizontal and vertical polarized reflectivity values. It is related to the shapes of the hydrometeors. Mathematically,

$$Z_{dr}(dB) = Z_h(dB) - Z_v(dB) \tag{2.10}$$

It is 0 for spherical, positive for oblate, and negative for prolate hydrometeors. K_{dp} is the range derivative of the Φ_{dp} along the radial. K_{dp} can be expressed mathematically as:

$$K_{dp} = \frac{\Phi_{dp}(r2) - \Phi_{dp}(r1)}{2(r2 - r1)} \tag{2.11}$$

Similarly to Z_{dr}, K_{dp} is positive for horizontal hydrometeors and negative for vertically oriented particles.

ρ_{HV} is the measurement of the similarity between Horizontal and Vertical pulses. It is a great tool for discriminating between meteorological and non-meteorological targets.

The most common signal processing algorithm in weather radar involves Pulse-Pair Processing based on the auto-correlation function. The unbiased estimate of the AutoCorrelation Function (ACF) of echo voltage $V(t, r)$ can be expressed as

$$\hat{R}(kTs) = \begin{cases} \frac{1}{N-|k|} \sum_{n=0}^{N-|k|-1} V_r^* V_r(n+k) & \text{for } |k| \leq N-1 \\ 0 & \text{otherwise} \end{cases} \tag{2.12}$$

The lag-0 of ACF is the power estimate. It can be expressed as:

$$\hat{P} = \hat{R}(0) = \hat{S} + \hat{N} \tag{2.13}$$

, where $\hat{S}$ is the signal power estimate and $\hat{N}$ is the noise estimate.

Similarly mean velocity is estimated from ACF at lag 1, expressed as

$$\hat{v}_r = -\frac{\lambda}{4\pi T_s} arg[\hat{R}(1)] \tag{2.14}$$

Since two pulses are required to have an estimate of power and the velocity, this process is referred to as Pulse-Pair Processing.

The above mentioned estimation techniques are time domain estimations. Similarly, the moments can also be estimated from the Gaussian power Spectrum.

$$\hat{S}(v) = \frac{S}{(2\pi)^{1/2}\sigma_v} e^{[-(v-v_r)^2/2\sigma_v^2]} + \frac{2NT_s}{\lambda} \tag{2.15}$$

If the spectral estimates obtained for velocites $v_1, v_2, ..., v_k$ (corresponding to the $f_1, f_2, ..., f_k$), the zeroth moment (power), first moment (velocity) and the second moment (spectrum width) can be estimated as:

$$\hat{P} = \sum_{k=0}^{N-1} \hat{S}(k) \tag{2.16}$$

$$\hat{v}_r = \frac{\sum_{k=0}^{N-1} v_k \hat{S}(v)}{\sum_{k=0}^{N-1} \hat{S}(v)} \tag{2.17}$$

$$\hat{\sigma}_v^2 = \frac{\sum_{k=0}^{N-1} (v_k - \hat{v}_r)^2 \hat{S}(k)}{\sum_{k=0}^{N-1} \hat{S}(k)} \tag{2.18}$$

2.1.3 Frequency Modulated Continuous Wave (FMCW) radars

In FMCW radars, a frequency chirp signal is transmitted. A chirp is basically a sinusoidal signal whose frequency varies linearly with time. If the frequency increase with time, it is called up chirp, and if it decreases with time, it is called down chirp. Most automotive radars use up-chirps as their transmit signals. Figure 2.3 shows the Amplitude Vs. Time plot of an FMCW chirp. It is easier to represent the chirp with a Frequency Vs. Time plot.

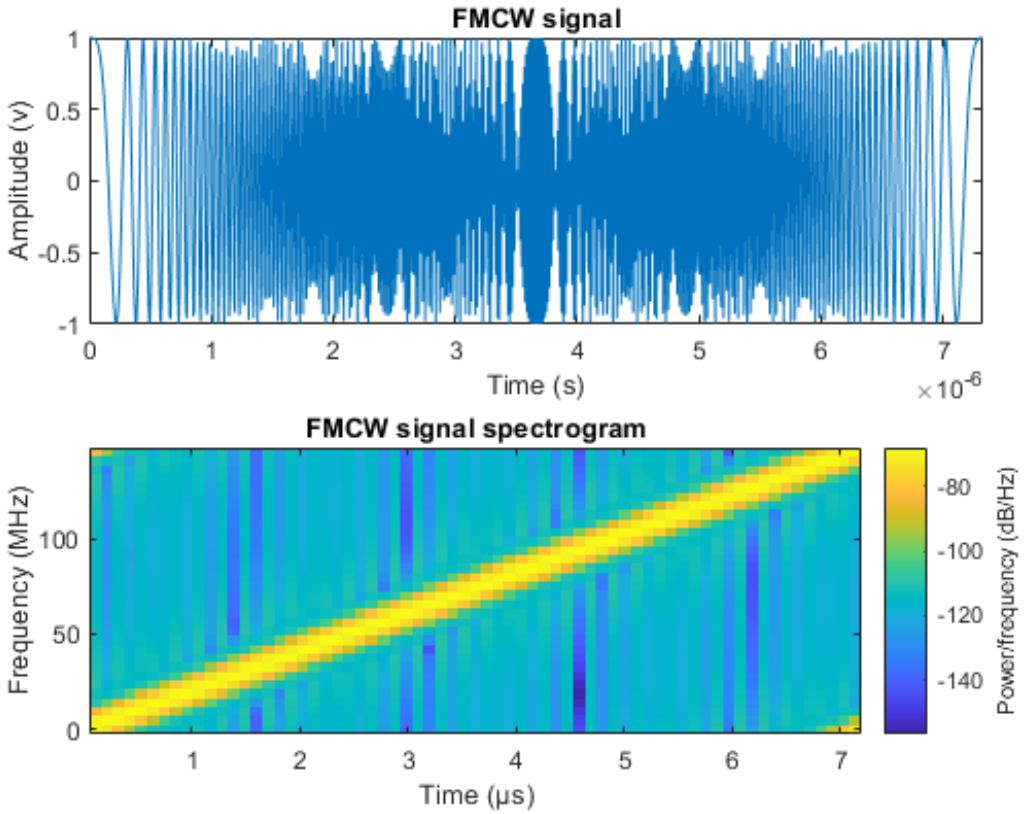


Figure 2.3: An example of FMCW chirp pulse, in time domain and frequency domain

The basic operation of an FMCW radar is shown in Figure 2.4. The synthesizer generates the FMCW chirp, which is transmitted by the TX antenna. The chirp gets reflected when it hits a target. The RX antenna receives the reflected chirp. A copy of the TX signal and the reflected signal gets mixed and generates the IF signal, whose

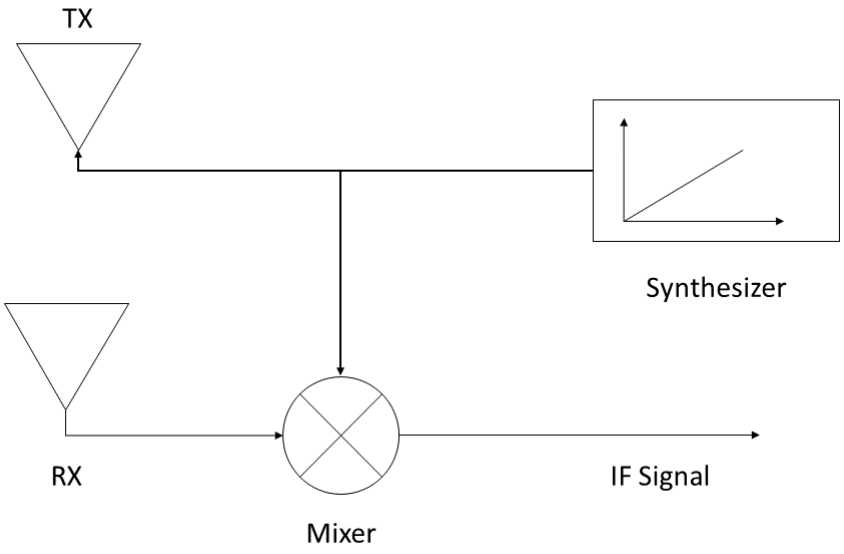


Figure 2.4: A basic block diagram of FMCW radar

frequency, also known as the beat frequency, depends on the range of the target. The beat frequency is equal to the instantaneous frequency difference between the TX and RX signal.

The received signal is a delayed version of the transmit signal. This can be represented in the Frequency-Time plot as shown in Figure 2.5. Here, T represents the two-way propagation delay. When this time-delayed version of the RX signal is mixed with the copy of the TX signal, the IF signal has frequency S_T, where $S_T = (S \times 2d)/c$. Here, S is the slope of the chirp. It is the rate at which the chirp ramps up and is obtained from the division of frequency bandwidth of the chirp and the chirp ramp time, i.e. $S = BW/T_C$. Similarly, d is the distance to the target and c is the propagation speed.

An ideal operation of the FMCW-automotive radar involves transmitting a number of chirps. The IF signal generated from every chirp is digitized with an Analog to Digital Converter (ADC) and generates a certain number of samples. Each chirp is received by each receiving element. The IF signal data cube that has IQ data for each sample, chirp, and the Rx-element is generated as illustrated in Figure 2.6.

The data cube then goes through a series of Fast-Fourier Transforms (FFTs) to estimate the range to the target, the velocity of the target, and the Angle of Arrival (AOA) of the target. FFT across the number of samples estimates the range as the frequency of the IF signal is dependent on the range to the target. A moving target introduces a phase difference across the chirps, so the FFT across the number of chirps estimates the velocity. Similarly, a phase difference is introduced in angular direction due to the RX-elements placement. So, the FFT across the number of Rx elements estimates the AOA. To estimate the AOA in both the Azimuth and Elevation directions, as in the case of 4D radars, a two-dimensional FFT is required across the horizontal and vertical RX elements.

Similar to pulsed-Doppler radar, parameters associated with FMCW radars are given as :

$$\Delta r = \frac{c}{2B} \tag{2.19}$$

$$r_a = \frac{F_s c}{2S} \tag{2.20}$$

$$\Delta v = \frac{\lambda}{2T_f} \tag{2.21}$$

$$v_a = \frac{\lambda}{4T_c} \tag{2.22}$$

$$\Delta\Theta = \frac{\lambda}{N \times dcos(\theta)} \tag{2.23}$$

$$\Theta_a = \sin^{-1}\left(\frac{\lambda}{2d}\right) \tag{2.24}$$

Here, Δr and r_a are range resolution and maximum unambiguous range respectively. Similarly, Δv and v_a represent the velocity resolution and maximum unambiguous velocity. Finally, $\Delta\Theta$ and Θ_a are angular resolution and maximum unambiguous AOA.

2.2 Specific Radars Used in This Dissertation

The first radar instrument used in this dissertation is CINRAD. CINRAD stands for China New Generation Weather Radar. The radar is manufactured by METSTAR Corporation. It is a dual-polarization S-band Doppler weather radar that incorporates the simultaneous transmission and reception of horizontal and vertical polarization (STAR mode). A passive power divider splits Klystron's high power output into two pathways, which are delivered to the Dual Polarization Feed horn on the antenna, which transmits both signals simultaneously, one with Horizontal Polarization and the other with Vertical Polarization. The Feed Horn divides the H and V polarization returns into two routes, which are subsequently fed to two digital receiver channels that are matched. The signal processor analyzes the data from the H and V channels to generate the dual-polarization parameters. It implements the volume coverage patterns (VCPs), similar to the NEXRAD radar systems.

The second radar instrument used in this dissertation is an airborne weather radar sensor originally from Garmin International, Inc.(shown in Figure 2.8), and is used as the basis for OU's PARADOX airborne radar. It is an X-band solid-state weather radar with a peak transmitter power of 40 watts. It uses a mechanically-scanning slotted waveguide array antenna to cover the field of view of 120° in azimuth and 60° in elevation.

The third radar instrument used in the dissertation is the Texas Instrument (TI)'s 77 GHz automotive radar sensor chip-set (AWR 1443) and evaluation system (AWR 1443-Boost), which is depicted in Figure 2.9. AWR 1443 is an FMCW transceiver that

operates at the frequency from 76 to 81 GHz, i.e., with a bandwidth of 4 GHz. It has two transmit channels, which can be used simultaneously, and four receive channels. The transmit power from each channel is 12 dBm. It is built with an integrated ARM processor and a hardware accelerator for radar data processing. This radar can be programmed for various applications like Short-range automotive radar, Mid-range radar, and Long-range radar sensing. AWR 1443 Boost is the evaluation system based on AWR 1443, which enables the capture of raw ADC data and performance evaluation. The RF front-end with appropriate antenna spacing, shown in Figure 2.10, allows for MIMO operations.

2.3 Machine Learning Techniques Used in this Dissertation

ML is a subset of AI that can be further divided into three categories: Reinforcement Learning, Supervised Learning, and Unsupervised Learning. In this dissertation, supervised learning is used for all three studies. Supervised learning performs the regression or classification by learning the functions that map the inputs to the outputs based on the input/output pairs. The learning is primarily done using the training data. To avoid any over-fitting, cross-validation data is used for the cross-validation. Then the performance is evaluated using the test sets. The functions can be a simple linear function to a complex neural network. In this dissertation, ANNs are used predominantly.

The foundations of the ANNs are based on the functioning of the human brain. As in a human's neural network, which consists of neurons, a neural network consists of nodes or perceptrons. A perceptron contains one or more inputs, weights for each input, a bias (or biases for each input), an activation function, and a single output.

$$Y = \sum_{i=1}^{i=n} (W_i X_i) + B \tag{2.25}$$

Here, Y is the output, W_i is the weight for the input X_i. B is the bias for the node. An ANN is a collection of interconnected perceptrons or nodes. An example of ANN is shown in Figure 2.12. Here the inputs are connected to the input layers, which are then connected to 5 nodes of a single hidden layer. The hidden layer gets its name because the feature inputs or outputs are not directly visible. Then the hidden layer is connected to an output layer. The training is usually based on the forward and backward propagation algorithms. Forward propagation calculates the outputs of each node based on the weights and biases assigned to them. In comparison, the backward propagation calculates the weights for each node for the next forward propagation based on some loss function. The goal of the training is to minimize the loss function through an iterative process until convergence is achieved. ANN can be deep or shallow. A shallow neural network has one hidden layer, while a deep neural network has two or more hidden layers. A special type of neural network is CNN. A CNN performs convolution on the set of interconnected features as in images or a signal to learn, making it very popular with applications such as image classifications (Jmour et al. 2018).

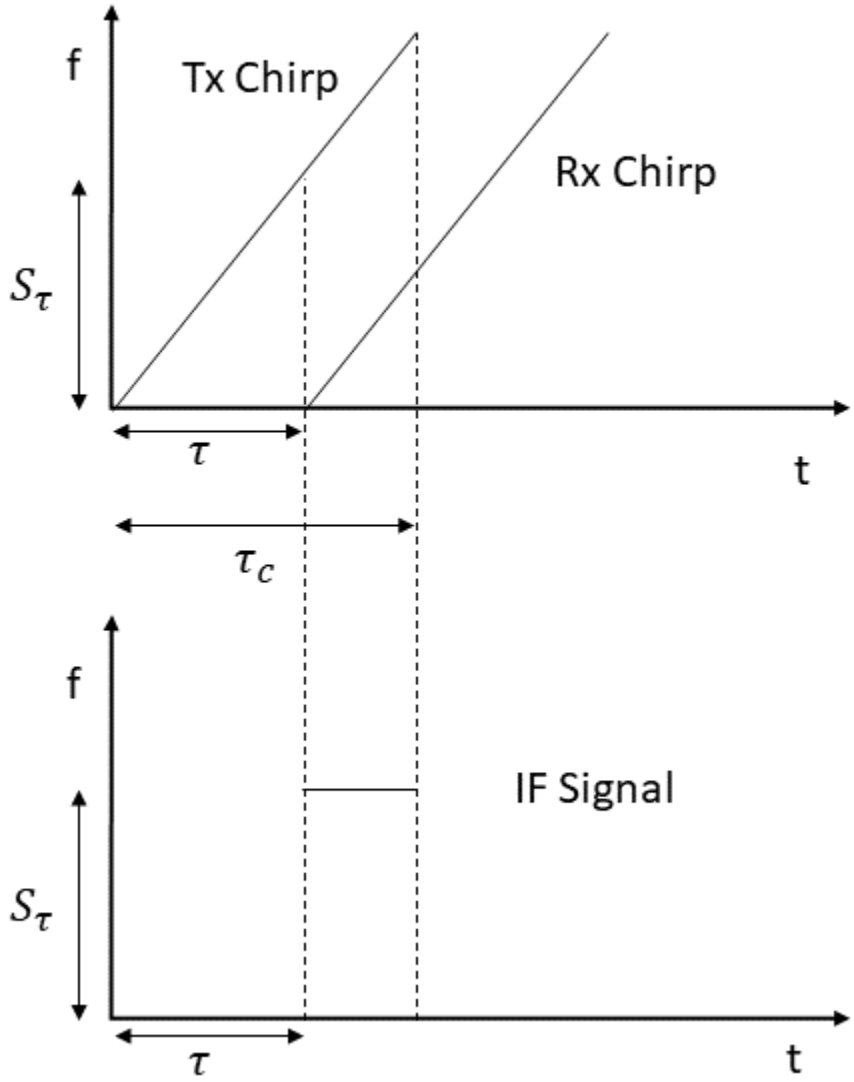


Figure 2.5: An illustration of FMCW radar operation.

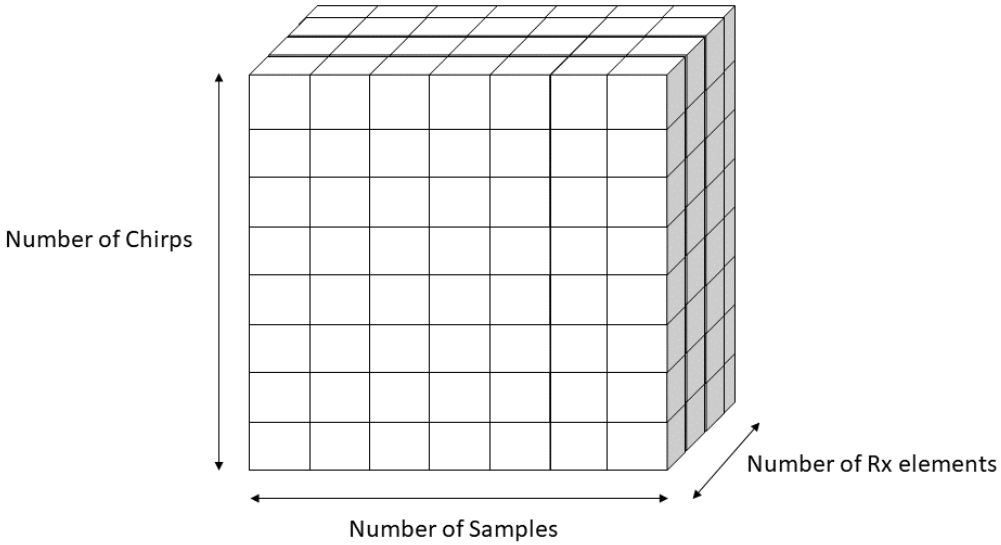


Figure 2.6: Radar data cube

Figure 2.7: A METSTAR Radar similar to the CINRAD.

Figure 2.8: GSX 70-PARADOX Airborne Weather Radar

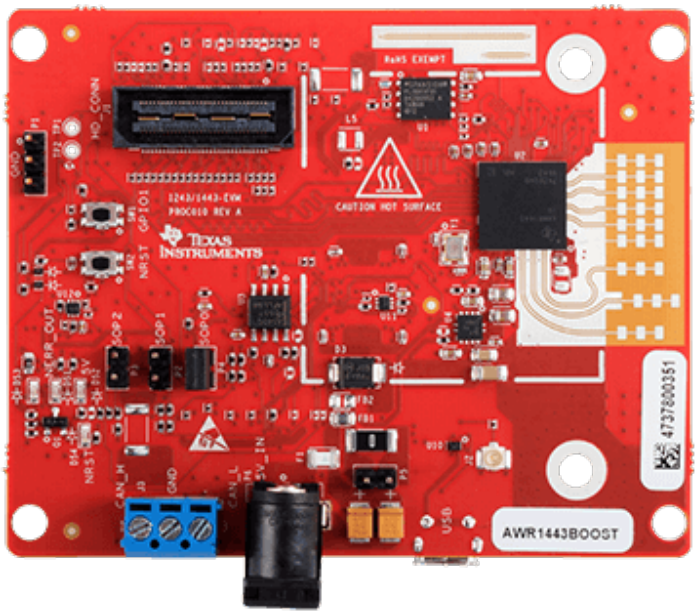


Figure 2.9: Texas Instrument 1443 Radar

Table 2.1: PARADOX2 System Parameters

Radar Parameters		Values
Mechanical dimensions	Antenna size	6 – 50 inches depending on platform requirements
	Transceiver diameter	8 inches
	Depth	6.3 inches
	Total Weight	9.5 lbs (including a 12 inch antenna, electronics and digital backend), <10 lbs without antenna
Operating Frequency		9.3 to 9.5 GHz
Antenna		Slotted Wavefuide Array and Mechanical scanning
FOV		±60 deg azimuth, 60 deg total elevation
Transmitter		Solid-state 40 watt peak power, support a wide range of waveforms and PRFs
Sensitivity		0 dBz@30 km Severe weather signatures up to 40 miles
Receiver		Real-time pulse compression receiver with optimized LFM and phase coding waveforms
Antenna beamwidth		Scalable: 12 inch panels: 5-8 ° az/el, enhanced to 2 ° through software processing, up to 50 inch panels –achieve 2 ° physical beamwidth
Scan speed		PPI 4 sec for one elevation, ∼3 sec RHI

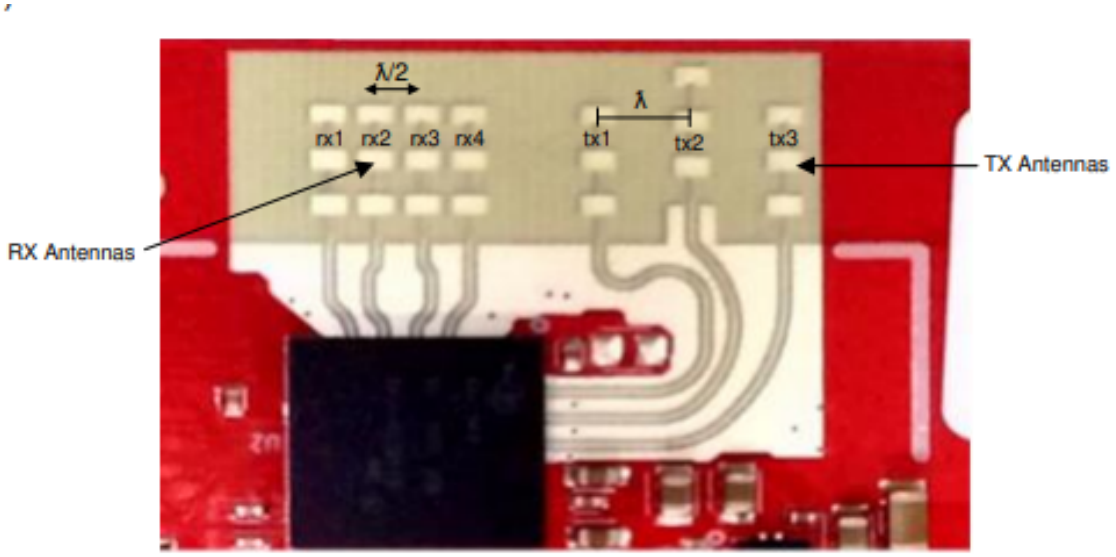


Figure 2.10: Antenna Configuration of Texas Instrument AWR 1443 Radar

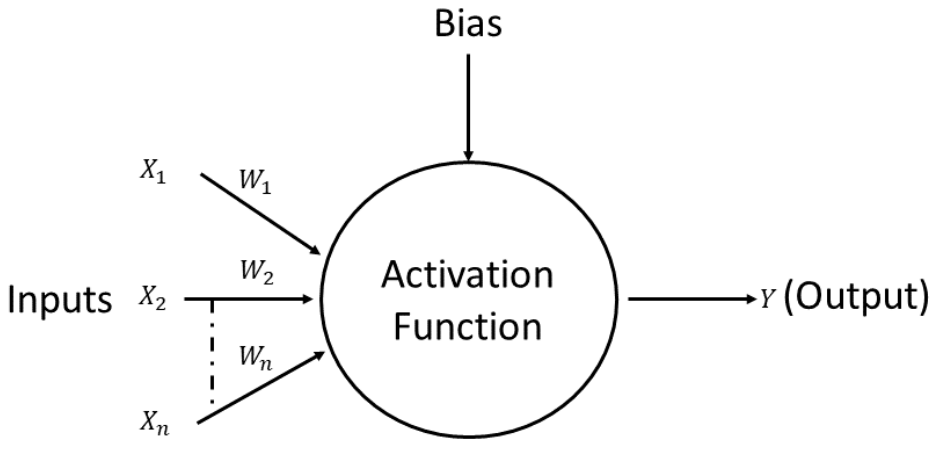


Figure 2.11: Structure of node in an ANN.

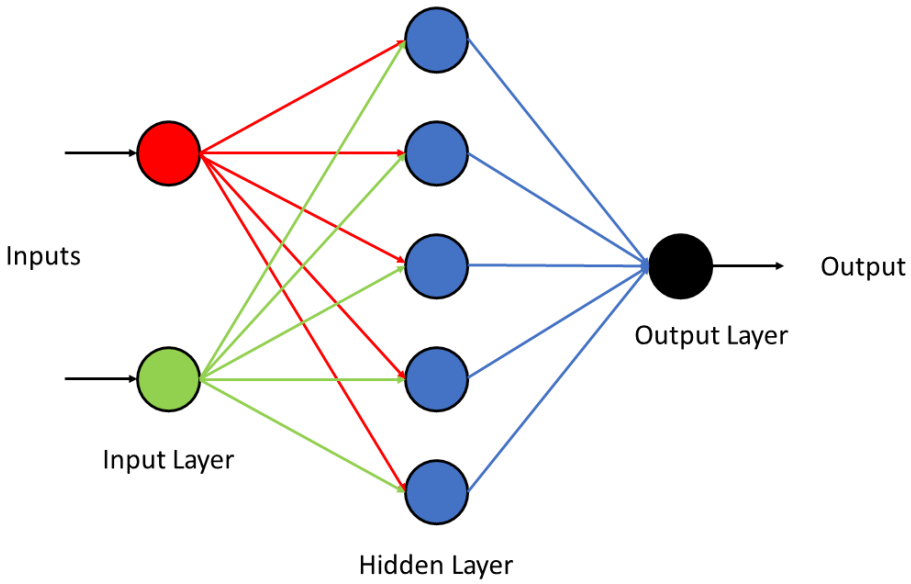


Figure 2.12: A 1-hidden layer ANN.

Chapter 3

Lightning Flash Rate Nowcasting Based on Ground-Based Polarimetric Radar Sensing and Machine Learning

3.1 Introduction

Timely detection and prediction of CG lightning strikes are essential for the safety of infrastructures, such as airports (Heitkemper et al. 2008; Steiner et al. 2012), and personal safety (Gomes 2017). Numerous previous studies have used various sensors and processing methods for detecting lightning and providing alerts (Carey and Rutledge 2000; Heitkemper et al. 2008; Hondl and Eilts 1994; Kohn et al. 2011; Lakshmanan and Smith 2009; Woodard et al. 2012). Weather radars (S-C-X band) have been used widely as one of the sensors for identifying the properties of the storm with lightning (Carey and Rutledge 2000; Hondl and Eilts 1994; Woodard et al. 2012). However, the applications are still limited in these aspects: (1) Polarimetric radar measurements are mainly used for classifying and detecting hydrometeors, not for predicting the flash rate. (2) The previous studies are limited to predicting the probability of lightning events instead of predicting the flash rates (flashes per unit time). (3) The usage of ML algorithms has been preliminary, primarily based on a few solutions, such as fuzzy logic (Kuk et al. 2012). (4) The lead-time for the prediction has been

very limited. The lead-time of up to 17 minutes has been demonstrated (Mosier et al. 2011; Wang et al. 2016). It is well-known that higher flash rates are usually associated with higher injuries and damage (Mazzetti and Fuelberg 2017). Flash rates can also be used for identifying conditions such as lightning jumps, which can help predict severe weather such as wind, hail, and tornados (Mazzetti and Fuelberg 2017; Williams et al. 1989). This work primarily focuses on using polarimetric radar measurements to predict the flash rate for lightning associated with thunderstorms. S-band polarimetric radar and lightning detection network datasets are combined for the training of ANN. The ELA is used in this implementation of the ML algorithm, for which a cluster of neural networks are trained, and their respective outputs are averaged to give a final prediction output (Smolyakov 2017). As part of supervised ML implementation, radar phenomenology knowledge and microphysical knowledge are incorporated into the feature selection process. A new group of features used in the ML algorithm is constructed, including general radar measurements and physical parameters outputs from storm-cell tracking algorithms. Representative and long-term data cases are used to verify the effectiveness of the proposed approach, and the prediction accuracy of lightning flash rates is evaluated up to 30 minutes of lead time.

