# The Classification and Simulation of Precipitating Convective Regimes over Darwin, Australia

by

Simon Caine

A Dissertation Submitted in

Fulfillment of the

Requirements for the Degree of

## Doctor of Philosophy

in

## Applied and Computational Mathematics

at

Monash University, Melbourne

December 2009

# General Declaration

**Monash University**
**Monash Research Graduate School**

## Declaration for thesis based or partially based on conjointly published or unpublished work

## General Declaration

In accordance with Monash University Doctorate Regulation 17/ Doctor of Philosophy and Master of Philosophy (MPhil) regulations the following declarations are made:

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes one original paper published in a peer reviewed journal. The core theme of the thesis is precipitation regimes over Darwin, Australia . The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the candidate, working within the School of Mathematical Sciences under the supervision of Steven T. Siems.

In the case of chapter 2 my contribution to the work involved performing all of the analysis and the writing up for publication.

| Thesis chapter | Publication title | Publication status | Nature and extent of candidate's contribution |
|---|---|---|---|
| 2 | Objective Classification of Precipitating Convective Regimes Using a Weather Radar in Darwin, Australia | Published | 90% |

I have renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Signed: ..... █████████ .........................................

Date: .........7|9|09.................

# Copyright Notices

## Notice 1

## Notice 2

# The Classification and Simulation of Precipitating Convective Regimes over Darwin, Australia

by

Simon Caine

A Dissertation Submitted in

Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in

Applied and Computational Mathematics

at

Monash University, Melbourne

December 2009

_____

Major Professor                                                          Date


_____

Graduate School Approval                                           Date

# Abstract

The regime nature of tropical convection occurring over Darwin Australia is explored in an attempt to condense the large variety of cloud types that occur over a tropical region in to a discrete and manageable number. The regimes are defined by their precipitation structure as observed by a CPOL radar and their links with known features of the tropical atmosphere are investigated.

Numerical simulations of tropical convection during the Tropical Warm Pool International Cloud Experiment were made using the Weather Research and Forecasting (WRF) model. The data used to provide the initial and boundary conditions for the model simulations was found to contain a large warm bias in the upper troposphere that detrimentally affected the simulated convection. The simulations were then evaluated against observations from a CPOL radar where it was found that the choice of microphysics scheme had a large impact on the quality of the simulations. One of the microphysics schemes used was found to have a significant problem simulating the precipitation coverage below the freezing level, while the other overestimated graupel coverage and underestimated snow.

The precipitation regimes previously defined were used to evaluate the model simulations. It was found that by themselves the regimes were of limited use. The technique used to originally define the regimes was then extended to include model and radar data, which proved to be a more useful (and objective) method for evaluating the statistical representation of precipitation in the simulations. It was found that the biggest problem with the WRF simulations was the representation of weak convective time periods. The second biggest problem with the WRF model is its ability to simulate periods of strong deep convection with large coverage of strati-

form precipitation, this was attributed to the incorrect forcing data.

This thesis is dedicated to Paul Caine, Janice Caine, and Sally Caine. Thank you for teaching me the value of knowledge and giving me the desire and ability to seek some for myself.

# Table of Contents

# List of Figures

ix

# LIST OF TABLES

# ACKNOWLEDGEMENTS

# Chapter 1

# Introduction

## 1.1 Tropical Convection

Tropical rainfall is largely controlled by convective processes. At the local scale tropical convection directly affects the environment though the generation of precipitation, providing the annual water supply for growing crops and human consumption for millions of people living in tropical regions. On much larger scales, numerous studies have shown (Orlanski, 1975; Rickenbach and Rutledge, 1998; Nesbitt et al., 2000, 2006; Houze et al., 1981; Houze and Churchill, 1984, 1987; Mapes and Houze, 1993; Mapes, 1993; Houze et al., 1980) that convection is often organized, varying from the local to meso-scale (e.g., topographically driven convection) to the large-scale in convectively coupled tropical waves (e.g., the Madden-Julian Oscillation). Tropical convection also affects the environment indirectly; detrainment from tropical convection leads to the formation of stratiform and cirriform clouds (Chou and Neelin, 1999). These clouds have much larger horizontal dimensions and lifetimes than the initial convection that lead to their generation, consequently they can have a significant impact on a region's radiation budget. At the global scale, deep convection is the primary mechanism for transporting water vapor and energy from the tropical boundary layer to the extra-tropics via the general circulation. The latent heat release associated with the stratiform component of tropical convection is likely to play an important role in the interaction of convection with large-scale dynamical features of the tropical atmosphere (Lin et al., 2004; Schumacher et al., 2004; Mapes and Lin, 2005; Tao et al., 2006).

While tropical convection directly and indirectly affects the environment from the local to the global scale, accurately representing tropical convection in numerical simulations has been long recognized as a difficult challenge. Before high resolution cloud resolving models were developed the only method for incorporating tropical convection in numerical models was via convective parameterisation (and it still is in many global, climate and large-scale models). Convective parameterisation attempts to incorporate the effects convection (water transport, energy transport, latent heat release etc) in models that are unable to resolve convection explicitly (due to low resolution of the numerical model). Although pameterisation has been shown to be a useful tool in numerical weather simulations (Arakawa, 2004), the complex nature of convection means that any given pameterisation will not work optimally under all meteorological conditions, and may not include important physical processes and feedback mechanisms.

One of the major problems with understanding, modeling, or parameterizing tropical clouds is that they come in many different forms. There are arguably infinite possible variations in the meteorological conditions that can lead to an infinite variation in clouds types (no two clouds are the same). In most cases these differences are likely not important to our understanding of tropical clouds or our ability to make accurate numerical predictions. The trick is to know when differences between clouds are important and when they are superficial. One methodology often employed to deal with this overwhelming complexity is define a number of regimes that contain or encompass clouds with similar properties. Previous research has shown that the wet season in Northern Australia (roughly November to March) can be roughly split into three regimes: the transition period, the monsoon period and the break period. The aptly named transition period describes a time of transition between the wet and dry seasons, typically between November - December (dry to wet) and Feburary-March (wet to dry). The active monsoon is defined by a deep,

westerly, lower tropospheric flow. Convection during the active monsoon often resembles that occurring over tropical oceans, producing large regions of stratiform precipitation and widespread rainfall (Keenan and Carbone, 1992; May et al., 2008). Within any given monsoon period the low level zonal winds temporarily revert to an easterly direction, these periods are referred to as break conditions. Convection during break conditions is typical of that occurring over continental areas, it is generally smaller in scale, more intense (compared with the active monsoon) and has a pronounced diurnal cycle with maximum convective activity occurring in the afternoon (Keenan and Carbone, 1992).

While the monsoon and break regimes have proven useful when studying convection over Northern Australia, an arguable downside of these regimes is that they are rather broad, all clouds are encompassed by only two regimes. Furthermore, the regimes are defined by the large-scale environment (lower tropospheric wind direction) rather than the properties of the clouds themselves. With the intent of objectively defining a number of regimes based on the properties of clouds, Jakob and Tselioudis (2003) and Rossow et al. (2005) used a simple clustering algorithm applied to data from the International Satellite Cloud Climatology Project (ISCCP) (Schiffer and Rossow, 1983) to show that tropical cloudiness on a 280x280 km$^2$ scale (the size of the ISCCP grid box) appears to be organized into six major regimes - three convectively active (and hence likely precipitating) and three suppressed. By their very nature the regimes identified in this fashion are based on the radiative signature of the tropical cloud systems, and further work (Jakob et al., 2005) established that the ISCCP-based regimes show distinct signatures in their radiative and thermodynamic properties. However, due to a lack of suitable data no conclusions on the precipitation characteristics of the ISCCP-based regimes could be drawn. More recently (Pope et al., 2009) applied a clustering algorithm to 49 seasons of radiosonde data from Darwin airport to define five regimes that occur over North

Australia during the wet season. The large-scale precipitation patterns associated with these thermodynamic regimes were then explored by linking each regime with the GPCP version 2 rainfall dataset (Adler et al., 2003).

Based on the success of these studies it is valid to ask the question if a similar regime character can be found using precipitation as the primary variable for discriminating between regimes. Previously Boccippio et al. (2005) used clustering analysis on vertical profiles derived from the Tropical Rainfall Measuring Mission (TRMM) precipitation radar to define a number of archetypal vertical structures for the entire tropics. The profiles used by Boccippio et al. (2005) were essentially vertical columns with a horizontal diameter of 4km, consequently very little spatial information entered into their regime definition.

An ideal tool to observe the spatial structure of convective cloud systems is the centimeter wavelength radar. Ground-based radars are routinely deployed by weather services worldwide (Demott and Rutledge, 1998; Kawashima et al., 2006) for research and forecasting purposes. The Australian Bureau of Meteorology has operated a polarimetric research C-band radar (C-Pol) (Keenan et al., 1998; May et al., 1999; Carey and Rutledge, 2000), in Darwin (Australia) since 1995. The combination of the availability of many years of radar data, and the ability to observe a variety of convective systems makes Darwin an ideal location to study the possible regime character of tropical precipitation on scales of a few hundred kilometers. The research of defining, describing, and comparing the regimes to known features of the tropical atmosphere is shown in chapter 2.

In addition to exploring the possible regime nature of convection in the Darwin region, this thesis will evaluate the Weather Research and Forecasting (WRF) model in a tropical environment. Evaluation of the WRF model to date has focused primarily on the extra-tropics (Cheng and Steenburgh, 2005; Done et al., 2004; Jankov et al., 2005). Where evaluation has taken place in the tropics, much of the research

has focused on tropical cyclones (Pattanayak and Mohanty, 2006; Rogers et al., 2007).

Darwin is the logical place to perform simulations in this study as it has a rich history of hosting intensive field campaigns, observational data gathered from field campaigns is extremely useful when evaluating a numerical model. During the 1986-1987 wet season the Australian Monsoon Experiment (AMEX) was run by the Australian Bureau of Meteorology Research Centre (BMRC) with "the aim of improving the understanding of the way that cumulonimbus convection and tropical cloud clusters interact with the larger-scale circulation; of air-sea interactions in the tropics; and of specific weather phenomena in the north Australian region." (Holland et al., 1986). The Marine Continental Thunderstorm Experiment (MCTEX) was held in between November and December 1995 with the aim of investigating the life cycle of island based thunderstorms that occur over the Tiwi Islands located north of Darwin (Keenan et al., 2000). The Darwin Area Wave Experiment (DAWEX) was held between October and December of 2001; the aim of this field campaign was to "study aspects of the atmospheric dynamics from the ground to the lower thermosphere, focusing on wave perturbations in the middle atmosphere in conjunction with detailed observations of tropical moist convection" (Hamilton et al., 2004).

The most recent field experiment to occur in Darwin was the Tropical Warm Pool - International Cloud Experiment (TWP-ICE), which ran between the 19th January and the 15 February 2006. TWP-ICE was a collaborative endeavor between the Australian Bureau of Meteorology (BoM), the U.S. Department of Energy - Atmospheric Radiation Measurement program (ARM) and a number of universities and research organizations worldwide. One of the primary aims of the experiment was "to examine convective cloud systems from their initial stages through to the decaying and thin high level cirrus and measure their impact on the environment"

(May et al., 2008). Over the course of the experiment radio sondes were launched every 3 hours from five ground-based sites and the Southern Serveyor, a research vessel patrolling the ocean around Darwin. Two radars were operating over the course of the experiment taking measurements of the cloud/precipitaiton structure and a number of flux sites were operating in the region. In addition to the ground-based observations a number of research aircraft were taking in-situ microphysical measurements.

As the WRF simulations presented in this thesis simulate convection that occurred during TWP-ICE, a brief overview of the synoptic conditions during the experiment is presented below, for further information about TWP-ICE a full description may be found in May et al. (2008). TWP-ICE began on the 19th of January 2006. During the first few days Darwin experienced typical monsoon conditions, with deep westerly winds, widespread convection and high cloud coverage. On approximately the 23rd of January a Mesoscale Convective System (MCS) developed in the region (May et al., 2008) dominating the local winds and precipitation. Between the 25th of January and the 4th of February a low pressure system to the south of Darwin (previously associated with the MCS) had a large effect on the weather in the TWP-ICE domain. The geostrophic winds about the low pressure system wrapped air from over the Australian continent around and into the experimental domain from the west (the classical definition of the Australian monsoon (Drosdowsky, 1996)). However, because the air-mass had continental origins it contained relatively low quantities of water vapor, leading to suppressed monsoon conditions with low cloud tops and relatively little precipitation. Towards the end of the experiment (approximately the 5th of February) Darwin experienced typical monsoon break conditions, with easterly winds and deep but relatively isolated convection, these conditions continued until the end of the experiment on the 12th Febuary.

Thus the TWP-ICE can be roughly split into 3 regimes or phases, the first being the active monsoon which occurred roughly between the 19th and 24th of January, the second being the suppressed monsoon between the 25th of January and the 4th of February and final phase between the 5th and the 12th of February where the region experienced typical monsoon break conditions. WRF simulations were chosen to coincide with the TWP-ICE field campaign to take advantage of the observational data collected during a variety meteorological conditions.
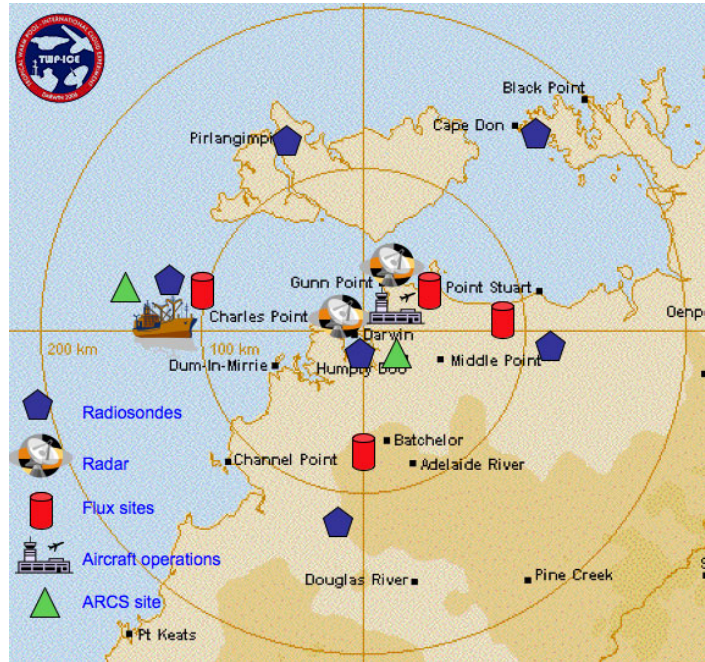


Figure 1.1: **TWP-ICE experimental setup**

## 1.2  Science Questions

The first aim of this thesis is to determine if rainfall over Darwin naturally divides into regimes. An attempt is made to define the regimes in an objective fashion, to base the regimes solely on the precipitation structures of clouds as seen by a CPOL radar, and to assume no a priori knowledge of the previously defined Monsoon and Break regimes. The research of defining, describing and comparing the regimes to

known features of tropical convection is shown in chapter 2.

A secondary aim of this thesis is to evaluate the performance of the WRF model in simulating the wide variety of clouds present in a typical wet season. The use of radar data to validate the WRF model will feature strongly in the analysis, and the models ability to simulate the four precipitation regimes defined in chapter 2 will be investigated. Further aims of this thesis are to explore additional ways that radar data can be used to evaluate the WRF model. As polarization data from the CPOL radar was not used in the definition of the precipitation regimes, an attempt to utilize the polarization information for model evaluation purposes will be shown in chapter 6.

# Chapter 2

# Precipitation Regimes

## 2.1   Introduction

The aim of this chapter is to determine if a number of regimes can be objectively defined from radar data and used to describe the rainfall/precipitation properties of convective systems occurring over long periods of time (e.g. seasons), for the region surrounding Darwin. In order to establish if tropical rainfall exhibits such regime character this study applies a cluster algorithm similar to that used in Jakob and Tselioudis (2003) and Rossow et al. (2005) to Frequency with Altitude diagrams (FADs) (Yuter and Houze, 1995; Neiman et al., 2005) of radar reflectivity derived from four wet seasons of radar data. The properties of the regimes identified by this algorithm are investigated and their relationship to the main circulation regime in tropical Australia - the monsoon - is established. The relationship of the radar-based regimes to the ISCCP-based tropical cloud regimes is also investigated.

## 2.2   Data Description and Method

The dataset used in this study is comprised of reflectivity values observed by the scanning C-band polarimetric (C-Pol) radar located near Darwin (Keenan et al., 1998). Figure 2.1 shows a map of the radar domain indicating land areas (gray) and the maximum coverage of the radar footprint (achieved at 3 km in the vertical) used in this study. The radar performs a full volume scan every ten minutes. Given the aim of studying possible precipitation regimes as they occur over long periods of

time it was necessary (because of computational limitations) to reduce the volume of information by using only one scan per hour. The hourly dataset used in this investigation is comprised of the four wet seasons (NDJFMA) of 1999/2000, 2001/02, 2002/03 and 2003/04, bringing the total number of sampled radar volumes to just over 13500.
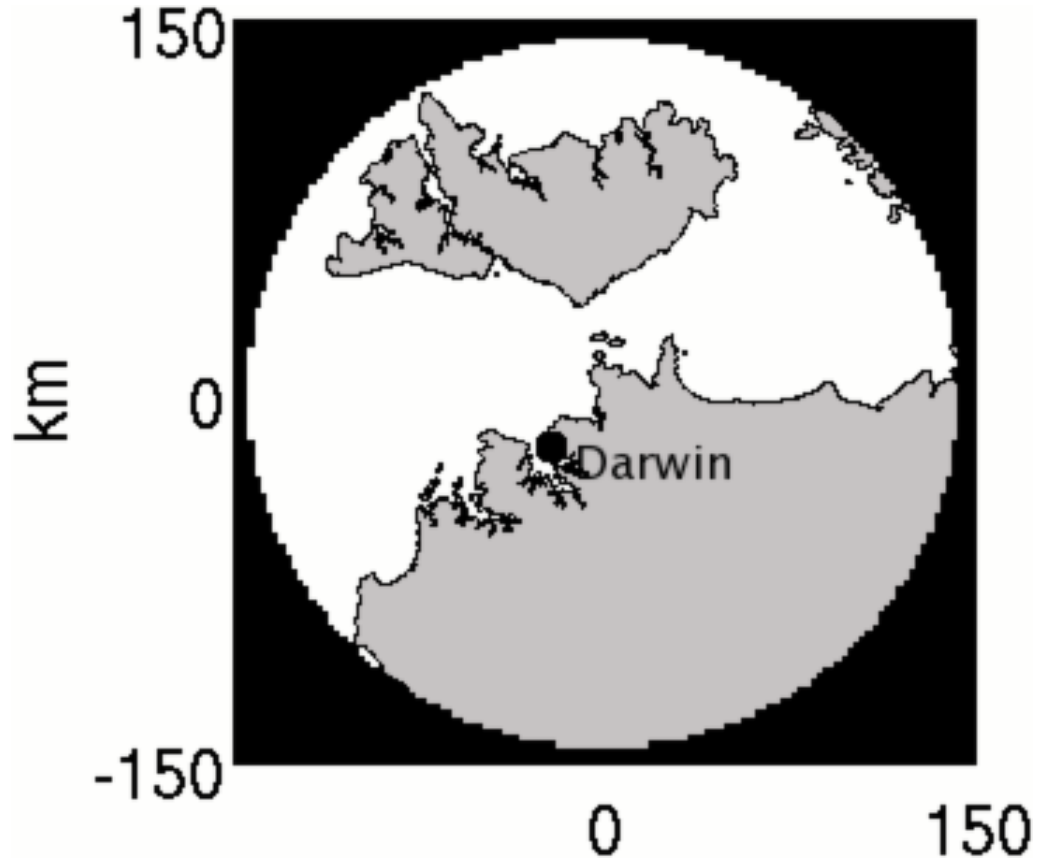


Figure 2.1: **Constant-Altitude Plan Position Indicator (CAPPI) showing the maximum coverage of the Gunn Point radar, near Darwin, Australia. CAPPI is valid at 3 km in the vertical. Land is shaded gray and oceans are white.**

As a first step in calculating frequency with altitude diagrams (FADs - Yuter and Houze (1995)) of radar reflectivity for each of the 13500 sampled volumes, the volume scans were interpolated into a cartesian space consisting of 40 vertical levels with a grid spacing of 500m, and 121 grid points in each horizontal direction with a grid spacing of 2.5 km. The change in the horizontal extent of the radar domain

with height is taken into account by applying masks of the actual observed area at each vertical level. The histograms forming the basis of the FADs were then created by iteratively searching through all valid reflectivity values at each vertical level and calculating the frequency with which these values fall into reflectivity bins of width 2 dBZ. Reflectivity values below 0 dBZ are excluded from the analysis as they are considered at the detection limit of the radar, especially at longer ranges. At each vertical level the relative frequency of occurrence of each reflectivity bin was then calculated by dividing the number in each bin by the total number of observed points defined by the radar mask at that level. Consequently, summation over the frequencies for each vertical level can be thought as a measure for the fractional coverage with radar signal (or precipitation) at that level. This deliberately deviates from the standard normalization of CFADS (the graphical representation of FADS i.e. Contoured Frequency with Altitude Diagrams) (Yuter and Houze, 1995), where each level is normalized by its fractional coverage. This choice was made to maintain area coverage information in the regime analysis that follows, as this provides a crucial distinction between regimes.

For illustrative purposes Figure 2.2 shows the mean FAD over all samples used in this study. An important summary measure used throughout the study is the total volume coverage (TVC) with radar echo. This quantity is calculated by summing the number of radar returns over all levels and reflectivity values then dividing by the total number of observable points. It therefore represents the fractional coverage of radar signal over the entire three dimensional domain. The TVC for the mean histogram is 0.04, indicating that on average approximately 4% of the volume scanned by the radar has a reflectivity value above the 0 dBZ threshold. It can also be seen that, on average, the maximum echo top height reaches approximately 16 km and reflectivity values greater than 40 dBZ are observed up to 8 km. A large fraction of the returns have reflectivity values less than 18 dBZ. As this is the cutoff

threshold for the TRMM radar, Figure 2.2 gives some indication of the number of hydrometeors the TRMM radar will not observe.
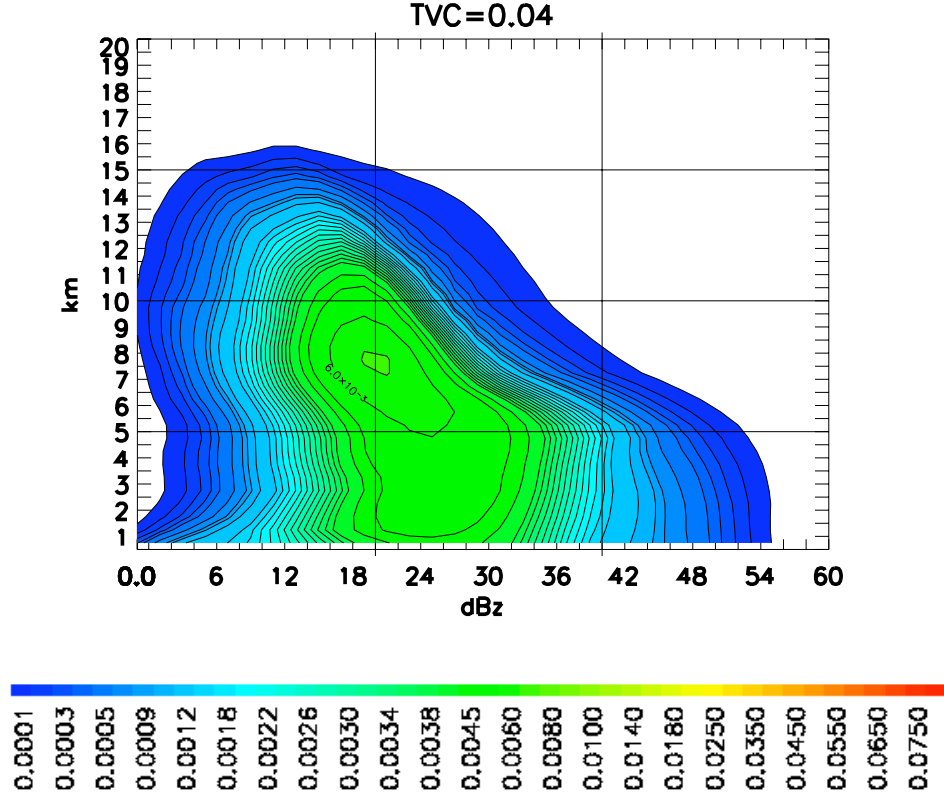


Figure 2.2: **Frequency with altitude diagram (FAD) showing the mean of all time periods used in the study (13508 histograms). The total volume coverage (TVC) is shown at the top of the diagram, indicating that on average 4% of the volume scanned by the radar is covered by hydrometeors.**

Each of the more than 13500 histograms provides an input into the definition of objectively derived regimes. In order to focus the clustering algorithm (see below) onto the major regimes, first all (trivial) null cases, i.e., those with little or no radar return anywhere in the volume were removed from the analysis. For this purpose any histogram which did not have at least 1% coverage in at least one of the vertical levels was defined as no-precipitation histogram and was removed. This eliminated

roughly 5000 histograms from the analysis. While these "no precipitation" cases do not enter the regime definition algorithm, they do nevertheless provide important physical information about the state of the atmosphere. Their high frequency of occurrence and their "trivial" nature justifies their removal to avoid them dominating the outcome of the cluster algorithm.

A K-Means clustering algorithm (Anderberg, 1973) was applied to the remaining histograms to identify possible reoccurring patterns in the FADs that could be indicative of recurring precipitation regimes. The choice of algorithm is largely driven by it's successful application in previous studies (Jakob and Tselioudis, 2003) and efficiency requirements due to the large sample size used here. As discussed in Jakob and Tselioudis (2003) and Rossow et al. (2005) a feature of this particular algorithm is that it requires the user to predefine the number of clusters searched for, $k$. This is overcome by repeating the analysis several times with varying values of $k$. Here $2 \leq k \leq 10$ are used. The basic concept of the iterative K-Means algorithm is as follows:

- randomly choose $k$ "seed vectors" from the dataset as initial cluster centroids,

- assign each input histogram to one cluster by finding the minimum Euclidean distance to any of the cluster centroids,

- calculate new cluster centroids by averaging over all histograms assigned to a cluster,

- use the newly calculated centroids as seed vectors, and

- iterate the algorithm.

With each iteration the difference between the old and new centroids decreases and the clusters are said to be stable when this difference is less than a prescribed threshold value. The histograms associated with each cluster are known to be closer

to the centroid of this cluster than to those of any of the others. Each regime is then characterized by its centroid or mean histogram. The regimes defined by the clustering algorithm should not depend on initial histograms chosen at random and tests were performed to ensure this was the case.

The advantage of using the K-Means algorithm is that it provides an objective method for defining regimes from large sets of data. However, as discussed above, the number of clusters, $k$, searched for in the algorithm needs to be prescribed. It is therefore necessary to define a strategy for choosing the optimal set of clusters from the application of the algorithm to values of $2 \leq k \leq 10$. Here the same "quasi-objective" method as in Rossow et al. (2005) is used. For each increase in cluster number the stability of the solution to the (randomly chosen) initial seeds is investigated. Furthermore, the emerging new cluster in the $k + 1$-analysis is compared to the existing $k$ clusters of the previous application. The optimal number of regimes in this study, $K$, is then defined as the smallest possible number of clusters for which i) the algorithm provides stable solutions with respect to initial seeds and ii) the addition of further clusters leads to regimes similar to the already existing ones. Here it is found that $K = 4$. Note, that since "no-precipitation" histograms have been excluded from the analysis to begin with, the total number of physical regimes defined in this study is five. Nevertheless, the rest of the paper will continue to refer to four precipitation regimes, treating the no-precipitation regime as a "zeroth" regime due to its trivial character. The main features of the four major precipitation regimes identified will now be described. It should be noted a particular storm event or mesoscale convective system (MCS) may and indeed does fall into multiple precipitation regimes. Transition of a storm event or MCS between regimes occurs when the temporal evolution of the event is captured within the domain of the radar. Transition between regimes may also occur when a storm event or MCS enters and leaves the domain of the radar.

## 2.3 Tropical precipitation regimes as identified in radar data

The application of a cluster algorithm to the FADs derived from hourly volume scans using the polarimetric C-Pol radar located near Darwin reveals four major precipitation regimes in that region. To determine the robustness of the precipitation regimes each wet season was analyzed separately. The regimes defined when using individual seasons (not shown) were found to be very similar in structure to the precipitation regimes defined here, although the RFO of the four regimes does vary from year to year.

The mean FADs for each of the four regimes are shown in Figure 2.3. The order of the panels in this figure is in terms of the relative frequency of occurrence (RFO) of each regime, which is shown in the top left-hand corner of each panel. Due to its trivial nature the no-precipitation regime is not plotted, however for completeness the RFO of this "zeroth" regime is displayed on the right hand side of each panel. It is evident that for the dataset used here - hourly data for four wet seasons (ND-JFMA) - this regime has the second highest frequency of occurrence (about 38%), highlighting the well-known fact that even on the scale of the radar footprint used here (about 300 km diameter) there are many times with no precipitation. Most of those occur in the transition season early and late in the wet-season definition used here (see Section 2.4). Another important regime characteristic displayed in Figure 2.3 is the total volume coverage (TVC) of the radar signal introduced in the previous section. The TVC provides a useful first indication for the differences between regimes. TVC increases steadily from 2% for the most frequent precipitation regime to 26% for the least frequent regime, indicating that in the Darwin region the tropical atmosphere exhibits regimes with small precipitation coverage significantly more frequently than those with large precipitation coverage.

Below follows a brief description of each of the regimes displayed in Figure 2.3. In describing the regimes, "low" reflectivities will refer to values between 0 and 20 dBZ, "medium" reflectivities will refer to values between 20 and 40 dBZ and "high" reflectivities will refer to values between 40 and 60 dBZ. Those ranges are indicated in the figures by vertical lines. To further aid conceptualization of the precipitation regimes Constant-Altitude Plan Position Indicators (CAPPI) examples of each regime are shown in Figure 2.4 for both 2.5 and 10 km in the vertical. These specific examples are representative of each regime and were chosen by finding those histograms with the minimum Euclidian distance from the cluster mean histograms displayed in Figure 2.3. These CAPPI examples confirm that the regimes are well differentiated by TVC and distribution of reflectivity values.
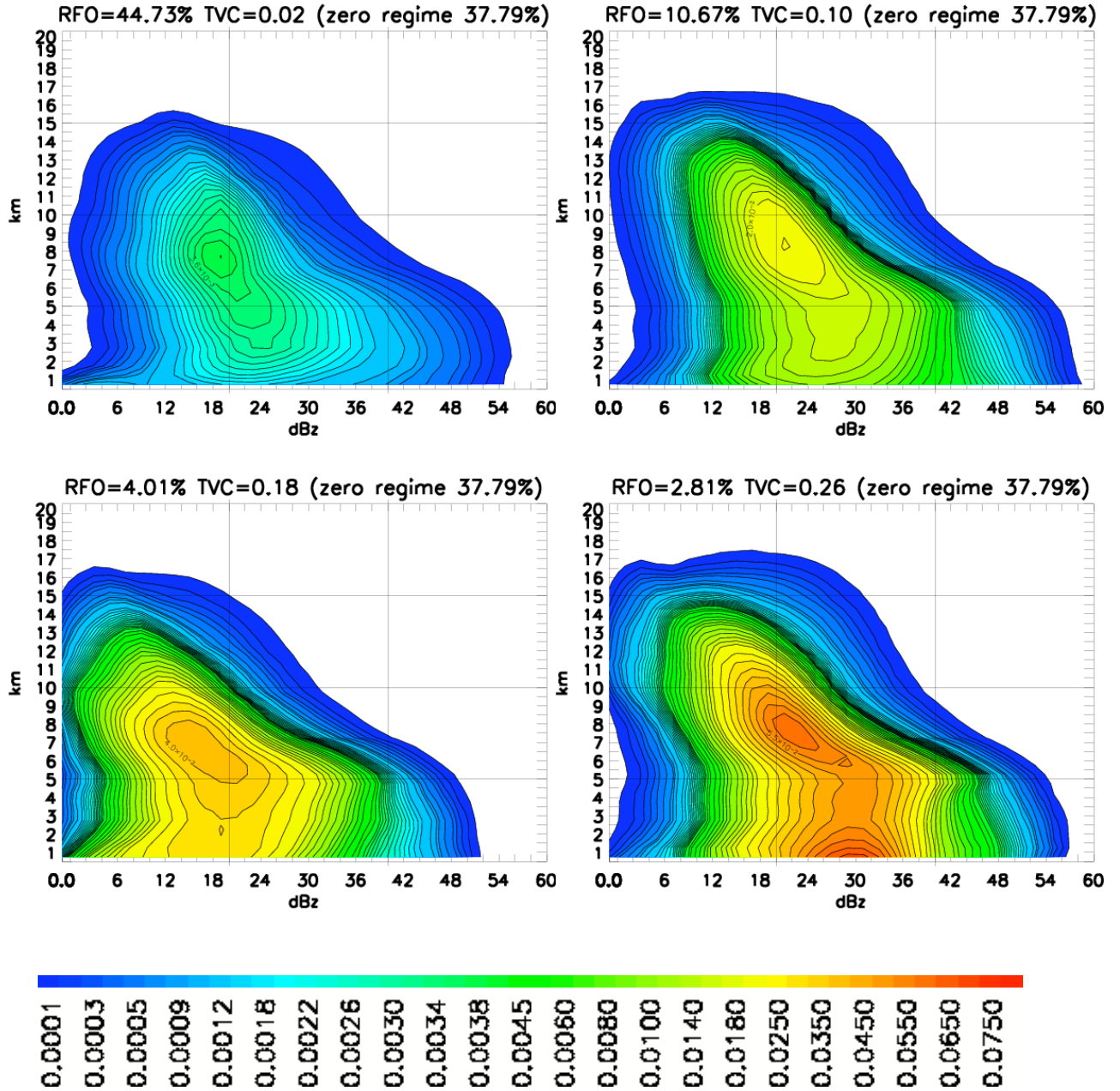
Figure 2.3: **The four precipitation regimes defined by K-means algorithm. Regimes are ordered (most to least) by their relative frequency of occurrence (RFO). The RFO of the "zeroth" regime, the regime containing time periods that have no precipitation over the entire radar domain, is shown for completeness.**
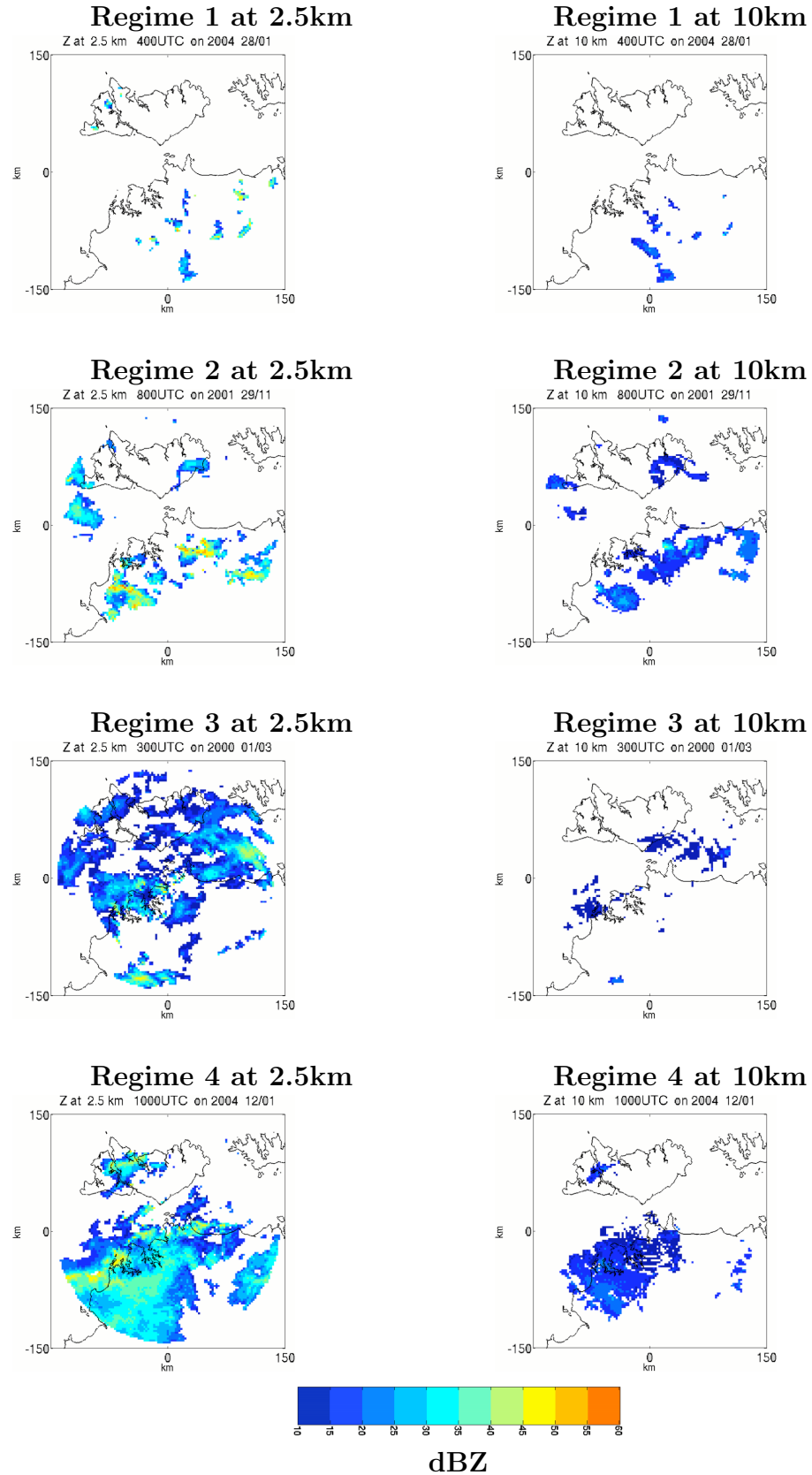
Figure 2.4: **CAPPI examples deemed representative of the four precipitation regimes, examples are given for both 2.5 and 10 km in the vertical.**

Regime 1 is the dominant weather condition occurring over Darwin during the four wet seasons analyzed. It occurs approximately 45% of the time. The precipitation associated with this regime is weak and shallow. It has the lowest frequency of reflectivity values above 40 dBZ and a maximum echo top height 2 km lower than the other regimes. The TVC for this regime is 2%, which is an order of magnitude smaller than the next closest regime. The low TVC value and absence of a bright band at the freezing level (approximately 5km) indicates that the convection has a low spatial/volume coverage and is therefore probably patchy and contains only a small proportion of stratiform precipitation. Average convective and stratiform rain contributions for each regime will be analyzed below to further investigate this hypothesis. On average higher rain rates occur over the continent than over the ocean in this regime (see Figure 2.5). This is an important feature of regime 1 and will be investigated further below.

Regime 2 has a RFO of approximately 10%. The TVC of this regime is an order of magnitude greater than in regime 1 indicating that the precipitation covers a much larger volume of the radar domain. In this regime the maximum echo top height extends to just over 16 km and there is a greater frequency of high reflectivity values at low levels, indicating that the convection is of greater intensity than was found for regime 1. It should be noted that regime 2 actually has the highest reflectivity values of all four regimes. Another defining feature of regime 2 is the increased occurrence of medium reflectivity values (20-40 dBZ) from the lowest level to about 13 km, with the greatest frequency occurring between 6 and 12 km.

As was the case in regime 1, the absence of a strong bright-band signal near 5 km is indicative of precipitation that is likely to be predominately convective, which will be further investigated below.

Regimes 3 has a RFO of approximately 4%. A significant increase can be seen in both TVC and echo coverage at low levels. The low occurrence of high reflectivity

values below 5 km and decreased echo top height indicates that the convective component of this regime is weaker than the convection associated with regime 2. Oceanic convection tends to be shallower and have lower maximum reflectivity than continental convection (Toracinta et al., 2002; Nesbitt et al., 2000), the possibility that regime 3 could be influenced by precipitating systems with oceanic origins will be investigated in section 2.4. The large increase in frequency of low and medium reflectivity values is likely due to the presence of significant areas of stratiform precipitation and would explain the relatively high TVC for this regime, evidence of stratiform precipitation can be seen in the form of a bright band at approximately 5km for medium reflectivity values. However, as the regime is an average of many time periods and contains both stratiform and convective precipitation, the bright band is less pronounced than would normally be the case for purely stratiform states. Further evidence of the stratiform component of this regime will be provided below.

Regime 4 is the rarest of all the regimes and has a RFO of under 3%. In contrast to regime 3 there is a large occurrence of high reflectivity values below 5 km. The maximum echo top height for this regime is higher than in regime 3 and is even slightly higher than was found for regime 2. The spread of reflectivity values at 15 km is larger than in regime 2 and extends to much higher reflectivity values than those found in regime 3. Regime 4 has the largest TVC and by far the greatest frequency of medium reflectivity values, indicating the presence of large amounts of stratiform precipitation with greater vertical extent than the stratiform precipitation seen in regime 3. The bright band in regime 4 is more pronounced than was seen in regime 3 and can be seen up to approximately 46 dBZ. Another defining feature of regime 4 is the large frequency of medium reflectivity at the lowest levels, particularly between reflectivity values of 24 and 36 dBZ.

Figure 2.5 shows the average rain rate in millimeters per hour for the four regimes. For each time period used in the study (approximately 13500 samples)

the 2.5 km reflectivity data was converted to rain rate maps using the relationship $Z = 305R^{1.36}$ (Bringi et al., 2001), where Z is the absolute reflectivity and R is the rain rate in mm/hour. Averaged rain rate maps for each regime were then created by summing the $n$ rain rate maps associated with each regime and dividing by $n$. $R_{total}$ is displayed in the top right hand corner of each figure, calculated by summing the rain rates over the entire radar domain and dividing by the number of valid grid points, thereby giving some measure (spatially averaged) of the rain rate of each regime. The rainfall distribution shown in Figure 2.5 a) shows that on average more rainfall occurs over land than over the ocean for regime 1, this is especially true over the Tiwi Islands. As expected this regime has the lowest rain rate of all the regimes, approximately 0.3 mm per hour averaged over the entire domain.