3.2 General Concepts

The flow chart for the flash rate prediction (or nowcasting) is shown in Figure 3.1. The data used for lightning nowcasting include both dual-polarized weather radar (CINRAD) and the LLIS data, which are unique in terms of the location around mountains in Hong Kong. The LLIS data combined with radar data is used for the training process, and the polarimetric radar data alone is used for prediction. CINRAD is an S-band, simultaneous dual-polarized radar that is identical to NEXRAD. The radar antenna has a gain of 45.7 dBi and beamwidth of 0.92° for both azimuth

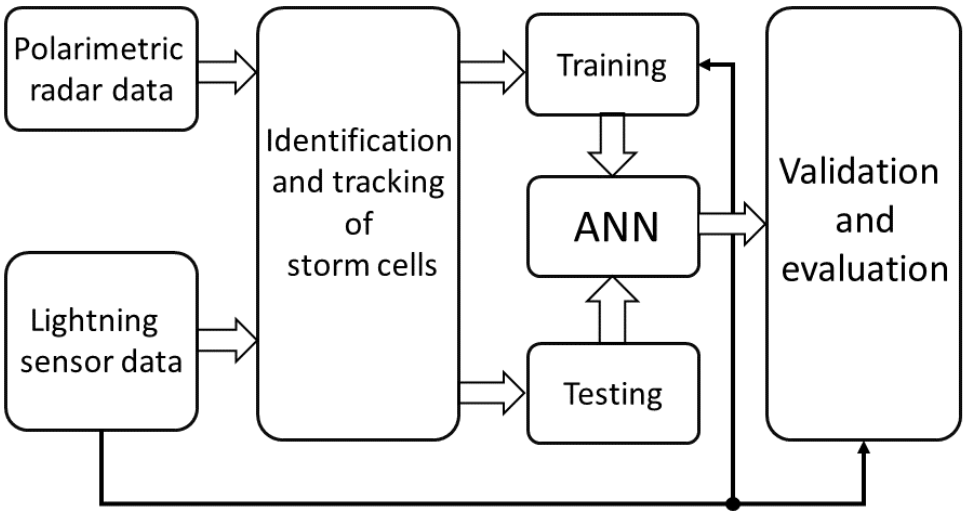


Figure 3.1: General (data processing) system operation and ML algorithm basic flow.

and elevation. The transmitter frequency is 2.92 GHz with a pulse width of 1.0/2.0 μs. Radar data from each Volumetric Coverage Scan (VCP) include reflectivity factor, velocity, spectrum width, Z_{dr}, Φ_{dp}, ρ_{HV} and K_{dp}. The detailed parameters on CINRAD are summarized in Table 3.1. On the other hand, the LLIS contains nine sensors manufactured by Vaisala, which have Detection Efficiency of up to 95% for CG lightning. Lightning data includes the time and location of the lightning strike, intensity, and the type of lightning (i.e., CG or intracloud (IC)). For each storm cell, "truth" flash rates are retrieved from the LLIS data, with the unit of flashes per minute per storm cell. Figure 3.2(b) shows the location map of the radar (the red mark) and the lightning detection network (black circles).

In Figure 3.1, firstly, storm cell identification and tracking are applied to the radar data. Tracking each storm cell is necessary for the study's approach because nowcasting is based on the "learned" behavior of individual storm cells. Also, the focus is on storm cell-based lightning nowcasting. For this approach, all the storm cells are uniquely identified and tracked throughout the VCPs, and predictions are made regarding the lightning flash rates of each storm cell, assuming cells, on average,

Table 3.1: Parameter of the CINRAD polarimetric weather radar system.

Parameter	**Description**
Year of implementation and operation	2014/15
Hardware supplier	Metstar, China
Software and radar processor	Sigmet IRIS
Polarization	Dual-Pol
Antenna height	588 m
Antenna diameter	8.54 m
Antenna gain	44 dB
Antenna beamwidth (AZ/EL)	0.92° for both azimuth and elevation
Elevation levels	0.07°, 0.86°, 1.82°, 2.66°, 3.58°, 5.38°, 9.95°, 14.96°, 21.95°, 33.95°
First sidelobe (down from the main lobe)	34 dB
Transmitter	Klystron
Peak power	650 kW
Pulse width	1.0/2.0 μs
Maximum range	256 km (every 6 minutes) 512 km (every 12 minutes)
Velocity dealiasing algorithm	Alternate PRF on ad
Nyquist velocity at lowest elevation scan	45.1 $\mathrm{m\,s^{-1}}$

(a) CINRAD radar.

(b) Locations of the radar and lightning network nodes (black circles).

Figure 3.2: Images of the sensors used in this study.

have the same characteristics. Different tracking algorithms are used in other studies (Dixon and Wiener 1993; Hu et al. 2019; Picel et al. 2018). For this research, TINT (TINT is not TITAN), which is based on the thunderstorm identification, tracking, analysis, and nowcasting (TITAN) algorithm, has been used for storm cell tracking (Dixon and Wiener 1993; Picel et al. 2018). Based on TITAN, storm cells are defined as contiguous regions with minimum reflectivity and volume of 30 dBZ and 25 km^3, respectively. The 30 dBZ threshold is applied for the plan position indicator (PPI) for all elevation levels. Polarimetric radar measurements for every radar resolution volume are used as input to TINT. Then, parameters and features required for tracking and ML are computed for each storm cell.

To analyze the radar features and validate initial ML studies, the entire data set containing a series of thunderstorm cases from 2017 to 2020, with each case containing 2-5 hours of observations, and covering different seasons, was used. These data sets are divided randomly, so training is done using some storm cases while testing is done using other storm cases. The operation of lightning nowcasting implemented a "sliding window" approach in that the usage of sensor data moves from the older/history datasets toward the newer/current datasets. For training, output parameters from

the storm tracking algorithm are combined with the number of flashes obtained from LLIS to form the training database. To address the operational requirements of specific aviation facilities, the training period and prediction lead time are set up to 30 minutes. For example, 30 minutes of current radar measurement data are used as training inputs. Then, the trained model will predict the lightning flash rates after up to 30 minutes. Meanwhile, the new radar measurements continue to arrive and become the new training data for further predictions. For validation, the trained neural network is tested using the test data up to 30 minutes later. For an example of small training data sets, three representative storm cases are shown in the following discussions. The first storm case was on 31 August 2017. The total number of unique storm cells identified was 248 for a duration of 3 hours. Approximately 40 cells were identified in each VCP scan. Similarly, the next storm case was on 22 August 2018. The total number of unique cells identified was 235 for a duration of 3 hours. Approximately 30 cells were identified in each VCP scan. Finally, the third storm case was on 20 April 2019. 574 individual cells are identified during a span of 5 hours. Approximately 30 storm cells are detected in each scan. In these storm cases, CG flash rates up to 2500 flashes per minute are observed in an area of about 11600 km^2. In addition, for an example of long-term training data set, the machine-learning algorithm was further tested for data obtained during 2020.

3.3 Selection of Features

Connections between dual-pol features and lightning have been studied before (Carey and Rutledge 1996, 2000; Caylor and Chandrasekar 1996; Hondl and Eilts 1994; Hu et al. 2019; Jameson et al. 1996; López and Aubagnac 1997; Mattos et al. 2016; Wiens et al. 2005). This study's selection of radar features has been based on previous studies of physical models and radar data analysis. During the physical process, a thunderstorm contains condensates in the form of supercooled water and graupels at

the lower level and ice crystals in the upper layer. During thunderstorm generation, these graupels and ice crystals collide through updrafts and downdrafts, transferring charges between them (Heitkemper et al. 2008). This Non Inductive (NI) charging causes the ice crystals and graupels to carry the electric charge of opposite polarity. Ice crystals carry the positive charge, while graupels, which reside at a lower height, carry the negative charge. This charging is the principal mechanism responsible for cloud electrification. The NI charging mechanism is also used to explain the tripolar and multipolar charge structure (Baker and Dash 1989; Bruning et al. 2014; Keith and Saunders 1990; Stolzenburg et al. 1998; Williams et al. 1989). The compositions, shapes, and orientations of these particles show distinctive features that are manifested through polarimetric radar measurements.

The reflectivity factor has been a useful feature of this microphysical process (Doviak et al. 2006). A storm cell with lightning usually has larger reflectivity values (>30 dBZ) above freezing altitude (Carey and Rutledge 1996, 1998). For a storm cell above the freezing altitude, water will be in the form of supercooled liquid, graupels, or ice crystals (Xu et al. 2010). The presence of ice crystals and graupels, which have lower reflectivity values and higher reflectivity values, respectively, in the upper layer (above freezing altitude) of the storm, is an important lightning initiating condition. In addition, reflectivity values between 30-35 dBZ between -10°C and -20°C isotherm is a good indicator of lightning initiation. Since there is a direct correlation between altitude and temperature, the height of the echo top has been related to the radiosonde temperature profile in this study.

Figure 3.3(a) shows one example of radar reflectivity PPI. As anticipated, the storm cells with a higher reflectivity have more flashes associated with them, as shown in Figure 3.3(b). Also, in 3.3(b), there are storm cells with comparatively higher reflectivity values but have fewer flashes, as shown in area 2. In contrast, area 1 has a considerably greater number of flashes. This can be explained by their vertical

profiles as given by Range Height Indicator (RHI) plots. Storm cell of area 1 has a higher echo top (10 km) than the storm of area 2 (8 km), shown in Figure 3.3(c). Higher echo tops were found to be associated with lightning. The higher probability of lightning for storm cells with higher echo top can be explained based on decreasing temperature with the increase of altitude in the troposphere, which is shown in Figure 3.3(d). Figure 3.3(d) gives a general idea of the attitude range for the temperature to be between -10 and -20 $^\circ$ centigrade. It can be inferred that the storm cells with an echo top greater than 7 km have the probability of lightning. Analyzing the dataset, it was found that the higher the altitude of the echo top of a storm cell, the greater number of flashes it produces. In Figure 3.3(b), area 2 has less number of flashes than area 1, which could be because the vertical extent of storm cell area 2 is shorter (8 km) than that of storm cell area 1 (10 km).

Z_{dr} has higher values for rainfall due to the oblate shape of waterdrops, while it is usually smaller for the ice crystals in the absence of any electric field due to their random orientation (Doviak et al. 2006). A sharp change in Z_{dr} around 0° isotherm indicates the presence of water and supercooled liquid and ice particles – the condition required by the lightning (Xu et al. 2010). Updrafts in convective storms have been shown to cause lightning. Typically positive Z_{dr}-columns and their vertical extent are well-correlated with updraft intensity and freezing of drops in these columns. These regions are likely sources of graupel embryos (Homeyer and Kumjian 2015; Kumjian et al. 2013). Therefore, the presence of a Z_{dr} column can be a useful feature of lightning.

K_{dp}, which is the differential phase shifts at two locations, has positive values for horizontally orientated hydrometeors and negative values for vertically orientated hydrometeors. Since ice crystals are presumed to be orientated randomly, K_{dp} values will be primarily close to 0 $^\circ\,\mathrm{km}^{-1}$ above freezing altitude, while having negative values sometimes since there are vertically orientated crystals affected by the strong

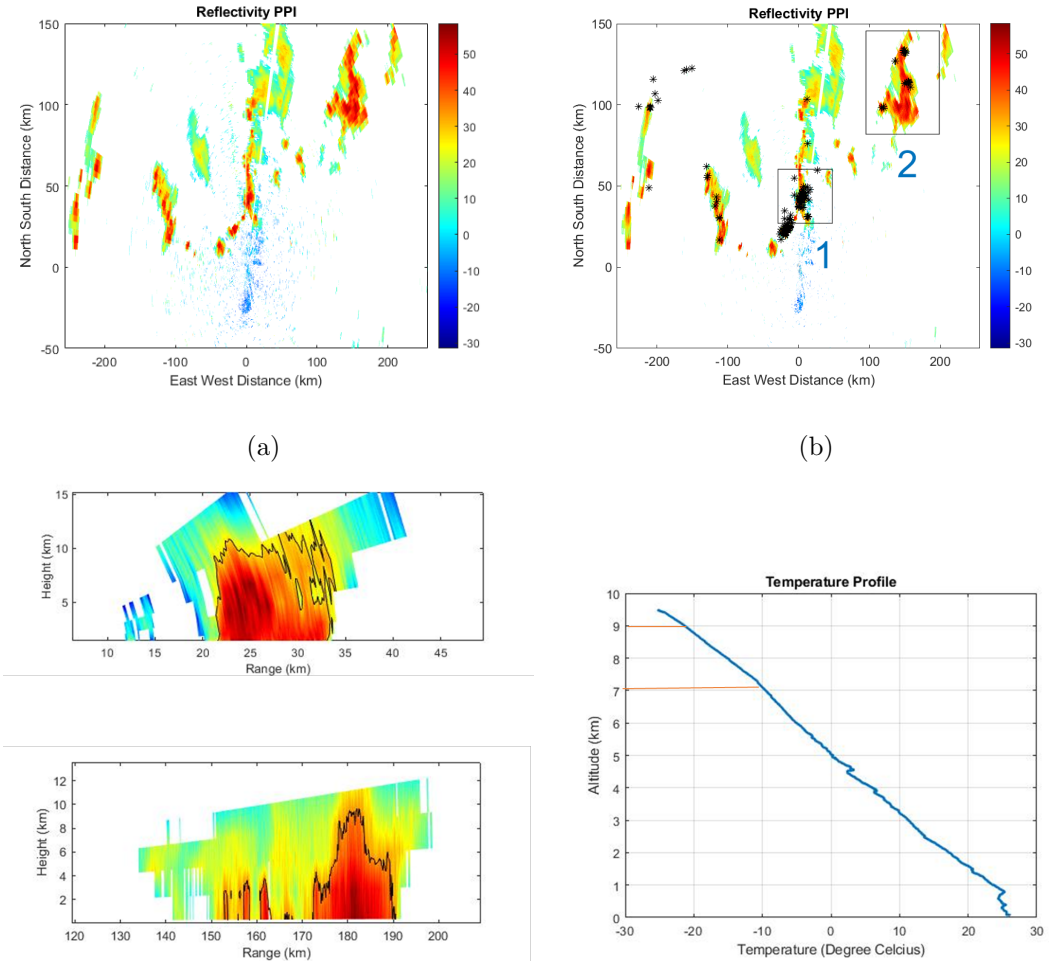
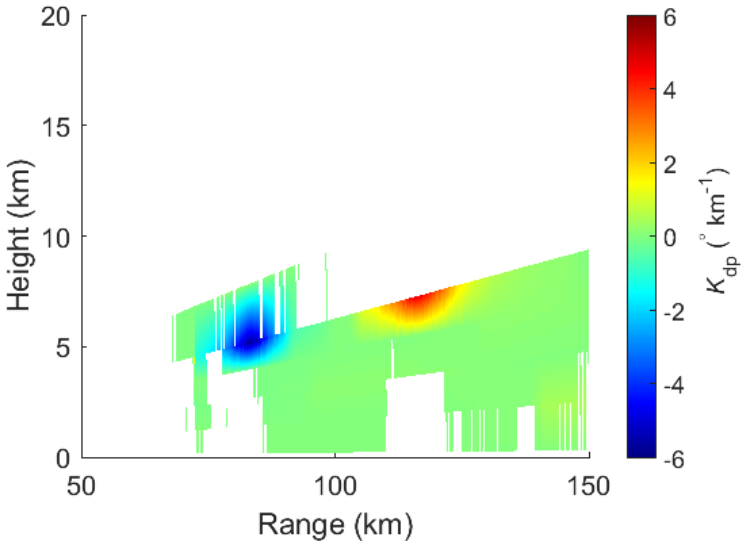


Figure 3.3: Dependency of the lightning on reflectivity and temperature.(a)Reflectivity PPI display from the CINRAD, (b)Lightning Location in Reflectivity PPI represented by blue scatter.1 and 2 represents the area with higher flash rate and lower flash rate respectively, (c) RHI plot of reflectivity for area 1 (top) and area 2(bottom), (d) Temperature profile of the storm on August 31st, 2017

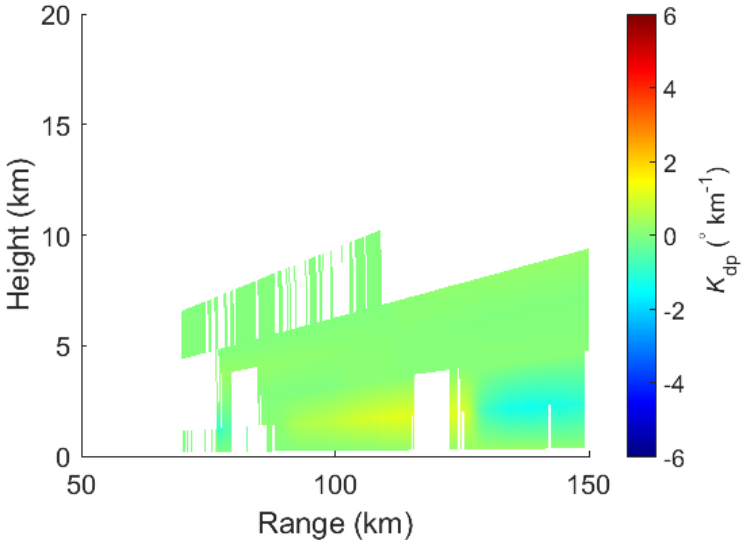
electric field (Preston and Fuelberg 2012). For example, Figure 3.4 shows the vertical profile of K_{dp} values before and after lightning initiation. In Figure 3.4(a), the presence of positive and negative K_{dp} values in the vertical profile observed at 12:00:08 is due to both horizontally and vertically orientated ice crystals. However, in the vertical profile observed at 12:06:09, as shown in Figure 3.4(b), which is after the lightning, K_{dp} values are more "normal" based on that ice crystals appear to be more randomly orientated.

ρ_{HV} has been primarily used to distinguish between meteorological and non-meteorological targets. It can also be used to study specific types of hydrometeors. Although the association of the (ρ_{HV}) with the lightning has been studied, a significant difference was not found in the distribution of ρ_{HV} values in the storm with and without thunderstorm and has not been used as features in this study.

Figure 3.5 shows the distribution of reflectivity (Figure 3.5(a)), Z_{dr} (Figure 3.5(b)), K_{dp} (Figure 3.5(c)) and ρ_{HV}(Figure 3.5(d)) collected from resolution volumes from the storm cells with and without lightning. The reflectivity distribution shows that storms with lightning have higher reflectivity values and a higher standard deviation of the reflectivity values. The distributions for both Z_dr and K_{dp} for storm cells without lightning are centered on 1 dB and 1 $^\circ\,\mathrm{km}^{-1}$. In contrast, for storm cells with lightning, the distributions are centered between 0 and 1, which confirms the possible changes in particle orientations. In terms of ranges of values, it was also found that Z_{dr} and K_{dp} values for storms with lightning can be as low as -4.94 dB and -8.27 $^\circ\,\mathrm{km}^{-1}$ while the storms without lightning have minimum Z_{dr} and K_{dp} values of -2.45 dB and -2.29 $^\circ\,\mathrm{km}^{-1}$, respectively. The presence of ice crystals and graupels above the freezing altitude (mixed-phase region) also attributes to the lower Z_{dr} values. Similar observations were made for the maximum value of Z_{dr} (5.63 vs 3.77 dB), and K_{dp} (8.31 vs 1.58 $^\circ\,\mathrm{km}^{-1}$) as storms with and without lightning were compared. This is because of the heavy rainfall that was observed during the



(a) K_{dp} values at 12:00:08.



(b) K_{dp} values at 12:06:09.

Figure 3.4: Vertical profiles of K_{dp} values before and after lightning strikes as seen in two consecutive scans (12:00:08 and 12:06:09) in a storm cell of storm case of 20 April 2019.

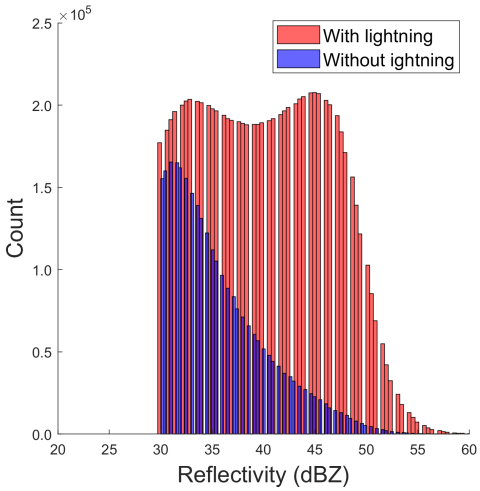
storm with lightning. This gave higher Z_{dr} values due to oblate water drops at a lower altitude. While significant statistical differences were observed for other radar variables, statistical measures were found to be similar when ρ_{HV} values of storms with and without lightning were compared.

Based on these observations, Z, Z_{dr}, and K_{dp} are the three most essential features for training ML algorithms. Specific statistics of these three measurements, such as maximum, minimum, and average values, are used, as part of the features of the algorithms. The detailed construction of the training feature vector is illustrated in Table 3.2, which combines features extracted from 5 consecutive VCPs to generate a training data set that covers a 30 minutes time span. For each VCP, 14 feature elements are extracted from radar data. The length of the feature vector reduces accordingly if a lead time shorter than 30 minutes is needed.

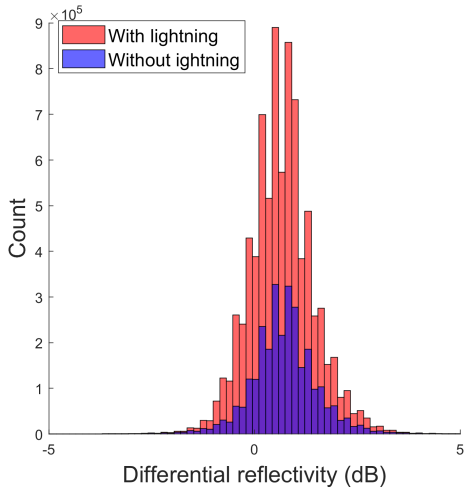
3.4 Training and Testing Procedure

Training and testing samples for supervised learning are generated from the storm cells by extracting the selected features at each scan. For a 30 minutes lead time, storm cells that have a "life span" of at least 60 minutes, or 10 VCPs, are used. The goal is to train the algorithm with up to 30 minutes of data and then estimate the flash rates of the storm cells up to 30 minutes later. The actual lead time can vary from 6 minutes to 30 minutes.

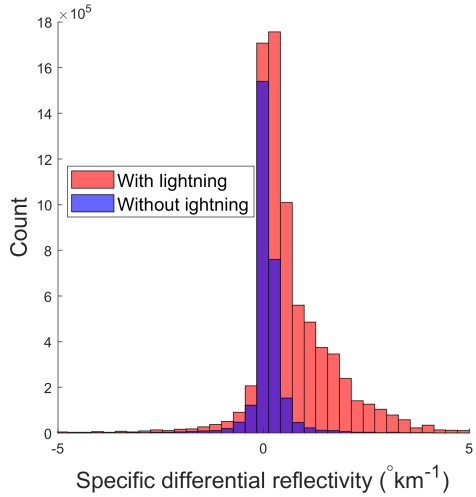
Figure 3.6 is used to clarify further the approach used for 30 minutes lead time. Since each VCP lasts for around 6 minutes, 5 consecutive VCPs, representing 30 minutes, can be used to predict the flash rate for the 10th VCP, i.e., after 24-30 minutes. Suppose the red and green storm cells are tracked consecutively over 11 VCPs; each cell creates two training or testing sample sequences. One of them uses storm parameters from VCP 1 to 10, and the other uses parameters from VCP 2 to 11. For VCP 1-10, the 6th VCP represents the current time. VCP 1 to 5 from both



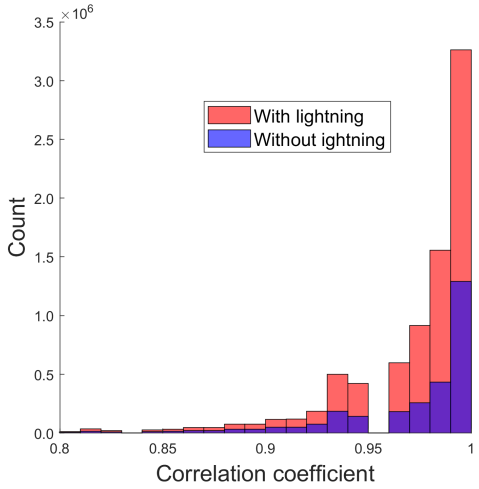
(a) Histogram of reflectivity.



(b) Histogram of Z_{dr}.



(c) Histogram of K_{dp}.



(d) Histogram of ρ_{HV}.

Figure 3.5: Histograms of the values of (a) Reflectivity in dBZ, (b) Differential Reflectivity in dB, (c) K_{dp} in $^\circ\,\mathrm{km}^{-1}$ and (d) ρ_{HV} of a thunderstorm for the radar measurement case of 20 April 2019. Storms with lightning have the Z_{dr} distribution with a slight negative shift and a positive shift of the K_{dp} distribution, suggesting horizontal orientations of particles.

Table 3.2: Feature parameter used for ML algorithm for 30 minutes prediction lead time. In total a 70-element feature vector is listed. Each parameter derived from Dual-Pol radar measurements are listed in each row. Each parameter is extracted from five consecutive scans, and each extraction results in an element in the feature vector, which is indexed by its location in the feature vector (a number in the second column).

Parameter	**Feature indices**
Storm area	1, 15, 29, 43, 57
Storm volume	2, 16, 30, 44, 58
Maximum reflectivity	3, 17, 31, 45, 59
Average reflectivity	4, 18, 32, 46, 60
Standard deviation of reflectivity	5, 19, 33, 47, 61
Maximum Z_{dr}	6, 20, 34, 48, 62
Minimum Z_{dr}	7, 21, 35, 49, 63
Average Z_{dr}	8, 22, 36, 50, 64
Standard deviation of Z_{dr}	9, 23, 37, 51, 65
Maximum K_{dp}	10, 24, 38, 52, 66
Minimum K_{dp}	11, 25, 39, 53, 67
Average K_{dp}	12, 26, 40, 54, 68
Standard deviation of K_{dp}	13, 27, 41, 55, 69
Number of CG flashes in six minutes	14, 28, 42, 56, 70

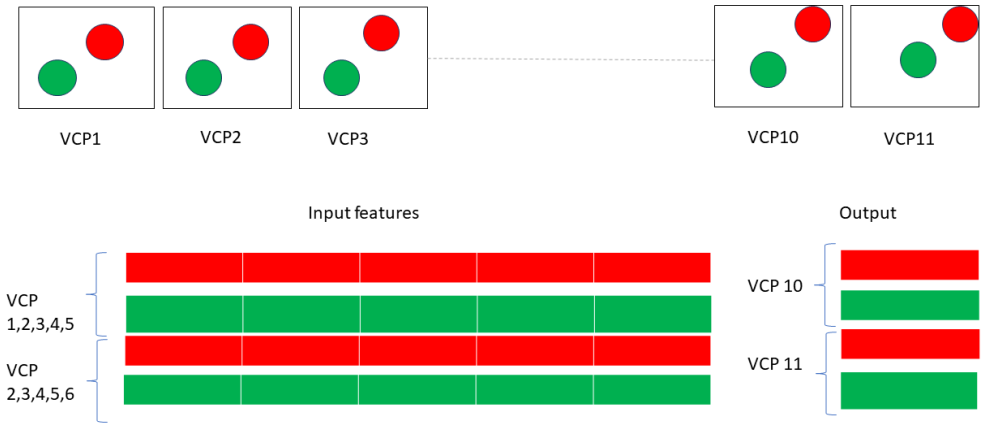


Figure 3.6: Illustrative description of creating training and test samples from radar VCPs.

the storms are used to create two input data of training/testing samples, and the prediction is made for the 10th VCP. For 2-11, the 7th VCP represents the current time, and VCP 2 to 6 forms the input data. The process continues in a "sliding window" approach. Depending on the lead time, different numbers of VCPs can be grouped into training or testing data with different spans. The "sliding window" approach is applied to all the storm cells in all VCPs for all storm cases, regardless of training or testing.

Two approaches are used for training and testing. In the first approach, the complete lightning datasets are randomly divided into training and testing datasets (70% training and 30% testing) based on specific storm cases. The second training and testing approach is to emulate the real-time nowcasting application; for this approach, during any stage of training, the system is retrained using additional training samples obtained from the storm case under test. The periods of these additional training samples are up to 30 minutes before the testing VCPs. Thus, the training data will have knowledge about storm parameters and flash rates of the storm under test, such as in the actual "sliding window" implementation. This approach is called "real-time" training in the following discussions. Further illustration is shown in Figure 3.7. The

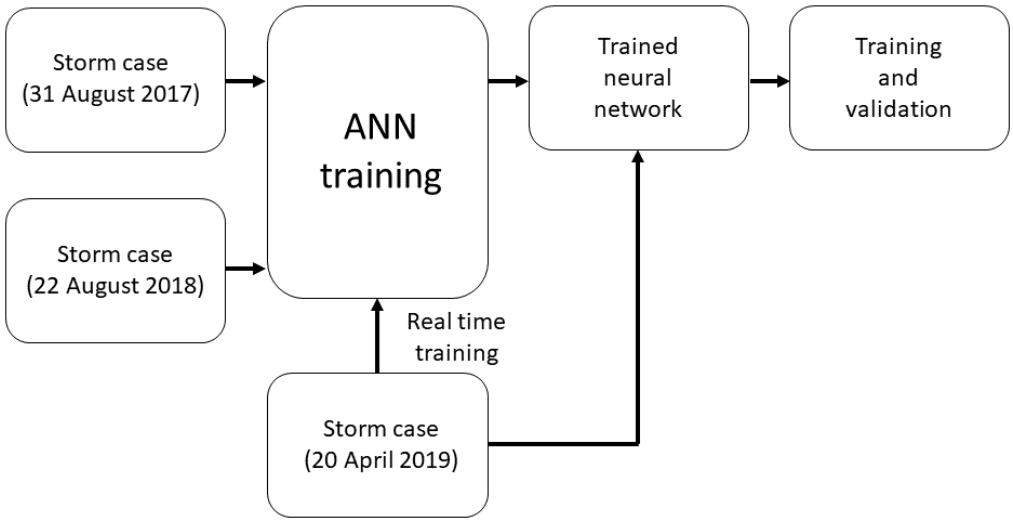


Figure 3.7: Illustrative description of an example of training process for nowcasting using multiple radar-measured storm cases.

study assumes that the training samples are created using storm cases of 31 August 2017 and 22 August 2018, and test samples are created using storm cases of 20 April 2019. The training process is then modified to implement "real-time training" by adding the information of the storm case under test on 20 April 2019 as part of the training samples. With the "real-time training", after an hour of initiation of the nowcasting process, storm parameters and true flash rates from the test storm case of 20 April 2019 are gradually ingested into the training samples. The model is then retrained with the updated training data before updated predictions are made.