Figure 2.5 b) supports the assertion that regime 2 contains convection initiated by convergence due to land / sea breezes, a slightly higher rain rate can be seen orientated parallel to the coast line of the mainland. Slightly higher rain rates can also be seen over the ocean in the west of the radar domain. Analysis of the diurnal cycle of rain rate for regime 2 (not shown) helps explain the patterns seen in figure 2.5 b). During the afternoon precipitation occurs preferentially over land areas, while during the night and early morning precipitation occurs over the ocean. The higher rain rates over the ocean in the west of the domain in figure 2.5 b) is due almost exclusively to precipitation occurring at night. Although $R_{total}$ of regime 2 is approximately 4 times that of regime 1 (1.1 compared with 0.3), its RFO is approximately 1/4 times that of regime 1 (approximately 11% for regime 2 compared with 45% for regime 1). Thus regime 1 and regime 2 contribute approximately equal amounts of rainfall to the region.
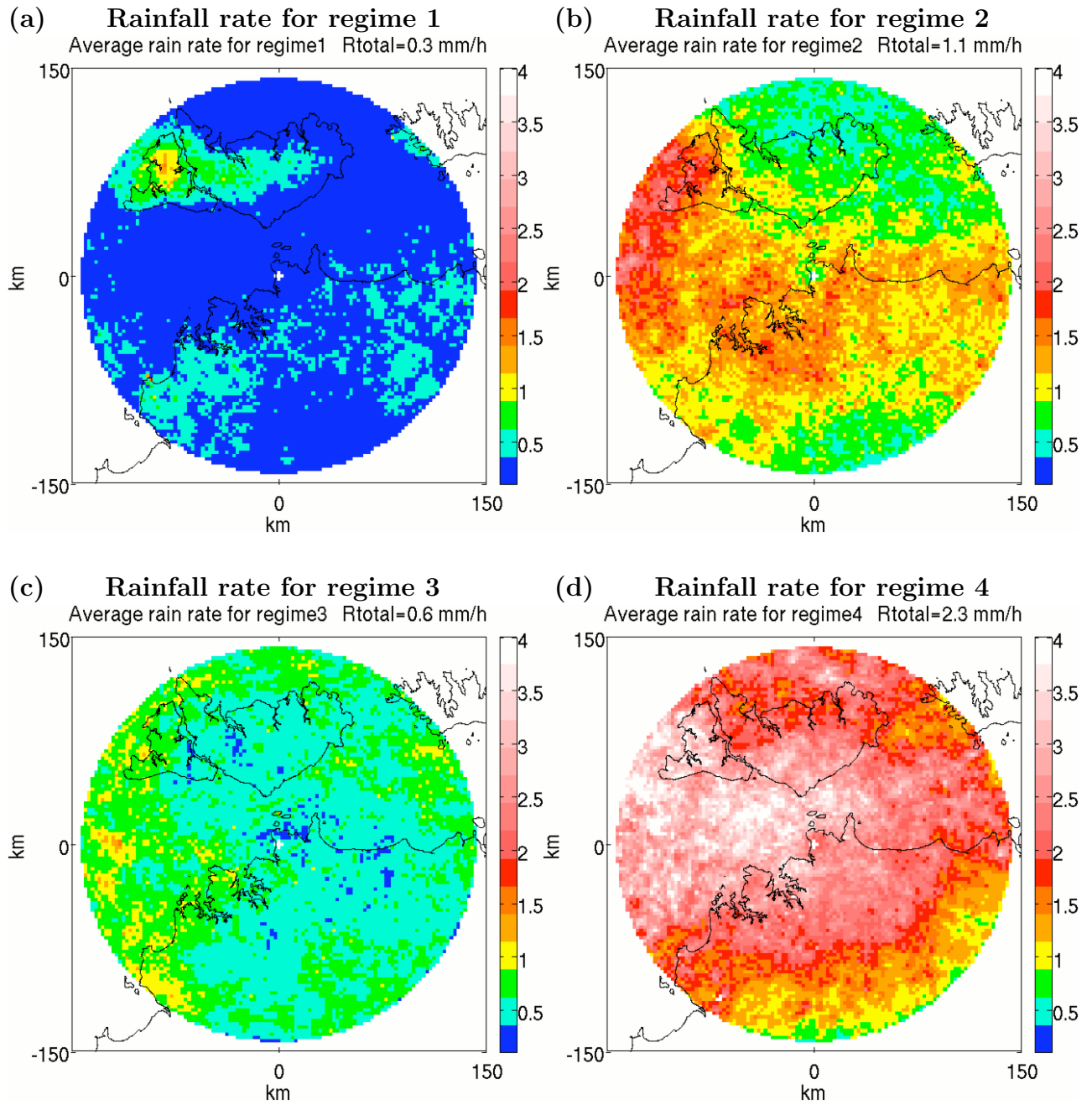
Figure 2.5: **Average rainfall rates for the four precipitation regimes in mm/hour**

|        | Regime 1 | Regime 2 | Regime 3 | Regime 4 |
|--------|----------|----------|----------|----------|
| Land   | 0.3518   | 1.0824   | 0.5844   | 2.0265   |
| Ocean  | 0.2252   | 1.1261   | 0.6656   | 2.6141   |

Table 2.1: **Rain intensity of each Regime split into land and ocean contributions.**

Figure 2.5 c) shows that in regime 3 rainfall is fairly constant throughout the domain of the radar, with perhaps a slight increase in rainfall over the ocean to the west. An interesting feature of this regime is that $R_{total}$ is much lower than was found for regime 2, despite the fact that the volume coverage of hydrometeors is much larger for regime 3 than regime 2. This might be explained by the finding that the convection associated with regime 3 is weaker than the convection associated with regime 2.

Figure 2.5 d) shows that in regime 4 high rainfall rates occur preferentially over the ocean. As expected $R_{total}$ for this regime is the largest of all four regimes, owing to the large amounts of stratiform precipitation and intense convection found within this regime.

To further test the spatial distribution of the precipitation within the regimes a land/ocean mask was applied and the rain intensities over land and ocean were calculated for each regime. The land (ocean) rain intensity is defined as the sum of the rainfall over the land (ocean) divided by the number of land (ocean) points. Table 2.1 shows that the rain intensity is greater over land in regime 1. Regime 2 has approximately the same rain intensity over land and ocean, as discussed above this regime contains contributions from convection occurring over land (during the afternoon) and over the ocean (during the night). Both regimes 3 and 4 have greater rain intensity over the ocean than the land.

Figure 2.6 shows the average rain rate for each regime split into convective and stratiform components. The rain fields measured with the C-Pol radar have been dis-aggregated into convective and stratiform areas using the algorithm described

by Steiner et al. (1995). The only modification to this was to utilise the polarimetric rain estimates obtained with the C-Pol radar using algorithms described by Bringi et al. (2001, 2004). These rain estimates were then converted to an effective reflectivity using a simple Z-R relation (Zeff=$305R^{1.36}$) in order to apply the Steiner et al. algorithms. This had the benefit of the classification being less sensitive to variations in measured reflectivity associated with drop-size distributions, attenuation of the sampled reflectivity, and hail contamination as well as being consistent simple mapping of the observed field. The convective and stratiform rain was then summed over the radar domain (a 300 km diameter circle) to calculate the contributions of convective and stratiform rain areas to the total rain accumulation. Figure 2.6 indicates that regime 1 is predominately a convective regime, with stratiform precipitation contributing 24% to the over all rain rate. Regime 2 is also predominately convective in nature, however the stratiform component of this regime has increased to 35%. Regime 3 contains the most stratiform precipitation relative to convective precipitation, 64% of the precipitation in regime 3 is stratiform in nature. Regime 4 also contains a large proportion of stratiform precipitation (53%), however the convective component of this regime is quite large, almost as high as the total rain rate for regime 2.

In summary, analysis of the CFAD's and convective/stratiform rain rates leads to the conclusion that regime 1 contains shallow, patchy and relatively weak convection. Furthermore, the average rain rate map for regime 1 and land/ocean rain intensities indicates that the convection associated with regime 1 preferentially occurs over land areas. Regime 2 was found to be a relatively strong convective regime that precipitates preferentially over the land during the day and over the ocean at night. The convection in this regime has greater intensity and scale than the convection in regime 1, with a greater proportion of the convection likely initiated by land/sea breezes. The convective nature of this regime was confirmed by the relatively large
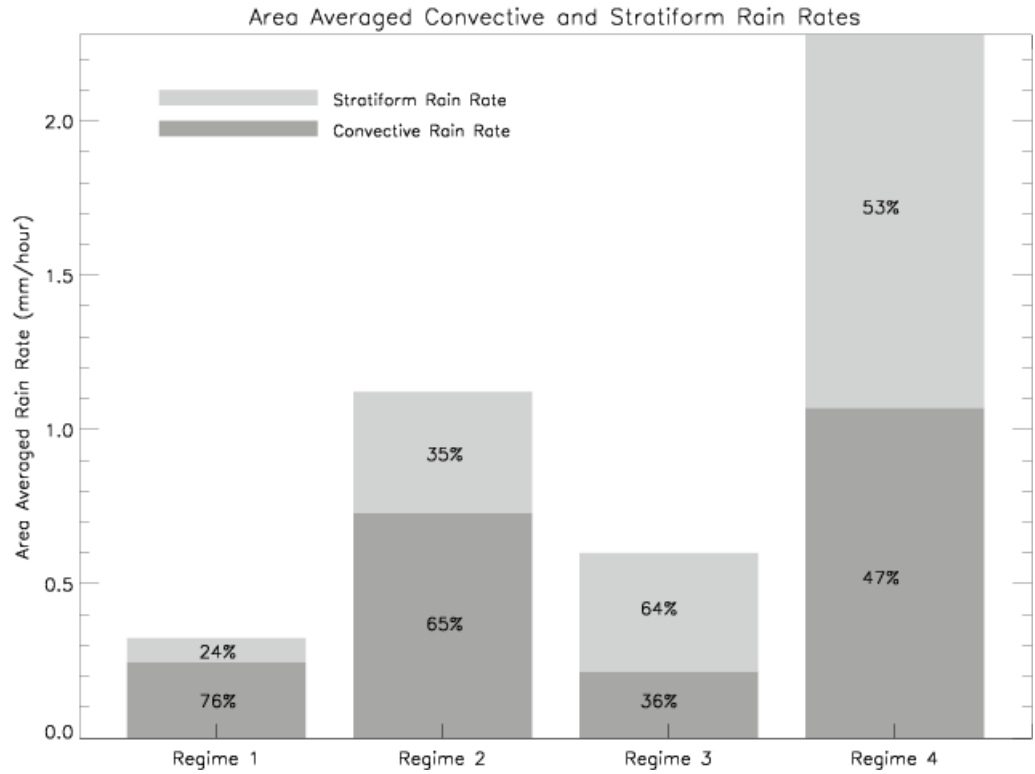
Figure 2.6: **Area Averaged stratiform and convective rain rates for the four precipitation regimes.**

proportion of convective rain as seen in figure 2.6 and absence of a bright band signature in Figure 2.3. Regime 3 was found to contain a mixture of relatively weak convection and large amounts of stratiform precipitation. The temporal evolution of regimes (not shown), indicates that regime 3 is likely to occur from maturing systems associated with regime 2 and 4. Regime 4 was found to be a mixture of relatively strong convection with large amounts of stratiform precipitation.

Having discovered some of the basic characteristics of the regimes the next section aims to shed further light on their characteristics by studying their seasonal and diurnal cycles, and their link to the monsoon.

## 2.4 Some characteristics of the precipitation regimes

The previous section has objectively identified four (five if the no precipitation is included) major convective regimes in the Darwin region. The four non-trivial regimes show characteristic distributions of radar reflectivity both in the horizontal and vertical on the scale of a typical weather radar footprint. While it was possible to link the regimes to known features of tropical convection, their link to large-scale features of the tropics as they affect the Darwin region remain to be elucidated. This section is aimed at establishing links between the regimes identified above and known features of the tropical atmosphere. First their response to the main forced modes of the atmosphere, namely the seasonal and diurnal cycles are investigated. This is followed by establishing the regime characteristics as a function of monsoon activity in the Darwin region.

### 2.4.1 The seasonal cycle

In the Darwin region the seasonal mode is dominated by the Australian summer monsoon, which dominates the region's weather usually from the second half of December into early March (Drosdowsky, 1996). Active monsoon periods are interlaced with monsoon breaks. The monsoon period is bordered by a build-up and decay period on the order of 1-2 months each, while the rest of the year has a distinct dry-season character.

Here the signature of the seasonal cycle on the convective regimes is investigated by studying the seasonal variation of their frequency of occurrence. Figure 2.7 shows the relative frequency of occurrence of the five regimes (including the no-precipitation regime) as a function of month. The no-precipitation regime (Regime 0) shows the expected seasonal cycle with a minimum in January, the month with the largest influence of the Australian summer monsoon, and maxima in the months of March and April, those in which the Darwin region transits into the dry season. The

RFO of the patchy convection regime (Regime 1), is remarkably constant through-out November through February with a decline in occurrence in March and April. The stronger convective regimes (Regimes 2-4) show an increase in their RFO from November to January, a decline in February and are virtually absent in March and April. The most notable change from November to January occurs for regimes 3 and 4, confirming the impression gained in the previous section that these regimes probably have a more monsoonal character (Steiner et al., 1995). This provides a further indication that the objectively identified precipitation regimes may repre-sent physically relevant states of the atmosphere around Darwin, a claim that will be further investigated below.
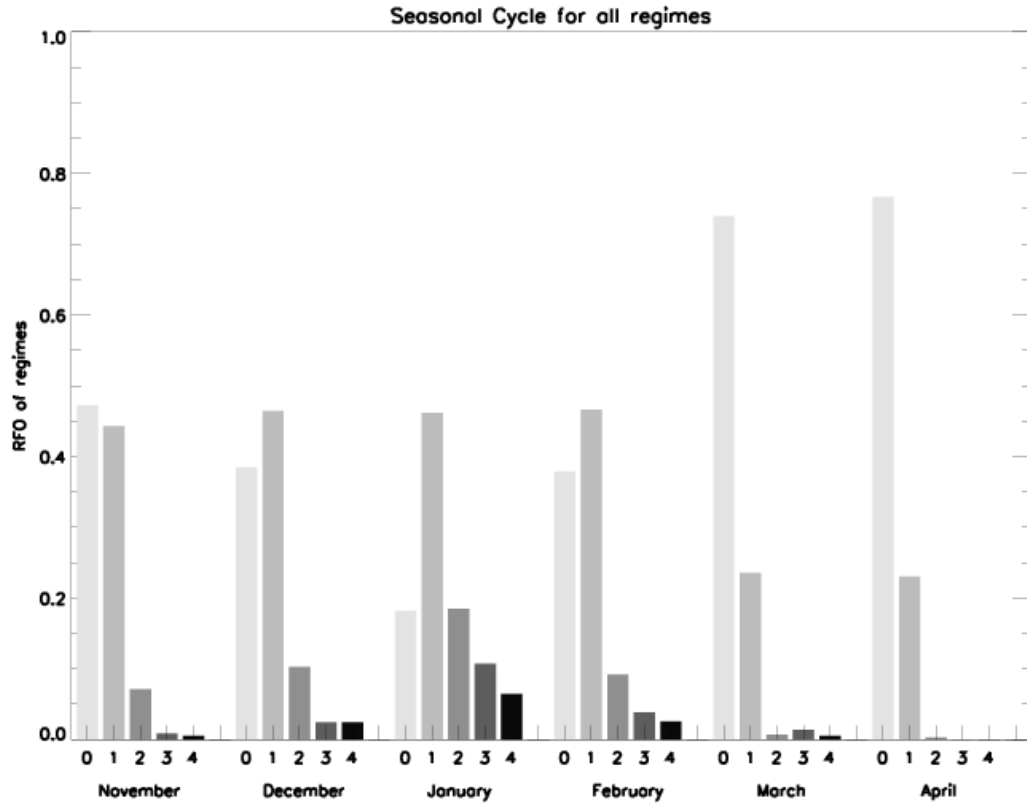


Figure 2.7: **Seasonal cycle of the precipitation regimes showing the relative frequency of occurrence of each regime, for the months November to April.**

### 2.4.2 The diurnal cycle

Figure 2.8 shows the diurnal cycle of the relative frequency of occurrence (RFO) of the four precipitating regimes. The RFO for each regime has been normalized including the no-precipitation regime. The diurnal cycles of the four precipitation regimes appear to be separated into two distinct patterns. Regimes 1 and 2 (top panels) have a peak in their RFO in the late afternoon, while regimes 3 and 4 (bottom panels) show a peak in the late at night and in the early hours of the morning. The two distinct patterns seen in the diurnal cycles may indicate a physical difference between the precipitation regimes. One possible explanation for this difference would be that the ratio of land to oceanic convection is different for the different regimes (as seen in table 2.1). Convection tends to peak during the late afternoon/early evening over land while over oceans convection tends to peak during the early morning (Yang and Slingo, 2001; Gray and Jacobson, 1977; Mapes and Houze, 1993). However, the nature of the convection associated with each regime (continental vs maritime) cannot be concluded solely from the diurnal cycle. Mapes and Houze (1993) have shown that the diurnal cycle of cold cloud clusters depends on the size of the cluster and Yang and Slingo (2001) showed that the diurnal signal over land often extends to the adjacent oceans. Furthermore Nesbitt and Zipser (2003) found that MCS rainfall over land peaks in the early morning. As regimes 3 and 4 both have high TVC values it is likely that MCS occurring over land are also associated with regimes 3 and 4.

Apart from the late afternoon maximum, regime 1 is equally likely to occur at any other hour of the day. This may be indicative for the fact that while land heating strongly influences regime 1, this relatively weak convection may well occur over the ocean and at other times of the day. Earlier it was postulated that regime 2 has a larger proportion of convection initialised by sea breezes. If this is the case, the minimum in the diurnal cycle for regime 2 (around 10 am) may well be due to

(a) **Regime 1**



(b) **Regime 2**
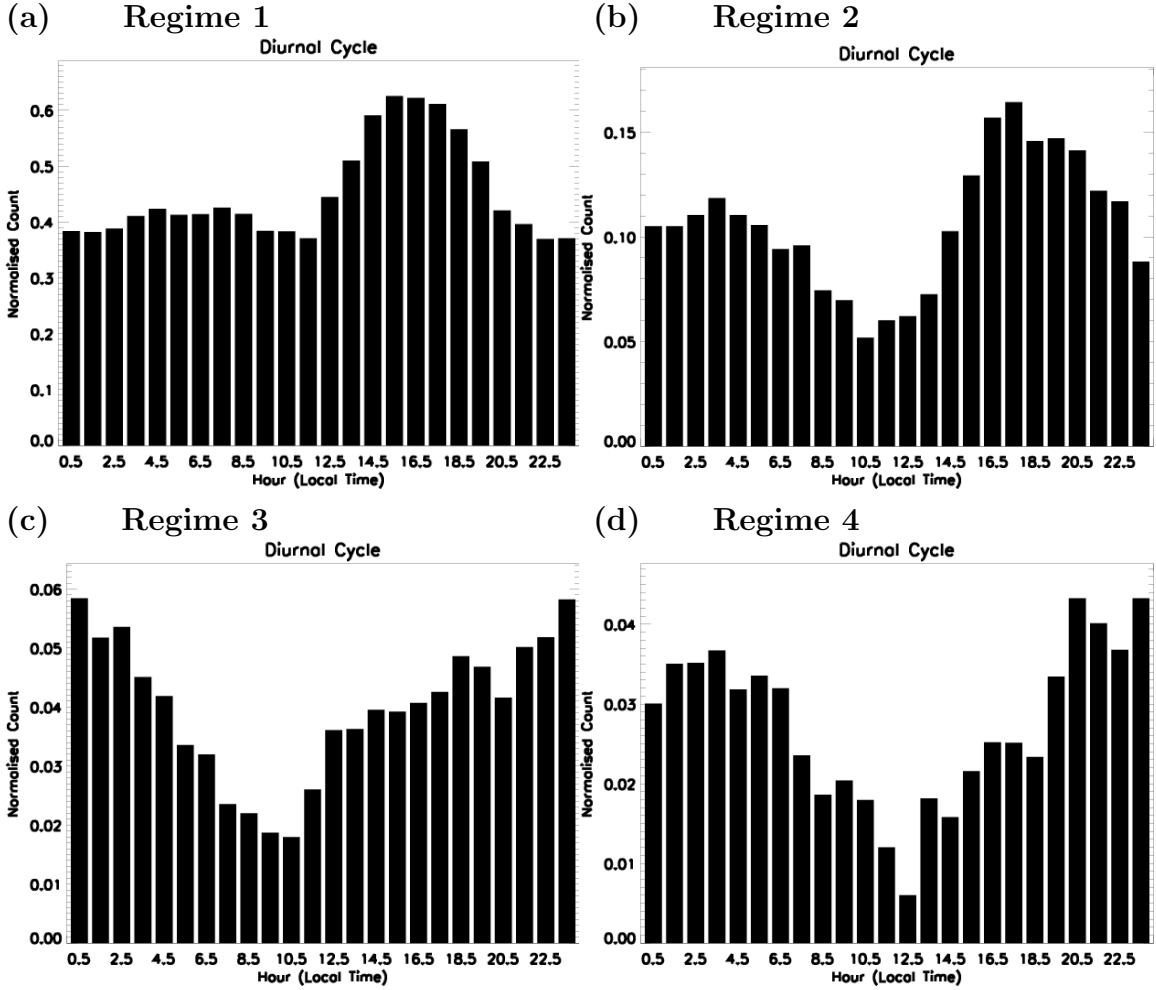


(c) **Regime 3**



(d) **Regime 4**



Figure 2.8: **Diurnal Cycle of the four precipitation regimes (in local time), the RFO has been normalised including the "zeroth" regime.**

the lack of the sea breeze at this time. Regime 2 also displays a secondary peak in the early hours of the morning, this is likely due the precipitation that occurs over the ocean in the western part of the domain.