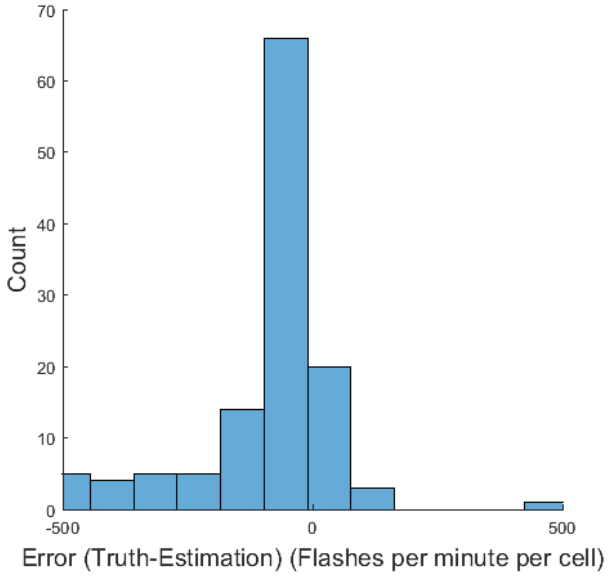
The ML model is an ANN with three hidden layers. Each hidden layer contains 20, 30, and 20 nodes, respectively. The selection of the neural network architecture is based on its accuracy and efficiency after performing experiments with different numbers of hidden layers and the number of nodes. The training function was selected as the Levenberg-Marquardt function, and the internal performance index is mean-squared-error. The training dataset is further divided into internal training and cross-validation datasets, which are used through an iterative process. First, a cluster of ten neural networks is trained, and then the final output from the system is calculated

as the average output values of the ten neural networks from the cluster. There is no difference among the neural networks regarding their basic structures. However, since they have different initial conditions and take different data inputs, an individual network may produce either overfitted or underfitted results, which is difficult to control. Such issues can be mitigated using multiple networks trained with randomly assigned initial weights.

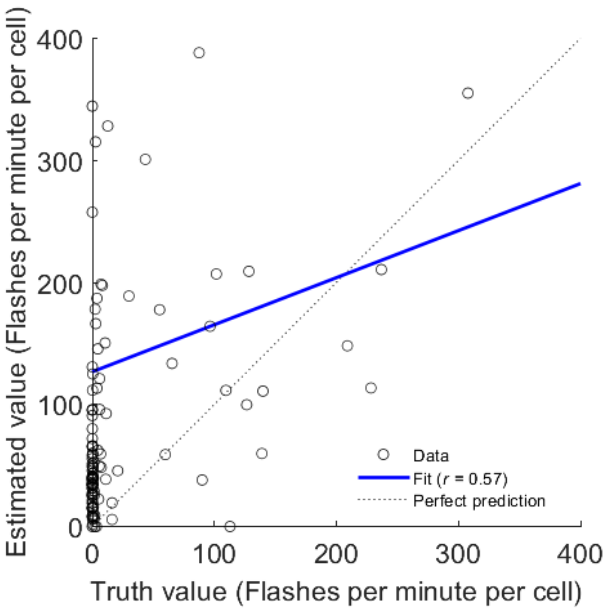
3.5 Results of Short-term Representative Cases

The ML-based solution's objective is to predict (or nowcast) the flash rates (defined as the number of CG flashes per minute per storm cell) associated with storm cells contained in radar VCPs up to 30 minutes ahead. Two performance metrics are used: Pearson's correlation coefficient (ρ), which measures the level of correlations between truth and predicted lightning flash rates, and the mean percentage error (MPE), which is defined as averaged percentage errors comparing predicted and true values of flash rates. The first measure is essential for general ML performance evaluation, while the second is helpful for the specific estimation problem.

Figure 3.8 illustrates the 30 minutes lead time performance of a single ANN based on randomly selected training and testing samples. The number of training and test samples were approximately 520 and 130, respectively. Although the distribution of error values are centred on zero in Figure 3.8(a), the ρ for the model was found to be 0.57, according to the regression plot in Figure 3.8(b). The ρ varies form about 0.5 to 0.8. Predictions from individual ANNs may result in various ρ values due to the random division of training and cross-validation sets during the iterative process. For the similar 30 minutes lead time test, Figure 3.9 shows how the overall accuracy of the prediction system improves by training and combining the cluster of the neural networks. The ρ is found to be 0.81 when the outputs from 10 ANNs are averaged.



(a) Distribution of errors (truth minus predicted).



(b) The corresponding regression/scatter plot.

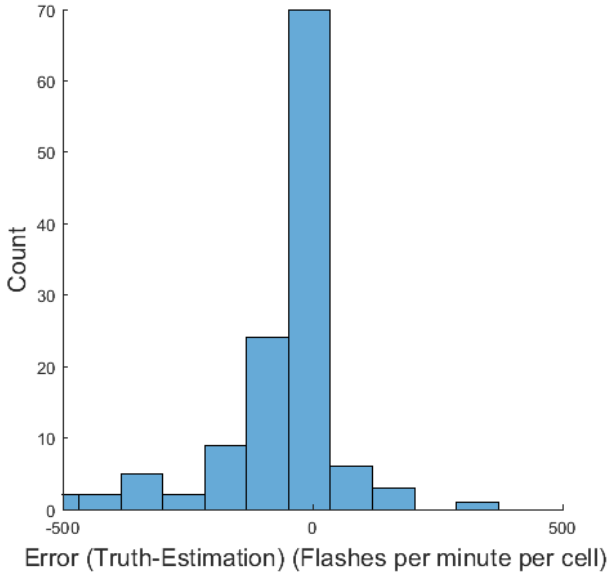
Figure 3.8: Prediction performances of 30 minutes lead-time based on error distribution and ρ for an individual ANN. Dual-pol features are used.

Table 3.3: Comparison of flash rate estimation accuracy when dual-pol features are used and not used in training. For both methods, 10 ANNs are trained.

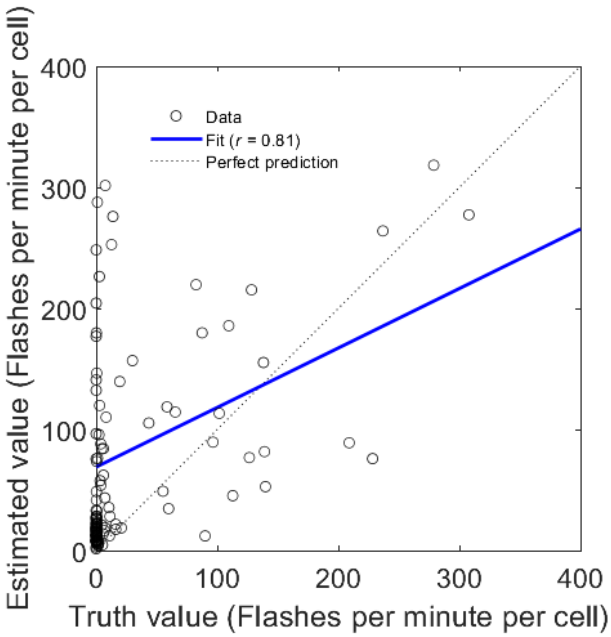
Metrics for comparison	**Dual-Pol**	**No Dual-pol**
Number of ANNs with ρ greater than 0.8	1	0
Number of ANNs with ρ between 0.7 and 0.8	4	0
Number of ANNs with ρ between 0.6 and 0.7	2	1
Number of ANNs with ρ less than 0.6	3	9
Mean ρ for all ANNs	0.82	0.68

The occurrence of error values centered on zero also increases because of ensemble training.

Even with ensemble training, when the dual-pol features are removed from the training dataset i.e., only reflectivity statistics are used, the performance of the now-casting degrades drastically. The distribution of error values is no longer centered on zero, and the ρ from the average neural network with the random selection of training and testing samples was found to be 0.49. Summary of the performance of networks when trained using dual-pol and without dual-pol features for the random division of training and test samples is shown in Table 3.3. It is clearly illustrated that the usage of dual-pol features leads to better flash rate estimation performance. The ρ is significantly larger when the training was done, including dual-pol features (0.82 vs. 0.68 for 30 minutes lead time). The differences between statistical metrics of dual-pol variables with and without lightning proved to be essential input features in the ML training.



(a) Distribution of errors (truth minus predicted).



(b) The corresponding regression/scatter plot.

Figure 3.9: Prediction performances of 30 minutes lead-time based on error distribution and ρ for 10 ANN clusters. The accuracy improves compared to Figure 3.8.

Table 3.4: Comparison of performance of different lead times (DP/no DP). 6th, 7th, 8th, 9th, 10th VCP correspond to 6 minutes, 12 minutes, 18 minutes, 24 minutes, and 30 minutes, respectively

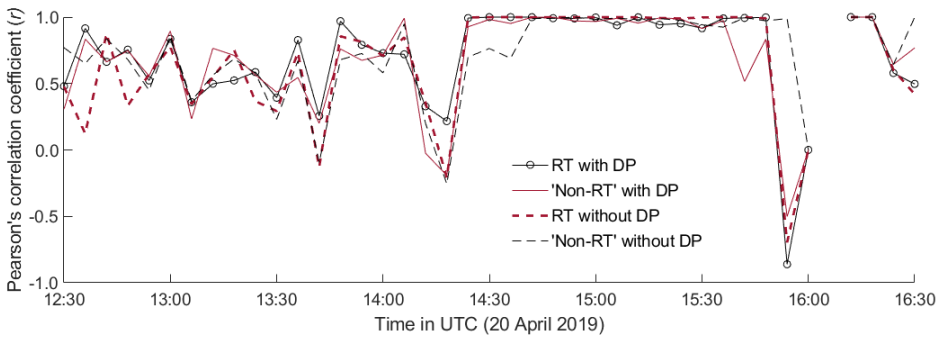
Metrics for comparison	**6th VCP**	**7th VCP**	**8th VCP**	**9th VCP**	**10th VCP**
Number of ANNs (ρ>0.8)	6/6	5/7	3/0	2/0	1/0
Number of ANNs (0.7<ρ<0.8)	4/2	4/3	6/0	4/0	4/0
Number of ANNs (0.6<ρ<0.7)	0/2	0/0	1/4	3/2	2/1
Number of ANNs (ρ<0.6)	0/0	1/0	0/6	1/8	3/9
Mean ρ for all ANNs	0.86/0.90	0.86/0.88	0.81/0.66	0.83/0.66	0.82/0.68

The next example shows the prediction performance for a 12-minute lead time. Figure 3.10 depicts the time-evolution of the Pearson's correlation coefficient (ρ) values and MPE values for the storm case of 20 April 2019, which is used for testing. The training is done using two storm cases: 31 August 2017 and 22 August 2018, and the real-time training starts around 14:00. From Figure 3.10(a), it is observed that when dual-pol features are included in real-time training, the performance is better in correlation. Between 16:00 and 16:30, ρ are comparatively low, and MPEs have high values. The storm tracking algorithm detected the merging of storm cells during this period. As a result, the ML algorithm cannot use the input features derived prior to the merging to make accurate predictions. Overall, models trained in real-time with dual-pol features have higher ρ for most durations, indicating a better ability to correlate the predicted and the true flash rates. This can also be observed in terms of MPE values in Figure 3.10(b). In summary, with small training data sets, ML models trained in real-time with dual-pol features have significantly lower overall MPE, in spite of possible bias during the course of the predictions.

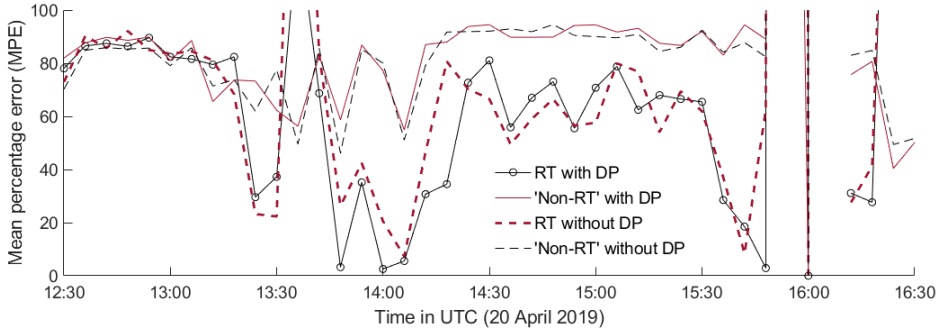
3.6 Results of Longer-Term Operation

The ML algorithm was further tested for data obtained for a longer-term operation. For this experiment, lightning data was collected from storms that occurred in the year 2020. In total, 20 thunderstorm cases were used. Although storm cases for all the months have been used to form the training data, storm cases from May and June constitute the most considerable portion. Naturally, thunderstorms are more frequent during the summer period of the year. Therefore, the training dataset consists of 1 case from March, 1 from April, 8 from May, 5 from June, 2 from August, and 3 from September. Each case contains 3-5 hours of training data.

Figure 3.11 shows the regression plots when the data obtained from the thunderstorms of 2020 was used in the ML experiment with a lead time of 30 minutes. The



(a) Pearson's correlation coefficient (ρ) for all 4 combinations

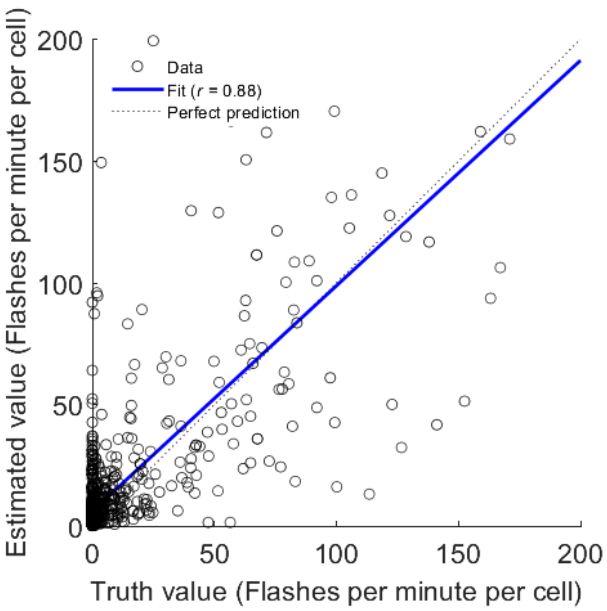


(b) Mean percentage error (MPE) for all 4 combinations.

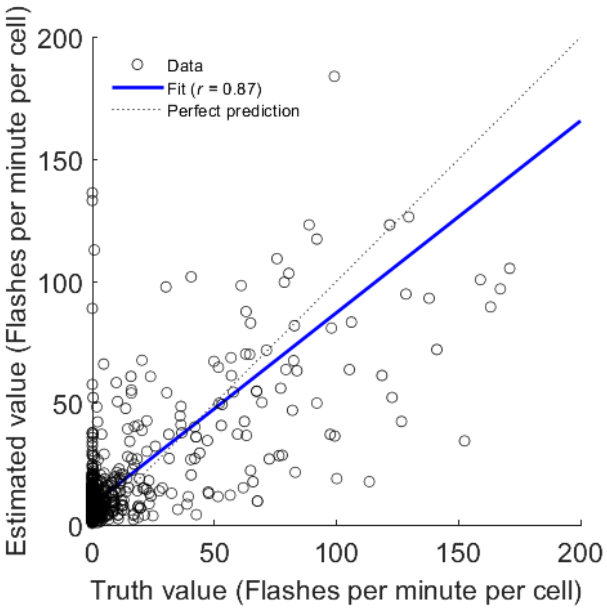
Figure 3.10: Performance of flash-rate prediction for 12 minutes lead time, using small training data sets, measured by Pearson's correlation coefficient (ρ) and mean percentage error (MPE) with and without real time (RT) and dual-pol features (DP).

total data were randomly divided into training and test sets. The number of training and test samples was approximately 4000 and 1000, respectively. As mentioned in previous sections, ten neural networks were trained, and the final output was calculated as the average value of the outputs from the ten individual networks. With and without the usage of dual-pol features, the Pearson's correlation coefficients (ρ) between true and predicted flash rates are found to be 0.88 and 0.87, respectively, as shown in Figure 3.11(a) and Figure 3.11(b). It is indicated that the proposed ML approach can be scaled for longer-term operational usage.

Besides the random division of data into training and testing sets, additional testing was done to emulate the lightning nowcasting operation. For this experiment,

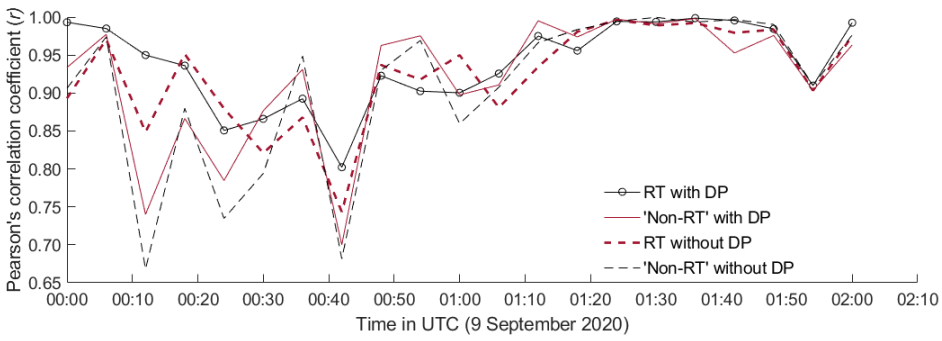


(a) Regression plot (With dual-pol features).

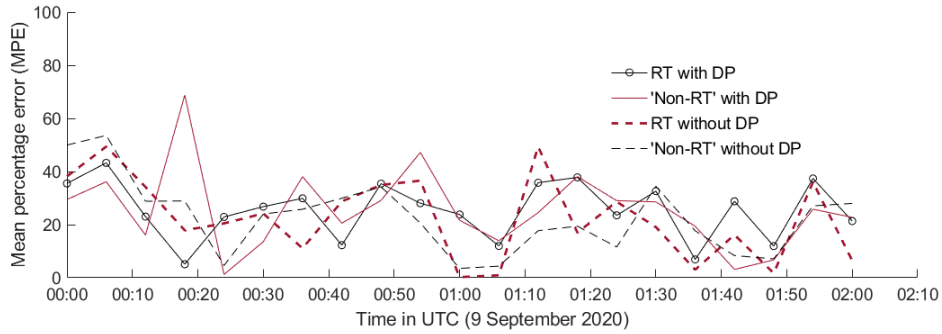


(b) Regression plot (Without dual-pol features).

Figure 3.11: Regression plots of the test samples obtained from the longer term ML regression experiment.



(a) Pearson's correlation coefficient (ρ) for all 4 combinations



(b) Mean percentage error (MPE) for all 4 combinations.

Figure 3.12: Performance of flash-rate prediction for 30 minutes lead time, measured by Pearson's correlation coefficient (ρ) and mean percentage error (MPE) with and without real time (RT) and dual-pol features (DP), based on long-term data training.

most of the 2020 lightning cases are used for training, and one case in September (9 September 2020) is used for testing. Figure 11 shows that for both metrics, Pearson's correlation coefficient (ρ) and MPE display good performance compared to the models trained from smaller data sets. Figure 11 also shows that even without using dual-pol features and real-time training, good prediction accuracy can be achieved, provided the ML model has been trained with larger datasets. This is possible because the larger data sets contain more information about the characteristics of individual storm cases. Similar performance can be observed when using other cases for testing and cases prior to the case for training, which further validated the operational value of the proposed algorithms.

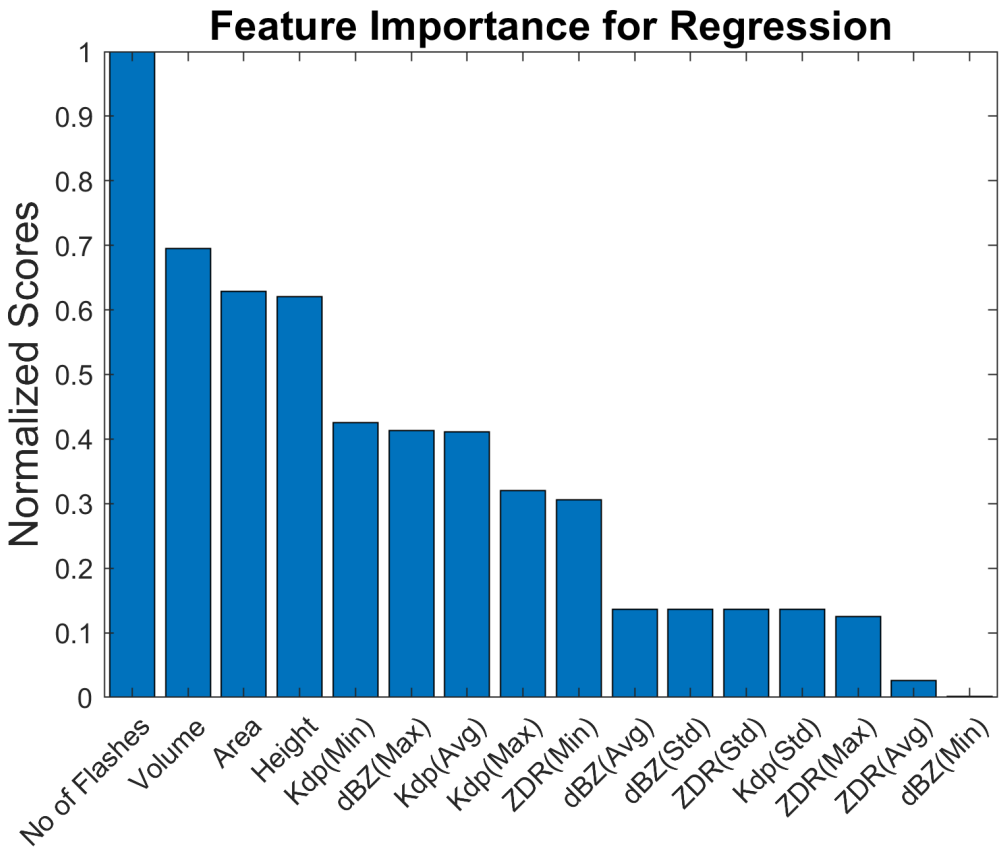


Figure 3.13: Feature importance metrics for all the input variables.

3.7 Analysis of Feature Importance

The importance of each feature can be determined by using a univariate F-test (Guyon and Elisseeff 2003). Each feature is examined using cross-correlation to determine the strength of the feature's relationship with the response variable. The cross-correlation is then converted into F-tests and p-values. The most important predictor has the smallest p-value of the test statistic. This technique is based on the statistical inference and does not depend on the ML algorithm as in the case of Permutation feature importance technique. Thus, the scores obtained using the F-test technique are consistent for different types of ML model. This can also be used to filter out irrelevant features before the training of ML models.

One input VCP (the first of the 5 VCPs) is examined to evaluate the importance of input features. It is clear from Figure 3.13 that the number of flashes in the previous scans has the maximum weight in predicting the future flash rate. Storm cell volume and area are also essential features since storm areas are expected to be directly proportional to lightning rates. The height of the storm, which is related to the temperature, is also of greater importance. The minimum value of K_{dp}, the maximum value of reflectivity, and the average value of K_{dp} values have approximately the same level of importance. The standard deviations have relatively minor importance, and the minimum value of reflectivity (which is about 30 dBZ) has the least importance.

3.8 Conclusion

In conclusion, this study has demonstrated that it is feasible to use polarimetric radar measurements to predict (or nowcast) lightning flash rates up to 30 minutes of lead time. The achieved prediction performance is based on the careful selection of feature variables as well as the ensemble technique in ML algorithm implementation. Also, it is shown that using dual-pol variables in training and using the "real-time training" technique can effectively improve the flash rate prediction performance when training is done using small, short-period data sets, although in some cases, other factors such as quality of training data and storm tracking results may affect actual results. On the other hand, when long-term training data is available, performance may be further enhanced using long-term data from different seasons, and the impacts of real-time training and dual-pol observations are less significant. The validations of the proposed algorithm are done using both representative cases from previous years as well as long-term data collection during the year 2020.

Chapter 4

Detection of High Ice Water Content Through Airborne Radar Sensing and Machine Learning

4.1 Introduction

HIWC represent a significant aeronautical hazard and a threat to jet engine operations at high altitudes (Dye et al. 2004) including commercial vehicles, rockets, super-sonic vehicles, and other space exploration mission platforms. The ice crystal causes jet engine power loss and engine rollbacks. The development of the next generation of airborne aviation weather radar has an urgent need to detect and monitor HIWC conditions. Data analysis studies, flight campaigns, and modeling studies have been ongoing to address the HIWC challenge in the aviation community (Harrah et al. 2019; Li et al. 2010). A physical knowledge-based model of radar observation of HIWC is needed for two reasons. First, it will guide the understanding and interpretation of the radar measurements from existing flight campaigns. Second, it will guide the industry and manufacturers in designing, developing, and certifying the new airborne radar products. There is numerous previous modeling effort related to airborne weather radar. For example, (Li et al. 2010) established the initial framework of using the Numeric Weather Prediction (NWP) model in the airborne aviation hazard estimation.

The extension of Airborne Doppler Weather Radar Simulation (ADWRS) for ice crystals was inveestigated, which would be part of the standard tools for industry developers (Harrah et al. 2019). The method of simulating ground-based polarimetric phased array radar using the advanced MATLAB tools was first introduced (Li et al. 2019, 2020). The airborne radar sensor simulation was extended to include antenna tilt control and various clutter effects (Gilliam 2020). None of these developments, however, have provided a validated simulation model and complete procedure for airborne forward-looking HIWC detection, especially for dual-polarized operations. On the other hand, there are a large amount of data from the previous flight campaigns (Harrah et al. 2019; Strapp et al. 2019, 2016b), and significant phenomenology analysis of these data (Hu et al. 2021). However, there is still a lack of connection between the scientific knowledge derived from these data and the application of the knowledge to actual radar designs. Further, lack of physical understanding and interpretation has hindered the acceptance of some practically appealing HIWC detection algorithms like Radar estimated IWC (RIWC) (Harrah et al. 2019) by the radar community since there is no model describing the origin of pulse-to-pulse variation of received signal amplitudes.

This study introduces an improved airborne weather radar simulation tool and modeling method. The modeling starts from validated microphysical models of ice particles and then uses them to simulate the stochastic behaviors of ice particles in a unit-sized, hypothetical radar resolution cell. Each radar pulse samples a Monte-Carlo realization of the particle states in the volume, according to the Probability Distribution Functions (PDF) between ice-water-content (IWC) and radar variables and the radar sensor parameters. The complete radar scans and I/Q signals can then be generated from the NWP truth field grid of IWC and the end-to-end system tool similar to (Li et al. 2010), while the airborne radar characteristics are included. The novel benefit of this approach is that a physics-based validated "radar observation model"

provides the connection between the IWC and the radar measurements, which incorporates the randomness in the measurement caused by ice crystals. Such randomness may be useful for HIWC detection but has not been fully understood. Also, the new simulation tool adds flexibility to the overall system evaluations since it provides a unified interface to read NWP IWC field and aviation encounter cases. Furthermore, for the first time, it can use 3-D data grid and antenna pattern, and then translate the samples to radar spherical grids. In addition, the application of the simulation model with ML based HIWC detection and classification cases is demonstrated, using realistic radar signals produced from the simulations.

4.2 Development Procedure

The initial modeling has been based on fight measurement data, including NASA's DC-8 campaign, NRC-Convair Campaign, and French Falcon Campaign. All the data collections were performed in 2015; while the NASA campaign was around the Gulf of Mexico, the latter two campaigns were in Cayenne, French Guiana (Hu et al. 2021). More details about the data sets used in these campaigns are listed in the Table. 4.1. Since NASA's flight campaign was the only one that included the forward-looking airborne weather radar data, the NASA flight data are used to verify the radar scan returns. However, data from the other campaigns are also very important for SCMC inputs. During NASA's flight campaigns in 2015, the instrument suite on the aircraft took measurements of the HIWC conditions associated with mesoscale convection systems along the coast and in the oceans, as well as tropical storms. Detailed descriptions of the flight test campaigns that acquired these in-situ probe data are available in (Harrah et al. 2019; Ratvasky et al. 2019).

The sensor data association supporting SCMC development is done among the measured radar reflectivity (can be extended to other radar parameters) using an

Flights	Temperature Range ($^{\circ}C$)	IWC Range (g/m^{-3})	# of PSD samples	% of Size Bins Set to Pure Ice Density	Typical AR values	Typical Canting Angles($^{\circ}$)
NASA DC-8	-58 to -15	0.001 to 3.2	22088	1.62	0.5	0
NRC Convair	-45 to -5	0 to 4.2	21598	28	0.5	35
French Falcon	-45 to -3	0 to 4.2	21614	30	0.7	20

Table 4.1: Comparisons of flight campaign datasets used in this study.

X-band airborne radar, as well as IWC and Particle Size Distribution (PSD) measured by in-situ probes. Spatial matching and correlation are done firstly; latitude, longitude, and elevation coordinates of each resolution volume were calculated based on the aircraft's latitude and longitude, altitude location data, and radar's azimuth, elevation, and range in PPI scans. Then these locations are compared with the header information obtained from the in-situ probe data files. The nearest location matching probe locations are recorded as the correlated radar resolution volume. Secondly, it was made sure that the difference in timestamps between radar data and IWC data was no greater than 15 minutes. It helps avoid the association with the resolution volume during the return flight path. Thirdly, timestamps from the IWC file and PSD file are compared and matched to incorporate PSD sensor data. The IWC values recorded with the IKP-2 probe are linked with the PSD data recorded with 2D-S and PIP microphysical imaging probes. Because of the in-situ measurements of both the IWC and PSDs, it was easier to correlate the sensor data based on the time. This sensor association process is summarized in Figure 4.1.

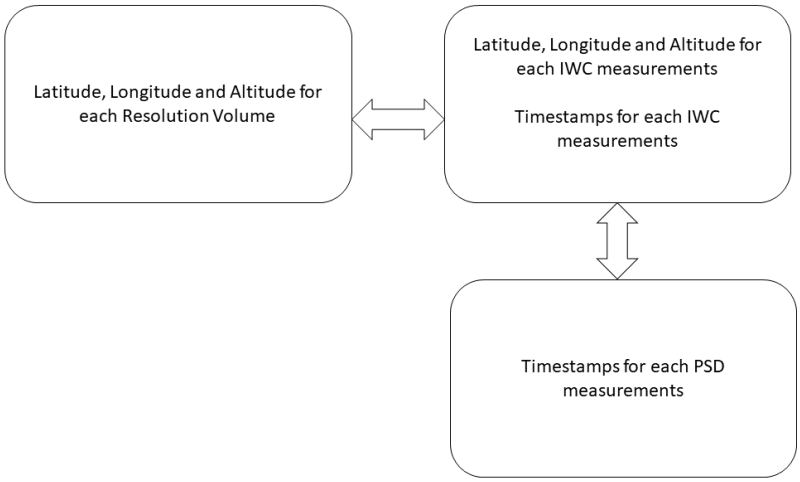


Figure 4.1: Data association procedure to correlate in-situ sensor measurement with radar measurements during 2015 flight test.

The overall development procedure of the simulation tool, as shown in Fig 4.2, emphasizes the relationships between the IWC values of clouds and the measured airborne radar observations. The study's approach establishes a simplified microphysical model of ice crystals while validating these models with the best possible knowledge extracted from measurement data. First, Monte-Carlo Method is used to address the uncertainty of the model parameters, and then a software package is developed associated with the previous system simulation tools. Eventually, the simulation tool may be helpful in the missions of the development of new airborne radar and aviation standards. (Shrestha et al. 2022).

4.3 Microphysical Properties of Ice Particles

This chapter aims to develop a physics-based radar sensing model for HIWC as an aviation hazard and then apply a ML-based processing framework to evaluate the HIWC hazard detection using airborne radars. Recent studies have examined the

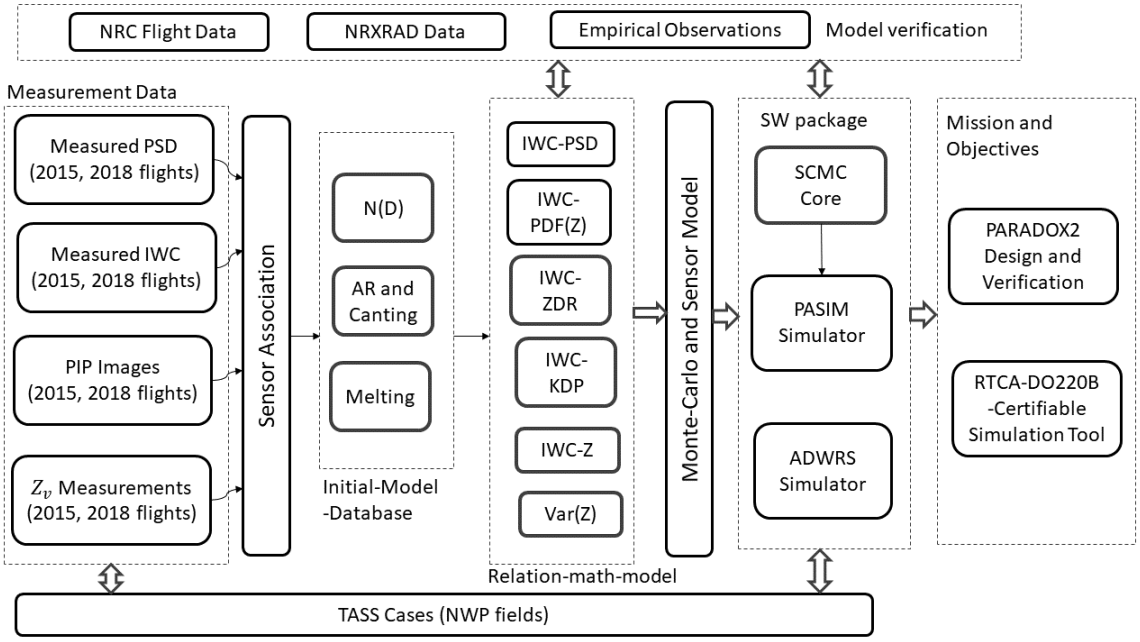
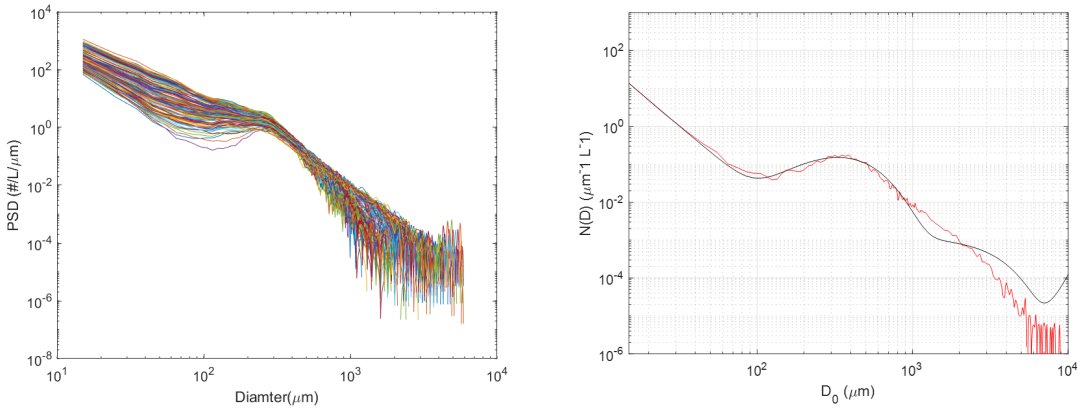


Figure 4.2: Organization of models and verification methods

microphysical properties of high-altitude ice. Convective storms lift a large amount of moisture to great heights where it freezes into tiny ice particles. Mostly, these ice particles are tiny, with reported mean mass diameters between 250 and 500 μm. Ice particles below 100 μm were reported (Leroy et al. 2017). Some common meteorological conditions are observed during high-altitude icing. These include high altitudes (11,000 feet to 40,000 feet), cold temperatures ($-10^{\circ}C$ to $-50^{\circ}C$), aircraft near convection clouds or thunderstorms, and the presence of turbulence (Grzych and Mason 2010).

Ice crystals' PSD is important for modeling and analyzing IWC and radar measurements. PSDs are number distribution functions given in $\#/L/\mu m$ that indicate the number of ice particles per unit volume in a liter, per unit bin size in μm. The PSDs are measured with 2D-Stereo for 10-1280 μm with a resolution of 10 μm. Similarly, for 100-6400 μm, Precipitation Imaging Probe (PIP) with the bin size of 100 μm was used. The example PSDs for the flight campaign is shown in Fig.4.3. Interestingly, it was found that all available measured PSDs are well-fitted with the



(a) (b)

Figure 4.3: Example of PSDs (a) collected from NASA 2015 HAIC-HIWC flight test campaign (for all the measured IWC values, during Aug 2015 flight).(b) Example of Cayenne, French Guiana flight collected HIWC PSD (red) and the fitted PSD (black) using the "double-Gamma" function and regenerated using the extracted function parameters.

"double-Gamma" function, which contains six parameters (b_1 to b_6), as shown in Equation (4.1), and illustrated in in Fig 4.3(b).

$$y = b_1 x^{b2} e^{-b3} + b_4 x^{b5} e^{-b6} \tag{4.1}$$

In addition to the size of the ice crystals, the mass-diameter relationship is of great importance. The mass-diameter relationship (m-D) is usually represented as a power-law $m = \alpha D^{\beta}$. Where m is a particle's mass, D is its diameter, α, and β are coefficients. The variability in mass with diameter also affects the effective density of ice particles (Heymsfield et al. 2004). Some of the relationships between the ice particle size and the ice crystals' effective density ρ are defined (Coutris et al. 2019).

4.4 Single-Cell Monte-Carlo (SCMC) Simulation Model

The first step in developing the aviation weather radar simulator is the single-cell Monte Carlo simulation (SCMC simulation). SCMC is essential in modeling the physical existence of hydrometeors in the atmosphere. SCMC also includes the calculation of scattering amplitudes, and radar variables can be calculated based on them. *Single Cell* in SCMC represents a unit-sized radar resolution volume in which the hydrometers are randomly distributed. The total number of hydrometeors and their shapes/sizes are defined by PSDs. Other microphysical parameters of these particles, such as the Axial Ratio (AR) of the hydrometeors and their Canting Angles (CA), are also modeled. Atmospheric parameters such as temperature, which can influence the scattering properties of hydrometeors, are also included. A novel aspect of the modified SCMC from the original codes (Li et al. 2010) is the inclusion of the ice crystal as a new type of hydrometeor and usage of the actual measured in-situ probe data in the modeling.

With the correlated sensor data, the SCMC uses random distribution assumptions to "fill-in" the other microphysical parameters that are unknown. For example, Static Air Temperature (SAT) measured by the DC-8 aircraft was used directly as the single radar resolution volume/cell temperature. AR and CA of ice particles are defined following uniform distributions in certain ranges. The dielectric constant of the ice crystal is estimated based on the measured temperatures. This dielectric constant is used together with the size of the ice crystals, the AR, and the CA to calculate the scattering amplitudes of the particles (Mishchenko et al. 1996). The AR and CA are assumed to follow uniform distribution, with recommended mean values in Table. 4.1. Default ranges of these values are based on studies in (Garrett et al. 2015), and existing PIP images. The actual canting angles would naturally depend on the wind

and electrification conditions. Therefore, the SCMC uses flexible configurations of these variables. SCMC calculates the scattering amplitudes from both the T-Matrix and the Rayleigh scattering. The scattering amplitudes are then used to calculate the radar observation variables such as reflectivity, Z_{dr}, K_{dp}, and ρ_{HV}.

4.4.1 Equations Used in the SCMC

For example, reflectivity, Z_{dr} and K_{dp} can be expressed as (Doviak et al. 2006)

$$Z_{h,v} = \frac{4\lambda^4}{\pi^4 \mid K_w \mid} \int_{Dmin}^{Dmax} \mid f_{hh,vv}(\pi, D) \mid^2 N(D)d(D) \tag{4.2}$$

$$Z_{dr} = 10log_{10}\frac{Z_h}{Z_v} \tag{4.3}$$

$$K_{dp} = 10^{-3}\frac{180}{\pi}\lambda Re\{\int_{D_{min}}^{D_{max}} \mid f_{hh}(0, D) - f_{vv}(0, D) \mid N(D)d(D)\} \tag{4.4}$$

Here, wavelength is the radar wavelength in millimeters (mm), and $N(D)$ represents the PSD in $mm^{-1}m^{-3}$. $f_{hh,vv}(\pi, D)$ and $f_{hh,vv}(0, D)$ are the backward and forward scattering amplitudes, respectively for co-polarized horizontal and vertical directions, respectively. K_w is the dielectric factor of water (Doviak et al. 2006).

The IWC in unit of $g\Big/m^3$ is related with PSD through

$$IWC = \frac{\pi}{6}10^{-3}\int_0^{D\max} \rho_e\,(D)D^3N\,(D)\,dD \tag{4.5}$$

In Equation (4.5), IWC is the key ice content parameter affecting aircraft safety. ρ_e is the effective density of ice crystals. When ρ_e has a unit of $g\Big/cm^3$ and equal-volume diameter of ice particle D has a unit of μm, a useful relationship based on flight measurement data (Dye et al. 2004) is:

$$\rho_e = 1.41 \times 10^{-10}D^{-1.1}, \;\; for \;\; D > D_t \tag{4.6}$$

Equation (4.6) is a modified version of the BF-95 model (Brown and Francis 1995). The density model in Equation (4.6) may need to be adjusted somehow to "fit" specific measurement, for example, ρ_e is often set as constant $0.917\,g\big/cm^3$ for particle diameter smaller than certain threshold, D_t. The actual value of D_t depends on specific temperature of the environment and environment conditions. The relationship is depicted in Fig 4.4. The dielectric properties of ice crystals, which affects the scattering properties is also dependent on the density of ice crystals (Ulaby et al. 2014).

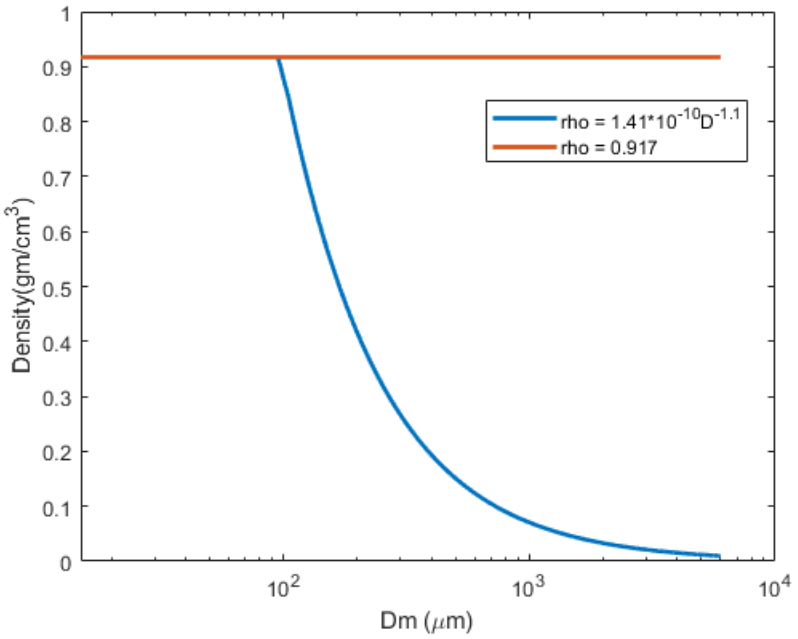


Figure 4.4: Effective Density vs equivolume diameter of ice crystals

Another useful variable, the total concentration of ice particles N_t, is the summation of numbers of concentration for all sizes, and is given by:

$$N_t = \int_{D_{min}}^{D_{max}} N(D)d(D) \tag{4.7}$$

Figure 4.6 shows the number of particles (N_t) recorded during the NASA flight campaign for different temperature ranges. Similar concentrations were observed in other

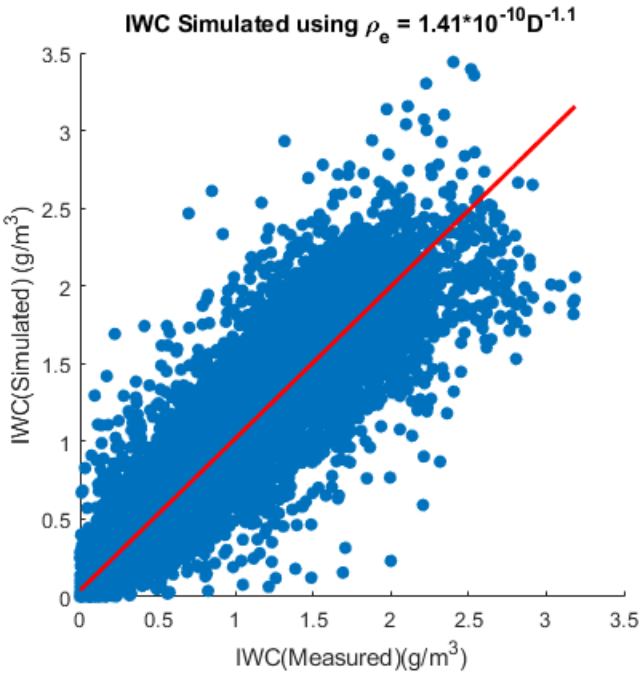
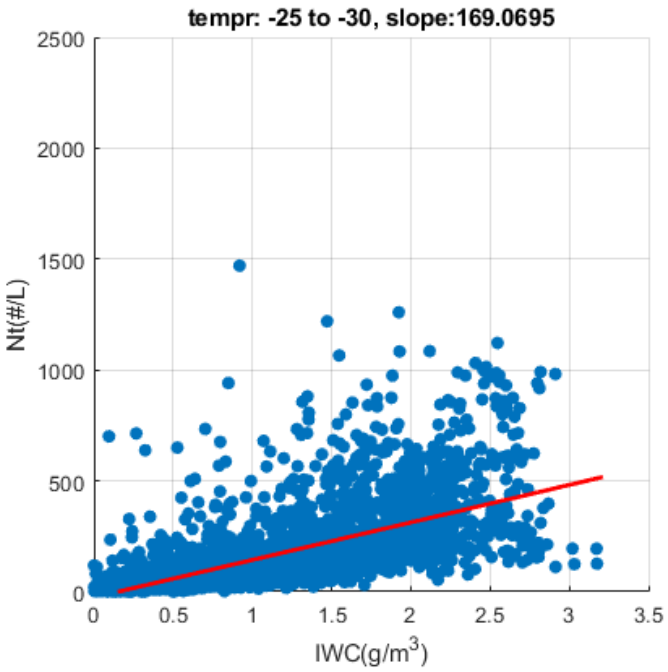


Figure 4.5: Measured IWC vs estimated IWC based on measured PSD and the modified effective density Model.

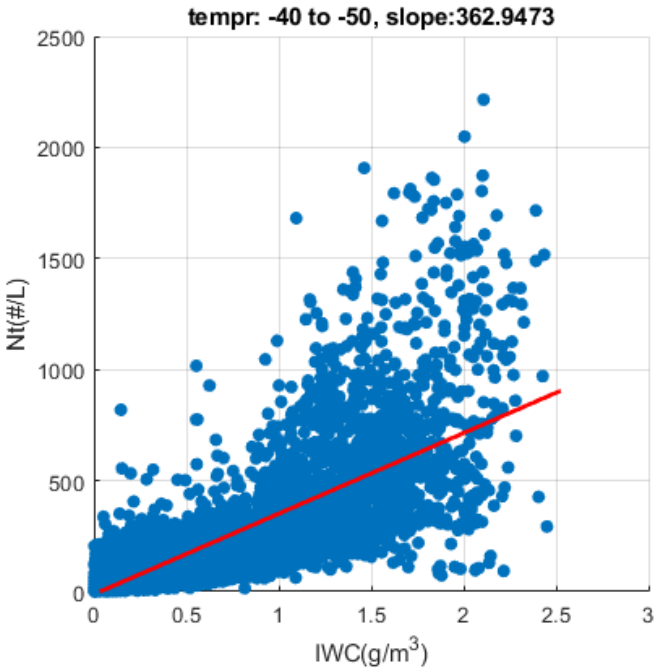
flight campaigns (Hu et al. 2021). This similarity helps to validate the data recorded to be used in the radar development process.

4.4.2 Mean and Variance of Reflectivity Factors

The pulse-to-pulse variation of radar return signal power, which is related to the variance of reflectivity factor, has been used in the "Swerling" algorithm developed by NASA team for HIWC detection (Harrah et al. 2019). This section discusses both the mean and variance of reflectivity factors as well as the RIWC parameter used in the "Swerling" algorithm, which have not been addressed before. Assuming M represents the voltage RCS resulting from N number of sources with individual RCS of σ_i and individual phase of ψ_i. Considering random phases with uniform distribution from 0 to 2π between N particles, we have



(a)



(b)

Figure 4.6: Comparisons of number of particles associated with IWC values for different temperature ranges (a) -25 to -30 $^{\circ}C$, (b) -40 to -50 $^{\circ}C$.

$$M = \sum_{i=1}^{i=N} \sqrt{\sigma_i} \exp^{j\psi_i} \tag{4.8}$$

$$M^* = \sum_{i=1}^{i=N} \sqrt{\sigma_i} \exp^{-j\psi_i} \tag{4.9}$$

The mean of the RCS can be written as (Jenn 2005):

$$Mean(\sigma) = E|MM^*| = \sum_{i=1}^{i=N} \sigma_i \tag{4.10}$$

Similarly, the variance of RCS can be written as

$$E[|MM^*|^2] = E[(MM^*)(MM^*)^*] = E[MM^*M^*M] = E[M^2(M^*)^2] \tag{4.11}$$

Since we have

$$(a_1 + a_2 + \ldots + a_N)^2 = \sum_{i=1}^{i=N} (a_i)^2 + 2 \times \sum_{\substack{1 \le j \le N \\ 1 \le k \le N \\ j \ne k}} a_j a_k \tag{4.12}$$

$$\begin{aligned} E[|MM^*|^2] = \\ E\left[\left(\sum_{i=1}^{i=N} \sigma_i \exp^{j2\psi_i} + 2 \times \sum_{\substack{1 \le j \le N \\ 1 \le k \le N \\ j \ne k}} \sqrt{\sigma_j}\sqrt{\sigma_k} \exp^{j(\psi_j + \psi_k)} \right) \right. \\ \left. * \left(\sum_{i=1}^{i=N} \sigma_i \exp^{-j2\psi_i} + 2 \times \sum_{\substack{1 \le j \le N \\ 1 \le k \le N \\ j \ne k}} \sqrt{\sigma_j}\sqrt{\sigma_k} \exp^{-j(\psi_j + \psi_k)} \right) \right] \end{aligned} \tag{4.13}$$

$$E[|MM^*|^2] = (\frac{1}{2\pi})^N \times$$
$$\int_0^{2\pi} \cdots \int_0^{2\pi} \Bigg[(\sum_{i=1}^{i=N} \sigma_i^{j2\psi_i} + 2 \times \sum_{\substack{1\le j\le N \\ 1\le k\le N \\ j\neq k}} \sqrt{\sigma_j}\sqrt{\sigma_k}^{j(\psi_j+\psi_k)}) \times$$
$$(\sum_{i=1}^{i=N} \sigma_i^{-j2\psi_i} + 2 \times \sum_{\substack{1\le j\le N \\ 1\le k\le N \\ j\neq k}} \sqrt{\sigma_j}\sqrt{\sigma_k}^{-j(\psi_j+\psi_k)}) \Bigg] d\psi 1 \ldots d\psi N \tag{4.14}$$

After the expansion of the itegrand of the above equation, any terms left with one or more exponential terms becomes zero as a result of the integration from 0 to 2π [$\int_0^{2\pi} e^{jk\theta} d\theta = 0$]. The only terms that do not have exponential terms are obtained as a result of the product between two conjugate terms. As a result,

$$E[|MM^*|^2] = \left(\frac{1}{2\pi}\right)^N \times (2\pi)^N \times \left(\sum_{i=1}^{i=N} \sigma_i^2 + 4 \times \sum_{\substack{1\le j\le N \\ 1\le k\le N \\ j\neq k}} \sigma_j \sigma k\right) \tag{4.15}$$

$$E[|MM^*|^2] = \left(\sum_{i=1}^{i=N} \sigma_i\right)^2 + 2 \times \sum_{\substack{1\le j\le N \\ 1\le k\le N \\ j\neq k}} \sigma_j \sigma k \tag{4.16}$$

Thus, the variance of RCS becomes

$$VAR(\sigma) = E[|MM^*|^2] - (E[|MM^*|])^2 \tag{4.17}$$

$$VAR(\sigma) = 2 \times \sum_{\substack{1\le j\le N \\ 1\le k\le N \\ j\neq k}} \sigma_j \sigma k \tag{4.18}$$

For spheroid particles, σ_i can be calculated using the Rayleigh-Gans scattering formula or T-Matrix theory (Mishchenko et al. 1996). The backscattering amplitude

f (Bringi and Chandrasekar 2001; Doviak et al. 2006) of each particle i is calculated and then related to RCS as:

$$\sigma_i = 4\pi \left| f_i \right|^2 \tag{4.19}$$

The reflectivity factor then can be defined as

$$Z = \frac{\bar{\sigma}\lambda^4}{V\pi^5 |K_w|^2} \tag{4.20}$$

Based on Equation (4.20), and (4.17) the variance of reflectivity factor is given by

$$VAR\left(Z\right) = \frac{VAR\left(\sigma\right)\lambda^8}{V^2\pi^{10}|K_w|^4} = \xi \cdot VAR\left(\sigma\right) \tag{4.21}$$

in which, $\xi = \lambda^8 \Big/ V^2\pi^{10}|K_w|^4$ is a coefficient related to wavelength.

Since the variance of RCS consists of mutual correlation terms $\sum_{i \neq j} \sigma_i \sigma_j$, and then based on Equations (4.19) and (4.20), the variance of reflectivity may be also expressed in terms of PSD and scattering amplitudes as:

$$VAR\left(Z_{h,v}\right) = \xi \left\{ \left[\int \left| f_{hh,vv}\left(\pi, D\right) \right|^2 N\left(D\right) dD \right]^2 - \left[\int \left| f_{hh,vv}\left(\pi, D\right) \right|^4 N\left(D\right) dD \right] \right\} \tag{4.22}$$

And the Radar-Derived IWC (RIWC) factor used in (Harrah et al. 2019) can be then expressed as:

$$RIWC \propto \log_{10}\left(\frac{VAR\left(Z\right)}{Z}\right) = \log_{10}\left[Z_{h,v} - \frac{\int \left| f_{hh,vv}\left(\pi, D\right) \right|^4 N\left(D\right) dD}{\int \left| f_{hh,vv}\left(\pi, D\right) \right|^2 N\left(D\right) dD} \right] \tag{4.23}$$

In Equations (4.22) and (4.23), the units of Z and $\mathrm{V}AR\left(Z\right)$ are $mm^{-6}m^3$ and $mm^{-12}m^6$, respectively.

4.4.3 SCMC Outputs and PDF of Radar Variables

Using probe-measured PSDs and the radar scattering models, SCMC outputs can be visualized in terms of histograms of radar variables regarding specific IWC values. The histograms are obtained by collecting all the reflectivity samples from SCMC that are within $\pm 0.1 \text{g} \Big/ m^3$ of specific IWC values. Fig. 4.7 compares the obtained histograms of two IWC values for the three flight campaigns. Interestingly, these histograms show multiple peaks for low IWC values, including a peak centered around 10 dBz. The histograms are more like Gaussian shapes for high IWC values. The extracted PDFs from the histograms using MATLAB are illustrated in Fig. 4.8. From these PDFs, we can observe the interesting trend of how the IWC affects the spreading of PDFs. On the other hand, this trend may show different behaviors for different flight campaigns in different regions. For example, the "spreading" of the PDFs may decrease with IWC for the French Guiana campaign results but increase with IWC for the Gulf of Mexico campaign, as shown in Fig. 4.8.

4.5 Verification of SCMC Model Outputs

Firstly, the SCMC simulation outputs are compared with the available flight test data and other open publications. The example outputs from Monte-Carlo simulations using the PSDs collected from the 2015 DC-8 flight campaign are shown in Figure 4.9 and 4.10, for the X band and S-band, respectively. Meanwhile, the similar simulation outputs from SCMC for the other two flight campaigns are shown in Figure 4.11. Comparing the simulated radar variables from the three different campaigns, it can be noticed that the Convair and Falcon flight measurement PSDs result in a higher range of reflectivity values than DC-8 data. The Z_{dr} values and distributions are different among the flight campaigns due to the particles' different AR and canting angles. The RIWC plots have similar trends among different flights, and they are in

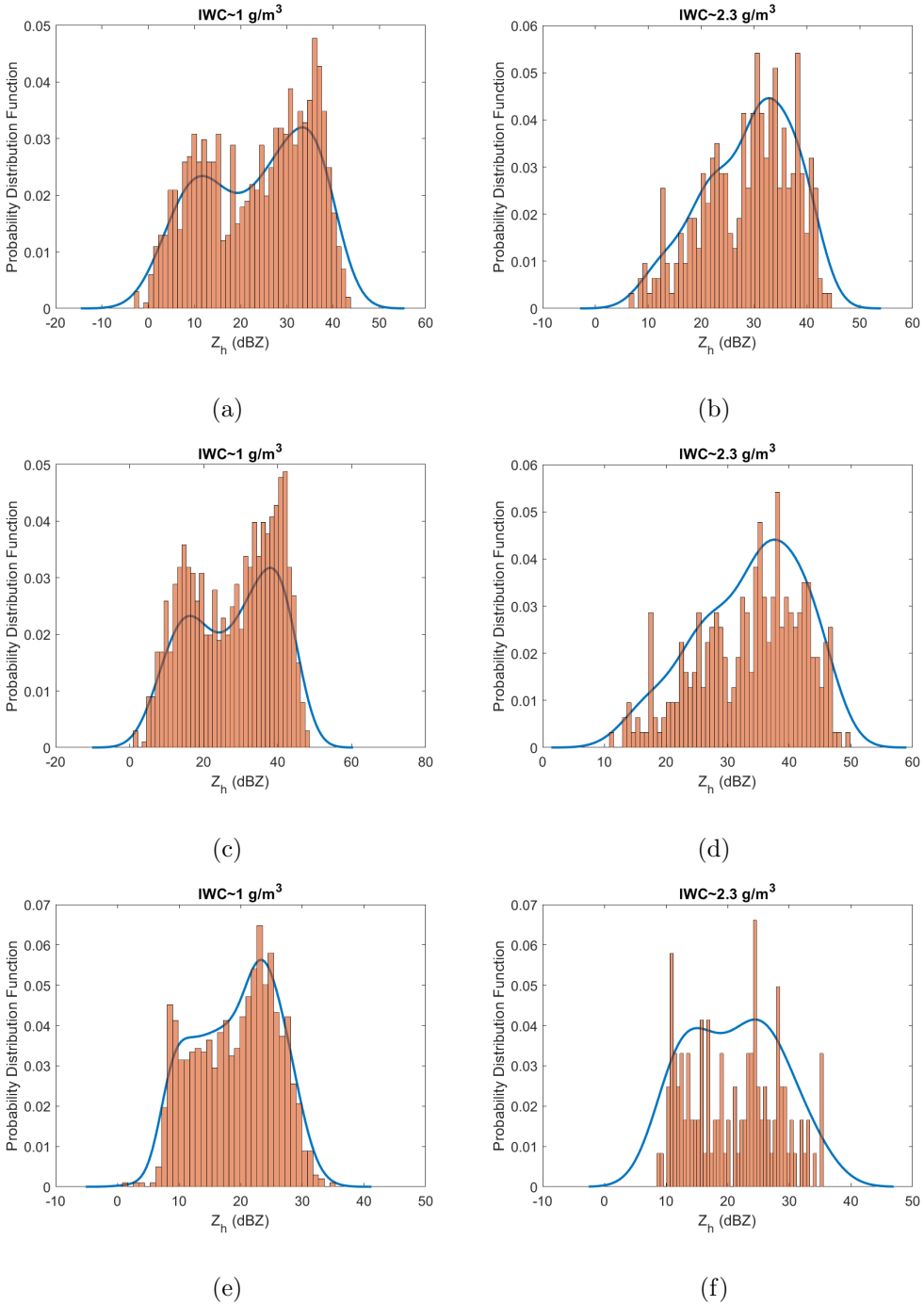
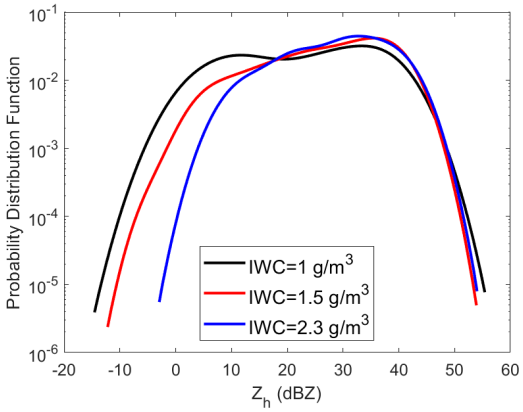
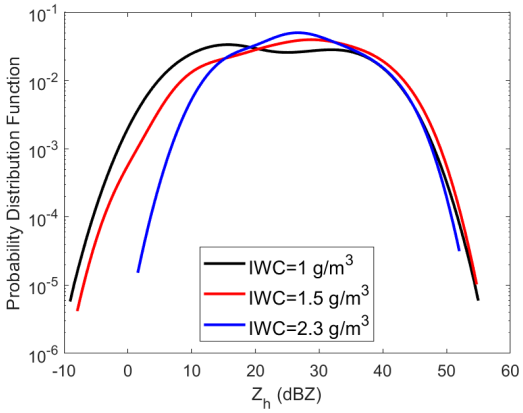


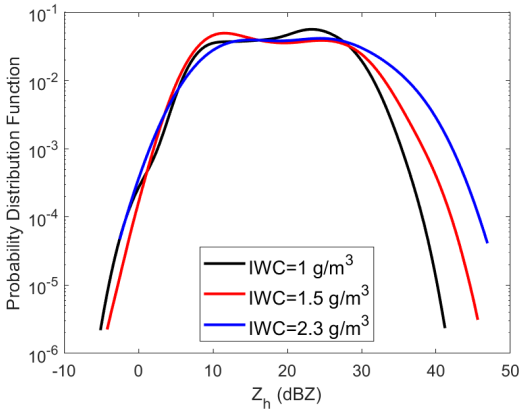
Figure 4.7: Normalized histograms and trend lines of horizontally polarized reflectivity factor values as outputs of SCMC. Vertically polarized reflectivity have similar distributions. The Left column is for IWC $\approx 1g/m^3$, right column is for IWC $\approx 2.3g/m^3$. (a) and (b) are for the Convair flight, (c) and (d) are the Falcon flight, and (e) and (f) are for the NASA DC-8 flight.