### 2.4.3 Regime occurrence and monsoon activity

Another well-known indicator for the character of convection in the Darwin area is the synoptic meteorology of the Australian summer monsoon. A simple but effective definition of monsoon activity is wind-direction in the lower troposphere, with westerly winds indicative for an active phase of the monsoon, while easterly winds

dominate the break as well as pre-and post monsoon periods (Drosdowsky, 1996). Convection tends to be shallower and more representative of that found over oceans during periods when a westerly wind places Darwin in an oceanic air mass. Easterly wind conditions bring air from North East Australia to the region and the convection is often more representative of that occurring over continents. Here the concept of an east/west split in wind direction that previously proved useful in other investigations e.g. (Keenan and Carbone, 1992) is utilized with the aim of determining if wind direction has any influence on the occurrence of the four precipitation regimes defined. The purpose of the analysis is not to find a clear delineation between the regimes, but rather to compare how the precipitation regimes found using cluster analysis compare to traditional methods. For the analysis, wind data from Darwin airport at 700 hPa is defined as easterly or westerly if the wind direction has any component in those directions. The frequencies of occurrence of wind direction within each regime are shown in Figure 2.9.
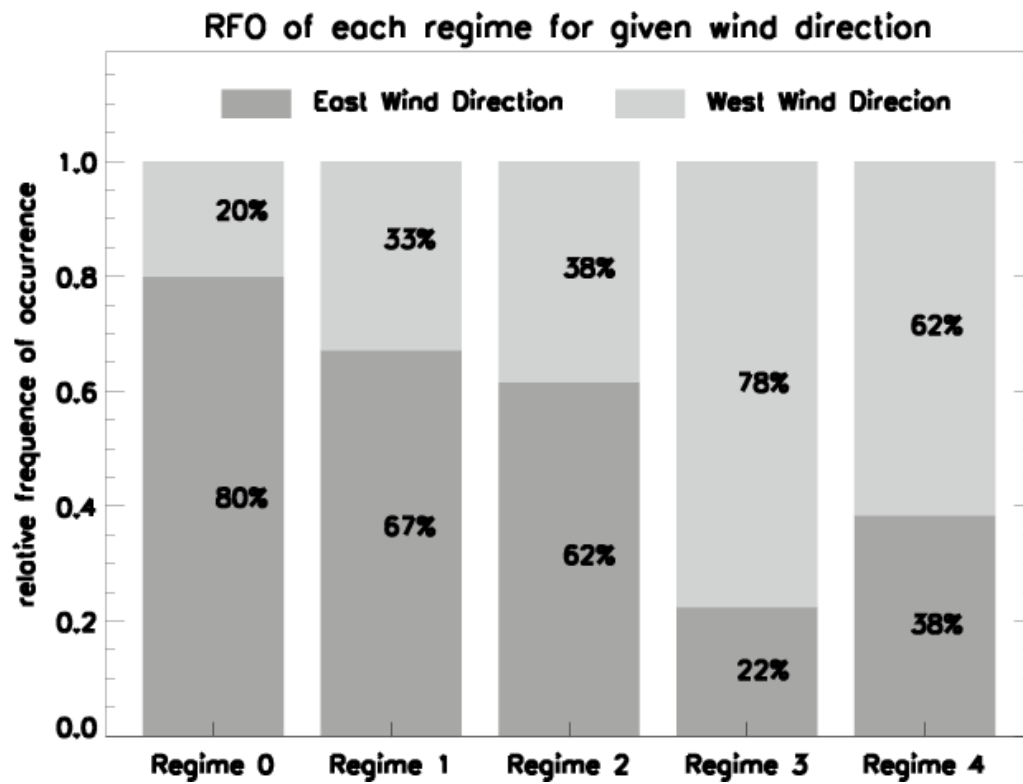
Figure 2.9: **Relative frequency of occurrence of the four precipitation and "zeroth" regimes for a given wind direction (defined as either easterly or westerly).**

It is evident from the figure that there are major differences between regimes in their relationship with wind direction and hence monsoon activity. Regimes 0 through 2 show a predominance in their occurrence during times when the wind is coming from the east (break conditions). During break conditions precipitating cloud systems are known to be typical of those with continental origin (Steiner et al., 1995). The no-precipitation regime shows the strongest association with easterly flow. As indicated in Figure 2.7 this regime occurs most frequently in the monsoon build-up and decay phases of the season. Although regime 2 is predominately an easterly regime, it contains more westerly flow than regime 1. This may help to explain why regime 2 was found to have to be a mixture of oceanic and land convection. In contrast, regimes 3 and 4 are most strongly associated with westerly conditions; westerly flow is typical for active monsoon periods, which have previously been shown to be associated with convection of maritime character (Keenan and Carbone, 1992).

While simple, analyzing the predominant wind direction in each of the precipitation regimes has further supported the findings of the previous sections, that the precipitation regimes identified in this study are physically sensible. While probably following intuition, this constitutes a non-trivial finding, as there is no physical a-priori information that supports that division entering into the cluster algorithm. The fact that the four (five) regimes could be objectively identified and match physical characteristics of convection rather well enables their use for further studies of convective systems in the Darwin area and possibly beyond.

## 2.5  Precipitation vs ISCCP-based cloud regimes

The previous section has established that the convective regimes identified in this study for the Darwin area exhibit two distinctly different patterns of their diurnal cycles and that the objectively defined regimes are able to naturally separate when

analyzing the synoptic scale winds. Furthermore, the strength of the convection and the presence and size of the stratiform precipitation component all constitute information contained in the convective regimes. This enables the use of these regimes for further studies of tropical convection and its interaction with the large-scale circulation. Given this and the fact that this work was partially motivated by the success of finding tropical regimes based on cloud properties retrieved from satellites, it appears worthwhile to investigate the relationship of the six ISCCP-based tropical cloud regimes of Rossow et al. (2005) with the "precipitation-based" regimes identified here. This will provide some insight into the precipitation structure (in a statistical sense) of the cloud regimes, which in turn have already been shown to have links to tropical circulation features (Rossow et al., 2005; Höglund, 2005).

To set the context for further discussion a brief summary of the six ISCCP weather states is given below, for a detailed discussion of the regimes the reader is referred to Rossow et al. (2005). Based on their cloud signature, three of the six regimes can be described as suppressed with respect to the occurrence of deep convection, whilst the other three have been identified as convectively active. Of the active regimes the deep anvil cloud (CD) regime exhibits a large coverage with optically thick clouds, most likely a mix of thick anvils and convective towers, while the convective cirrus (CC) regime is dominated by a large coverage with cirrus clouds of small to medium optical thickness. Probably the weakest (in terms of convection) of the three convectively active regimes is the mixed cloud (MIX) regime, termed because it likely consists of a mixture of shallow, congestus and deep convection without significant anvil and cirrus coverage. The suppressed thin cirrus (STC) regime is categorized as a suppressed regime and is dominated by a large coverage with thin cirrus clouds. As shown by Rossow et al. (2005) this regime occurs close to convection without containing significant amounts of convective clouds in the area over which the ISCCP histogram is calculated. The suppressed shallow cloud

regime (SSCL) regime has high frequencies of occurrence for clouds with high cloud top pressure (low tops) and is interpreted as consisting in most part of suppressed shallow clouds with a low total cloud cover as is typically found in shallow cumulus cloud fields. The suppressed shallow cloud regime with high cloud cover (SSCH) is also dominated by low clouds, but generally shows a much higher total coverage with clouds indicative of stratocumulus clouds. Stratocumulus is almost never observed in the Darwin area resulting in an extremely low frequency of occurrence of the SSCH satellite regime in the study region. To avoid sampling biases this regime was excluded from further analysis.

As in previous studies (Höglund, 2005) one ISCCP weather state per day is defined for each ISCCP grid-box. As Darwin straddles several ISCCP grid-boxes, sensitivity studies to the choice of grid point were carried out, but showed little sensitivity with regard to the main conclusions. Hence the analysis presented here uses four ISCCP grid boxes each having an area of 280x280 km$^2$ located roughly between 10°-15°S and 128.5°-133°E. The relationship of the ISCCP-based cloud regimes to the precipitation regimes of this study, which are identified hourly, is investigated by calculating the mean frequencies of occurrence of each precipitation regime on all days that fall into a particular ISCCP regime. Note that the C-Pol radar detects precipitation sized particles and is unlikely to observe cirrus and non-precipitating clouds.

Figure 2.10 shows the RFO of the five regimes of this study for each ISCCP-based cloud regime. The cloud regimes are sorted as in Rossow et al. (2005) from what is considered the most convectively active regime (CD) to the most suppressed regime (SSCL). Several noteworthy features are evident in Figure 2.10. The occurrence of no precipitation (Regime 0) increases steadily from CD to SSCL. This supports the intuitive interpretation used in earlier work defining the ISCCP-based cloud regimes as convectively active vs suppressed. Of the suppressed regimes, SSCL clearly shows

the least occurrence of deep convective regimes (2-4) and is mostly characterized by the relatively shallow precipitation regime 1. There is a steady increase in the deep convective regimes (2-4) for the MIX and CC regimes, mostly at the expense of the no-precipitation regime. Nevertheless, both these cloud regimes remain dominated by the regimes with low precipitation coverage (regimes 1 and 2). The only cloud regime with a significant occurrence of the regimes with high precipitation coverage (regimes 3 and 4) is the CD regime. In this regime the frequency of occurrence of no precipitation at all has dropped to around 10 % indicating the frequent occurrence of precipitation in that regime. In earlier studies (Rossow et al., 2005; Höglund, 2005) it was found that the ISCCP-based CD regime occurs predominantly in active phases (i.e., convergence zones) of large-scale tropical circulation features such as the Madden-Julian Oscillation (MJO) or the Australian summer monsoon. This and the identification of the CD regime as the only regime containing high frequencies of regimes 3 and 4 leaves the intriguing possibility that what is thought of as "typical" maritime precipitating convection is in fact a convective type that is enabled and supported only by larger-scale circulation features.

An in depth study into the precipitation characteristics of the ISCCP regimes is beyond the scope of this investigation, however for completeness a simplistic calculation is performed to determine the average rain rate per hour of each ISCCP regime. As the rain rates for the ISCCP regimes have been derived from data obtained from the region surrounding Darwin, they are not intended to be applied to the tropics as a whole. The rain rate for each of the ISCCP regimes was calculated by multiplying the rain rate for each precipitation regime (Figure 2.5) by the frequency with which the radar regimes occur within the given ISCCP regime (Figure 2.10). These results are displayed in Table 2.2 and show that the rain rates for the ISCCP regimes is highest for the most convective regime CD, and decreases corresponding to convective intensity, with the suppressed regimes having the lowest rain rates. It
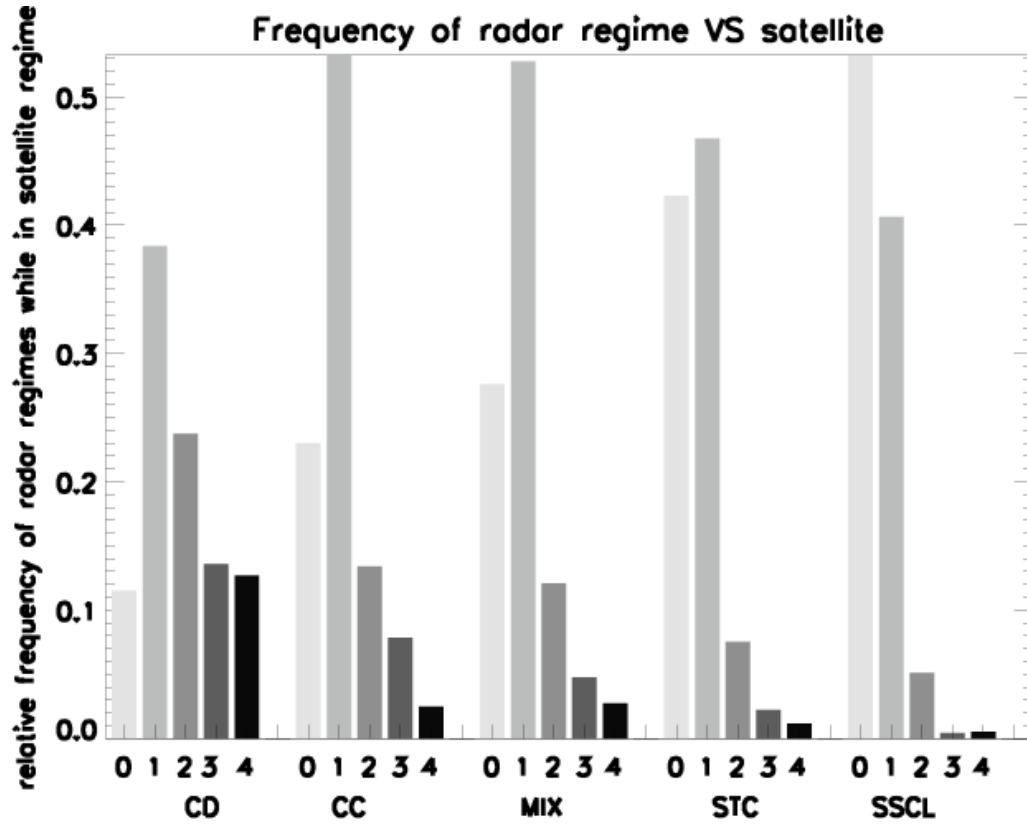
Figure 2.10: **Comparison between satellite and radar regimes. Relative frequency of occurrence as a function of radar regime, for the satellite regimes defined by Rossow et al. The SSCH satellite regime has not been included due to an extremely low frequency of occurrence.**

is interesting to note that the CC and the MIX regimes both have approximately the same rain rates. Although simplistic, associating an average rain rate with each of the ISCCP regimes can be thought of as a first step in creating/assessing parameterizations for numerical modeling. For example Figure 2.10 and Table 2.2 show that from a precipitation point of view regimes CC and MIX are quite similar. Therefore it may be useful to create parameterisations for tropical weather conditions based on the ISSCP regimes (which represent distinct and different states of the tropical atmosphere that can distinguish the four main classes found here), using the radar regimes to estimate of the total rainfall and latent heating associated with each regime.

| CD Regime | CC Regime | MIX Regime | STC Regime | SSCL Regime |
|-----------|-----------|------------|------------|-------------|
| 0.8007 mm/h | 0.4325 mm/h | 0.3981 mm/h | 0.2729 mm/h | 0.1994 mm/h |

Table 2.2: **Calculated average rain rate of each ISCCP regime in mm/hour.**

## 2.6 Summary

This chapter has investigated if the three dimensional structure of precipitation in the Darwin area as measured by a scanning polarimetric weather radar can be objectively classified into regimes. It has furthermore posed the question if those regimes have physical connections to known features of the tropical atmosphere in the region. A cluster algorithm was applied to two-dimensional histograms of reflectivity with height derived from quasi-instantaneous radar volumes. After excluding a regime with virtually no radar returns in the volume (a "no-precipitation" regime) the analysis revealed the existence of four precipitation regimes in the Darwin region:

- **Regime 1:- a patchy convective regime of medium intensity and low area coverage. This regime occurs most frequently in the afternoon and during break conditions.**

- **Regime 2:- a strong convective regime with relatively small area coverage. This regime has a peak in the diurnal cycle in the afternoon and a secondary peak in the early morning.**

- **Regime 3:-a weak convective regime with large area coverage and large stratiform regions. This regime occurs most frequently during the late night / early morning and during monsoon conditions.**

- **Regime 4:- a strong convective regime with large area coverage and large stratiform regions. This regimes has a peak in the diurnal cycle in the late night/early morning and has links to monsoon conditions.**

The overall statistics of the occurrence of these regimes once again highlights that even during convectively active periods such as the Australian Summer Monsoon, the spatial coverage with precipitation is small overall (cf. Figs. 2.2 and 2.3).

Linking the regimes identified by the algorithm to known features such as the seasonal and diurnal cycle and a simple description of monsoon activity through wind direction has confirmed their different physical character. Regime 1 and 2 show signs of a strongly land-induced diurnal cycle with an afternoon peak and occurs throughout the wet season. They are largely associated with winds from the East, typical for monsoon build-up and decay as well as monsoon break conditions, during which convection is often observed to be of continental character. Regimes 3 and 4 have their largest frequency of occurrence in January - the peak of the monsoon season. Their diurnal variation peaks at night or in the early morning and they are largely associated with westerly winds, as is typical for monsoon conditions.

Combining the precipitation regimes with tropical cloud regimes previously identified from ISCCP data reveals interesting relationships between the two. Arguably there exist probably three (maybe four) broad precipitation regimes across the five ISCCP-based regimes present at Darwin. The suppressed shallow cloud regime (SSCL) is dominated by the no-precipitation and the patchy convection regimes. The stronger convection regimes, in particular the maritime ones, do not feature strongly in this cloud regime. Whilst still dominated by the patchy convection regimes, the convective cirrus (CC) and mixed (MIX) cloud regimes and to a lesser extent the suppressed thin cirrus (STC) regime show larger frequencies of occurrence than the stronger precipitation regimes. The fraction of maritime versus continental precipitation regimes shifts steadily to more and more maritime with increasing convective character as identified in the cloud regimes themselves. By far the most active and most "maritime" regime is the convective deep anvil cloud regime (CD). About 50% of the time in this cloud regime one of the three strong convective pre-

cipitation regimes is present and this is the only regime where the strong maritime regime occurs with almost equal frequency to its weaker counterpart. The occurrence of no precipitation or patchy convection in this cloud regime are significantly reduced compared to any other cloud regime, showing the significantly different character of this regime. Given the aforementioned strong links of the CD cloud regime to large-scale circulation features (Rossow et al., 2005; Höglund, 2005) further investigations into the precipitation characteristics using the regimes definitions derived here seem warranted.

The objective identification of precipitation regimes from radar data and the establishment of some of their physical characteristics opens several avenues for further research. Their simplest application is in the objective identification of the overall character, i.e., mostly maritime versus mostly continental, and intensity, i.e., strong versus weak. Such characterization obviously already exists and is used regularly, but has so far largely been based on subjective criteria and/or the use of wind direction to identify convective character. The use of radar data as suggested here is both objective and more directly based on the precipitation features themselves. An obvious question for future research is how general the regimes found at Darwin are for the wider tropical region. Applying the techniques proposed here to other radar locations in the tropics, such as Kwajalein (Cetrone and Houze, 2006), would be a natural extension of this study. Another possibility would be the application to data collected by the TRMM precipitation radar. It would also be interesting to investigate if regimes defined in a similar way to those used here could helpful in objectively determining some of the "building blocks" used in the Stretched Building Block conceptual model described by Mapes et al. (2006).

The regimes identified here represent the statistical characteristics of convection at the spatial scale of the radar footprint (roughly 300 km$^2$). Studying the regime identification as a function of scale is therefore another possible extension of this

work. As has been highlighted here, combining the precipitation regime information with data from different types of instruments is another promising avenue of research. The brief example of the ISCCP-based cloud regimes emphasizes the potential of such investigations. The presence of a large suite of instruments deployed by the US Department of Energy's ARM program (Ackerman and Stokes, 2003) as well a recently conducted field studies such as TWP-ICE (May et al., 2008) make Darwin the ideal location for such studies. On a longer time-scale objectively defined precipitation regimes may also prove useful when studying local changes in rainfall patterns due to changes in synoptic conditions and seasonal weather patterns (e.g. El Niño). A similar analysis could be performed over a much larger time scale to determine how or if the precipitation regimes change with time and determine any possible links with climate change. Finally it would be very interesting to investigate if models, in particular those that resolve deep convection, are able to reproduce the observed regimes when forced with realistic large-scale conditions. Recent field data will enable such studies in the near future.

# Chapter 3

# The Weather Research and Forecasting (WRF) Simulations and Thermodynamic Evaluations

In this chapter the numerical model used to simulate tropical convection over Darwin is introduced, the model domains and physical packages used in each simulation are detailed, and the five simulated periods within TWP-ICE are defined. Data from TWP-ICE is then used to evaluate the forcing data used to initialise and provide the boundary conditions for the mesoscale simulations.

## 3.1 The WRF model

One of the aims of this thesis is use a numerical weather prediction model to simulate the wide variety of cloud types present during a typical wet season over Darwin, Australia. The model simulations will be evaluated primarily on the precipitation associated with these cloud systems, taking advantage of data from the TWP-ICE field campaign and the Bureau of Meteorology's (BoM) research radar. However, ultimately we wish to use the precipitation regimes defined in chapter 2 in the evaluation process.

The WRF model is a publicly available non-hydrostatic model that can be run either as a research model, or as an operational numerical weather prediction model. It has been designed to simulate the atmosphere on a variety of length scales ranging from meters to hundreds of kilometers, and comes with multiple physics modules that can be chosen to suit the needs of a researcher/forecaster in their specific application. When used for mesoscale numerical weather prediction WRF requires

data from another model (usually a global model or another mesoscale model) to provide the initial and boundary conditions. The WRF model has the option of two-way nesting, where information from the high resolution domain is passed back to the low resolution domains, or one-way nesting, in which information flows only from the low resolution domains to the high resolution domain. All simulations presented in this document were made using WRF version 2.2.1.

### 3.1.1   Model Domains

The domains used in all simulations are shown in Figure 3.1. The outermost domain (d01) has a grid spacing of 34 km and each successive domain is nested (one-way) with a grid spacing ratio of 3:1 to its parent domain. Thus the grid spacing for d02, d03 and d04 is $\sim 11.3$, $\sim 3.8$ and $\sim 1.3$ km respectively. Domain four (d04) has been chosen to coincide with the domain of the BoM research CPOL radar used in chapter 2.