(a)



(b)



(c)

Figure 4.8: Comparisons of extracted probability density functions of reflectivity factor values for different IWC values.The PSDs associated with temperature range $\leq$ -25°C are used. (a) Convair data, (b) Falcon data, (c) NASA DC-8 data.

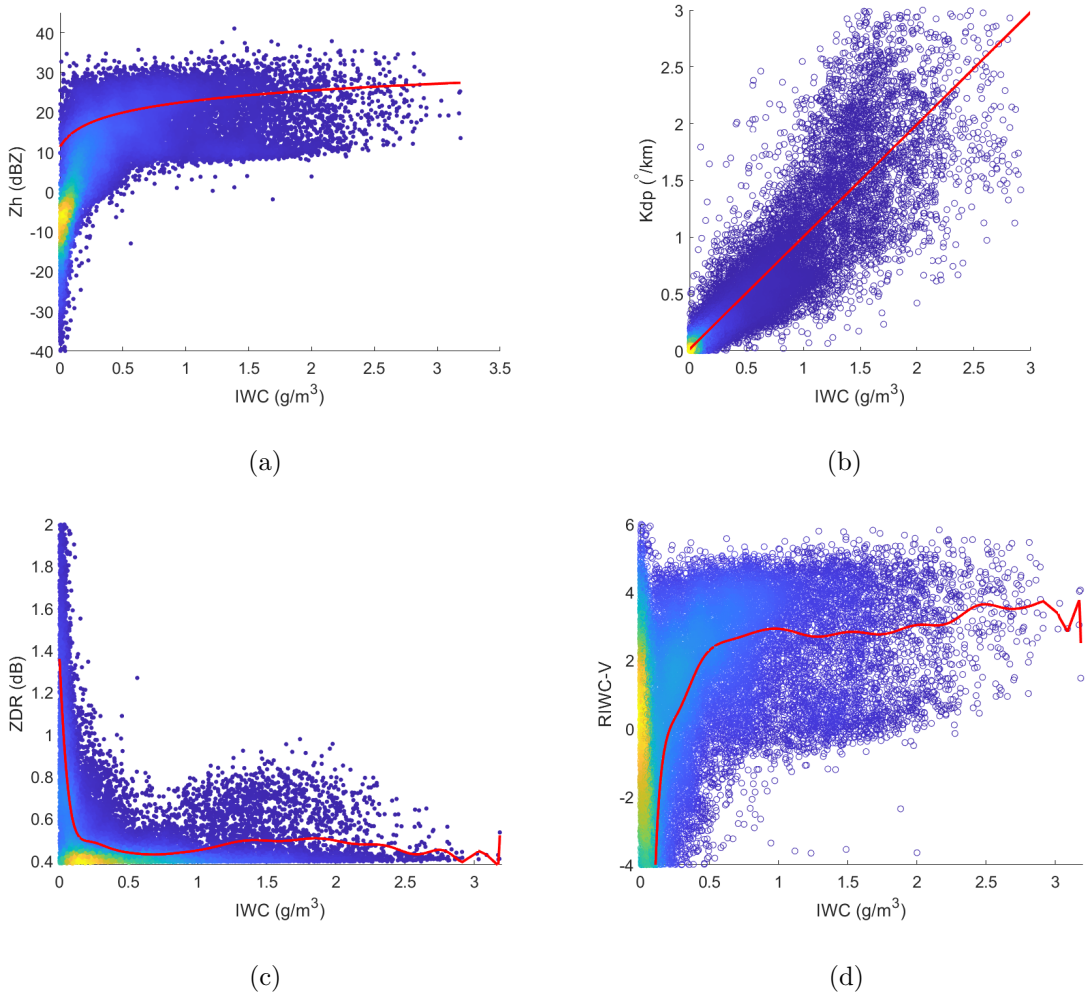


Figure 4.9: SCMC output in form of scatter-plots. Radar frequency is 9.41 GHz. Dots with density colors represent outputs from 23400 Monte-Carlo runs using 2015 DC-8 flight campaign data, and red lines represent regression trends. (a) IWC vs mean horizontal reflectivity factor Z_h, (b) IWC vs K_{dp}, (c) IWC vs Z_{dr}, (d) IWC vs RIWC for vertical polarization (Equation (4.23)).

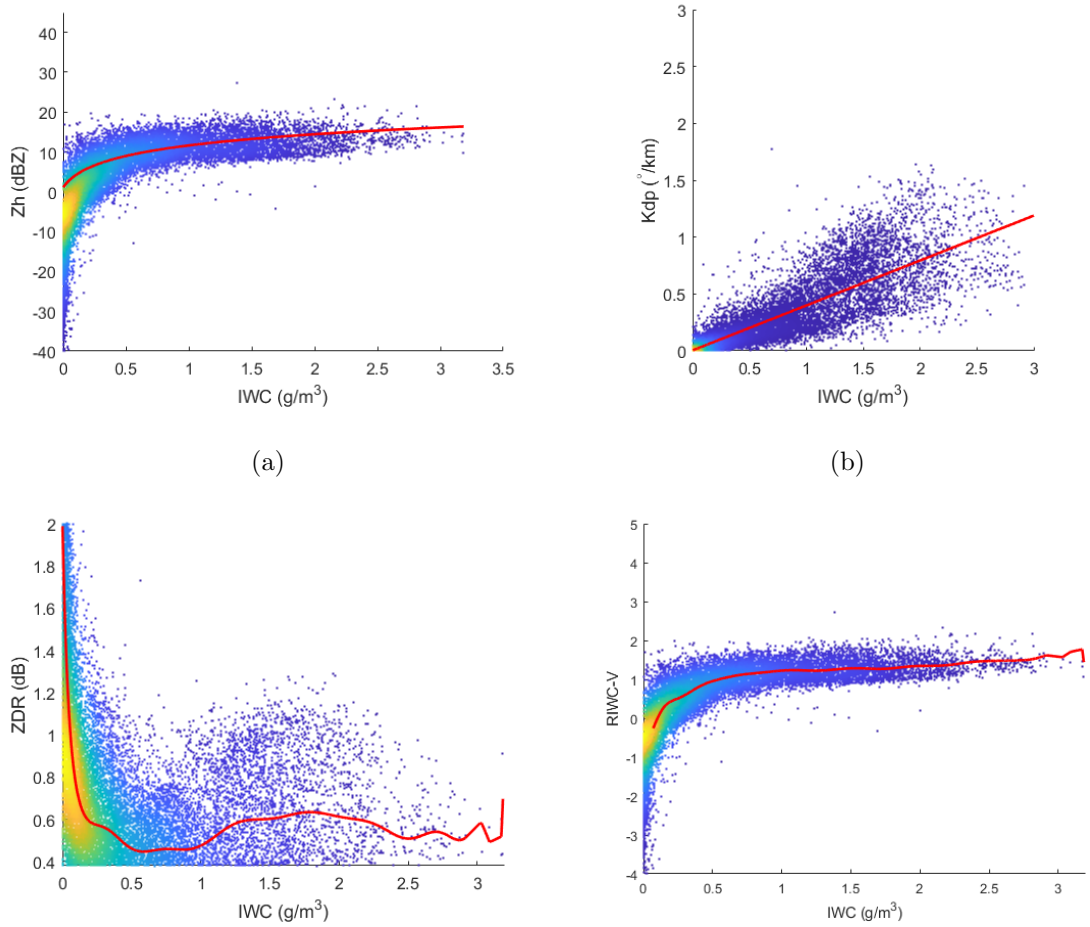


Figure 4.10: SCMC output in form of scatter-plots. Radar frequency is 3 GHz. Dots with density colors represent outputs from 23400 Monte-Carlo runs using 2015 DC-8 flight campaign data, and red lines represent regression trends. (a) IWC vs mean horizontal reflectivity factor Z_h, (b) IWC vs K_{dp}, (c) IWC vs Z_{dr}, (d) IWC vs RIWC for vertical polarization (Equation (4.23)).

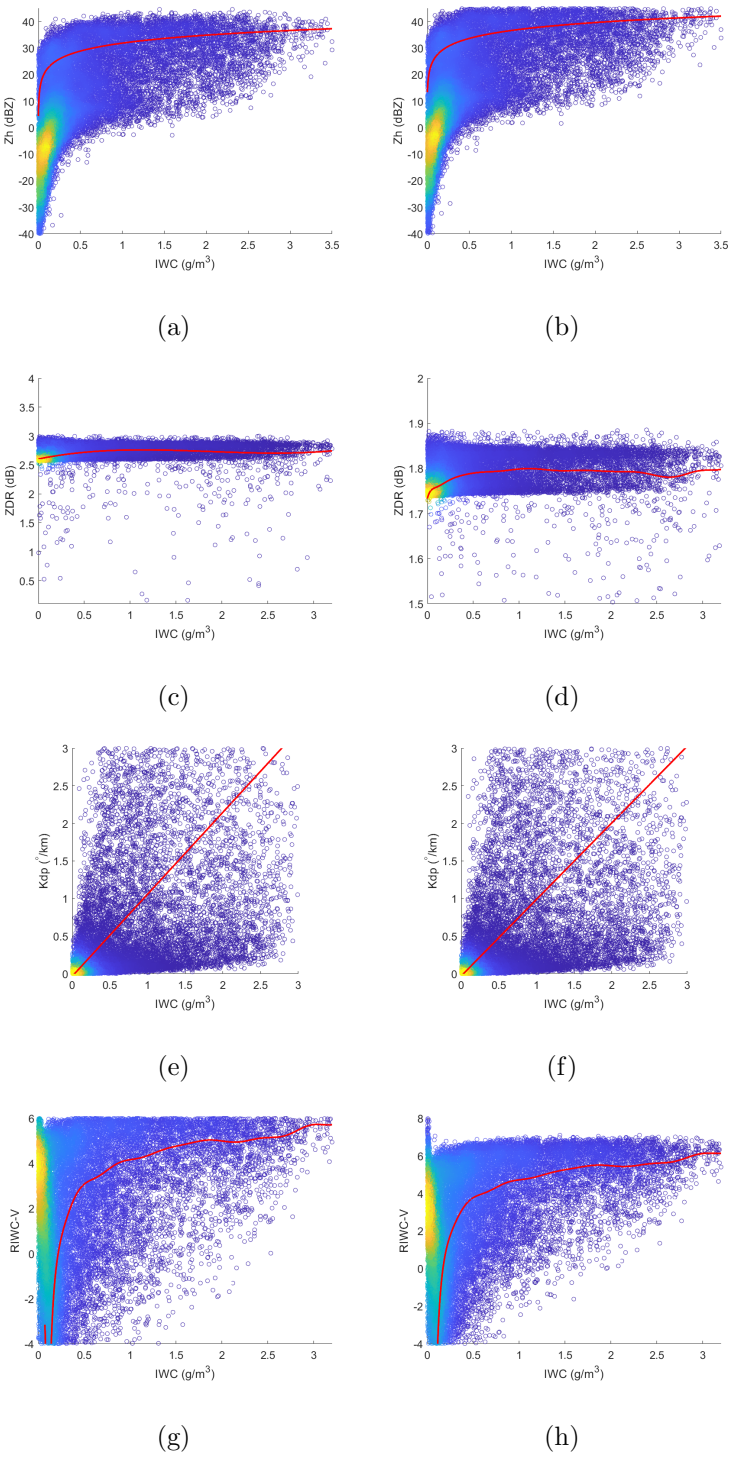


Figure 4.11: SCMC output, using the PSD data collected from Convair (Left column) and Falcon (Right column) campaigns. (a) and (b): IWC vs Z_h, (c) and (d): IWC vs Z_{dr}, (e) and (f): IWC vs K_{dp}, (g) and (h) IWC vs RIWC for vertical polarization

the correct range of the IWC values, similar to the results in (Harrah et al. 2019). However, the RIWCs do not show a linear relationship with the IWC values. The most interesting radar variable is the K_{dp}, which offers a consistent statistically linear relationship between IWC and K_{dp}. A similar relationship is introduced in (Ryzhkov et al. 1998).

As part of the validation of SCMC outputs with measured radar data, Figure 4.12 (a) shows a comparison between the SCMC simulated IWC-Z curve and the measurement curve derived from the 2015 flight campaign measurement. The figure compares the reflectivity values generated by the SCMC with the measured average reflectivity against specific IWC values. A mean value of the reflectivity values assigned to an IWC value ($\pm 0.1\ g\Big/m^3$) was used for comparison. For example, for the IWC value of 1 $g\Big/m^3$ in Figure 4.12 (a), all reflectivity values that correspond to IWC values between 0.9 and 1.1 $g\Big/m^3$ are considered for the average calculation. The reflectivity values start at around 22 dBz for IWC of 1 $g\Big/m^3$ and increase to around 24 dBz for the IWC of 2.5 $g\Big/m^3$, for both simulations and measurements. From these comparisons, we can see that the changes in reflectivity values vs. IWC are rather insignificant.

Similarly, RIWC from both the NASA flight campaign and the outputs from the SCMC simulation versus different IWC values are compared (Figure 4.12 (b)). The RIWC for the NASA data can be computed based on the mean magnitude (μ) and the variance of the magnitude (σ) of the I and Q signals of the return pulses (Harrah et al. 2019):

$$\mu = \frac{1}{n}\sum_{i=1}^{n}(I_i^2 + Q_i^2) \tag{4.24}$$

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(I_i^2 + Q_i^2) - \mu^2 \tag{4.25}$$

Finally, RIWC is calculated by

$$RIWC = log_{10}(\frac{\sigma^2}{\mu}) \tag{4.26}$$

The trend is similar to the trend of the reflectivity. The NASA-RIWC estimates, however may introduce an error of about $\pm 0.5g\Big/m^3$ (see discussion in (Harrah et al. 2019)). A similar margin of errors is observed with the RIWC estimations simulated from SCMC.

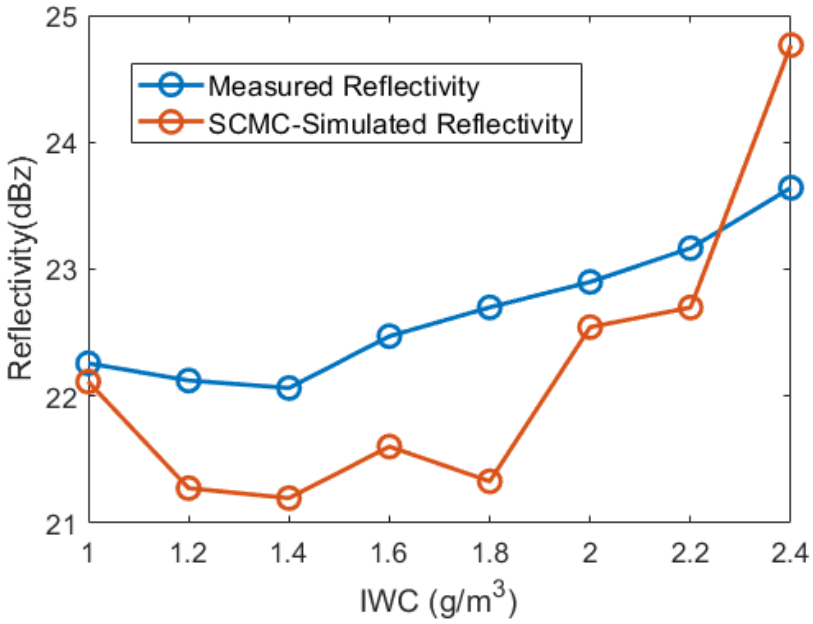
4.6 Airborne Weather Radar System Simulation

4.6.1 Background

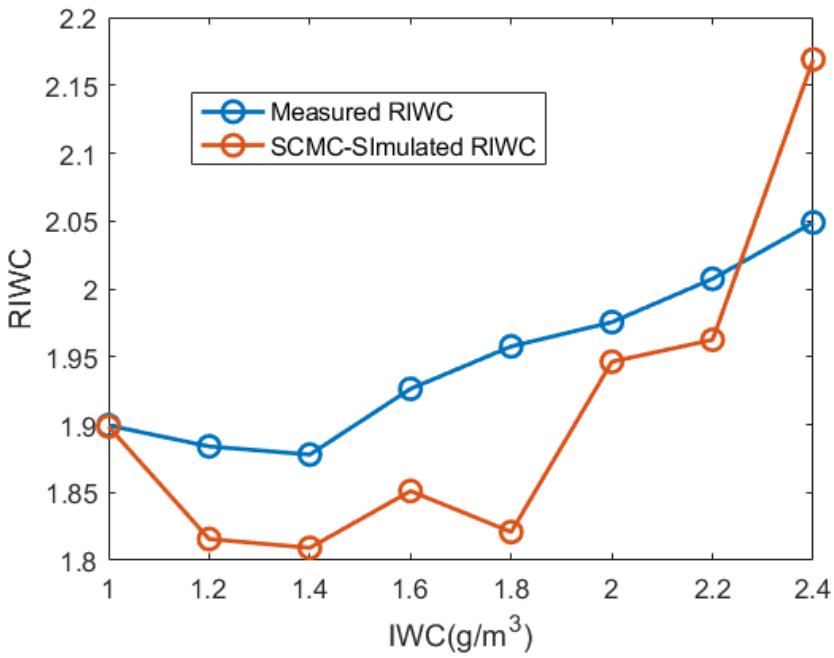
The output of the SCMC simulation can be used in an end-to-end system simulator to generate the in-phase and quadrature (I/Q) signals. A time-domain, 3D scanning region airborne radar simulator is developed based on MATLAB, and the simulation flow is depicted in Figure 4.13.

PASIM (Li et al. 2019, 2020) is a weather radar simulation tool operating in time domain. It was originally developed for ground-based weather radar system simulations, while it is extended to airborne weather radars. The expansion of the simulator, which is called "airborne PASIM", uses the same methods to generate I/Q signals from radar moment data while supporting airborne motions, geometries, and coordinate system translations. The simulator takes "truth" weather radar variables as inputs. It generates the radar received IQ signals, which take account of antenna patterns, waveforms, and RF transceiver characteristics. The radar system developers then use the simulated I/Q signals to validate the HIWC detection and IWC estimation algorithms.

The 3D "truth" weather field is derived from the simulated data from NASA's Terminal Area Simulation System (TASS) (Ahmad and Proctor 2011; Proctor 1987). The TASS data field includes three-dimensional fields of IWC as part of its outputs.



(a)



(b)

Figure 4.12: Comparison of reflectivity (dBz) and RIWC from 2015 DC-8 flight measurement and SCMC outputs, for different IWC values. (a) Averaged reflectivity values, (b) Averaged RIWC values.

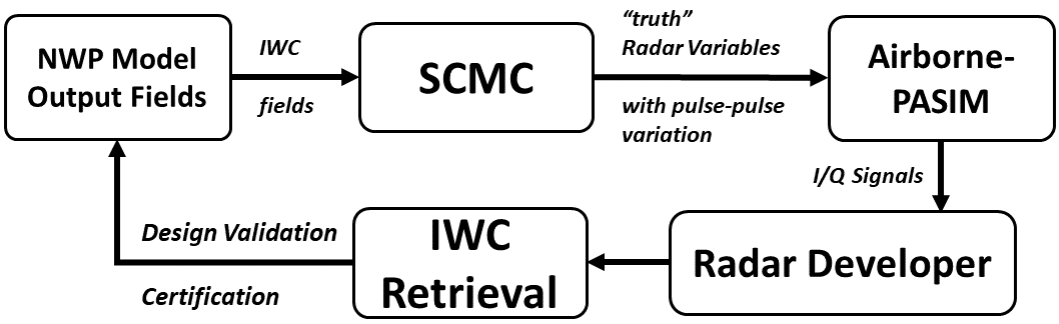


Figure 4.13: Structure and configuration of SCMC, and its relation with the airborne radar simulation tool.

It also includes secondary output products such as radar reflectivity factor, graupel, wind velocity, and other variables that are arranged into the Cartesian coordinate grid. The first step is to map these data to a radar-based polar grid system, which can be achieved using neighbor (NN) interpolation or finding the intersecting volume between radar resolution and the TASS Cartesian grid volume. Then an average is taken of all the IWC values of these enclosed grid volumes. For each IWC value in the polar radar grid, the SCMC run produces the polarimetric weather radar variables. These variables are used as inputs to the modified airborne radar simulator.

First, based on the reflectivity factor values, the scattering amplitudes $|S_{hh}|$ and $|S_{vv}|$ are computed for each resolution volumes using the following equation (Doviak et al. 2006):

$$\left|S_{hh,vv}\right| = Z_{hh,vv}\frac{0.93\pi^6 R^2 c\tau\theta^2}{16\log(2)\,\lambda^4} \tag{4.27}$$

,

In Equation 4.27, Z_{hh}, Z_{vv} are the reflectivity factors (linear scale), θ is the antenna's 3 dB beam width, τ is signal pulse width. Next, these scattering amplitudes are applied in MATLAB's phased array systems toolbox and are used to create I/Q data for a single pulse over all the resolution volumes within the field of view. To generate the I/Q data for multiple pulses, the auto-correlation function is utilized.

Randomness is induced in the I/Q data based on sampling the SCMC output. More randomness is introduced to explore the pulse to pulse power variation for HIWC detection. Then the mean velocity and spectrum width are included in the following auto-correlation equation for horizontal and vertical polarizations (Bringi and Chandrasekar 2001; Doviak et al. 2006).

$$R_{hh,vv}(mT_s) = |S_{hh,vv}|e^{-8(\frac{\pi\sigma_v mT_s}{\lambda})^2}e^{-j\frac{4\pi v_m T_s}{\lambda}} + N\delta_m \tag{4.28}$$

In Equation 4.28, $S_{hh,vv}$, v_m and σ_v are the average signal power, mean velocity, the spectrum width, respectively. T_s is the pulse repetition time, n is the lag index of the auto-correlation function. N is noise power. The dual-pol radar variables, including Z_{dr}, ρ_{HV}, and Φ_{dp} are implemented using the auto-correlation function and Pulse-Pair Processing (PPP) (Doviak et al. 2006):

$$R_{hh}(mT_s) = R_{hh}(mT_s) \tag{4.29}$$

And,

$$R_{vv}(mT_s) = [\rho_{HV}R_{hh}(mT_s) + \sqrt{1-\rho_{HV}^2}R_{vv}(mT_s)]\frac{e^{j\Phi_{dp}}}{\sqrt{Z_{dr}}} \tag{4.30}$$

An example output of the SCMC-airborne radar system simulation is shown in Figure 4.14. Figure 4.14 (a) is the IWC field generated from the TASS model (which is treated as the de facto "truth IWC field"). The IWC filed data in the Cartesian grid have been converted to a polar grid to be consistent with the radar field of view. Figure 4.14 (b) shows the reflectivity output of SCMC. Simulated reflectivity scans from the airborne radar are shown in Figure 4.14 (c) and (d), respectively, using different antenna beamwidth values of a similar aperture (dish type, similar to existing radar products). In Figure 4.14 (c), the 3 dB beamwidth of the antenna was assumed to be 1° (which can be achieved using angular super-resolution processing (Wang et al. 2012)). The effective antenna beamwidth is 3.5° in Figure 4.14 (d). The

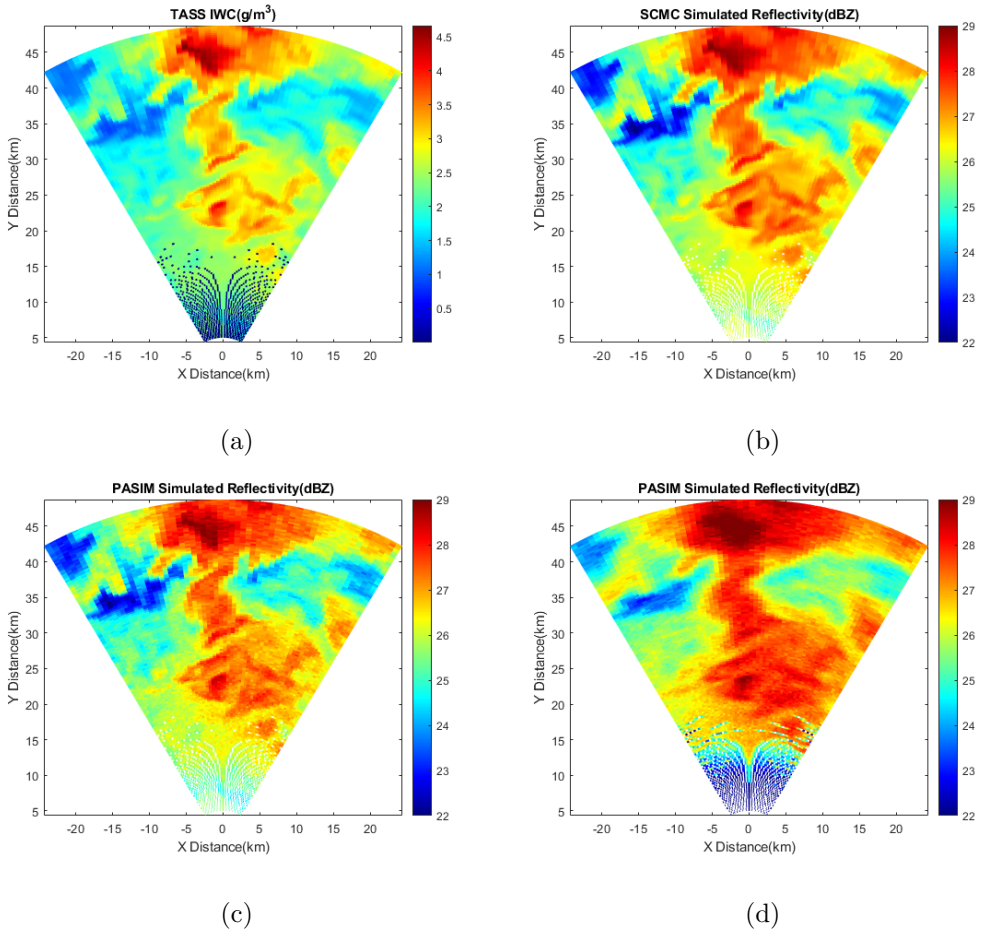


Figure 4.14: End-to-end airborne radar system simulation (PPI scans) (a) IWC field from the NWP model, (b) Reflectivity field from the SCMC, (c) Reflectivity field from Airborne Radar System Simulator (1° beamwidth), (d) Reflectivity field from Airborne Radar System Simulator (3.5° beamwidth).

effect of the antenna on spatial resolution can be seen clearly, which will, in turn, affect HIWC detection performance.

Next, simulated PPI scans of all the dual-polarized radar variables were generated, as shown in Figure 4.15. The K_{dp} plots show that the relationship between IWC and K_{dp} is almost linear and can be expressed as $K_{dP} \approx IWC$, for the X-band radar system. This relationship was also verified from the SCMC outputs. The values of Z_{dr} are plausible based on the axial ratios defined in the SCMC simulation (i.e., 0.3 to 0.7). However, the Z_{dr} values are distributed in a small range between 0.44 dB

and 0.45 dB; thus, it is not convenient to infer HIWC conditions from Z_{dr} values alone.

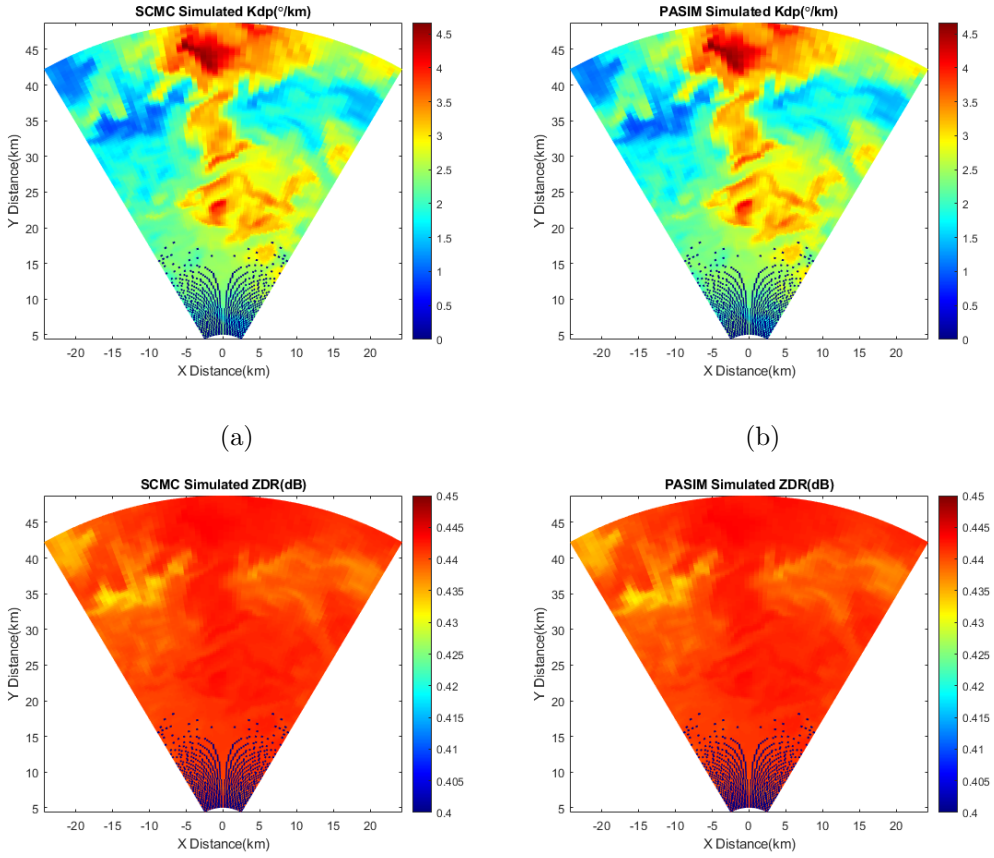


Figure 4.15: End-to-end simulations of PPI scans of polarimetric radar outputs, (a) A snapshot of K_{dp} field output from SCMC, (b) K_{dp} field estimation from Airborne Radar System Simulator, (c) A snapshot of Z_{dr} output field from SCMC, (d) Z_{dr} field estimation from Airborne Radar System Simulator (Assuming 1° beamwidth for system simulations).