### 3.1.2   Model Physics

The WRF model (version 2.2.1) comes with 7 microphysics schemes, 3 cumulus parameterization schemes, 2 surface layer schemes, 3 land surface models, 3 planetary boundary layer schemes, 2 longwave radiation schemes and 3 shortwave radiation schemes. A list of the physics modules common to all simulations presented in this thesis is provided below, for further information see Skamarock et al. (2005). The surface layer scheme used in all simulations is the **Similarity theory (Eta)** scheme (option 2). The Land-Surface Model used in all simulations is the **Noah LSM** (option 2). The Planetary Boundary Layer scheme used is the **Mellor-Yamada-Janjic** scheme (option 2). The long wave radiation scheme used is the **Rapid Radiative Transfer Model (RRTM)** (option 1). The short wave radiation scheme used is the **Goddard Shortwave** (option 2). Cumulus Parameterization is used only in
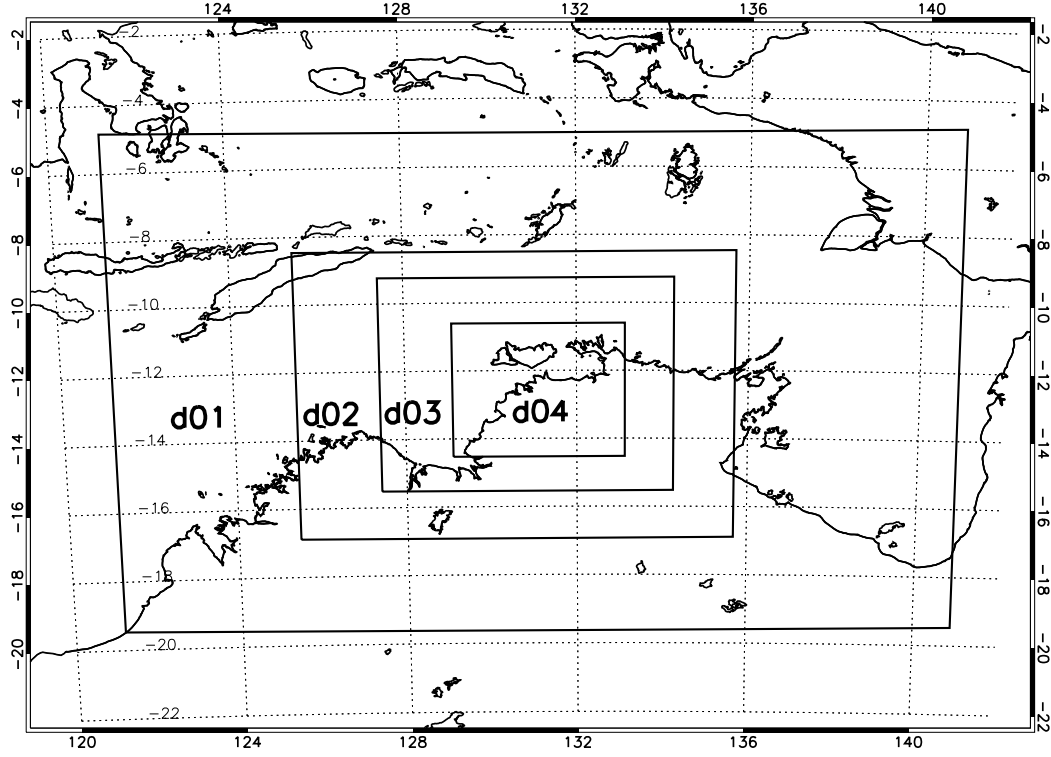
Figure 3.1: **Domains used in all simulations. The grid spacing in domain 1 (d01) is 34 km, the grid spacing in domain 2 (d02) is $\sim$ 11.3, the grid spacing in domain 3 (d03) is $\sim$ 3.8 and the grid spacing in domain 4 (d04) is $\sim$ 1.3 km.**

domains 1 and 2 where the grid spacing is too large to resolve convective updrafts. The cumulus Pameretization scheme used in domains 1 and 2 is the **Betts-Miller-Janjic** scheme (option 2).

Each of the simulations presented in this thesis have been run using two different microphysics schemes to evaluate how sensitive the simulations are to the microphysics module used, all other physics modules have been kept static and are listed above. The two microphysics schemes used are the **Purdue Lin** scheme (Lin et al., 1983; Chen and Sun, 2002) (hereafter referred to as Lin) and the **Thompson et al.** scheme (Thompson et al., 2004) (hereafter referred to as Thompson).

The Lin microphysics scheme is a single-moment microphysics scheme with six classes of water including water vapor, cloud water, rain, cloud ice, snow, and graupel. The Lin microphysics uses a Marshall-Palmer distribution to represent the drop size distribution of the precipitating categories, which is dependent on the mixing ratio of each category and a number of set parameters such as the densities and intercept parameters for each category.

The Thompson scheme is a more sophisticated microphysics scheme that was developed primarily to improve forecasts of supercooled liquid water, and has been tuned to the results of the bin microphysics scheme of Rasmussen et al. (2002). The Thompson scheme includes the same six water classes as the Lin microphysics. However, in the Thompson scheme the size distributions for each category are specified by a generalized gamma distribution and the intercept parameters for each distribution change as a function of temperature or mixing ratio. Furthermore, a number concentration for cloud ice is also included.

### 3.1.3 Simulations

Five simulations during the TWP-ICE period were run so that the broad range of meteorological conditions that occur during a typical wet season could be simulated. The simulations have been named according to the meteorological conditions during the time of the simulation. Each simulation was initialized and ended at 12 UTC on the date listed in table 3.1. From table 3.1 it can be seen that there is some overlap between the simulations e.g. the monsoon ends on the 24th of January and Monsoon/Suppressed simulation begins on the 23rd of January. Whenever results are displayed in a time series spanning multiple simulations, data from the end of a simulation is chosen in preference to data from the beginning of a simulation (to remove spin-up period).

|  | Monsoon | Monsoon/Suppressed | Suppressed | Suppressed/Break | Break |
|---|---|---|---|---|---|
| start date | 18/01/06 | 23/01/06 | 28/01/06 | 03/02/06 | 09/02/06 |
| end date | 24/01/06 | 29/01/06 | 03/02/06 | 09/02/06 | 15/02/06 |

Table 3.1: **Start and end dates of the 5 simulations during the TWP-ICE period**

## 3.2   Initialisation data

The initial and boundary conditions used to force the WRF simulations are provided by the National Centers for Environment Prediction (NCEP) Global Forecast System (GFS) Final Analysis (FNL), hereafter referred to as forcing or input data. The skill of any forecast or simulation is dependent on the quality of the initial and boundary conditions used to drive the model. Therefore, before assessing the WRF model in detail the data used to force the model is compared with observations to determine if any biases exist that could lead to inaccuracies in the simulations. The observational data used for comparison is called the Xie-Klein Forcing dataset; this dataset was created using the constrained variational analysis technique described in Zhang and Lin (1997) and Zhang et al. (2001) for the purpose of forcing single column models and for model evaluation Xie et al. (2009). The observational data used to create the Xie-Klein dataset comes from the radiosondes, flux sites and radars depicted in Figure 1.1. The valid domain of the Xie-Klein dataset is technically defined as the area bounded by the five outer radiosonde sites (pentagram region). However, for the analysis described below it is assumed that the forcing applies to the entire region and observations from the Xie-Klein dataset are compared directly with average quantities calculated over domain 4.

Figure 3.2 shows the results for a) input zonal wind (red westerly) b) input meridional wind (red southerly) c) observed zonal wind, d) observed meridional wind, e) the WRF zonal wind and f) WRF meridional wind over the entire TWP-ICE period. Figures 3.2 a), b), e) and f) are a composite of five simulations, where the

black lines on all figures indicate the end of a simulation. Where there is an overlap between simulations data from the first 24 hours of the overlapping simulation is excluded to minimize any spin up effects.

Figures 3.2 indicates that the winds provided by the NCEP FNL agree qualitatively well with the observational data. There are only a few times during which there is an obvious difference between the NCEP and observed winds. Some of these differences include an underestimation of the low level easterly winds during the passage of the MCS (23rd-24th). An overestimation of the low level easterly winds during the end of the experiment (7th-12th) and an underestimation of the easterly winds in the stratosphere, especially during the suppressed monsoon period (25 January to the 2nd of February).

The WRF simulated zonal and meridional winds, Figure 3.2 e) and f), are quite similar to the NCEP winds; perhaps the most interesting feature of note is that after the passage of the MCS (approximately the 23rd of Jan), the simulated zonal winds (easterly) are much weaker than observed. Even though the zonal wind in the NCEP data is weaker than the observations during this time, the WRF model has further weakened the zonal wind during the course of the simulation.

Figure 3.2 illustrates the three regimes or phases that occurred during TWP-ICE. Phase 1) is described as the active monsoon and occurs between roughly the 18-21th of January (first simulation) during which the lower tropospheric winds are westerly (positive numbers, red). Towards the end of the simulation (22nd -24th) a MCS dominated the large-scale wind flow changing the average winds over the regime from westerly to easterly. Phase 2) occurs between roughly the 25th of January to the 5th of February, during this time a low pressure system was located to the south of Darwin. Due to the circulation around the low pressure system the low level average winds over the domain were westerly (the classical definition of the active monsoon), however the air had continental origins and was consequently

very dry, which lead to suppressed conditions. Phase 3 occurs between roughly the 6th-12th of February during which the low level tropospheric winds have reverted to an easterly direction typical of a monsoon break period.
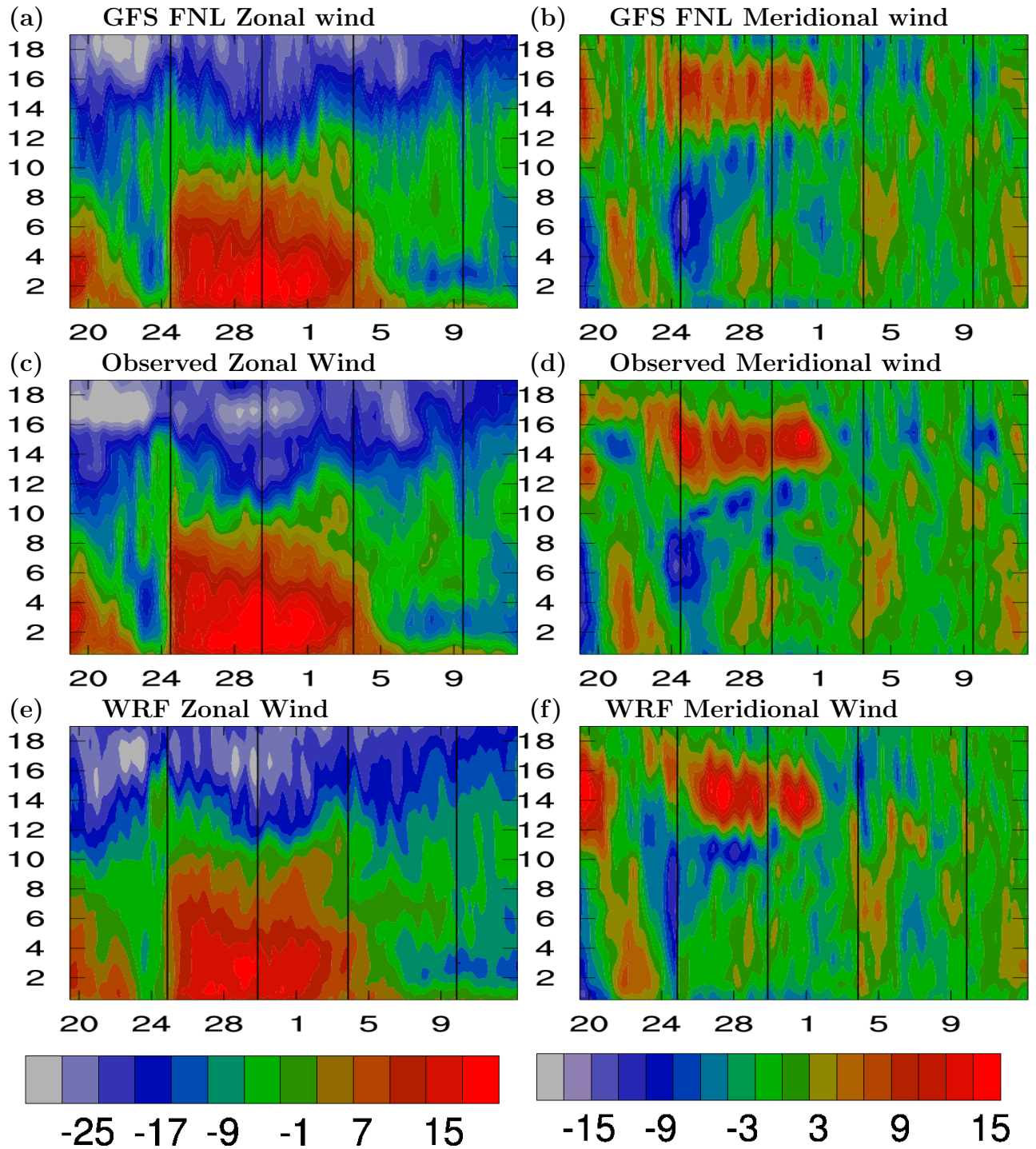
Figure 3.2: **Domain-averaged vertical profiles of Zonal and Meridional Winds derived from a) and b) the GFS FNL data. c) and d) the Xie Klein forcing dataset and e) and f) the WRF simulations**

Figure 3.3 shows the temperature profiles for a) the input temperature c) the observed temperature and e) the WRF temperature for the entire TWP-ICE period. Figure 3.3 b) shows the difference between the input and observations temperature over the TWP-ICE period and Figure 3.3 d) shows the difference between the WRF simulated and observed temperature over the TWP-ICE period.

Figure 3.3 b) shows that the NCEP data has two layers at different altitudes that have a large positive temperature bias, one centered at approximately 15 and the other at 17 km. The bias at 17 km level (stratosphere) reaches as much as 5 degrees and is consistently present throughout the TWP-ICE period. The temperature difference between the input data and observations is approximately 1-2 degrees at 15 km during the first half of the experiment and decreases towards the end of the experiment (although it is still present). Given the magnitude of these temperature differences it is expected that any simulation forced with these data will be adversely effected. For example during times of deep convection it would be expected that the temperature bias would limit updrafts or perhaps even cap convection.

Figure 3.3 d) shows the temperature difference between WRF simulations and observations for the TWP-ICE period. As was the case for the NCEP data there is a large difference between the simulated and observed temperatures at 17 km. However the temperature bias has spread vertically in the simulation and ranges from approximately 15.5 km to 18.5 km, compared with 16 km to 18 km in the NCEP data. The temperature bias in the NCEP data at 15 km is not present in the simulations at that altitude. There is however a large temperature bias at approximately 11-12 km at the beginning of the simulation that decreases in altitude and spreads vertically as the simulation progresses, by the 24th of January the bias is centered at approximately 8 km and ranges from around 5 to 12 km.

One of the most interesting features of the temperature bias can not be seen in Figure 3.3 d) because the initial spin-up period for the WRF model is not shown

in this figure (the simulation was started on the 18th of January at 12 UTC, while the Xie-Klein forcing dataset begins on the 19th of January at 0 UTC). In order to assess the temperature bias during the initial spin-up period a plot showing the difference between the observations and a simulation that was initialized on the 21st of January at 12 UTC is shown in Figure 3.3 f). During the first 12 hours of the simulation the temperature bias initially centered around 15 km (as in the input data) is seen to rapidly decrease in altitude to approximately 12 km and spread vertically. Figure 3.3 f) indicates that the temperature bias at 12 km is linked with the NCEP bias at 15 km and is likely not an indication of a problem with the WRF model, but rather a consequence of incorrect forcing. As the magnitude of the temperature bias is quite large (2-3 degrees) it is expected that the simulations will be negatively affected by the incorrect temperature forcing, making it hard to determine whether the WRF model itself is performing poorly due to inaccurate physics modules or if any inaccuracies stem from incorrect forcing. It would be useful to run further simulations that have been forced with data from another source, perhaps the Australian Bureau of Meteorology's Tropical Cycle Limited Area Prediction Scheme (TCLaps) or ECMWF, however such an analysis will be left to future work.[1]

---

[1]Some preliminary results of a simulation initialised with the ECMWF interim re-analysis data is provided in the appendix. The simulation shown was run using WRF version 3.0.1 for the monsoon period. This simulation was run purely for interest and at this stage no analysis has been performed (the simulation has not been quality checked either)
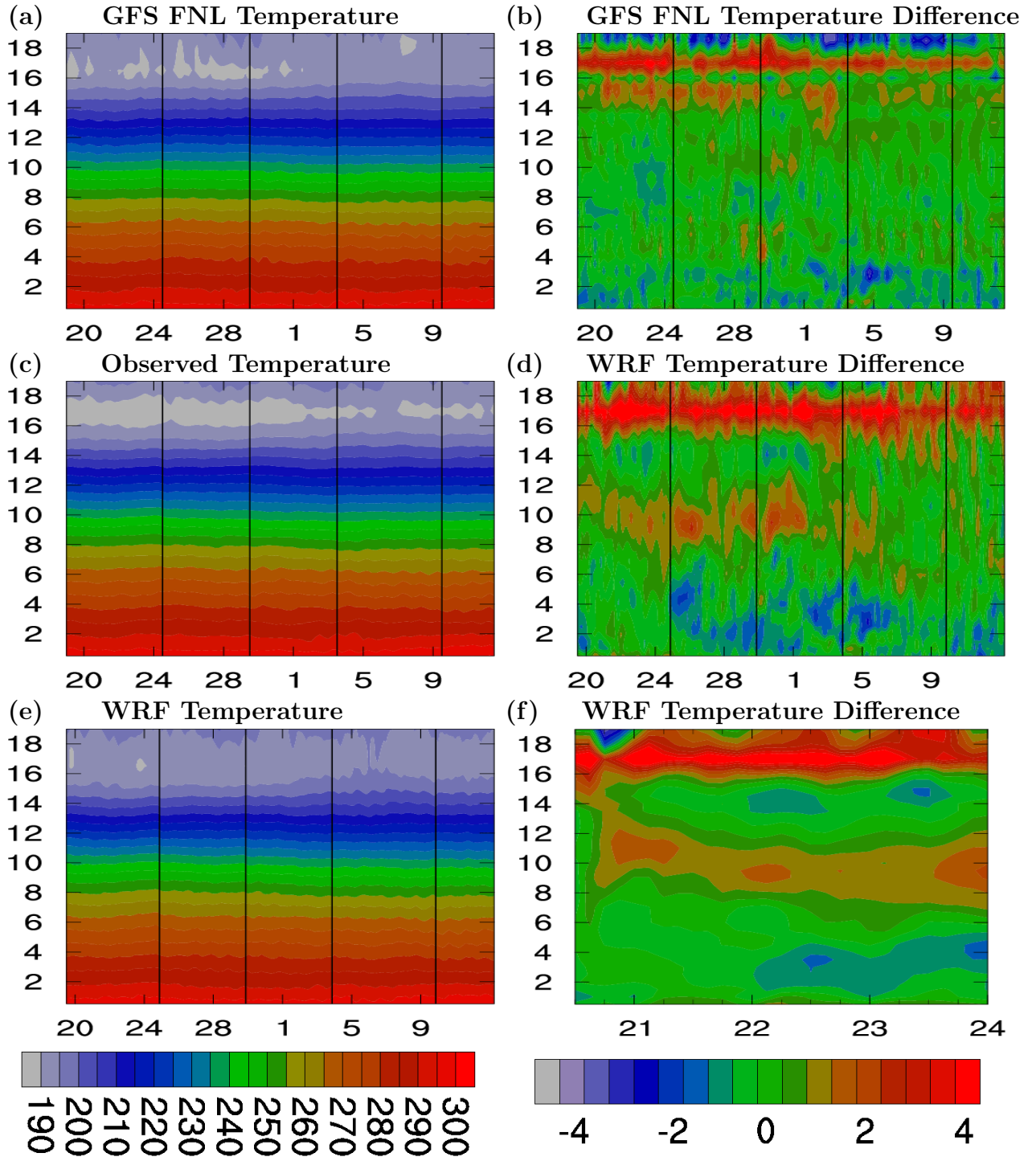
Figure 3.3: **Domain-averaged Temperature profiles over the entire TWP-ICE period for a) the GFS FNL data, c) the Xie-Klein forcing dataset, e) the WRF simulation. Differences in temperature between the Xie-Klein data and b) the GFS FNL data, d) The WRF simulations over the entire TWP-ICE period and f) a WRF simulation starting on the 21st of January.**

## 3.3   Vertical Velocity

Figure 3.4 shows the vertical velocity profiles for both the WRF model and observations for the entire TWP-ICE period. In Figures 3.2 and 3.3 the difference in winds and temperature between simulations using Lin and Thompson microphysics were so small that only the results for the Lin microphysics scheme were displayed. Vertical velocity is the exception to this rule, therefore vertical velocity profiles for a) Lin microphysics b) Observed vertical velocity and c) Thompson simulation are displayed in Figure 3.4. These figures have been included to give an overview of the domain-averaged vertical motions over the entire TWP-ICE period, the monsoon and break period will be assessed in more detail below. In general it appears from Figure 3.4 that the simulations using Lin microphysics perform better at simulating the strength of the observed domain-averaged upward motion. Domain-averaged upward vertical motion is consistently underestimated throughout the TWP-ICE period in the Thompson simulation, especially during the active monsoon.