4.7 Detection of HIWC using Simulated Airborne Radar Data

The previous sections established an end-to-end framework for evaluating the algorithms for detecting and classifying the HIWC conditions through airborne radar sensor measurements. In this section, an illustrative example of algorithm evaluation based on simple ML (i.e., ANN) is provided. The combination of flight test data and simulation data provides sufficiently large amounts of training and testing data for the evaluation. The biggest challenge is the selection and the application of the features and usage of these features for quantitative predictions. Even though the measured radar variables show different relations with IWC levels, there is a need to use the measurable and relevant radar variables in the ML algorithm as features. These features include reflectivity, Z_{dr}, K_{dp}, and RIWC for both polarizations. By doing so, it is convenient to better focus on the optimal detection and the achievable estimation results. One of the issues is the usage of temperature as a feature variable, which is further investigated through the following experiment.

ML algorithms are applied for classifying different levels of IWC (e.g., high and low IWC regions) and/or estimation of IWC values.A regression supervised learning method based on ANN is once again used for IWC predictions based on radar inputs. Similarly, the classification of the HIWC (based on threshold value of either 1 or 1.5 $g\Big/m^3$) can be implemented. Using the SCMC simulated radar reflectivity, temperature, and the IKP probe estimated IWC as input feature variables, a simple Neural Network with 2 hidden layers (16 nodes each) is trained for the regression and the classification. The dataset is divided randomly to generate the training (70%), testing (15%) and the cross validation (15%) datasets. The regression results and classification's confusion matrix are shown in the Figure 4.16 and Figure 4.17, respectively. The correlation coefficients of around 0.94 or more were achieved for the

training, cross validation and the test data sets. Similarly, a classification accuracy of better than 86% is achieved (In the confusion matrix, 1 represents the HIWC levels (IWC $> 1\, g/m^3$), and 0 represents IWC values of $1\, g/m^3$ or less). The input features used in this ML experiment include SCMC outputs: Reflectivity, Z_{dr}, K_{dp}, and RIWC in both Horizontal and Vertical Polarization. In addition, the temperature measured by the aircraft's temperature sensor is used.

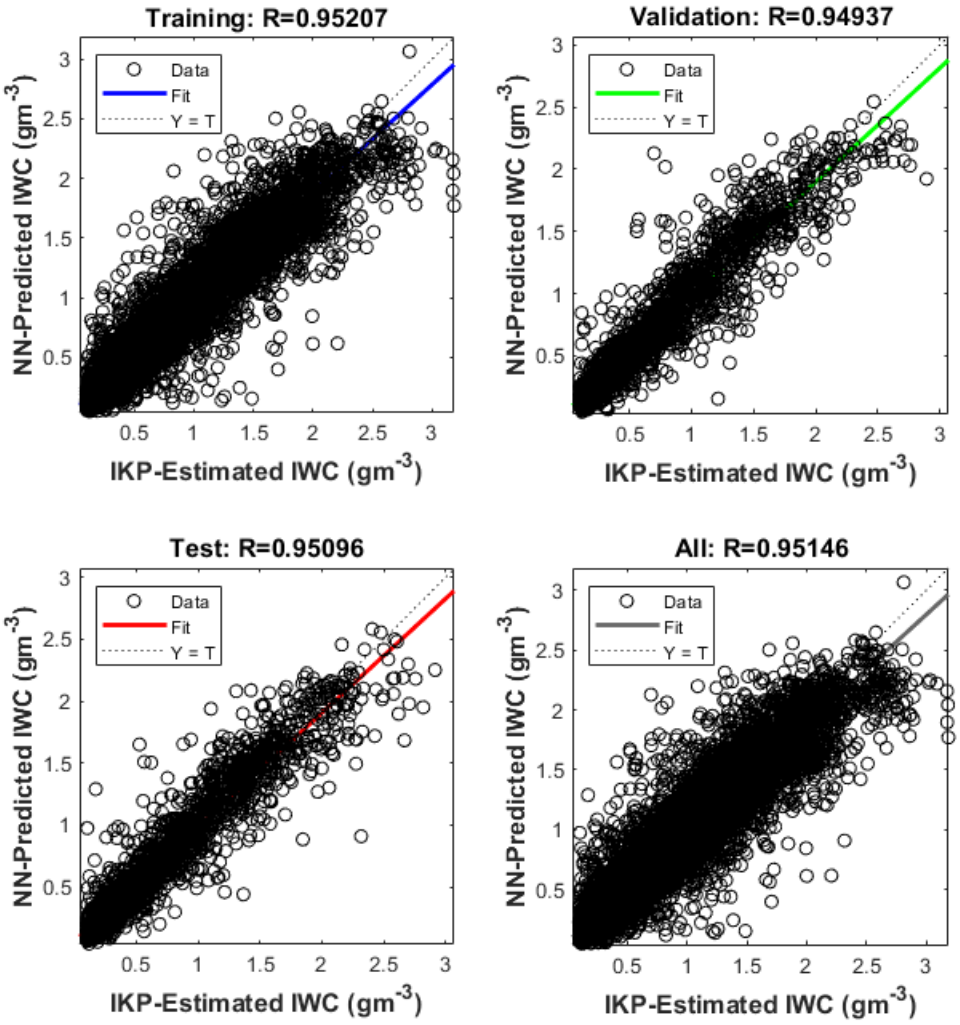


Figure 4.16: Regression plots for HIWC detection using temperature as a feature

Figure 4.17: Confusion Matrices for HIWC detection using temperature as a feature

Although the temperature has been shown to correlate with the IWC in previous studies, "radar-only" detection and estimation of HIWC are crucial. As the temperature measuring instruments in a flight provide in-situ measurements and radar cannot estimate the temperature, it is not convenient nor accurate to associate the aircraft measured temperatures like SAT or TAT with the radar PPI scans. Figure 4.18 and Figure 4.19 shows the regression analysis and classification performance of the ML algorithm without the use of temperature. Only the radar variables from the SCMC outputs are used as the inputs to the supervised learning approach. The same

neural network architecture was used with the same division ratio of the training, testing, and cross-validation datasets. The correlation coefficient of 0.92 or more and the classification accuracy of 85% or more were achieved. The results verified that temperature is not a critical feature variable for HIWC detection, even though it is important as a pre-condition for HIWC in general. Similar tests can be performed for different combinations of feature variables to attain further optimizations of ML algorithms.

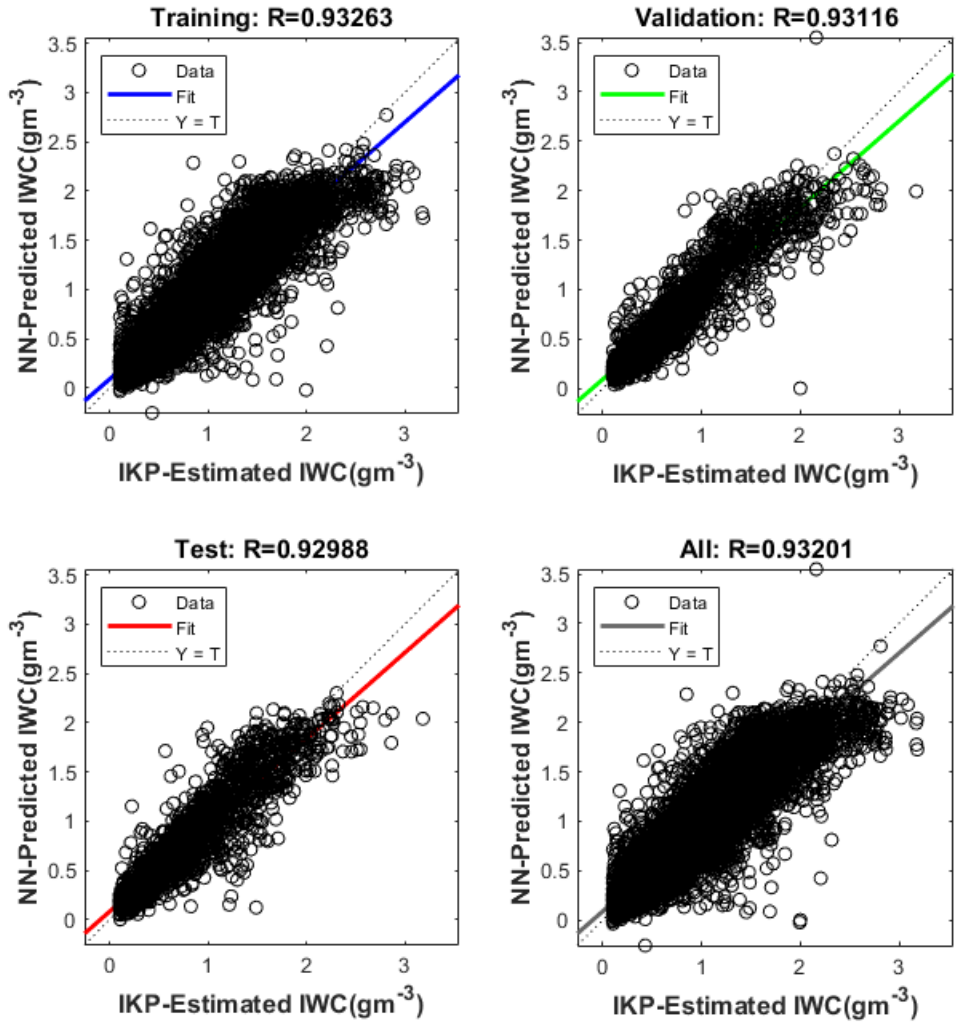


Figure 4.18: Regression plots for HIWC detection, without using temperature as an input feature.

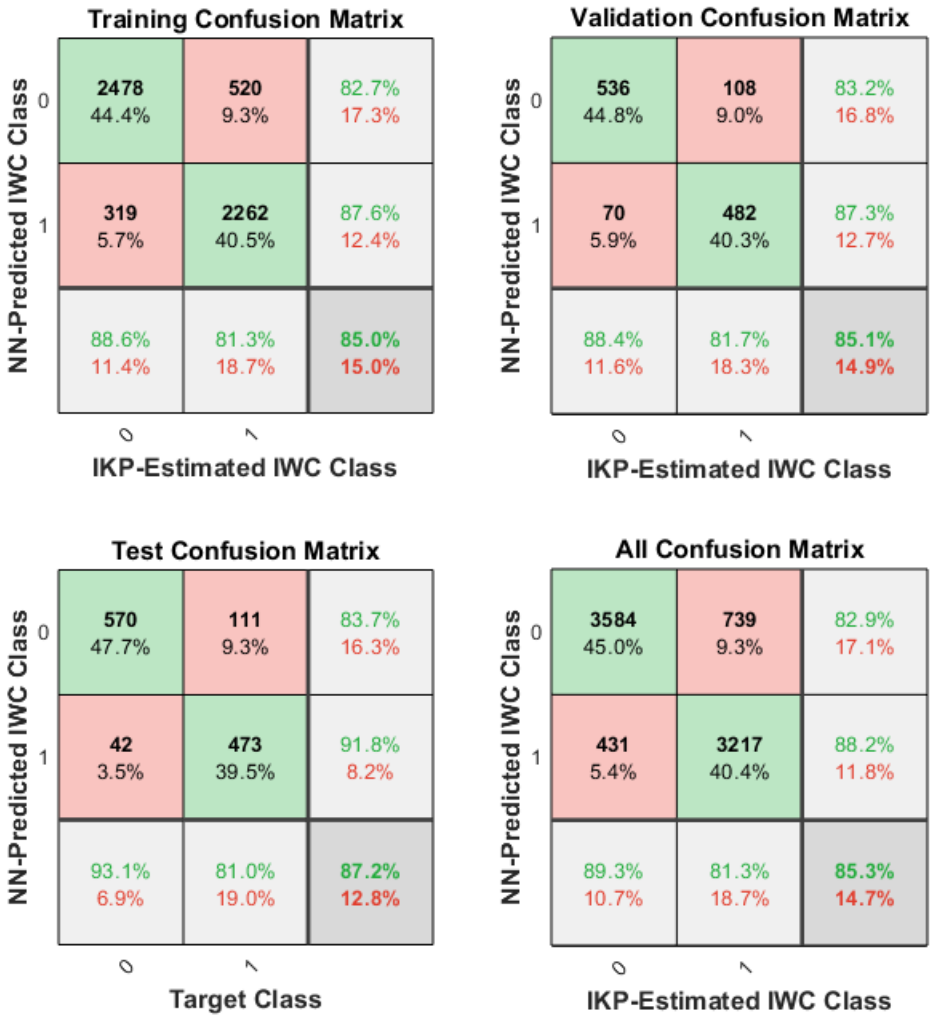


Figure 4.19: Confusion Matrices for HIWC detection, without using temperature as an input feature.

Figure 4.20 shows the ranking of the features used according to their importance. The ranking is based on the univariate F-test (Guyon and Elisseeff 2003). It can be seen that the K_{dp} is the most crucial feature for estimating IWC. It is due to the linear relationship between IWC and K_{dp} as seen in the SCMC outputs. The mean reflectivity values (Z_h and Z_v) are of secondary importance, while Z_{dr} and the variances of the reflectivity values (which include RIWCs) are less significant. This signifies the importance of K_{dp} and the potential values of incorporating polarimetric airborne weather radar variables for HIWC detection and estimation.

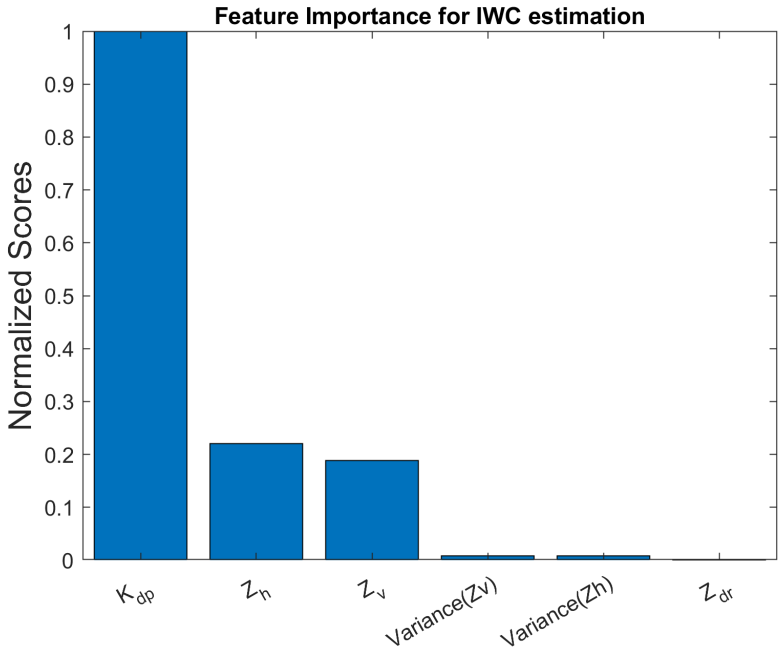


Figure 4.20: Confusion Matrices for HIWC detection, without using temperature as an input feature.

To investigate the effectiveness of dual-polarized radar variables as features and provide examples of using the airborne-PASIM, Z_h, K_{dp} and Z_{dr} is used as input feature variables for the same NWP weather field as in Figure.4.14 as the test data set. Z_h, K_{dp}, and Z_{dr} from the SCMC are used as the training dataset. ANN is designed and optimized herein as a predictor-classifier for IWC values. Again, the truth IWC field for airborne-PASIM (the output variable for the test dataset)

is taken from the TASS output IWC grid instead of actual field measurements. For each radar resolution cell spatially distributed over a PPI scan, "truth" radar variables are computed for it by averaging SCMC outputs, and I/Q signal samples of it are generated for both H and V channels based on an airborne encounter geometry. Z_h, K_{dp}, and Z_{dr} are estimated using these I/Q data, and then they are fed to the ANN (trained using SCMC outputs) to predict the IWC values. The visual illustration of the ML-based IWC estimation process is shown in Figure 4.21. The results of IWC prediction are shown in Fig. 4.22 (b). Compared to Fig. 4.22 (a), a small bias of IWC estimation can be observed, while the overall result is promising and validates the potential usage of polarimetric radar measurement for IWC *estimation*.

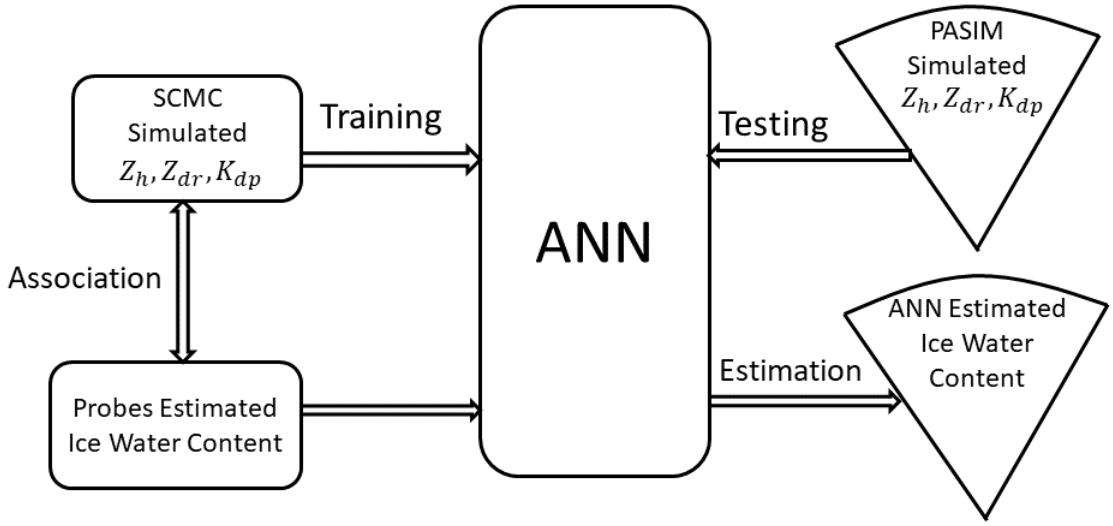
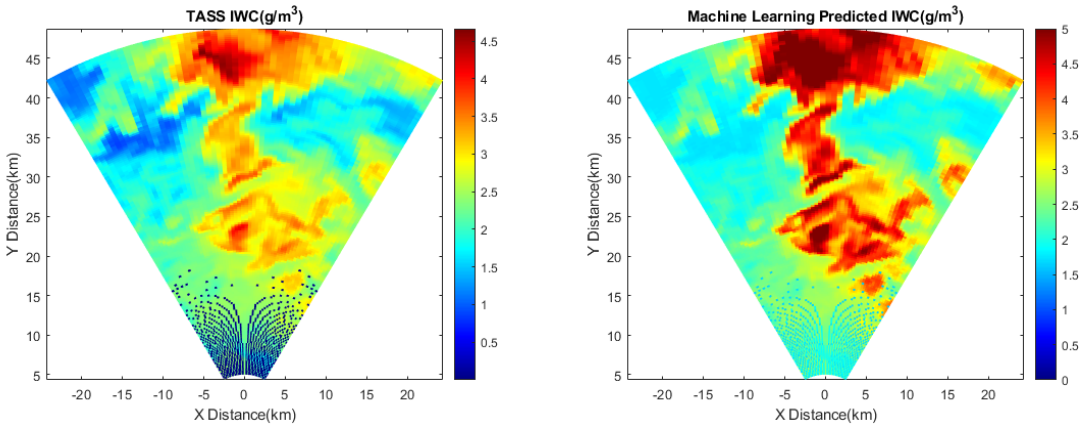


Figure 4.21: Illustration of the ML Based IWC Estimation Process.



(a) (b)

Figure 4.22: Prediction of IWC levels using (K_{dp} and Z_{dr}) feature vector. (a) IWC values from TASS output field (b) IWC values predicted by ML algorithm (trained ANN using SCMC outputs)

4.8 Discussion and Conclusion

In this chapter, the HIWC was introduced as an aviation hazard that impacts jet-engine operation. The relationship between the IWC and the radar variables was studied based on the various flight campaigns. In-situ measured PSDs of ice crystals are used in the SCMC simulation to generate the radar variables using T-matrix or Rayleigh scattering theories. After generating the radar variables, an end-to-end airborne weather radar simulator is developed. The simulator incorporates the radar system parameters to generate realistic I and Q data. ML experiments are also performed using the simulated radar variables for detecting and estimating IWC values. The validations of the SCMC outputs require comparison with the radar data (remote) collected during the flight tests. Since the SCMC outputs are based on the in-situ measured PSDs, temporal as well as spatial (Radar PPI scans vs. flight path in-situ data) correlations between radar and probe measurements are performed. More insights into the HIWC measurements are obtained through this comparison and validation, while errors in the sensor correlation process may still be an issue. The

current NASA flight campaigns do not have dual-polarized radar measurements, so the validations of polarimetric variables as features are mainly based on simulations. Future flight campaigns with dual-polarized radar measurements will be necessary for further system validations.

Chapter 5

Environment Classification Based on Vehicle-Mounted Radar and Machine-Learning

In this chapter, a challenge associated with automotive radar operations is addressed. One of the issues of the automotive radar is its low resolution, which affects its target classification performance. This chapter introduces techniques to improve angular resolution and classification accuracy. First, a method called MIMO is implemented for angular resolution improvement. Secondly, it is a huge issue for practical operations of automotive radar in actual street environment for the discrimination between pedestrians and moving vehicles. Naturally, micro-Doppler signatures of different types of radar targets are used for target classification here, while the previous studies are very limited in terms of classification accuracies. A new way to incorporating the micro-Doppler features into the ML algorithms is used in this chapter.

5.1 Improvement of Angular Resolution via MIMO

5.1.1 Introduction

Autonomous vehicles have always been a field of considerable research interest. Past research have demonstrated achievements assuring that self-driving cars are, in fact the future of mobility. Self-driving cars have been made possible by sensor fusion

technique, which incorporates sensors, including camera and radar. Cameras have the best resolution. Nevertheless, their ability to sense may be affected in weather or night conditions. Radars are not affected by these conditions but lack the resolution when compared with other optical sensors (Engels et al. 2021; Gamba 2020; Patole et al. 2017).

Most automotive radars are FMCW radars whose range resolution depends on the bandwidth of the FMCW chirp, and angular resolution depends upon the number of the receiving antennas. Having a higher number of receiving antenna elements will improve the angular resolution. Instead of increasing physical receiving antennas, it is possible to generate virtual receiving antennas by adding transmitting antennas, commonly known as the MIMO technique (Li and Stoica 2008; Rao 2018). MIMO requires orthogonal signals in multiple transmitting antennas. Previous research has explored the possibility of using MIMO techniques in an FMCW radar. Various techniques to implement orthogonality in an FMCW radar have been discussed in the previous article (De Wit et al. 2011). Commercial automotive radars have implemented the capability of MIMO using TDM and BPM in 2Tx and 4Rx systems (Rao 2018). Although the angular resolution is improved, the maximum unambiguous velocity is reduced by half. This problem will only exacerbate when more Tx are added to improve the angular resolution. The maximum unambiguous velocity reduces by the factor equal to the number of Tx elements. Thus, it becomes necessary to explore other multiplexing techniques. This study proposes the FDM technique to achieve orthogonality. The theoretical concepts of FDM-based MIMO radar have been presented in articles (De Wit et al. 2011; Zwanetski et al. 2013). Frequency modulated signal with different starting frequencies for two Tx antenna is used to create 8 Rx virtual channels. However, the limitation of FDM is that it usually requires an increment in sampling frequency of Analog to Digital Converter (ADC), to accommodate for the two frequencies. In this study, the choice of design parameters like the two

starting frequencies is discussed to eliminate the drawbacks of FDM-MIMO identified in prior research. There have been limited studies on the FDM-MIMO-based automotive radar. In this study, a full radar system has been simulated in MATLAB environment, which shows the possibility of using FDM in automotive radars. FDM is shown to achieve the same improvement in angular resolution as TDM-MIMO.

5.1.2 Theory of Operation

An automotive radar transmits FMCW chirps. When a chirp transmitted by FMCW radar is reflected from a target, it returns to the radar receiver. The received signal is the delayed version of the transmitted signal. The delay depends upon the distance of the target from the radar. When mixed with the transmit reference signal, baseband, or IF signal, the received signal is generated with the frequency corresponding to the target range ($f = S2d/c$). Here, S is the slope of the chirp, d is the target's distance, and c is the propagation speed of light. The signal is sampled by an ADC and received for multiple chirps and Rx antenna elements. FFTs and advanced techniques like digital beamforming are applied to estimate the range, velocity, and AOA (Patole et al. 2017; Richards 2005).

The angular resolution, the ability to resolve two targets in the angular domain, depends upon the number of receiving antenna elements in an array system (Gamba 2020; Rao 2018). Angular resolution of a general antenna system is given by Equation 5.1.

$$\Delta\Theta = \frac{\lambda}{Ndcos\theta}, \tag{5.1}$$

Here, λ is the transmitter wavelength, d is the spacing between elements, N is the number of elements, and θ is the pointing angle. For antenna spacing of $\lambda/2$ and $\theta = 0^{\circ}$ (boresight), the equation becomes $\Delta\Theta = 2/N$. Thus, it is clear that the number of receiving antenna elements should be increased to have a better angular

resolution. One intuitive way is to increase the number of receiving antenna elements. However, due to the size constraints of automotive radars, it is not always pragmatic to increase the number of receiving antenna elements. Another way is to add transmitting elements. For two transmitting antenna elements separated by 2λ, as shown in Figure 5.1.2, it is possible to create 8 virtual antennas. The phase difference in each four Rx antenna elements will be the same as if 8 physical Rx antenna elements were used. The signal transmitted from the first transmitting element will create the first four virtual receiving array elements. The signal transmitted from the second transmitting element will create the remaining four virtual receiving array elements.

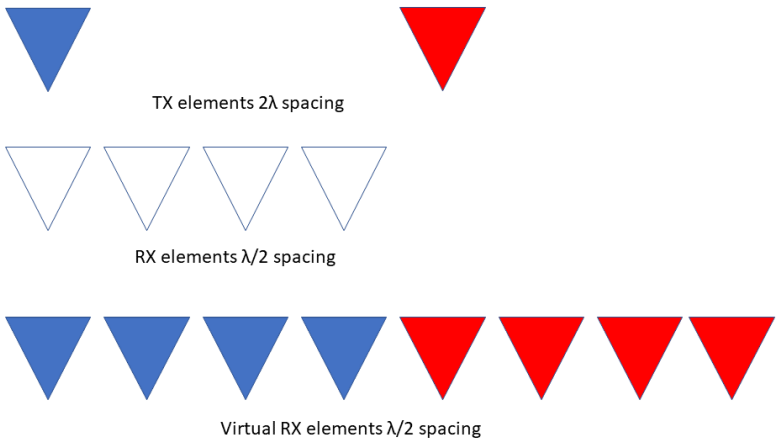


Figure 5.1: Virtual array demonstration for the case of two transmitter and four receiving array, which generates eight virtual receiving array elements.

The practical implementation of MIMO requires orthogonal waveforms transmitted from at least two antennas. One of the simplest methods is to use TDM as used in the previous report (Rao 2018). In this technique, two transmitting antenna transmits the signal alternatively. Despite being the easiest to implement, the maximum

velocity that can be measured is reduced. To overcome this issue, FDM can be implemented. Instead of transmitting signals alternatively, signals with different carrier frequencies are transmitted, thus creating an offset in the beat frequencies for two signals corresponding to two antennas. In the case of a system containing two transmit antennas, if the frequencies of the transmit signals are f and $f \pm \delta f$, and then the received signal is mixed with the reference signal of frequency f, the baseband signals of frequency $S2d/c$ and $S2d/c \pm \delta f$ are generated which can be separated digitally in frequency domain (De Wit et al. 2011).

The IF signals are sampled by ADCs (analog-to-digital converters), whose sampling rate plays an important role in determining the maximum range the radar can measure. According to the Nyquist Sampling criteria, the maximum frequency that can be sampled using an ADC with a sampling rate of f_S is $f_S/2$. The maximum range that can be measured is $Dmax = (c \times f_S)/(2 \times S)$ (Gamba 2020). Adding offset to frequency will require an increment in sampling rate by δf to sample signals transmitted by two antennas in the receiver chain. This study demonstrates a technique to eliminate the need for adding the sampling rate. If the frequencies of the transmitted signals are offset by half of the sampling rate ($f_S/2$), with the techniques of IQ demodulation, the same targets can be resolved in two halves of the spectrum after doing a range FFT. In Figure 5.2, a target located at 8 meters is estimated to be at 8 and -16 meters simultaneously. The negative range does not have any physical significance and is only the predetermined offset value in range measurement. By considering this offset in the digital backend processing, the actual range can be estimated.

5.1.3 Simulation

A practical automotive radar was simulated using MATLAB. MIMO was achieved using two Tx and four Rx antennas, respectively. The simulator can perform the

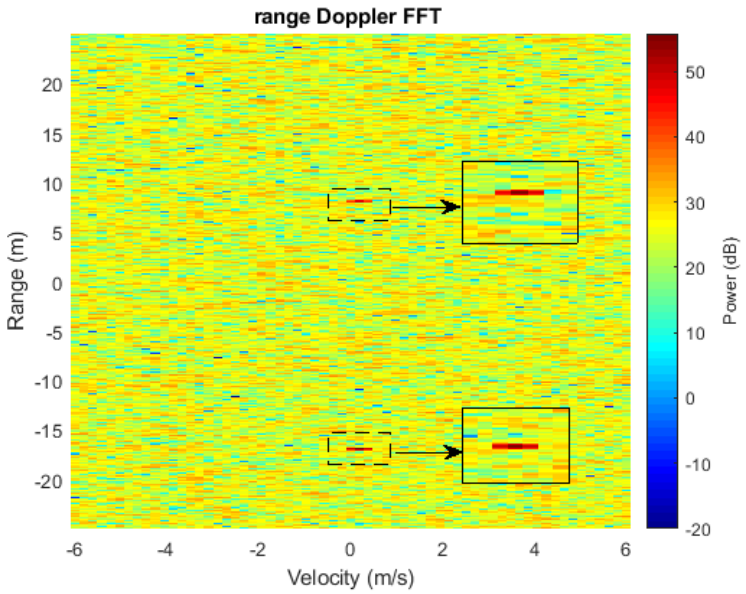


Figure 5.2: Range Profile demonstration whrn FDM MIMO is used.

TDM as well as FDM MIMO simulations. For TDM, two Tx antennas transmit the signals alternatively. Similarly, for FDM, the two Tx antennas transmit the signals of two different frequencies of 77 GHz and 76.995 GHz, respectively. The sampling rate of the ADC is 10 KiloSamples Per Second (KSPS). The two transmit frequencies are offset by half of the sampling rate. Basic radar parameters are shown in Table 5.1.

Based on the information about the target provided to the simulator, it can generate simulated I/Q time series output. There can be various targets at different locations with different velocities. The simulator is based on the time domain signal modeling. The range information is encoded into the signal as time delay, while velocity and angle information are encoded as phase shifts. The phase shift due to a moving target can be observed in signals from different chirps, while the phase shift due to the angular location can be observed in signals from receiving antenna elements. A datacube is generated based on the I/Q signals containing complex in-phase and quadrature data arranged in range, chirp, and antenna element directions. Simple FFT techniques are used to extract range, velocity, and angular information.

Table 5.1: Radar Simulation Parameters

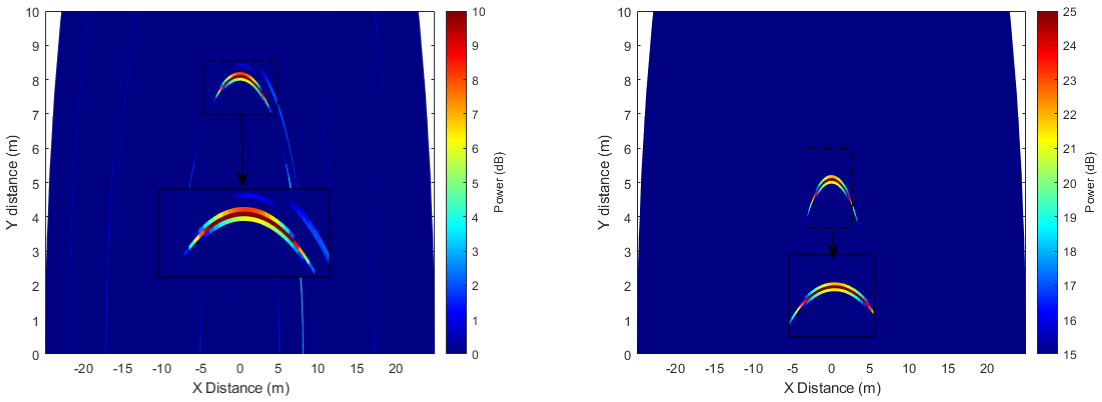
Radar Parameters	Values
Center Frequency	77 GHz
Bandwidth	1.798 GHz
Frequency Slope	29.982 MHz/μs
Number of Chirps	64
Number of Tx Antenna	2
Number of Rx Antenna	4
Idle Time	100 μs
Tx power	12 dBm
Maximum Unambiguous Range	24.99meters
Maximum Unambiguous Velocity	6.08 ms^{-1}
Spatial Resolution with MIMO (Boresight)	14.323 $^{\circ}$

FFT along range sample gives the range of the target. FFT along chirp (slow time) gives velocity, and FFT along the antenna element direction gives the angular information. As for comparisons, different Digital Beamforming techniques can also be used to extract the AOA.

5.1.4 Experimental Results

Multiple experiments with various target scenarios were performed to validate the performance of the MIMO mode for TI's automotive radar. Two representative target scenarios are presented here. In the first experiment, a single stationary target at 8 meters at boresight was considered. In the second experiment, three targets at 5 meters were considered. They were assumed to be moving with the radial velocities of 2.5 ms^{-1}, 5 ms^{-1}, and 7.5 ms^{-1}, respectively. Their angular locations were

-18°, 0°, and 18°, respectively. Figure 5.3 shows the range-AOA profile when a single Tx antenna was considered, i.e., no MIMO was implemented. Figure 5.3 (a) and Figure 5.3 (b) depicts the range-AOA profile for single target consideration and three targets consideration, respectively. Without using the MIMO technique, the three targets are not resolved spatially.



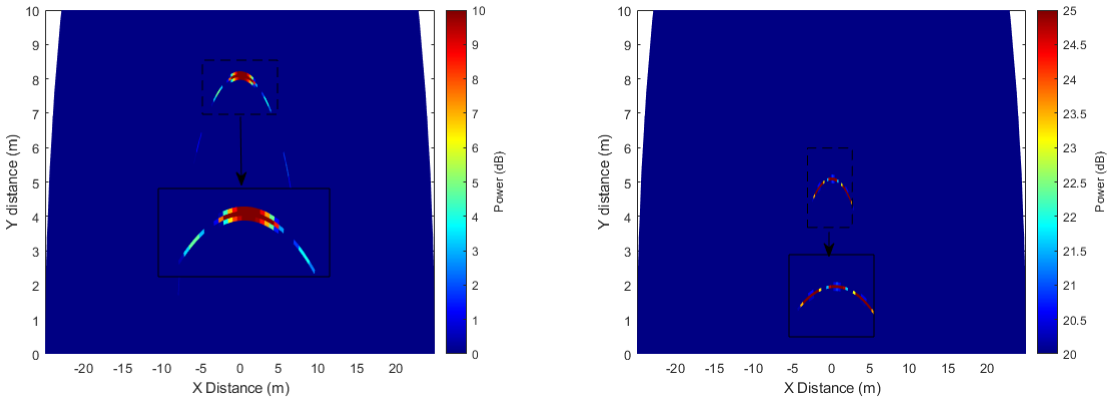
(a) (b)

Figure 5.3: Range-Angle of Arrival Map without using MIMO. (a) Single target (b) Three targets

For TDM- based MIMO, each TX antenna was enabled alternatively. The range-AOA profiles for the two experimental simulations are shown in Figure 5.4. Figure 5.4 (a) shows that the use of TDM MIMO provides better angular information about the target. Also, the three targets are resolved, as shown in Figure 5.4 (b).

Similar results were observed from the results based on FDM MIMO. Figure 5.5 shows that compared to non-MIMO, the angular resolution is improved significantly for both target scenarios.

Figure 5.6 compares the AOAs obtained using different MIMO schemes for the three target scenarios. When MIMO was not used, only one peak at 0° is observed. For both FDM and TDM MIMO, three targets (-18°, 0°, and 18°) are resolved. This is because the resolution has improved by a factor of 2 (28.64° to 14.32° at boresight).



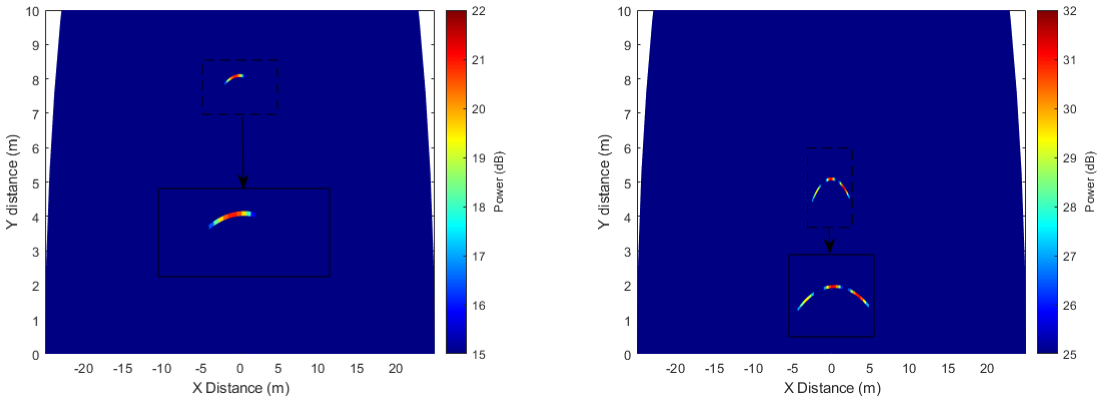
(a) (b)

Figure 5.4: Range-Angle of Arrival Map using TDM MIMO. (a) Single target (b) Three targets

There are some biases in the estimation. This can be corrected by increasing FFT points or by using digital beamforming techniques.

Although both TDM and FDM can achieve the orthogonality required by a MIMO system, the use of TDM for MIMO affect the maximum velocity that can be measured without any ambiguity. The maximum unambiguous velocity is given by Equation 2.22.

In Equation 2.22, the maximum unambiguous velocity (v_a) depends on the time duration between two chirps (T_C). Since, in TDM, the effective duration between two chirps transmitted by a TX antenna increases, the maximum unambiguous velocity decreases. In this simulation, the maximum unambiguous velocities for TDM and FDM-based MIMO are 3.04 and 6.08 ms^{-1}. Out of three targets moving with velocities of 2.5, 5, and 7.5 ms^{-1}, the TDM range-Doppler map shown in Figure 5.7 (a) indicates only one velocity value (2.5 ms^{-1}) is estimated accurately. But, the FDM range-Doppler map shown in Figure 5.7 (b), two velocity values (2.5 and 5 ms^{-1}) are estimated accurately.



(a) (b)

Figure 5.5: Range-Angle of Arrival Map using FDM MIMO. (a) Single target (b) Three targets

5.1.5 TDM MIMO Implementation using TI's AWR 1443 Radar

In addition to simulation results, Texas Instruments AWR 1443 Automotive radar was used to implement TDM-MIMO. The specification of the radar is similar to the simulator used and is shown in Table 5.1. The radar used and the experimental setup is shown in Figure 5.8. A stationary sedan vehicle at boresight was used as a target. The range-AOA map for the target at the range of 13 meters is shown in Figure 5.9. As observed in the TDM and FDM-based MIMO simulation, the angular resolution is improved when MIMO feature is enabled in TI radar. The angular resolution at boresight decreases from 28.64° to 14.32° in both MIMO simulations and when TI radar is used for measurement, thus validating the simulations.

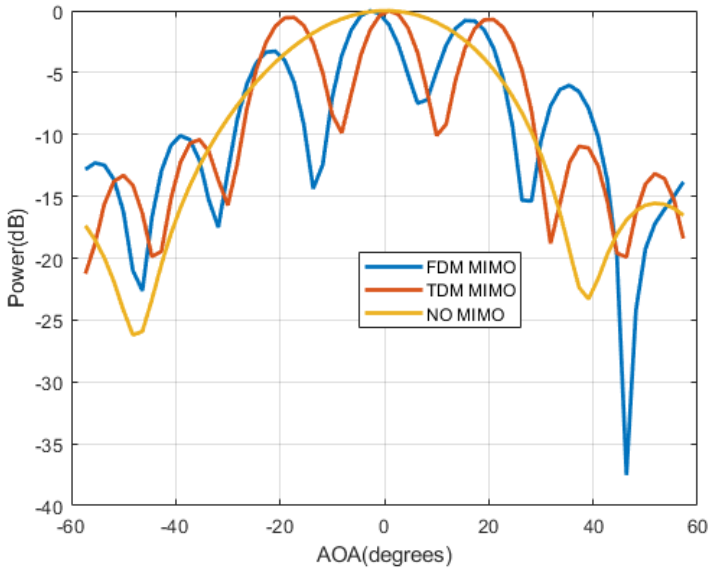
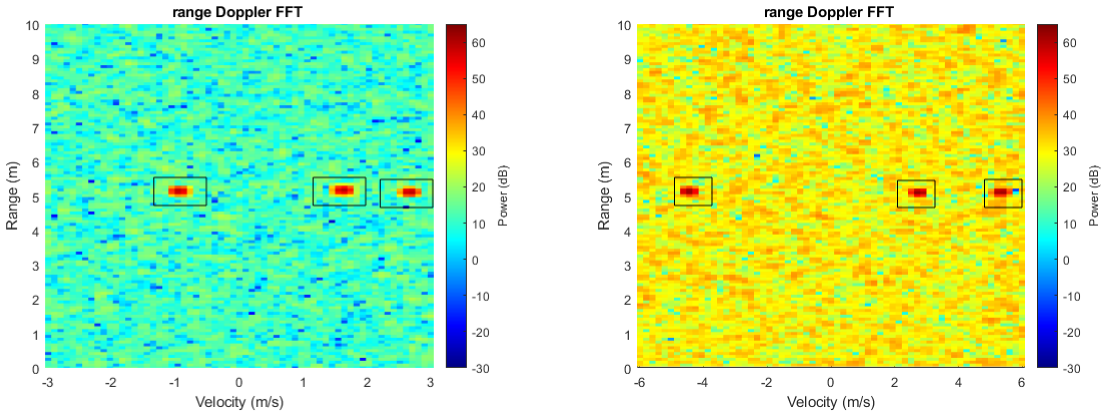


Figure 5.6: AOA comparison for different MIMO schemes.

5.1.6 Conclusion

To summarize, an end-to-end automotive radar with MIMO capability was simulated and partially verified, and the results were partially validated with actual measurements. The use of MIMO in radar systems can improve angular resolution. In this section, orthogonality of MIMO channels is achieved using both TDM and FDM. TDM is easier to implement but reduces the maximum unambiguous velocity. This problem can be overcome by using FDM. FDM may require ADCs with a higher sampling rate. Here, with IQ demodulation and using correct choices of frequencies in two transmit channels, there is no need to increase the receiver sampling rates. However, the implementation of FDM techniques can add some complexity to the radar system design. The complexity increases with the increment in the number of Tx elements. For future studies, other orthogonality techniques like phase division multiplexing or code division multiplexing can be implemented to achieve MIMO

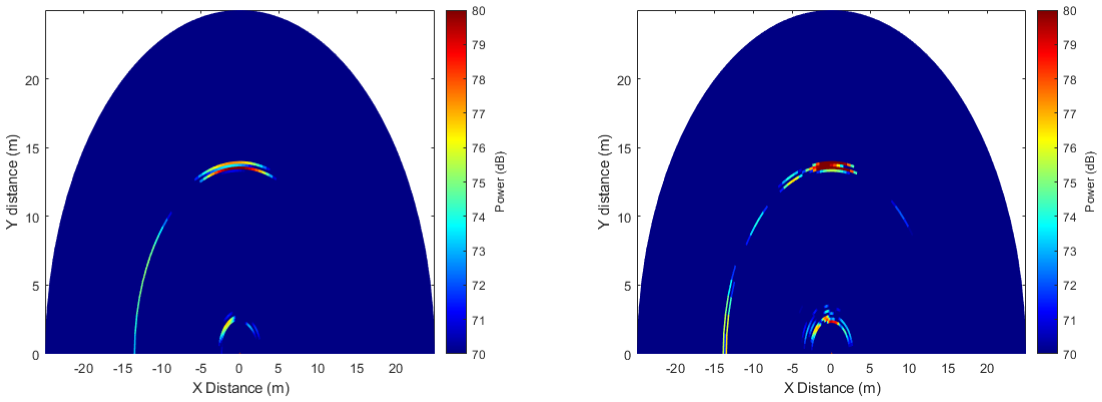


(a) (b)

Figure 5.7: Range Doppler map for MIMO radars. (a) TDM (b) FDM

(a) (b)

Figure 5.8: Measurement using TI radar (a) Radar used for the measurement (b) Experimental Setup



(a) (b)

Figure 5.9: Range-Angle of Arrival Map of a single car target, measured using TI AWR1443 radar. (a) Without MIMO (b) With MIMO

(Alaee-Kerahroodi et al. 2019; Rambach and Yang 2017). In addition, further studies about the feasibility of implementing these techniques in actual automotive radar products are recommended.

5.2 Micro Doppler based Classification of Pedestrian and Vehicle

5.2.1 Introduction

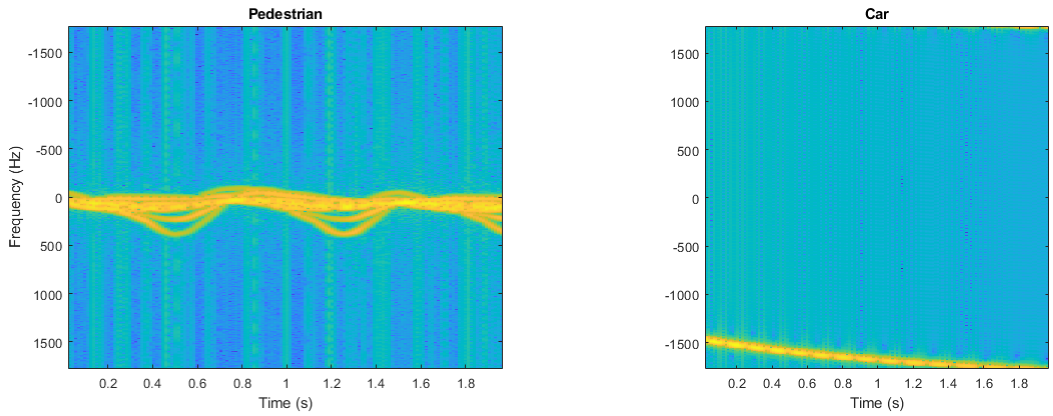
In addition to the radar's capability to detect the targets' bulk movement, it can also detect small oscillations within the target. These oscillations are referred to as Micro-Doppler. The source of Micro-Doppler can be swinging arms and legs in humans and rotation of the wheels in a vehicle. A periodic Micro-Doppler modulation can induce the modulation in Doppler frequency as the side-band shifts around the Doppler frequency (Chen et al. 2014).

This study focuses on the classification of a pedestrian and a vehicle. These two classes of targets represent a simple traffic condition. The classification of the targets is necessary for an autonomous system to make an optimum decision. Initial classification strategies studied the classification of stationary and moving targets based on the Doppler filtering (Hyun et al. 2017). But with the advancement in autonomous vehicle technology, studies about the classification of the types of moving targets are becoming popular. Target classification based on the Radar Cross Section (RCS) has been popular (Mansukhani et al. 2021). However, the classification using RCS may be affected by various factors like the size, and orientation of the target and it may be computationally inefficient (Kim et al. 2018). Micro-Doppler based classification has been implemented using the feature extraction on a target cluster of a range-Doppler map (Prophet et al. 2018). However, target clustering might be difficult. In this study, micro-Doppler is represented in the form of a spectrogram and is used directly with CNNs for classification. This eliminates the need for feature extraction and avoids any errors associated with the clustering and the manual feature extraction process.

Figure 5.10 shows an example of the difference in the Micro-Doppler characteristics of a pedestrian and a vehicle. The plots are simulated using MATLAB's radar toolbox. The simulation assumes that the 77 GHz FMCW radar is used. These features of the different targets can be the basis of classification using machine learning algorithms.

5.2.2 Signal Processing for Pedestrian Detection

Figure 5.11 shows a high-level signal processing chain for the micro-Doppler analysis. First, the base-band ADC data goes through a Range-FFT (FFT across the fast time samples) and then the Doppler FFT (FFT across the chirps in a frame). From Range-Doppler Maps, the targets are detected. Peak detection can be used with only one target in a frame, or a Constant False Alarm Rate (CFAR) detection can



(a) (b)

Figure 5.10: Micro Doppler characteristics of a Pedestrian (a) and a Vehicle(b), Generated using the Matlab's Radar Toolbox

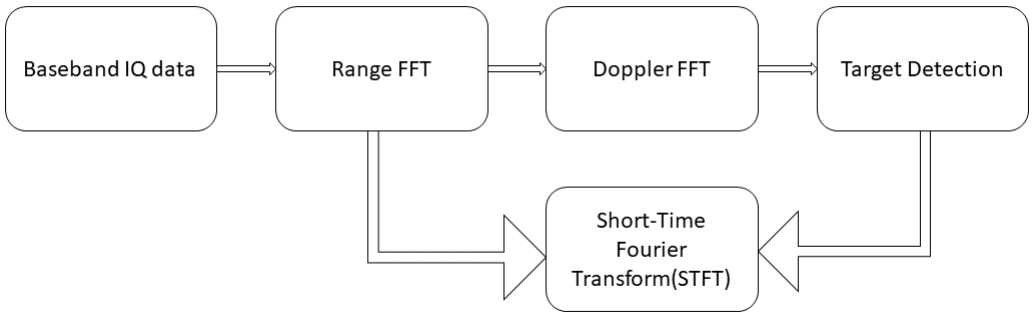
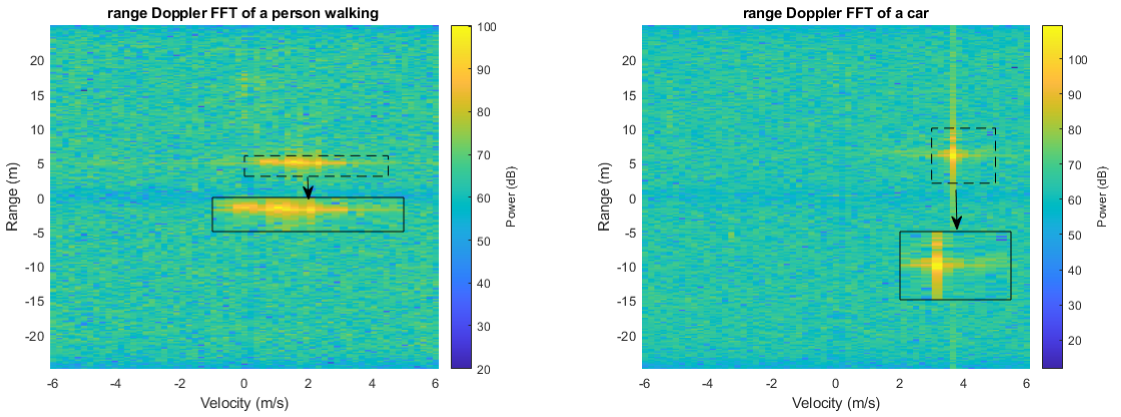


Figure 5.11: Signal Processing chain for the micro Doppler analysis.

be applied to detect multiple targets in the Range-Doppler Map. Then for each detected target, Short-time-Fourier Transform (STFT) is performed using the Range-FFT data. STFT determines the variation of the frequency spectrum over time or the target's micro-Doppler characteristics.

TI 1443 FMCW radar was used to collect the data for a pedestrian and a car, respectively. The start frequency of the radar was 77 GHz, with a frequency slope of 29.982 $MHz/\mu s$. 64 chirps were collected for each frame, and the chirp repetition time was 160 μs. The influence of the pedestrian's gait can be clearly observed in the Range-Doppler map as shown in Figure 5.12. For a pedestrian, the target velocity is smeared across multiple Doppler bins. This is because a walking person, in addition to the bulk Doppler phase information, exhibits the Doppler properties induced due to the swinging arms and legs and the rotating torso. The rotation of wheels in the car does not have notable Micro Doppler signatures. To interpret these signatures, it is often convenient to use the time-frequency properties of the return signal. For this, STFT is used. The spectrogram analysis of STFT generates the visual representation of the frequency variation with time. The data for 64 chirps were divided into smaller segments of 8 chirps using a repetition factor of 7. 1024 points FFT was performed for each segment. This resulted in the frequency resolution of 781 Hz. As a result, spectrogram plots of the size 1024×57 were used.

Figure 5.13 shows the STFT spectrogram of a pedestrian and a car together, collected using the 77 GHz Radar. The gait of a pedestrian can be identified in the spectrum plots. The periodic frequency variations on the STFT spectrogram of a pedestrian are due to the periodic movement of body parts. Such variations are absent in the STFT spectrogram of a moving car. These differences can be used as features for the target classification.



(a) (b)

Figure 5.12: Range Doppler Map of a Pedestrian(a) and a Vehicle(b), Collected using TI Radar

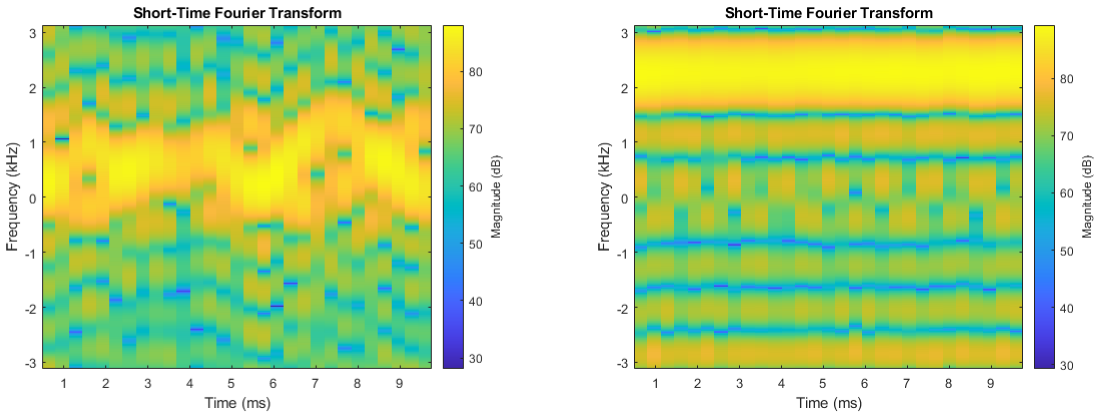
5.3 Classification of Person and Car Using CNN

In section 5.2.2, it is shown that the spectrograms of pedestrians and the vehicle in front of the radar exhibit different micro-Doppler characteristics. The spectrogram plots can be used with image classification algorithms like CNNs for the target classification. In this study, a 15-layers CNN was used for classification. The details of the 15 layers are shown in Table 5.2. The layer starts with the input spectrogram layer, which goes through a series of CNNs, Batch Normalization Layers, ReLUs, and Max-Pooling Layers before being unfolded into fully connected layers. Then with the help of Softmax and cross-entropy functions, the final classification output is generated.

The data collected using the TI 1443 radar were first converted into their spectrograms and labeled appropriately (car vs. pedestrian). A car was driven towards and away from the radar. Similarly, a person was made to walk towards and away from the radar. Around 2142 micro-Doppler images (approximately 50% for each class) were used for training and testing. Out of these images, 1200, 400, and 542 were selected randomly for training, validation, and testing purpose, respectively. Figure 5.14 shows the performance of the micro-Doppler-based classification using CNN in

Table 5.2: CNN Parameters

Layer number	Layer type	Description
1	Input layer	Size of spectrogram (1024*57)
2	2D Convolution Layer	8 layers with size (3,3)
3	Batch Normalization Layer	Batch Normalization
4	ReLU	Rectified Linear Activation
5	Max Pooling	(2,2) max Pooling with stride 2
6	2D Convolution Layer	16 layers with size (3,3)
7	Batch Normalization Layer	Batch Normalization
8	ReLU	Rectified Linear Activation
9	Max Pooling	(2,2) max Pooling with stride 2
10	2D Convolution Layer	32 layers with size (3,3)
11	Batch Normalization Layer	Batch Normalization
12	ReLU	Rectified Linear Activation
13	Fully Connected Layer	2 fully connected layers
14	Softmax	Softmax
15	Classification Output	Cross Entropy



(a) (b)

Figure 5.13: Micro Doppler characteristics of a Pedestrian(a) and a Vehicle(b), Collected using TI Radar

the form of confusion matrix. As shown in the plot, an accuracy level of 94% or higher was achieved for training, testing, and cross-validation sets, respectively. This establishes the foundation of target-classificaion using a radar.

5.3.1 Discussion and Conclusion

This section explored a classification strategy based on the micro-Doppler information. The target classification using the standard visual approach in a range angle domain may fail because of the lower radar resolution. As a result, the different characters between two or more types of targets may not be learned by a machine learning algorithm. The use of micro-Doppler information gives a distinct character. These characters are not only easy to visually interpret but can form the features for the target classification. The use of micro-Doppler for the target classification showed encouraging results.

The limitation of this study, however can be the extensive computations required by the CNNs. An automotive radar system has to be cost effective and adding processors to support the computations required by a CNN can make the entire

radar system expensive. Another limitation of this study is the use of only two targets for classification. In future, more target scenarios need to be experimented. For examples, to create a road-like scenario, different types of vehicles like motorcycle, cars, trucks, etc. and different movement of pedestrians like walking, running, etc. can be included. Similarly, targets moving at different angles in front of the radar can also create a realistic road like target scenarios.