Figure 3.5 shows the domain-averaged vertical velocity profiles for the monsoon simulation only. During the first part of the monsoon simulation (19th to the 21st) both the Lin and Thompson simulations have similar vertical velocity profiles, both show subsidence between 11 km and 16 km that is not seen in the observations and both simulations underestimate the upward motion below 11 km. On the 21st of January the simulations begin to differ; the Lin simulation has stronger upward motions (that extend higher) during the first half of the 21st followed by stronger subsidence between 11 and 15 km during the second half of the 21st. The timing of the upward motion in both simulations is out of phase with the observations, on the 21st the observations show strong upward motion around 0 UTC while the simulations have weak upward motion approximately 6 hours later.

While in general the Lin simulation seems to perform better than the Thompson simulation during the first half of the monsoon period, it does suffer some of the

same problems i.e. subsidence in the 11-16km levels that is not present in the observations and a shift in the timing of convection. However, the strength of the domain-averaged upward motion in the Lin simulation is closer to the observed values.

On the 23rd of January the Thompson scheme shows strong subsidence between 11 and 16 km that is not seen in either the observations or the Lin simulation. The strength of the upward motion below the subsidence layer is significantly weakened and the large-scale ascent during the passage of the MCS is not well captured in the Thompson simulation. During the passage of the MCS the Lin microphysics performs much better in simulating the observed updraft strengths, however the large-scale ascent occurs over a longer period of time (between roughly the 23rd to the end of the simulation), whereas the observations show more intense upward motion between 12 UTC on the 23rd to 0 UTC on the 24th.

Figure 3.5 b) d) and f) show the domain-averaged vertical velocity for the break simulation only. The main difference between the Lin and Thompson simulations during this period is the strength of the vertical motion; as was the case in the monsoon simulation the Lin microphysics produces stronger vertical motion that is closer to the observed values. Both the Lin and the Thompson simulations have a significant problem in the timing of the events throughout the break period. On the 9th of February both simulations completely miss the convection seen in the observations. However, this period falls into the first 12 hours of the simulation and is therefore affected by the initial spin-up period for the model. On the 10th and the 11th of February both simulations show convection occurring around 6 UTC (approximately 3 p.m. local time), the observations do show some vertical ascent on the 10th of February and a hint of upward vertical motion between 5 and 12 km on the 11th, however agreement between model and observations is poor. Both simulations completely miss the nocturnal ascent on the 10th (low level),

11th and 12th seen in the observations between 18 and 0 UTC (3-am -9-am local time). Instead the simulations show subsidence between 10 and 15 km during this time (after the afternoon convection), this may be responsible for suppressing any nocturnal convection.

## 3.4   Summary

In summary the zonal and meridional winds provided by the GFS FNL data correlate well with the observed winds from the Xie-Klein dataset. The temperature data on the other hand differers by as much as 5 degrees in the stratosphere, and between 2-3 degrees at approximately 15 km. The WRF simulations forced with this temperature data were also found to have a large temperature bias, however the bias had shifted to approximately 12 km. The height and magnitude of this temperature bias is expected to have a detrimental affect on the WRF simulations.

Lin simulation seems to perform better than the Thompson simulation when compared with the observations of vertical velocity (though Lin simulation still underestimates the strength of the upward vertical motion). Both simulations show strong subsidence between 10 and 16 km (especially during the active monsoon) that is not seen in the observations. Subsidence at this level is likely responsible for the large temperature difference seen at these altitudes, while also suppressing ascent at the lower levels. Differences in the vertical velocity profiles between the Lin and Thompson simulations cannot be attributed to any problems with the forcing data, because both use the same model setup. The fact that such large differences are produced by changing only the microphysics scheme indicates that the microphysics schemes are behaving differently enough to feed-back into the dynamics of the model. Differences in the latent heating or the vertical profiles of latent heating are the likely cause of these dynamical differences. Chapters 4 and 6 will look more closely at the microphysics of the simulations.
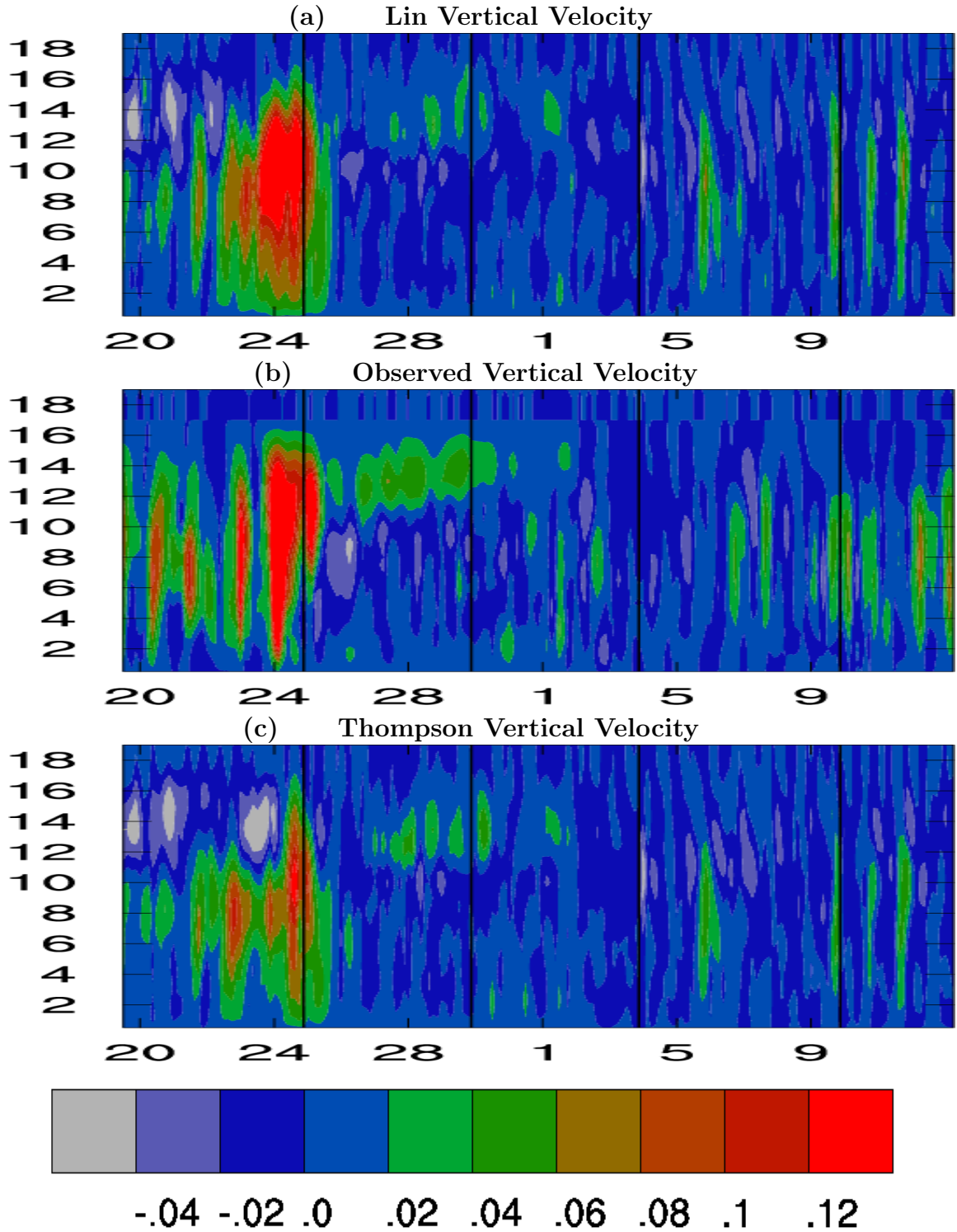
Figure 3.4: **Domain-averaged vertical velocity profiles over the entire TWP-ICE period for a) the Lin simulations, b) the Xie-Klein forcing dataset and c) the Thompson simulations.**
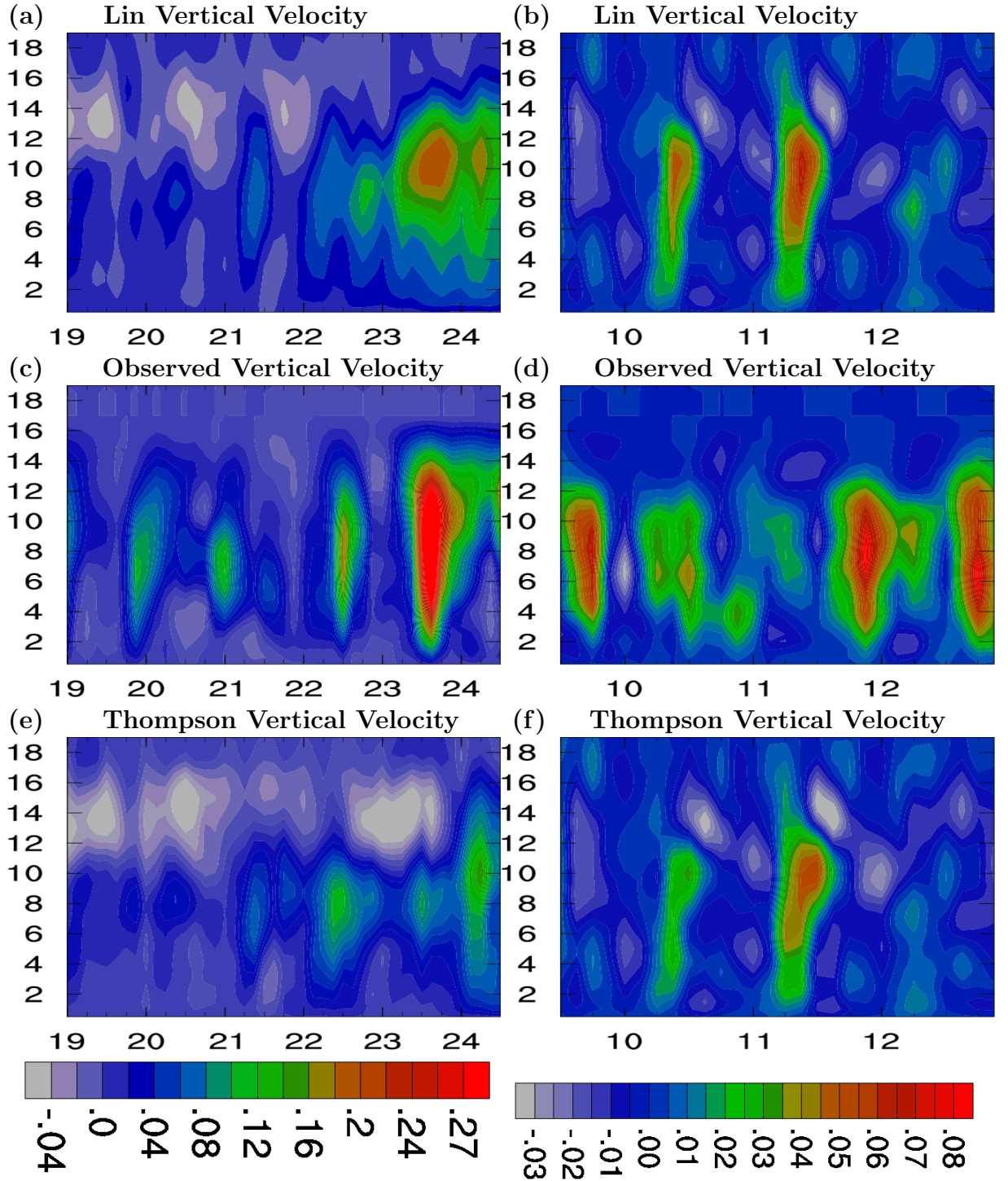
Figure 3.5: **Domain-averaged vertical velocity profiles over the Monsoon period for a) the Lin simulations, c) the Xie-Klein forcing dataset and e) the Thompson simulations. Domain-averaged vertical velocity profiles over the Break period for b) the Lin simulations, d) the Xie-Klein forcing dataset and f) the Thompson simulations.**

# Chapter 4

# Model Simulated Radar Reflectivity

## 4.1   Introduction

Ground-based radars are able to observe precipitation/hydrometeors over a large spatial scale (up to 150 km horizontally and 20 km vertically) allowing observations of precipitation over a significant proportion of a typical cloud resolving domain. While point observations such as rain gauges will always be useful, especially in dense networks, the sheer quantity of information made available from a weather radar makes its use an attractive choice when validating a numerical model. This chapter describes the how data from the WRF model are converted into a form that can be compared with observations from a weather radar. A simple analysis (domain-averaged results) is provided to demonstrate the effect of the conversion algorithm. More complicated comparisons between the radar and model are performed in chapters 5, 6 and 7.

## 4.2   Simulated Reflectivity

To compare radar reflectivity data (in dBZ) to the WRF simulations an algorithm provided by the Read/Interpolate/Plot (RIP) graphical package (developed at NCAR) was used to convert the model microphysical data into simulated radar reflectivity. The algorithm assumes that all particles are spheres of constant density, and for single moment bulk microphysics schemes (such as the Lin scheme) that the particle size distribution for each of the precipitating category (rain, graupel and

snow) follows an exponential function such as that depicted in equation 4.1.

$$N(D) = N_0 e^{(-\lambda D)} \tag{4.1}$$

where N(D) is the number concentration of particles with diameter D, $N_0$ is the intercept parameter and $\lambda$ is the slope parameter which is defined as

$$\lambda_{cat} = \left(\frac{\pi N_0 \rho_{cat}}{\rho_{air} q_{cat}}\right)^{\frac{1}{4}} \tag{4.2}$$

$\rho_{air}$ is the density of dry air and $\rho_{cat}$ and $q_{cat}$ are the density and mixing ratio of the hydrometer category.

The equivalent reflectivity factor for rain $(Z_e)$ is taken as the 6th moment of the size distribution. Thus

$$Z_e = \int D^6 N_0 e^{-\lambda D} \, \mathrm{dD} \tag{4.3}$$

Assuming the rain droplets are perfectly spherical and after some calculation (see Appendix) the equivalent reflectivity for rain becomes

$$Z_e = \frac{720(\rho_{air} q_r)^{\frac{7}{4}}}{N_r^{\frac{3}{4}}(\pi \rho_r)^{\frac{7}{4}}} \tag{4.4}$$

Equation 4.4 describes the equivalent reflectivity factor for rain in terms of variables available from standard numerical model output. The equivalent reflectivity factors for both snow and graupel are derived by modifying 4.4 with the assumption that the snow and graupel particles are spheres of solid ice. Equation 4.5 relates

the diameter that a solid ice sphere would have to obtain the same reflectivity as a particle of snow (or graupel) that does not have constant density.

$$D_{solid} = D_s \left(\frac{\rho_S}{\rho_I}\right)^{\frac{1}{3}} \tag{4.5}$$

where $\rho_I$ is the density of ice and $\rho_s$ is the density of snow.

Taking the 6th power of equation 4.5 and including a factor that accounts for ice having a different reflective capacity than water (0.224) we have for snow

$$Z_e = \frac{720(\rho_{air}q_s)^{\frac{7}{4}}}{N_s^{\frac{3}{4}}(\pi\rho_s)^{\frac{7}{4}}} \times 0.224\left(\frac{\rho_s}{\rho_I}\right)^2 \tag{4.6}$$

and for graupel

$$Z_e = \frac{720(\rho_{air}q_g)^{\frac{7}{4}}}{N_g^{\frac{3}{4}}(\pi\rho_g)^{\frac{7}{4}}} \times 0.224\left(\frac{\rho_g}{\rho_I}\right)^2 \tag{4.7}$$

The reflectivity factors for rain, snow and graupel (equations 4.4, 4.6 , 4.7) can then be added to obtain the total reflectivity factor for each model grid cell. As reflectivity factors are typically measured in $mm^6 m^{-3}$ the total reflectivity factor is multiplied by $10^{18}$ to convert it from $m^6 m^{-3}$ to the standard units used for radar data. Finally, to convert the reflectivity factor measured in dB to the standard dimensionless unit (dBZ), the logarithm of the total equivalent reflectivity factor is multiplied by 10.

$$Z_e \text{ (in dBZ)} = 10log_{10}[Z_e(\text{in } mm^6 m^{-3})] \qquad (4.8)$$

The dBZ conversion algorithm has been designed to accept output from the WRF model (originally designed for MM5, the predecessor of WRF) and can be used to create simulated radar reflectivity fields for simulations using different microphysics scemes. When applied to simulations using Lin microphysics the intercept parameters $N_0$ must be provided for each precipitation category ($N_s = 3 \times 10^6$ $N_g = 4 \times 10^6$ $N_r = 8 \times 10^6$). When applied to simulations using Thompson microphysics the intercept parameters for each category must be calculated. In the Thompson scheme the snow intercept is a function of temperature (equation 4.9) and the rain and graupel intercepts are a function of the rain and graupel mixing ratios respectively (equations 4.10 and 4.11).

$$N_s = min\{3 \times 10^8, 2 \times 10^6 \times exp[-0.12min(-0.001, T_c)]\} \qquad (4.9)$$

$$N_g = max\{1 \times 10^4, 2.38(\tfrac{\pi \rho_g}{\rho_{air} q_g})^{0.92}\} \qquad (4.10)$$

$$N_r = (\tfrac{1 \times 10^{10} - 8 \times 10^6}{2})tanh[\tfrac{4(1 \times 10^{-4} - q_r)}{1 \times 10^{-4}}] + (\tfrac{1 \times 10^{10} + 8 \times 10^6}{2}) \quad (4.11)$$

After applying the conversion algorithm to both the Lin and Thompson simulations the simulated radar reflectivity data was interpolated to 20 equally spaced vertical levels (every 500m) and averaged horizontally (taking care to do this in

absolute reflectivity) so that the horizontal and vertical resolution of the simulated reflectivity data was comparable to the gridded radar data (see chapter 2 for more details).

The process of converting the microphysical data to simulated reflectivity described above is not fully consistent with the particle size distribution for snow within the Thompson microphysics scheme. The Thompson microphysics scheme uses the sum of two gamma functions to model the particle size distribution for snow particles. However, it has been discovered that the version of the Thompson microphysics code used in the simulations presented in this thesis contains a coding bug that produces unphysical results. Given that results derived from the Thompson simulations will remain unphysical even if the correct reflectivity conversion algorithm is used, the figures and text in the following chapters will continue to be based on results derived from reflectivity data generated using particle size distributions represented by the exponential distribution shown in Equation 4.1. In the summary section at the end of each chapter an attempt will be made, where possible, to outline how the results presented in this thesis are modified when the correct particle size distribution for snow is used in the reflectivity conversion algorithm. A new reflectivity conversion algorithm (received from Greg Thompson via private communication) has been used to recalculate figures for the Monsoon period only. Please note that as well as including the correct particle size distribution for snow, the new reflectivity conversion algorithm also contains a more sophisticated treatment of melting snow and graupel. Updated Figures for the monsoon period have been included in the Appendix.

## 4.3 Domain Averaged Results

As a first check to determine how the conversion algorithm was performing, maximum reflectivity and total precipitation coverage plots were created for both model

simulated and radar reflectivity data. As reflectivity is proportional to the diameter of the hydrometeors to the power of six, maximum reflectivity plots give an indication of how the model is performing during the most intense convective events. Maximum reflectivity plots were created by searching through the domain (either model or radar) and finding the maximum reflectivity value at each level. The total precipitation coverage plots were created by calculating the fraction of the domain (at each level) that was covered by a reflectivity value greater than 0 dBZ.

### 4.3.1   Monsoon Simulation

Figure 4.1 shows the maximum reflectivity and total precipitation coverage during the monsoon simulation for the radar a) and b), the Lin microphysics scheme c) and d) and the Thompson microphysics scheme e) and f). Besides some notable differences, the model simulated maximum reflectivity compares quite well with the radar derived maximum reflectivity profiles for both the Lin and Thompson simulations. During the first half of the simulation the drop-off of high reflectivity values (50-60 dBZ in red) above the freezing level (approximately 5km) is captured well for both the Lin and the Thompson simulations. In general the Lin simulation produces larger reflectivity values at higher altitudes than observed by the radar; during the passage of the MCS the Lin simulation has reflectivity values between 50 and 60 dBZ (red) all the way up to 9 km whereas the radar observes these values up to approximately 7 km. Reflectivity values between 40-50 dBZ (orange) can be seen all the way up to 15 km in the Lin simulation compared with 12 km for the radar.

The Thompson scheme performs better than the Lin scheme for maximum reflectivity values between 40-60 dBZ (red to orange), but starts to differ from the observations in the 30-40 dBZ range (yellow), underestimating the height of these values/particles by as much as 6 km. At lower reflectivity values (below 30 dBZ i.e.

blueish colors) the Lin scheme seems to perform quite well in terms of the height that hydrometeors with these reflectivity values reach.

Figures 4.1 b) d) and f) highlight some important difference between the Lin and the Thompson simulations that cannot be seen in the maximum reflectivity plots. The Thompson simulation has a drastic drop-off of precipitation coverage below 5 km that is not present in either the radar observations or the Lin simulation. Furthermore, coverage above the freezing level is actually higher in the Thompson simulation on the 24th of January than occurs at any stage during the Lin simulation. As the freezing level is at approximately 5 km the drop-off in coverage is likely an indication of a problem with the conversion between snow and rain/graupel in the Thompson scheme. However, at this stage all that can be said is that the Thompson simulation appears to have a significant problem representing the coverage of precipitating particles below the freezing layer.

In general both simulations underestimate the precipitation coverage during the first half of the monsoon period (specifically the 20th-22nd). In chapter 3 the vertical velocity profiles were found to be much too weak during this time period (possibly due to subsidence in the 12-18km layer due to incorrect forcing data). Hence, it is not unexpected that precipitation coverage would be underestimated during this time. During and after the passage of the MCS the Lin simulation produces precipitation coverage profiles that are in general agreement with the observations. There are however some notable differences: the observed increase in precipitation coverage just before the 23rd is present in the Lin simulation at the correct time, but drops off too quickly above the freezing level and is significantly underestimated by the 7 km level. During the first 12 hours on the 23rd of January the radar observes relatively little precipitation coverage (approximately 10-20% except at the lowest levels). While the Lin simulation shows a slight drop-off of coverage for a few hours, it soon begins to increase again and never reaches the values that are observed by

the radar.

In summary the Lin simulation overestimates the coverage in time (it starts too early) and underestimates the precipitation coverage with height, especially at higher altitudes. The underestimation of precipitation at higher altitudes is likely due to the incorrect forcing in the GFS data. The overestimation in time may also be linked to the bias, a decrease in the convective strength due to capping of the convection by the warm layer may lead to a decreased rain efficiency. Precipitation in the Thompson simulation appears to suffer from both the bias in the large-scale forcing and an inherent problem within the microphysics scheme itself. The coverage of precipitation appears to be sandwiched between 5 and 10 km, the temperature bias above 10 km and a problem with rain below the freezing level.
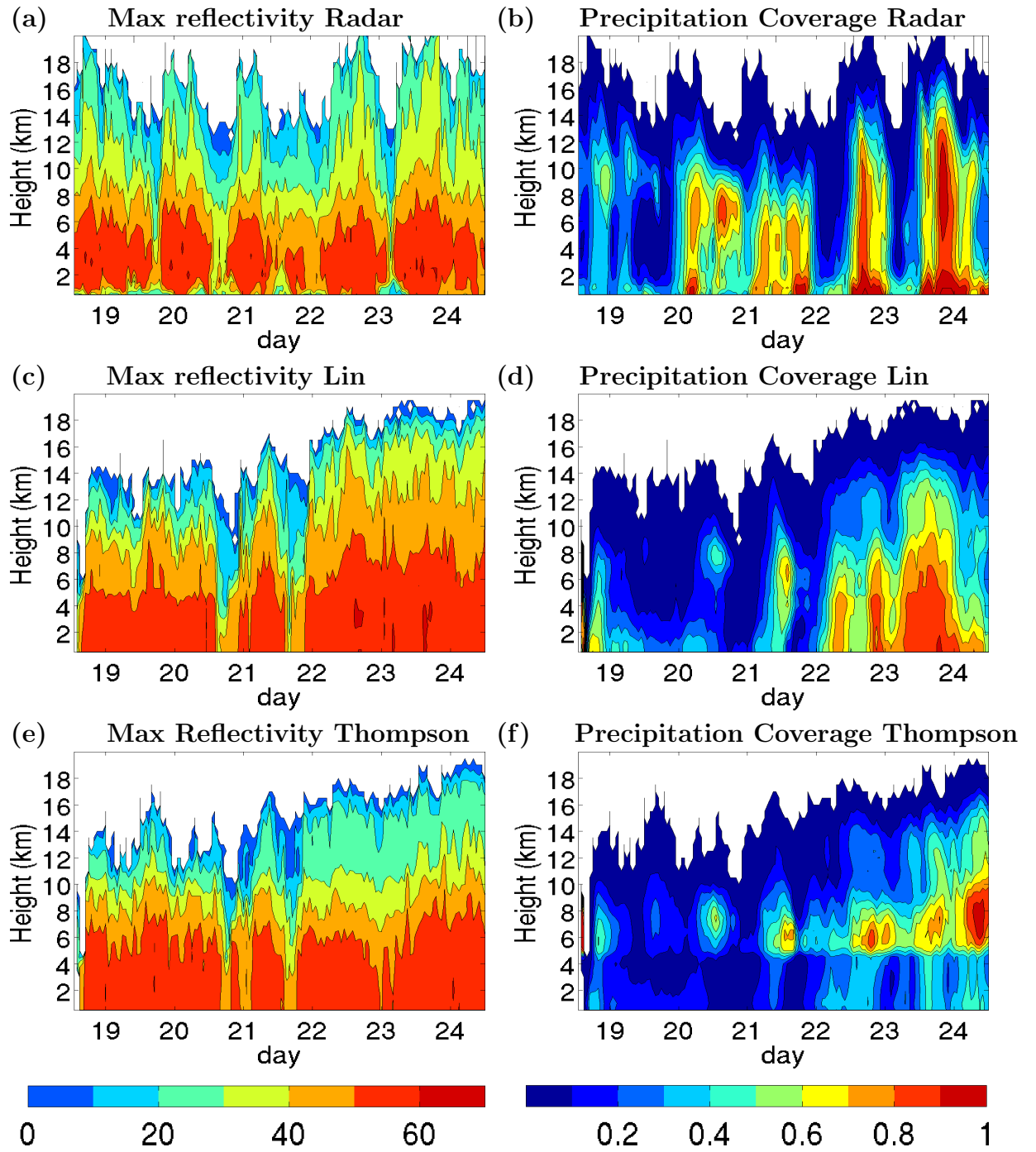
Figure 4.1: **Maximum Reflectivity and Precipitation Coverage profiles for the Monsoon period derived from the radar a) and b, the Lin simulation c) and d) and the Thompson simulation e) and f)**

### 4.3.2 Monsoon/Suppressed Simulation

Suppressed monsoon conditions occurred during TWP-ICE when the MCS that spun up during monsoon period moved to the south of Darwin and transitioned into a tropical low. Circulation about the low wrapped dry air from over the continent into the domain, suppressing much of the convection typically present during the westerly wind conditions.

Figure 4.2 shows the maximum reflectivity and total precipitation coverage during the suppressed monsoon simulation for the radar a) and b), the Lin microphysics scheme c) and d) and the Thompson microphysics scheme e) and f). This simulation was initialised on the 23rd of Jan (12 UTC) as the MCS was passing through the model domain. As a consequence of the model spin-up the precipitation event is delayed somewhat in this simulation and the precipitation associated with the MCS does not finish until the 25th, see figure 4.2 b) and e). As the domain enters the suppressed monsoon period the Lin microphysics captures the low level (below 2km) precipitation coverage well.

Both the Lin and Thompson microphysics overestimate the highest maximum reflectivity values (red), that regularly extend all the way to the freezing level, a feature rarely observed by the radar (except at the start of the simulation when the MCS is still in the TWP-ICE domain). As with the monsoon simulation the Thompson microphysics has major problems below the freezing level during the first two days of the simulation, however the problem is much less obvious in the suppressed conditions. Precipitation events during the suppressed conditions are likely to be very different to those occurring during the monsoon conditions. During suppressed conditions it is expected that precipitation is generated primarily through weak convection, typical of that associated with regime 1 in chapter 2. Whereas in monsoon conditions the precipitation generated is likely to be similar to regimes 3 and 4. It is difficult to tell from figure 4.2 f) whether the Thompson microphysics is

underestimating precipitation below the freezing level, chapters 5 and 6 will investigate the possibility that the problem below the freezing level is dependent on the synoptic conditions or the types of clouds present.

Figure 4.2: **Maximum Reflectivity and Precipitation Coverage profiles for the Monsoon/Suppressed period derived from the radar a) and b, the Lin simulation c) and d) and the Thompson simulation e) and f)**

### 4.3.3  Suppressed Simulations

Figure 4.3 shows the maximum reflectivity and total precipitation coverage between the 28th of January to the 9th of Feburary for the radar a) and b), the Lin microphysics scheme c) and d) and the Thompson microphysics scheme e) and f). The results shown in Figure 4.3 are derived from two simulations, the first initialized on the 28th of January and the second initialized on the 3rd of February; the black vertical line denotes the begining/end of the second/first simulation.

Throughout the entire period (28th January - 9th of February) the radar observes relatively high precipitation coverage capped at 2 km, on 3 occasions (1st, 6th and 9th of Feb) deeper convection can be seen breaking through the 2 km level and then past the freezing layer. During the first 5 days the Lin simulation does better than the Thompson simulation at producing the low level precipitation coverage, however the convection often reaches 3-4 km rather than being capped at 2 km.

On the 1st of February the radar observes convection breaking through the freezing level leading to significant precipitation coverage (for suppressed conditions). The Lin simulation does show an increase in precipitation coverage at the correct time, but the coverage drops off way too quickly above the freezing layer. The Lin and the Thompson simulations capture well the transition from the suppressed monsoon conditions to the "clear" regime on approximately the 3rd of February, but fail to stay in the clear regime as long as is observed. Relatively deep convection (past 10 km) occurs in both simulations on the 5th of February, a day earlier than observed. It is interesting to note that while the problem in the Thompson physics below the freezing level is still present, it is much less obvious in these simulations. This will be further investigated in chapter 6.

Figure 4.3: **Maximum Reflectivity and Precipitation Coverage profiles for the Suppressed and Suppressed/Break period derived from the radar a) and b, the Lin simulation c) and d) and the Thompson simulation e) and f)**

### 4.3.4   Break Simulation

Figure 4.4 shows the maximum reflectivity and total precipitation coverage during the break simulation for the radar a) and b), the Lin microphysics scheme c) and d) and the Thompson microphysics scheme e) and f). The maximum reflectivity plots show that the convection is very intense during this period, the radar observes high reflectivity values (red) up to 7 km daily. On the 10th and the 15th of February high reflectivity values (red) reach 8 km and reflectivity values between 40-50 dBZ (orange) can be seen all the way up to 15 km. These reflectivity values extend much higher than any other period in this study, which is consistent with break period convection being more intense than convection occurring in active monsoon conditions. The simulated maximum reflectivity values (red and orange) also extend to higher altitudes than in previous simulations. However, the height that the highest reflectivity values (red) reach is actually overestimated in both simulations by as much as 2-3 km.

The timing of the convection in the model simulations is particularly poor during the break period. The spin-up period for the model is likely the reason that the convection observed by the radar on the 9th is not present in either simulation. However, the model simulations are out of phase with the radar observations until the 14th of February. The model simulations produce convection during the afternoon when the radar observes no deep convection, then the radar observes convection during the night which the simulations fail to capture. The reason the WRF model is unable to simulate the nocturnal precipitation during this period will be explored in Chapter 5. If the timing issue of the convection is ignored the Thompson simulation seems to be producing more realistic precipitation coverage above the freezing level than the Lin simulation.

## 4.4   Summary

This chapter has explored how data from the WRF microphysics schemes can be converted into simulated radar reflectivity. Domain-averaged results have highlighted a number of differences between the WRF simulations and radar observations. Some of these differences include an underestimation of precipitation coverage at higher altitudes in both the Lin and the Thompson simulations, an overestimation of maximum reflectivity values at higher altitudes in the Lin simulations, and a significant problem in the Thompson scheme at simulating precipitation coverage below the freezing level. Some of these differences are likely due to inaccuracies in the WRF simulations themselves, rather than an artifact of the dBZ conversion algorithm. For example the underestimation of precipitation coverage at higher altitudes can be explained as a consequence of the inaccurate forcing data used to drive the model simulations. Therefore, despite the differences stated above, the general agreement between the observed and simulated reflectivity data provides enough incentive to continue using the simulated reflectivity data for model evaluation.

Results for the Thompson simulation when the correct particle size distribution is used in the reflectivity conversion algorithm are shown in Figure 5 of the appendix. These figures show that while some details differ, the conclusions of this chapter remain largely unchanged. Specifically, the under-representation of precipitation below the freezing level is still evident. Some notable changes that do occur include decreased precipitation coverage at higher altitudes (above approximately 10 km) and increased maximum reflectivity at and above the freezing level. The increased maximum reflectivity values around the freezing level is likely due to the more sophisticated treatment of melting snow and graupel in the new reflectivity conversion scheme, while the decreased precipitation converge at higher altitudes is a consequence using the correct particle size distribution for snow.
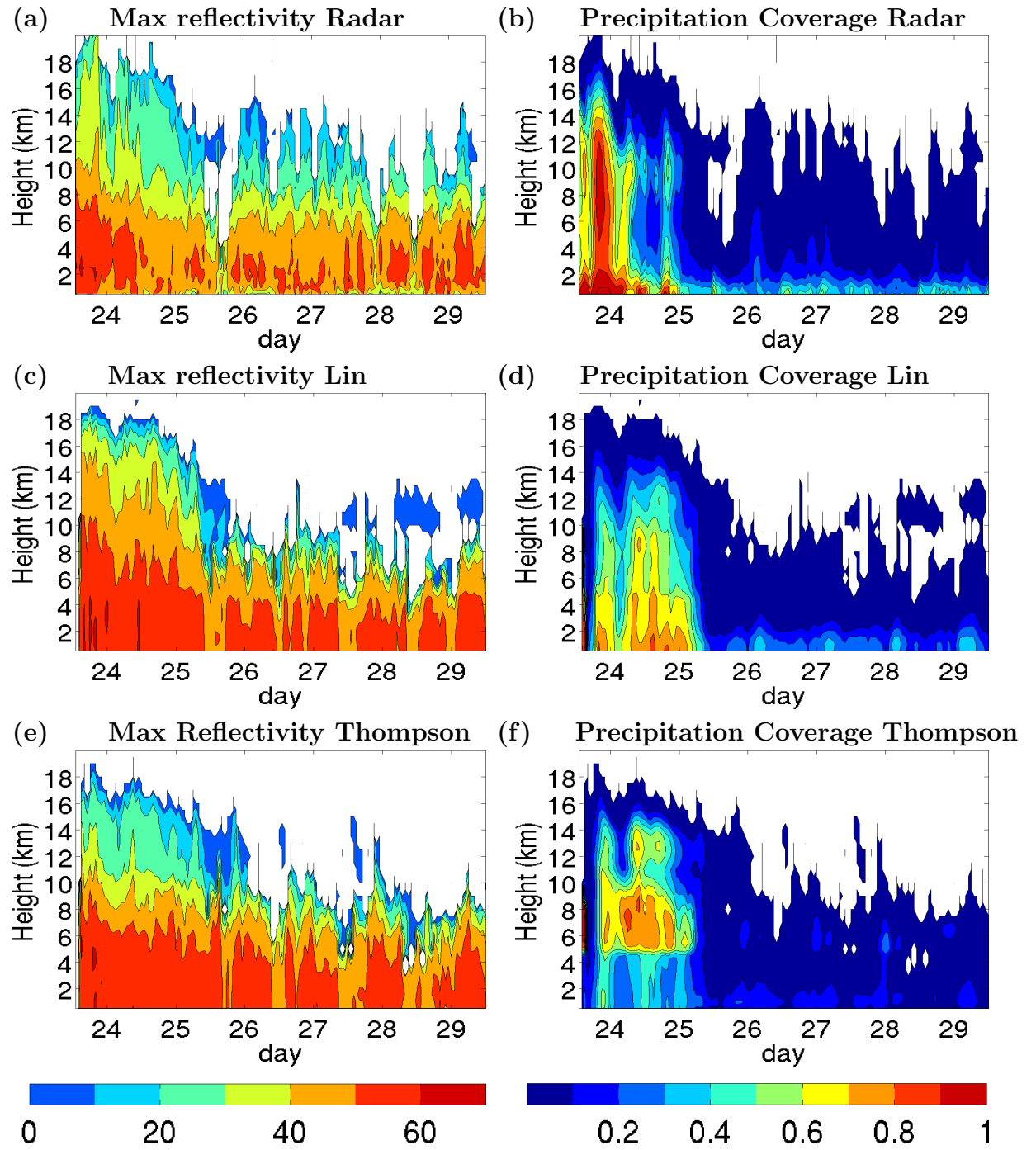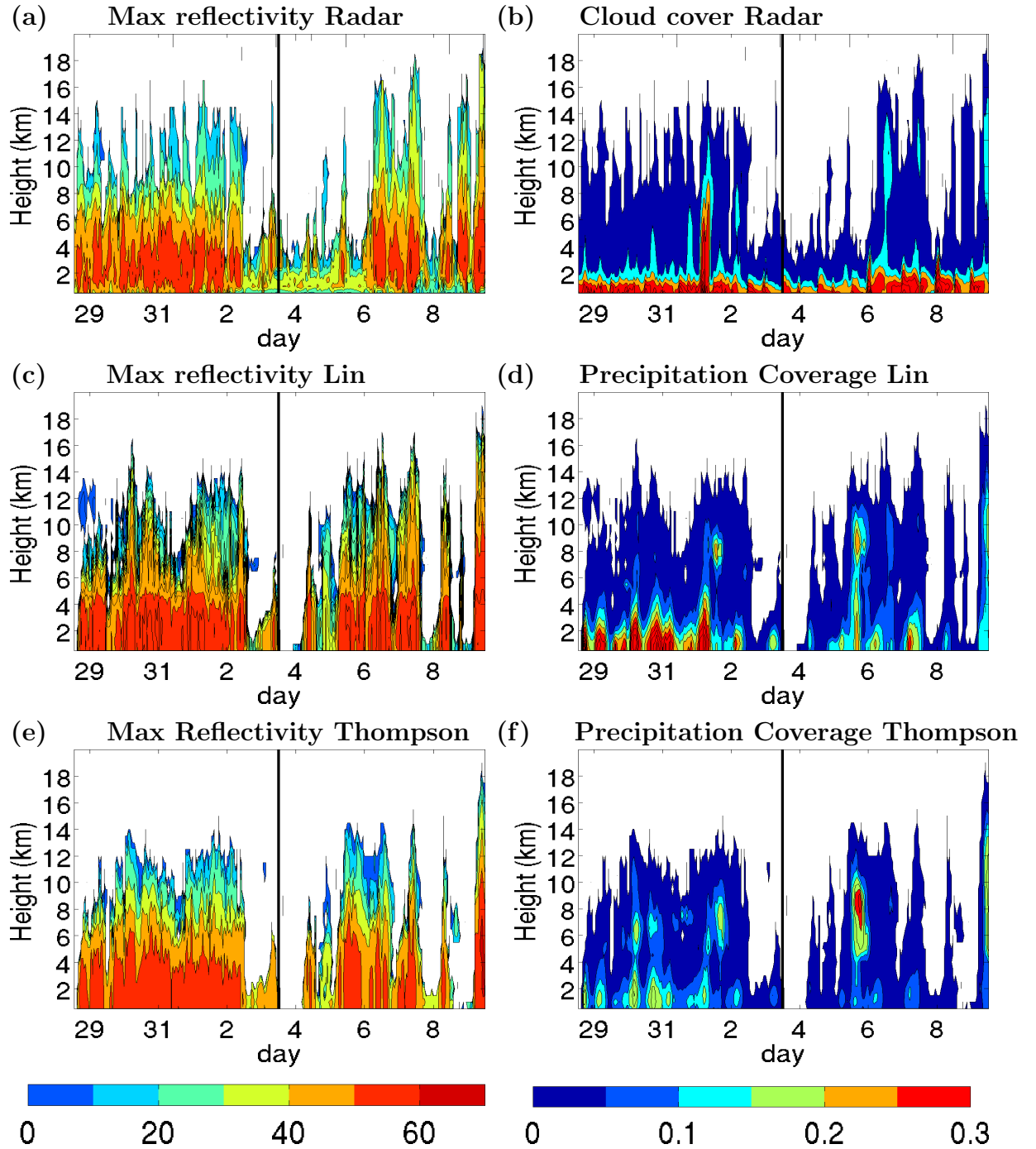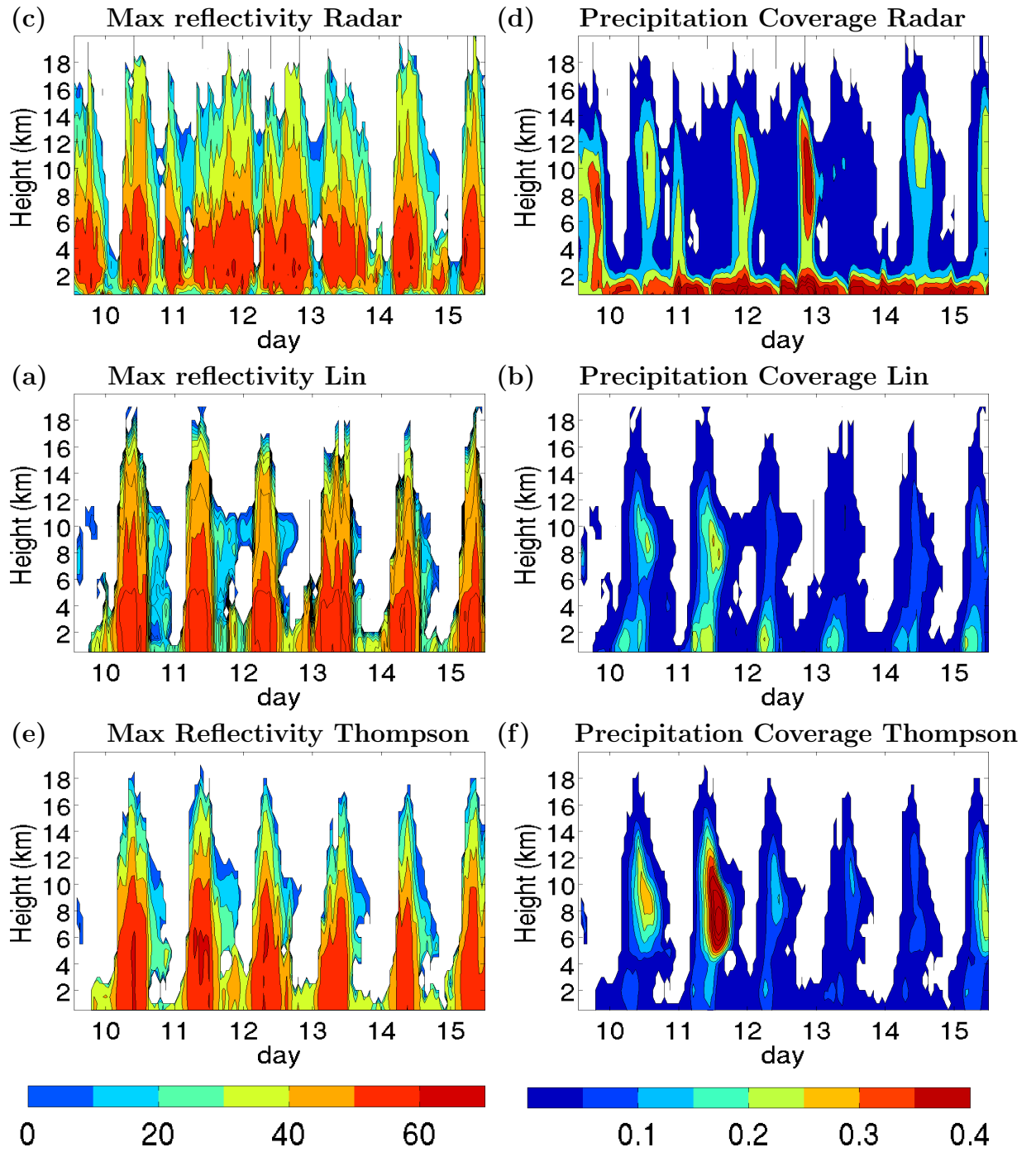
Figure 4.4: **Maximum Reflectivity and Precipitation Coverage profiles for the Break period derived from the radar a) and b, the Lin simulation c) and d) and the Thompson simulation e) and f)**