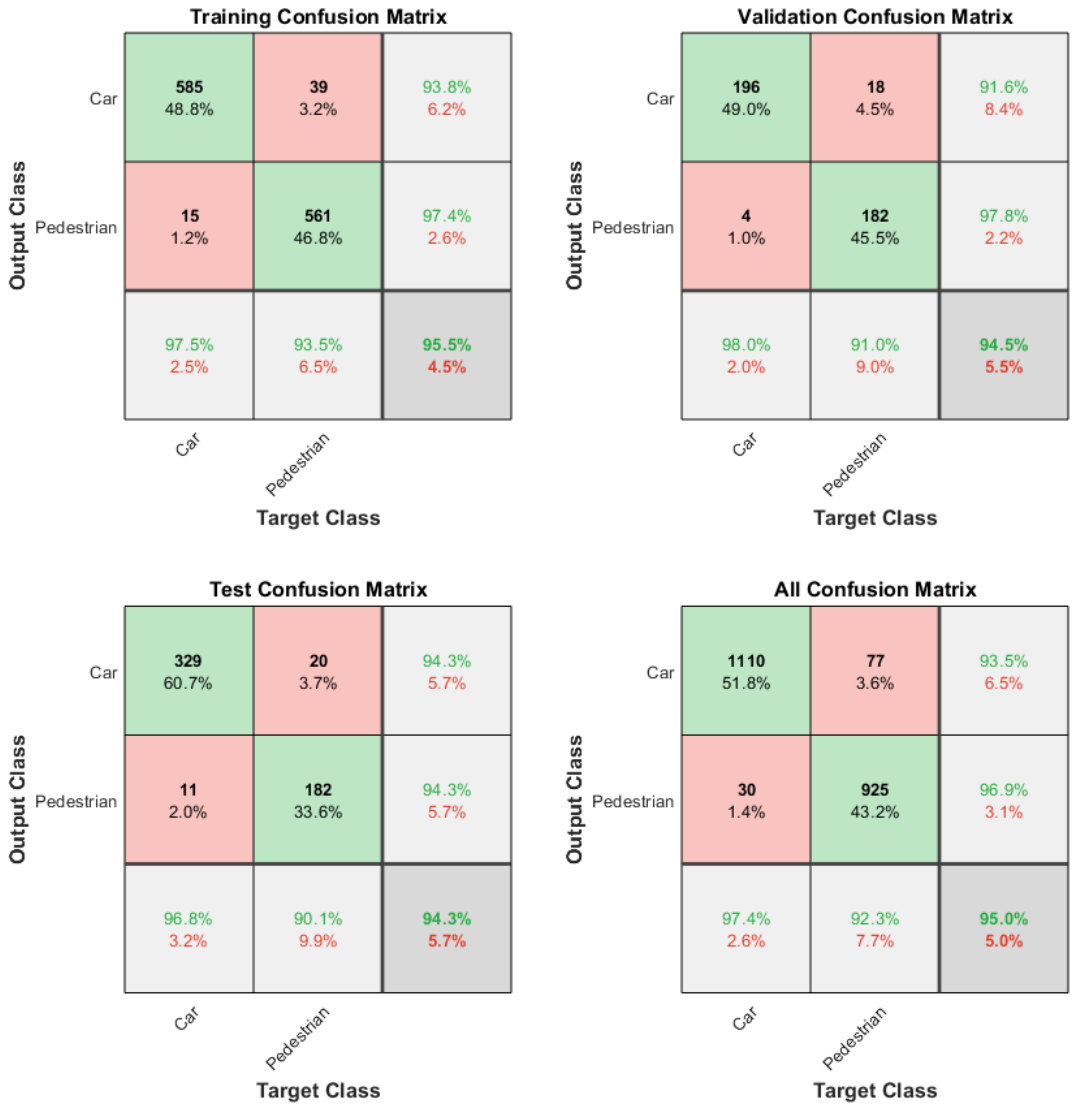


Figure 5.14: Confusion Matrix for the Classification of the Pedestrian and a Car Using Micro-Doppler Spectrum Data

Chapter 6

Conclusions and Future Work

This chapter concludes this dissertation by summarizing the main research findings and discussing their value and contribution. It also reviews the study's limitations and suggest opportunities for future research.

6.1 Summary of Contributions

This dissertation explores ML-based studies in radar meteorology and also in radar algorithm development. Similar research has been done in the past but these studies had some limitations. The dissertation studies seek to address some of those limitations. For example, past research utilized ML technology to predict the occurrence of lightning. However, there is a gap in the research as studies to quantify lighting events are limited. In this study, utilizing ANN and feature variables derived from polarimetric radars, accurate prediction of flash rate up to 30 minutes was achieved. Similarly, for HIWC estimation, there is a need for a radar simulator to accurately characterize the ice crystals. A novel microphysics-based, end-to-end radar simulator was developed for this mission based on data collected from NASA flight campaigns. Additionally, using data generated from this simulation, the ANN achieved 85% accuracy in HIWC hazard detection, which is the first time to be reported in the science community. Lastly, there is a need for the accurate classification of objects in ground

traffic to improve the safety of an autonomous driving system. For this, the angular resolution has to be improved. In this dissertation, an innovative FDM-based MIMO solution has been proposed to improve the angular resolution. Also, with the novel method of assimilating time-spectrum features, the initial classification accuracy of better than 85% between pedestrians and vehicles has been achieved using measured radar data. In general, the applications of this dissertation study have the potential to reduce airborne accidents, and automotive casualties and improve overall weather forecasting. The contributions of each mission are further discussed in detail in the paragraphs below.

6.1.1 Prediction of Lightning Flash Rate Using Ground-based Weather Radar Observations

This study used ML with the ground-based weather radar to predict the flash rate. The lightning flash rate estimation is a novel application in itself, as previous studies focused mainly on detecting lightning. Quantifying the lightning event aids in better conveying the severity of lightning in the weather forecasts. It can also be an indicator of other weather phenomena, such as tornado events. As a result, it provides insights to weather forecasters and researchers alike. In addition, the nowcasting lead time was increased to 30 minutes, compared to 15 minutes in the previous literature. This helps to provide additional time to take necessary precautions in lightning events. In addition, the importance of dual-pol features, like the presence of negative K_{dp}, is demonstrated in the lightning studies. This information can be significant for weather radar researchers in further studies of lightning events.

6.1.2 Detection of HIWC Conditions Using Airborne Weather Radar

In this study, a complete radar-based characterization of the HIWC was introduced. The study started with the study of HIWC's microphysical properties like PSD and density, followed by the development of the radar variables for the HIWC, a full Radar system simulation, and the estimation of HIWC using ML algorithms. The previous studies primarily focused on the radar reflectivity factor and the temperature, while studies using the dual-pol features are limited. To mitigate this gap, the dual-pol features are used in this study which led to the observation of the significance of the dual-pol features, particularly the K_{dp}. Although the temperature has significance for HIWC conditions, the temperature dependency on HIWC is difficult to utilize in a practical scenario. So the "radar-only" detection of HIWC, as demonstrated in the study, will be benefited from the dual-pol radar measurement features. An end-to-end radar simulator is developed. The simulator can assist other radar developers and researchers to characterize the HIWC conditions and to generate the radar variables for the development of IWC retrieval algorithms. The outcome of this study may be beneficial to FAA and RTCA to develop future airborne weather radar standards.

6.1.3 Target classification With Automotive Radars

In this study, the use of automotive radar for classification purposes has been discussed. The radar has been deemed inferior when compared to a camera in terms of resolution, which affects the classification capabilities. The resolution can be improved using MIMO technology. TDM-based MIMO, despite affecting the velocity performance, has dominated the automotive radar segment. In this study, FDM-based MIMO was introduced as an alternative MIMO technique. FDM-based MIMO does not reduce the maximum unambiguous velocity. The results from the simulation encourage the radar developers to explore the possibility of implementing FDM

MIMO. Similarly, the difficulties in the visual classification techniques using radar can be overcome by the micro-Doppler characterizations of the targets. Micro-Doppler spectrogram is shown to be used directly with a CNN without the need of extracting features manually. Thus, the classification should be done on the different(like micro-Doppler) domains. This should encourage the researchers not to solely rely on the techniques used for image classification in the visual domain, as in the case of cameras.

6.2 Future Work

The use of ML techniques in radar technologies has shown impressive results. However, there are some limitations. The applications of ML (in any domain) can generate expected outcomes, with or without the knowledge of the domain. However, the knowledge of the domain is fundamental as well. Additional care should be taken while using ML as a research tool to fulfill scientific objectives. The domain knowledge can be beneficial with the data pre-processing steps. The other limitation could be the availability of data. Although the amount of data available for ML experiments is increasing, the specific data required for a specific scientific study can be expanded.

6.2.1 Lightning Detection Application

The lightning flash rate estimation using radar and ML showed some encouraging results. However, the current study only used radar data for estimating lightning flash rate. Since the radar data cannot provide the temperature estimation directly, the temperature information was inferred from altitude. This was done by using the echo top height as an input feature. The further usage of measured temperature may improve the accuracy of the flash rate estimation. The input features can be modified

to include only the region above freezing altitude, as this region is more significant for the generation of lightning strikes. In addition, the inclusion of temperature will also help in extracting more features like the presence of Z_{dr} column. Also, this study only focused on one geographical region of Hong Kong. Similar studies in other geographical regions with different climatology would help to expand it to a more generalized solution.

6.2.2 Monitoring HIWC as an Aviation Hazard

The HIWC estimation has used the simulated radar data in the ML implementation, as there are limited direct radar measurement data. The use of measured data is crucial for the accurate validation of the results. More HIWC flight campaigns using the dual-polarized airborne weather radar are necessary. In addition, there are some limitations in using flight campaign data. The association of the in-situ data and remote-sensing radar data collected by different instruments (microphysics probe vs. radar) may introduce errors as it is difficult to achieve exact temporal and spatial data associations between these instruments. Future work would need to address this data correlation challenge better.

6.2.3 Further Development of Automotive Radar

Different MIMO schemes like CDM, QPSK, or DCM can be explored to further investigate angular resolutions' impacts on ML processing. Future automotive radars' classification performance and spatial resolution can benefit from imaging 4D multi-domain radar sensing (range, Doppler, azimuth, and elevation).

Bibliography

Ahmad, N. and F. Proctor, 2011: Simulation of benchmark cases with the terminal area simulation system (tass). *49th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*, 1005.

Alaee-Kerahroodi, M., B. S. MR, M. Naghsh, and B. Ottersten, 2019: Cdm-mimo in next-generation mmwave automotive radar sensors. *Proceedings of URSI AP-RASC*.

Baker, M. and J. Dash, 1989: Charge transfer in thunderstorms and the surface melting of ice. *Journal of crystal growth*, **97 (3-4)**, 770–776.

Branch, W. D. T., 2012: Dual-polarization radar principles and system operations. *NOAA.[Available online at http://www. wdtb. noaa. gov/courses/dualpol/documents/DualPolRadarPrinciples. pdf]*.

Bringi, V. N. and V. Chandrasekar, 2001: *Polarimetric Doppler weather radar: principles and applications*. Cambridge university press.

Brown, P. R. and P. N. Francis, 1995: Improved measurements of the ice water content in cirrus using a total-water probe. *Journal of Atmospheric and Oceanic Technology*, **12 (2)**, 410–414.

Bruning, E. C., S. A. Weiss, and K. M. Calhoun, 2014: Continuous variability in thunderstorm primary electrification and an evaluation of inverted-polarity terminology. *Atmospheric Research*, **135**, 274–284.

Carey, L. and S. Rutledge, 1996: A multiparameter radar case study of the microphysical and kinematic evolution of a lightning producing storm. *Meteorology and Atmospheric Physics*, **59 (1-2)**, 33–64.

Carey, L. D. and S. A. Rutledge, 1998: Electrical and multiparameter radar observations of a severe hailstorm. *Journal of Geophysical Research: Atmospheres*, **103 (D12)**, 13 979–14 000.

Carey, L. D. and S. A. Rutledge, 2000: The relationship between precipitation and lightning in tropical island convection: A C-band polarimetric radar study. *Monthly weather review*, **128 (8)**, 2687–2710.

Caylor, I. and V. Chandrasekar, 1996: Time-varying ice crystal orientation in thunderstorms observed with multiparameter radar. *IEEE transactions on geoscience and remote sensing*, **34 (4)**, 847–858.

Chase, R. J., D. R. Harrison, A. Burke, G. M. Lackmann, and A. McGovern, 2022: A machine learning tutorial for operational meteorology, part i: Traditional machine learning. *arXiv preprint arXiv:2204.07492*.

Chen, V. C., D. Tahmoush, and W. J. Miceli, 2014: *Radar micro-Doppler signatures*. Institution of Engineering and Technology.

Coutris, P., A. Schwarzenboeck, D. Leroy, A. Grandin, F. Dezitter, and J. W. Strapp, 2019: Uncertainty of the ice particles median mass diameters retrieved from the haic-hiwc dataset: A study of the influence of the mass retrieval method. *SAE International Journal of Advances and Current Practices in Mobility*, **2 (2019-01-1983)**, 140–150.

De Wit, J., W. Van Rossum, and A. De Jong, 2011: *Orthogonal waveforms for FMCW MIMO radar*. IEEE.

Dixon, M. and G. Wiener, 1993: TITAN: Thunderstorm identification, tracking, analysis, and nowcasting—A radar-based methodology. *Journal of atmospheric and oceanic technology*, **10 (6)**, 785–797.

Doviak, R. J. et al., 2006: *Doppler radar and weather observations*. Courier Corporation.

Dye, J. E., S. Lewis, M. Bateman, D. M. Mach, F. J. Merceret, J. G. Ward, C. A. Grainger, et al., 2004: Final report on the airborne field mill project (abfm) 2000-2001 field campaign. Tech. rep.

Engels, F., P. Heidenreich, M. Wintermantel, L. Stäcker, M. Al Kadi, and A. M. Zoubir, 2021: Automotive radar signal processing: Research directions and practical challenges. *IEEE Journal of Selected Topics in Signal Processing*, **15 (4)**, 865–878.

Gamba, J., 2020: *Radar Signal Processing for Autonomous Driving*. Springer.

Garrett, T. J., S. E. Yuter, C. Fallgatter, K. Shkurko, S. R. Rhodes, and J. L. Endries, 2015: Orientations and aspect ratios of falling snow. *Geophysical Research Letters*, **42 (11)**, 4617–4622.

Gilliam, M., 2020: All-weather sense and avoid (saa) radar clutter modeling and control.

Gomes, C., 2017: Lightning related human risks and risk management. *American Journal of Management Science and Engineering*, **2 (5)**, 65.

Grzych, M. L. and J. G. Mason, 2010: Weather conditions associated with jet engine power loss and damage due to ingestion of ice particles: What we've learned through 2009. *14th Conf. on Aviation, Range and Aerospace Meteorology*.

Guyon, I. and A. Elisseeff, 2003: An introduction to variable and feature selection. *Journal of machine learning research*, **3 (Mar)**, 1157–1182.

Harrah, S. D., et al., 2019: Radar detection of high concentrations of ice particles-methodology and preliminary flight test results. Tech. rep.

Heitkemper, L., R. F. Price, and D. B. Johnson, 2008: *Lightning-warning systems for use by airports*. 04-02, The National Academies Press, Washington, DC.

Heymsfield, A. J., A. Bansemer, C. Schmitt, C. Twohy, and M. R. Poellot, 2004: Effective ice particle densities derived from aircraft data. *Journal of the atmospheric sciences*, **61 (9)**, 982–1003.

Hoang, L., 2019: High power radio frequency solid-state amplifiers and combiners for particle accelerators: From module to system design approach. Ph.D. thesis, Acta Universitatis Upsaliensis.

Holle, R. L., 2016: A summary of recent national-scale lightning fatality studies. *Weather, Climate, and Society*, **8 (1)**, 35–42.

Homeyer, C. R. and M. R. Kumjian, 2015: Microphysical characteristics of overshooting convection from polarimetric radar observations. *Journal of the Atmospheric Sciences*, **72 (2)**, 870–891.

Hondl, K. D. and M. D. Eilts, 1994: Doppler radar signatures of developing thunderstorms and their potential to indicate the onset of cloud-to-ground lightning. *Monthly Weather Review*, **122 (8)**, 1818–1836.

Hu, J., et al., 2019: Tracking and characterization of convective cells through their maturation into stratiform storm elements using polarimetric radar and lightning detection. *Atmospheric Research*, **226**, 192–207.

Hu, Y., et al., 2021: Dependence of ice microphysical properties on environmental parameters: Results from haic-hiwc cayenne field campaign. *Journal of the Atmospheric Sciences*, **78 (9)**, 2957–2981.

Hyun, E., Y.-S. Jin, and J.-H. Lee, 2017: Moving and stationary target detection scheme using coherent integration and subtraction for automotive fmcw radar systems. *2017 IEEE Radar Conference (RadarConf)*, IEEE, 0476–0481.

Jameson, A., M. Murphy, and E. Krider, 1996: Multiple-parameter radar observations of isolated florida thunderstorms during the onset of electrification. *Journal of Applied Meteorology*, **35 (3)**, 343–354.

Jenn, D., 2005: *Radar and laser cross section engineering*. American Institute of Aeronautics and Astronautics, Inc.

Jmour, N., S. Zayen, and A. Abdelkrim, 2018: Convolutional neural networks for image classification. *2018 international conference on advanced systems and electric technologies (IC_ASET)*, IEEE, 397–402.

Keith, W. and C. Saunders, 1990: Further laboratory studies of the charging of graupel during ice crystal interactions. *Atmospheric research*, **25 (5)**, 445–464.

Kim, S., S. Lee, S. Doo, and B. Shim, 2018: Moving target classification in automotive radar systems using convolutional recurrent neural networks. *2018 26th European Signal Processing Conference (EUSIPCO)*, IEEE, 1482–1486.

Kohn, M., E. Galanti, C. Price, K. Lagouvardos, and V. Kotroni, 2011: Nowcasting thunderstorms in the Mediterranean region using lightning data. *Atmospheric Research*, **100 (4)**, 489–502.

Kuk, B., H. Kim, J. Ha, H. Lee, and G. Lee, 2012: A fuzzy logic method for lightning prediction using thermodynamic and kinematic parameters from radio sounding observations in south korea. *Weather and forecasting*, **27 (1)**, 205–217.

Kumjian, M. R., A. P. Khain, N. Benmoshe, E. Ilotoviz, A. V. Ryzhkov, and V. T. Phillips, 2013: The anatomy and physics of zdr columns. *Journal of Applied Meteorology and Climatology*, **2014**.

Lakshmanan, V. and T. Smith, 2009: An algorithm to nowcast lightning initiation and cessation in real-time. *Proceedings of the Fourth Conference on the Meteorological Applications of Lightning Data, American Meteorological Society, Phoenix, AZ, paper*, Vol. 6.

Leroy, D., et al., 2017: Ice crystal sizes in high ice water content clouds. part ii: Statistics of mass diameter percentiles in tropical convection observed during the haic/hiwc project. *Journal of Atmospheric and Oceanic Technology*, **34 (1)**, 117–136.

Li, J. and P. Stoica, 2008: *MIMO radar signal processing*. John Wiley & Sons.

Li, Z., S. Perera, Y. Zhang, G. Zhang, and R. Doviak, 2019: Phased-array radar system simulator (pasim): Development and simulation result assessment. *Remote Sensing*, **11 (4)**, 422.

Li, Z., Y. Zhang, L. Borowska, I. R. Ivić, D. Mirković, S. Perera, G. Zhang, and D. S. Zrnić, 2020: Polarimetric phased array weather radar data quality evaluation through combined analysis, simulation, and measurements. *IEEE Geoscience and Remote Sensing Letters*, **18 (6)**, 1029–1033.

Li, Z., Y. Zhang, G. Zhang, and K. A. Brewster, 2010: A microphysics-based simulator for advanced airborne weather radar development. *IEEE transactions on geoscience and remote sensing*, **49 (4)**, 1356–1373.

López, R. E. and J.-P. Aubagnac, 1997: The lightning activity of a hailstorm as a function of changes in its microphysical characteristics inferred from polarimetric radar observations. *Journal of Geophysical Research: Atmospheres*, **102 (D14)**, 16 799–16 813.

Mansukhani, J., D. Penchalaiah, and A. Bhattacharyya, 2021: Rcs based target classification using deep learning methods. *2021 2nd International Conference on Range Technology (ICORT)*, IEEE, 1–5.

Mason, J., W. Strapp, and P. Chow, 2006: The ice particle threat to engines in flight. *44th AIAA Aerospace Sciences Meeting and Exhibit*, 206.

Mason, J. G. and M. Grzych, 2011: The challenges identifying weather associated with jet engine ice crystal icing. Tech. rep., SAE Technical Paper.

Mattos, E. V., L. A. Machado, E. R. Williams, and R. I. Albrecht, 2016: Polarimetric radar characteristics of storms with and without lightning activity. *Journal of Geophysical Research: Atmospheres*, **121 (23)**, 14–201.

Mazzetti, T. O. and H. E. Fuelberg, 2017: An analysis of total lightning flash rates over Florida. *Journal of Geophysical Research: Atmospheres*, **122 (23)**, 12–812.

McGovern, A., K. L. Elmore, D. J. Gagne, S. E. Haupt, C. D. Karstens, R. Lagerquist, T. Smith, and J. K. Williams, 2017: Using artificial intelligence to improve real-time decision-making for high-impact weather. *Bulletin of the American Meteorological Society*, **98 (10)**, 2073–2090.

Minhas, S., Z. Khanam, S. Ehsan, K. McDonald-Maier, and A. Hernández-Sabaté, 2022: Weather classification by utilizing synthetic data. *Sensors*, **22 (9)**, 3193.

Mishchenko, M. I., L. D. Travis, and D. W. Mackowski, 1996: T-matrix computations of light scattering by nonspherical particles: A review. *Journal of Quantitative Spectroscopy and Radiative Transfer*, **55 (5)**, 535–575.

Mosier, R. M., C. Schumacher, R. E. Orville, and L. D. Carey, 2011: Radar nowcasting of cloud-to-ground lightning over houston, texas. *Weather and Forecasting*, **26 (2)**, 199–212.

Patole, S. M., M. Torlak, D. Wang, and M. Ali, 2017: Automotive radars: A review of signal processing techniques. *IEEE Signal Processing Magazine*, **34 (2)**, 22–35.

Picel, M. H., S. Collis, B. Raut, S. Carani, R. Jackson, M. Van Lier, and A. Fridlind, 2018: TINT—TINT Is Not TITAN. Easy-to-Use Tracking Code Based on TITAN—Details and Uses. *Eighth Symposium on Advances in Modeling and Analysis Using Python*, American Meteorological Society 98th Annual Meeting, Austin, Texas.

Preston, A. D. and H. E. Fuelberg, 2012: The use of polarimetric radar data in determining lightning cessation. *22nd International Lightning Detection Conference*, 45th International LLightning Meteorology Conference, Broomfield, Colorado, USA.

Proctor, F., 1987: The terminal area simulation system. volume 1: Theoretical formulation. Tech. rep., NASA.

Prophet, R., M. Hoffmann, A. Ossowska, W. Malik, C. Sturm, and M. Vossiek, 2018: Pedestrian classification for 79 ghz automotive radar systems. *2018 IEEE Intelligent Vehicles Symposium (IV)*, IEEE, 1265–1270.

Ragonese, E., G. Papotto, C. Nocera, A. Cavarra, and G. Palmisano, 2022: Cmos automotive radar sensors: mm-wave circuit design challenges. *IEEE Transactions on Circuits and Systems II: Express Briefs*.

Rambach, K. and B. Yang, 2017: Mimo radar: time division multiplexing vs. code division multiplexing.

Rao, S., 2018: Mimo radar - application report. Texas Instrument.

Ratvasky, T., D. Duchanoy, J.-F. Bourdinot, S. Harrah, W. Strapp, A. Schwarzenboeck, F. Dezitter, and A. Grandin, 2015: Preliminary flight deck observations during flight in high ice water content conditions.

Ratvasky, T. P., et al., 2019: Summary of the high ice water content (hiwc) radar flight campaigns. *International Conference on Icing of Aircraft, Engines, and Structures*, GRC-E-DAA-TN66897.

Richards, M. A., 2005: *Fundamentals of radar signal processing*. Tata McGraw-Hill Education.

RTCA, 2019: Sc-230, airborne weather detection systems: https://www.rtca.org/sc-230/. URL `https://www.rtca.org/sc-230/`, URL `https://www.rtca.org/sc-230/`.

Ryzhkov, A. V. and D. S. Zrnic, 2019: *Radar polarimetry for weather observations*. Springer.

Ryzhkov, A. V., D. S. Zrnic, and B. Gordon, 1998: Polarimetric method for ice water content determination. *Journal of Applied Meteorology*, **37 (2)**, 125–134.

Schreurs, D. and M. O'Droma, 2009: *RF Power Amplifier Behaviorial Modeling*. Cambridge University Press, New York, 289 pp, 14-15 pp.

Schultz, C. J., W. A. Petersen, and L. D. Carey, 2009: Preliminary development and evaluation of lightning jump algorithms for the real-time detection of severe weather. *Journal of Applied Meteorology and Climatology*, **48 (12)**, 2543–2563.

Shrestha, Y., Y. Zhang, J. Fusselman, G. M. McFarquhar, W. Blake, and S. D. Harrah, 2022: Potential application of paradox (polarimetric airborne radar operating at x-band) to high ice water content (hiwc) monitoring. *2022 IEEE Radar Conference (RadarConf22)*, IEEE, 1–6.

Smolyakov, V., 2017: Ensemble learning to improve machine learning results. *Stats and Bots*.

Steiner, M., W. Deierling, and D. Johnson, 2012: Lightning safety at airports - material for thunder. Proceedings in 3rd WMO/WWRP International Symposium on Nowcasting and Very Short Range Forecasting. Rio de Janeiro, Brazil.

Stolzenburg, M., W. D. Rust, B. F. Smull, and T. C. Marshall, 1998: Electrical structure in thunderstorm convective regions: 1. mesoscale convective systems. *Journal of Geophysical Research: Atmospheres*, **103 (D12)**, 14 059–14 078.

Strapp, J., et al., 2016a: The high ice water content study of deep convective clouds: Report on science and technical plan. *Tech. Rep. DOT/FAA/TC-14/31*.

Strapp, J., et al., 2019: An assessment of cloud total water content and particle size from flight test campaign measurements in high ice water content, mixed phase/ice crystal icing conditions: Primary in-situ measurements. *Rapp. Tech. Contract DTFACT-15-C-00011, Tasks*, **6 (6.3)**.

Strapp, J. W., L. E. Lilie, T. P. Ratvasky, C. R. Davison, and C. Dumont, 2016b: Isokinetic twc evaporator probe: Development of the ikp2 and performance testing for the haic-hiwc darwin 2014 and cayenne field campaigns. *8th AIAA Atmospheric and Space Environments Conference*, 4059.

Ulaby, F. T., et al., 2014: *Microwave radar and radiometric remote sensing*, Vol. 4. University of Michigan Press Ann Arbor, MI, USA.

Waldschmidt, C., J. Hasch, and W. Menzel, 2021: Automotive radar—from first efforts to future systems. *IEEE Journal of Microwaves*, **1 (1)**, 135–148.

Wang, J., S. Zhou, B. Yang, X. Meng, and B. Zhou, 2016: Nowcasting cloud-to-ground lightning over nanjing area using s-band dual-polarization doppler radar. *Atmospheric Research*, **178**, 55–64.

Wang, S., Z. Li, and Y. Zhang, 2012: Application of optimized filters to two-dimensional sidelobe mitigation in meteorological radar sensing. *IEEE Geoscience and Remote Sensing Letters*, **9 (4)**, 778–782, doi:10.1109/LGRS.2011.2181484.

Wiens, K. C., S. A. Rutledge, and S. A. Tessendorf, 2005: The 29 june 2000 supercell observed during steps. part ii: Lightning and charge structure. *Journal of the atmospheric sciences*, **62 (12)**, 4151–4177.

Williams, E., M. Weber, and R. Orville, 1989: The relationship between lightning type and convective state of thunderclouds. *Journal of Geophysical Research: Atmospheres*, **94 (D11)**, 13 213–13 220.

Woodard, C., L. Carey, W. Petersen, and W. P. Roeder, 2012: Operational utility of dual-polarization variables in lightning initiation forecasting. *Electronic Journal of Operational Meteorology*, **13 (6)**, 79–102.

Xu, W., E. J. Zipser, C. Liu, and H. Jiang, 2010: On the relationships between lightning frequency and thundercloud parameters of regional precipitation systems. *Journal of Geophysical Research: Atmospheres*, **115 (D12)**.

Zhang, G., 2016: *Weather Radar Polarimetry*. 1st ed., CRC Press, Inc., USA.

Zwanetski, A., M. Kronauge, and H. Rohling, 2013: Waveform design for fmcw mimo radar based on frequency division. *2013 14th International Radar Symposium (IRS)*, IEEE, Vol. 1, 89–94.

Appendix A - List Of Symbols

c	Propagation Speed
ΔR	Range Resolution
$\Delta\Theta$	Angular Resolution
Δv	Velocity Resolution
K_{dp}	Specific Differential Reflectivity
λ	Wavelength
P_r	Radar Receiver Power
P_t	Radar Transmit Power
Φ_{DP}	Differential Phase Shift
r	Range to the target
ρ	Pearson's Correlation Coefficient
R_a	Maximum Unambiguous range
r_a	Range ambiguity
$\hat{R}$	Auto-Correation Function
ρ_{HV}	Correlation Coefficient
Rx	Receiver
Tc	Time period between 2 FMCW chirps
Ts	Pulse Repetition Frequency
Tx	Transmitter
v_a	Maximum Unambiguous Velocity
Z_{DR}	Differential Reflectivity

Appendix B - List Of Acronyms and Abbreviations

2D-S	2D-Stereo
ACF	Auto-Correlation Function
ADC	Analog to Digital Converter
ADWRS	Airborne Doppler Weather Radar Simulation
AI	Artificial Intelligence
ANN	Artificial Neural Network
AOA	Angle of Arrival
AR	Axial Ratio
BPM	Binary Phased Modulation
CG	Cloud to Ground
CINRAD	China New Generation Weather Radar
CMOS	Complementary Metal Oxide Semiconductor
CNN	Convolutional Neural Network
CNN	Convolutional Neural Network
DP	Dual-Polarization
ELA	Ensemble Learning Approach
FAA	Federal Aviation Administration
FDM	Frequency Division Multiplexing
FFT	Fast Fourier Transform
FMCW	Frequency Modulated Continuous Waveform
FOV	Field of View
HIWC	High Ice Water Content

HKO	HongKong Observatory
HPA	High Power Amplifier
IC	Intra-Cloud
IQ	In-phase and Quad-phase
IWC	Ice Water Content
LLIS	Lightning Location Information System
MIMO	Multiple Input Multiple Output
ML	Machine Learning
mmWave	Millimeter wave
MPE	Mean Percentage Error
NASA	National Aeronautics and Space Administration
NEXRAD	Next Generation Radar
NI	Non-Inductive
NN	Neural Network
NWP	Numeric Weather Prediction
PARADOX	Polarimetric Airborne Radar Operating at X-band
PASIM	Phased Array radar SIMulator
PDF	Probability Density Functions
PIP	Precipitation and Imaging Probe
PPI	Plan Position Indicator
PPP	Pulse Pair Processing
PSD	Particle Size Distribution
ReLU	Rectified Linear Unit
RHI	Range Height Indicator
RIWC	Radar Estimated IWC
RTCA	Radio Technical Commission for Aeronautics

SAT	Static Air Temperature
SCMC	Single Cell Monte Carlo
STALO	Stable Local Oscillator
STAR	Simultaneous Transmission and Reception
STFT	Short-Time Fourier Transform
SW	Spectrum Width
TASS	Terminal Area Simulation System
TAT	Total Air Temperature
TDM	Time Division Multiplexing
TI	Texas Instrument
TINT	TINT Is Not TITAN
TITAN	Thunderstorm Identification, Tracking, Analysis, and Nowcasting
VCP	Volume Coverage Pattern