# Chapter 5

# Precipitation

## 5.1   Introduction

Using rain gauges to evaluate the precipitation in a numerical model is of limited value as individual rain gauges cannot be resolved in a numerical model, and using the closest model grid-point to the rain gauge location is not ideal; a point by point correlation between model simulations and reality is well beyond the capacity of current numerical models. Furthermore, to accurately represent the spatial distribution of precipitation over a large region, particularly during convective showers, large dense rain gauge networks are required. One of the major advantages of radar is the large spatial extent over which hydrometeors can be detected. By exploiting Z-R relationships (relationships between absolute reflectivity and rain rate) it is possible to derive precipitation rates on the ground (in mm/h). Many authors of have investigated Z-R relationships at different locations and used large dense rain gauge networks to calibrate and evaluate the performance of the statistical relationships used (Gunn and Marshall, 1958; Fujiyoshi et al., 1990; Brandes et al., 2002; Crosson et al., 1996; Bringi et al., 2001). In this study the 2.5 km level radar reflectivity values are converted to rain rates on the ground using the formula $Z = 305R^{1.36}$, where Z is the absolute reflectivity and R is the rain rate in mm/hour. This formula was calibrated using rain gauge and data from the Gunn Point radar in Darwin (Bringi et al., 2001) and should therefore be ideal for use in this study.

This chapter explores whether the spatial distribution of precipitation averaged over each simulation can be used to evaluate the WRF simulations, as well as high-

light any differences between the Lin and Thompson microphysics schemes. By linking the spatial distribution of precipitation averaged over each simulation with the maps derived for the four precipitation regimes in chapter 2, a first attempt at linking the simulations with these regimes is provided. Links between the simulations and precipitation regimes will be explored in more detail in chapter 7.

## 5.2   Results

### 5.2.1   Monsoon Simulation

Figure 5.1 a) shows the time series of domain-averaged precipitation rates (left axis) and accumulated precipitation (right axis) for the radar as well as the Lin and Thompson simulations. In Chapter 4 the Thompson simulation was found to underestimate precipitation coverage below the freezing level, consequently precipitation rates derived from the 2.5 km model data are also likely to be underestimated. Any differences between the Lin and Thompson precipitation rates may provide valuable insight into the representation of precipitation below the freezing level in the Thompson microphysics scheme.

During the first half of the monsoon period there is very little difference between the two simulations. Both drastically underestimate the precipitation produced on the 20th of January and the precipitation event on the 21st of January is delayed by approximately 3-4 hours in both simulations. The main difference between simulations occurs after the 22nd of January. The Thompson simulation underestimates the peak rain rates on the 22nd, but adequately captures the observed decrease in precipitation on the 23rd. The Lin simulation overestimates the peak precipitation rates on the 22nd and fails to capture the full extent of the decreased precipitation observed on the 23rd. During the passage of the MCS the Lin simulation fails to reach the peak rainfall rates observed by the radar. However, higher precipitation

rates over the 24 hour period between 12 UTC on the 22nd and 12 UTC on the 23rd lead to an overestimation of the accumulated precipitation by as much as 50 mm. The Thompson simulation on the other hand drastically underestimates the precipitation generated during the passage of the MCS, leading to an underestimation of the accumulated precipitation by approximately 70 millimeters. That the Thompson simulation underestimates precipitation rates comes as no surprise when considering the deficiency in precipitation coverage highlighted in chapter 4. However, precipitation rates do not appear to be severely affected during the first 4 days of the simulation. Only during the passage of the MCS do significant differences between the precipitation rates in the Thompson and Lin simulation occur.

It therefore seems plausible that the problem in the Thompson microphysics may be linked with a physical property or process associated with the precipitation generated by the MCS. During the passage of the MCS it was seen that precipitation coverage (in particular that of stratiform precipitation) was significantly higher than any other time in the simulation (see Figure 4.1 f)), while precipitation coverage was a minimum during the first 4 days of the simulation. At this stage there is not enough information to determine whether the problem with the Thompson microphysics is primarily associated with stratiform precipitation. It will be treated as a working hypothesis and left to later chapters to explore.

Figure 5.1 b), c) and d) show the spatial distribution of precipitation averaged over the entire simulation for the radar, Lin and Thompson simulation respectively. Domain-averaged precipitation rates are shown on the top of each figure. The Lin simulation is able to roughly simulate the spatial distribution of precipitation, correctly capturing the high precipitation rates observed over the ocean in the west of the domain (although the region of high precipitation rate extends too far north). In general precipitation is over estimated everywhere in the Lin simulation, leading to an overestimate of the domain-averaged rain rate (2.1 mm/hour compared

with 1.7 mm/hour observed by the radar). Low precipitation rates observed by the radar over land in the east of the domain are not present in the Lin simulation. The Thompson simulation shows little skill in predicting the spatial distribution of precipitation, there is no sign of the observed precipitation maximum in the west of the domain and erroneously high precipitation rates occur to the north of the Tiwi Islands.

Comparing figures 5.1 b), c) and d) with the rain rate maps found for the four precipitation regimes (Figure 2.5) it can be seen that the spatial distribution of the radar and Lin precipitation appears quite similar to the precipitation pattern for regime 4. In Chapter 2 it was found that regime 4 was comprised of strong convection with large quantities of stratiform precipitation. Assuming the hypothesis that the Thompson physics has problems simulating stratiform precipitation is correct, it naturally follows that the higher precipitation rates seen in the west of the domain in the Lin and radar figures are due to stratiform precipitation.

Figure 5.1: **a) Time series of domain-averaged precipitation rates and accumulation over the monsoon period for the radar, Lin and Thompson simulations. b) Spatial distribution of precipitation rates averaged over the monsoon period for the radar c) the Lin simulation and d) the Thompson simulation.**

## 5.2.2   Monsoon / Suppressed Simulation

Figure 5.2 shows the average precipitation rates derived from radar and model simulated reflectivity over the monsoon / suppressed monsoon simulation. Once again the Lin simulation produces more precipitation than the Thompson simulation. However, during this period the domain-averaged precipitation rate in the Thompson simulation is much closer to the radar derived rain rates. This does not necessarily mean that the Thompson microphysics is simulating precipitation in a more realistic fashion than the Lin microphysics scheme. One possible cause for the close agreement is compensating errors. Figure 5.2 a) shows that the Thompson simulation underestimates the precipitation associated with the event on the 23rd, but this deficiency is made up by the precipitation generated between the 24th and 25th of January. The high precipitation rates between the 23rd and the 24th are likely artifacts of the model as they do not correspond to precipitation events observed by the radar. It is probably good fortune that the underestimation of precipitation on the 23rd (most likely due to the problem in precipitation below the freezing level) is made up the precipitation during the 23rd to the 24th (which is not observed and would likely be even be higher if not for the problem below the freezing level).

Figures 5.2 b) c) and d) show that the model simulations do not agree well with the observed spatial precipitation pattern. The simulations fail to capture the observed decrease in precipitation in the eastern half of the domain and a line of high precipitation rates orientated along the coast (more obvious in the Lin simulation) is not observed by the radar. To investigate the differences between the model and radar precipitation patterns in more detail Figure 5.3 shows precipitation maps for the radar and model simulations over different time periods.

Figures 5.3 a), d), and g) show precipitation maps for the 12 hour period between the start of the simulation and 0 UTC on the 24th of January, during which the largest precipitation event occurs. Figures 5.3 b), e) and h) show the precipitation

maps for the 36 hour period between 0 UTC on the 24th and 12 UTC on the 25th, during which both model simulations produce unobserved precipitation events and Figures 5.3 c), f) and i) show the precipitation maps between 12 UTC on the 25th and the end of the simulation. From these figures it can be seen that the high precipitation rates orientated along the coast-line occur primarily during the first 12 hours of the simulation. Normally the first 12 hours of a simulation would be defined as the spin-up period for the model and the data would be rejected. However, in this simulation the unrealistic convection generated during the spin-up period (which has the appearance of a land/sea breeze effect) appears to affect the precipitation for the next 36 hours. After 12 UTC on the 25th both simulations appear to perform quite well, both capture the patchy nature of convection throughout the domain and the higher precipitation rates in the south west of the domain. During this time the Thompson simulation actually produces more precipitation than the Lin simulation. Given that stratiform precipitation has little to no effect on precipitation rates during suppressed conditions, this result appears consistent with the previous stratiform hypothesis.

During suppressed conditions it would be expected that regime 1 would be the most prevalent of the four precipitation regimes, regime 1 was found to be the weakest of the convective regimes and had low precipitation coverage and low domain-averaged precipitation rates. Indeed the spatial distribution of the radar and both model simulations appears similar to that found for regime 1 (see Figure 2.5) for the period between 12 UTC on the 25th and the end of the simulation.

In conclusion the first 48 hours in these simulations are greatly affected by the initial spin-up period in the model. Given that the model was initialized during the passage of the MCS is is not surprising that the model would be affected by the spin-up period. However, the duration of the spin-up effect is well beyond what would normally be expected for the WRF model. After the first 48 hours both

model simulations perform well and are able to simulate the spatial distribution of precipitation with some skill. Similarities between the spatial patterns of precipitation and regime 1 indicate that stratiform precipitation contributes little to the overall rain rate during the latter half of the simulations, and during this period the Thompson simulation does not seem to be adversely affected by the problem below the freezing level highlighted in chapter 4.
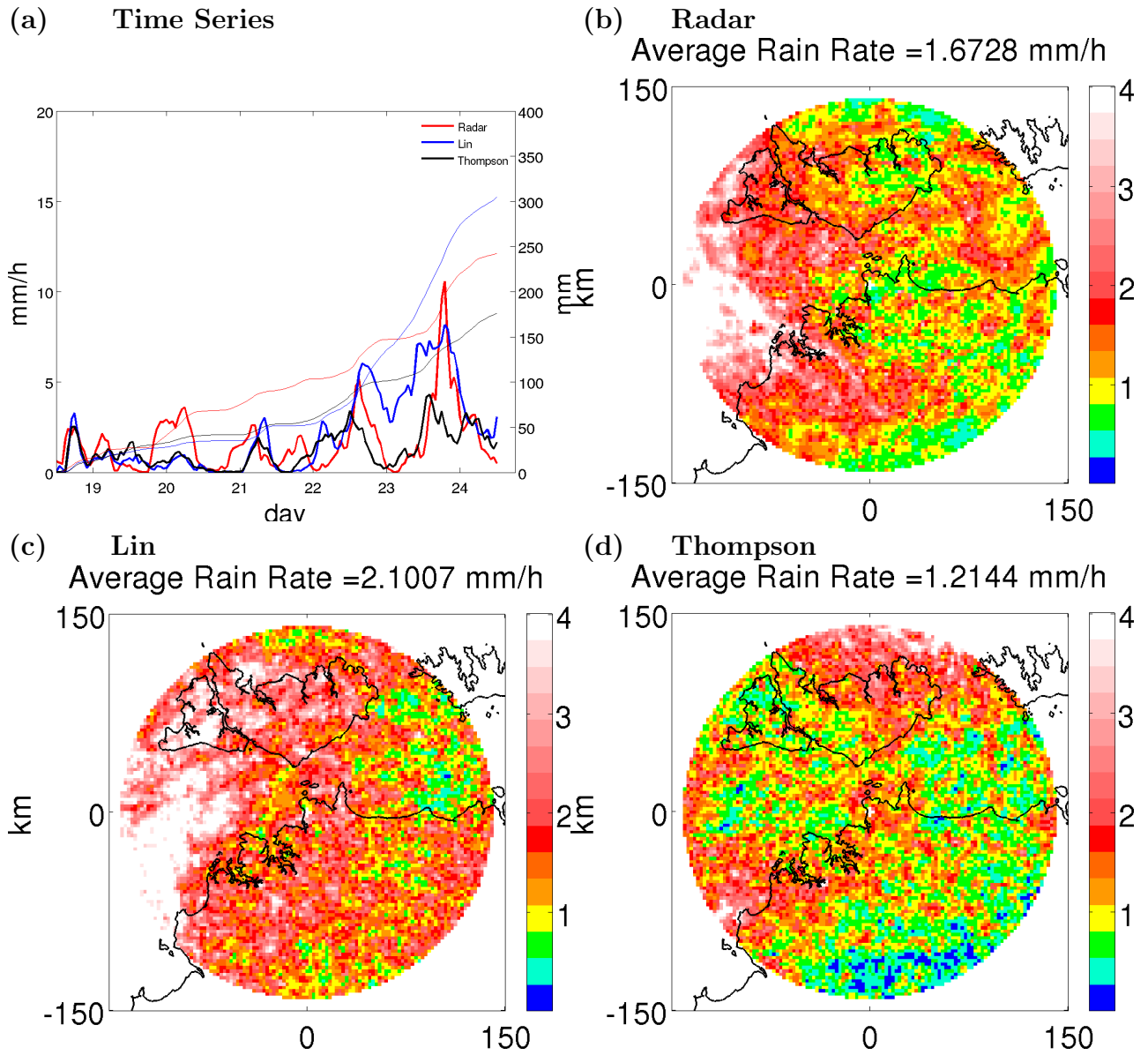
Figure 5.2: **a) Time series of domain-averaged precipitation rates and accumulation over the monsoon/suppressed period for the radar, Lin and Thompson simulations. b) Spatial distribution of precipitation rates averaged over the monsoon/suppressed period for the radar c) the Lin simulation and d) the Thompson simulation.**
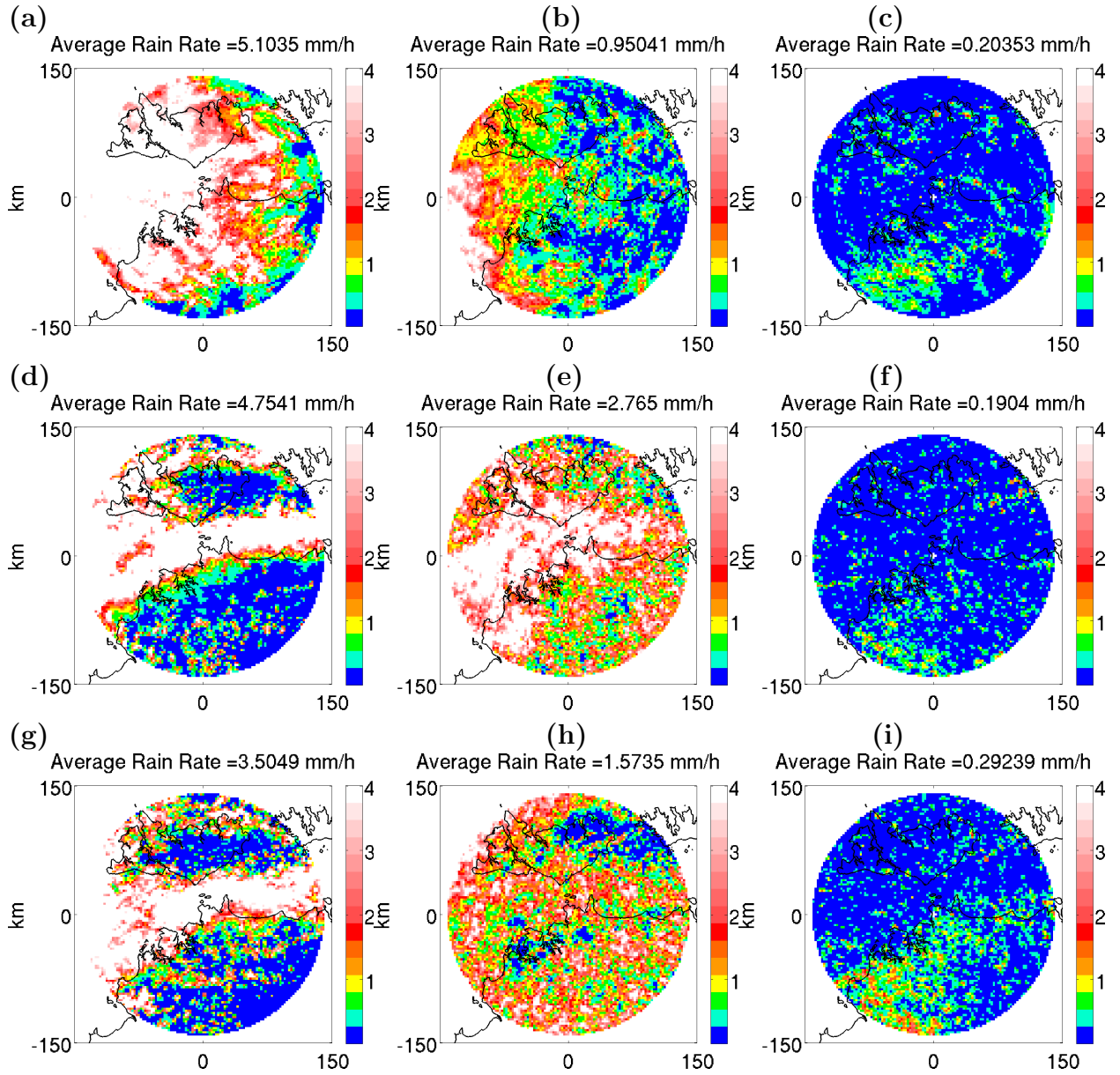
Figure 5.3: **Spatial distribution of precipitation for the first 12 hours of the monsoon/suppressed monsoon simulation for a) radar, d) Lin simulation g) Thompson simulation. Spatial distribution of precipitation occuring between the 0 UTC on the 24th to 12UTC on the 25th of January for b) radar, e) Lin simulation h) Thompson simulation. Spatial distribution of precipitation occuring between 12UTC on the 25th till the end of the simulation for c) the radar, f) Lin simulation i) Thompson simulation.**

### 5.2.3 Suppressed and Suppressed/Break Simulations

Figure 5.4 shows the precipitation rates derived from radar and model simulated reflectivity averaged over the suppressed monsoon to break period (two simulations spanning the 28th of January to the 9th of February). In contrast to the previous periods investigated, both the Thompson and Lin microphysics schemes produce similar precipitation rates throughout the entire period. However, both overestimate the precipitation generated when compared with the radar. Accumulated precipitation is greater in the Thompson simulations from the 28th of January till the 6th of February, but during the last three and a half days higher precipitation rates in the Lin simulation leads to a higher total accumulation by the end of the period. In general both microphysics schemes capture the timing of the precipitation events well, with their biggest error being the overestimation of precipitation between the 30th and 31st of January. Following this period of overestimated precipitation both models completely miss the observed precipitation event that occurs just before the 1st of February. During the last 12 hours of this period the Lin simulation over estimates the precipitation rate by a factor of two. The Thompson simulation also over estimates the precipitation during this time, but not by as much as the Lin simulation. Both the Thompson and the Lin simulation miss the precipitation event during the night on the 8th of Feb.

The spatial patterns of precipitation show that the Lin and the Thompson simulations overestimate the precipitation on the mainland. The Lin simulation clearly shows the region of high precipitation centered around the Tiwi Islands (in the north of the domain), but slightly overestimates the precipitation rates there. Given that there is a preference for precipitation over land and precipitation rates are quite low, it appears that the majority of the precipitation produced during this period is diurnally driven weak convection. The spatial distribution of precipitation in both models and the radar resembles that found for regime 1. Given that regime 1 was

found to be a convective regime with very little stratiform precipitation, the similar precipitation rates/accumulations between the Lin and Thompson simulations seem to indicate that stratiform hypothesis has some merit. During times where there is relatively little stratiform precipitation the Lin and Thompson simulations are most similar.
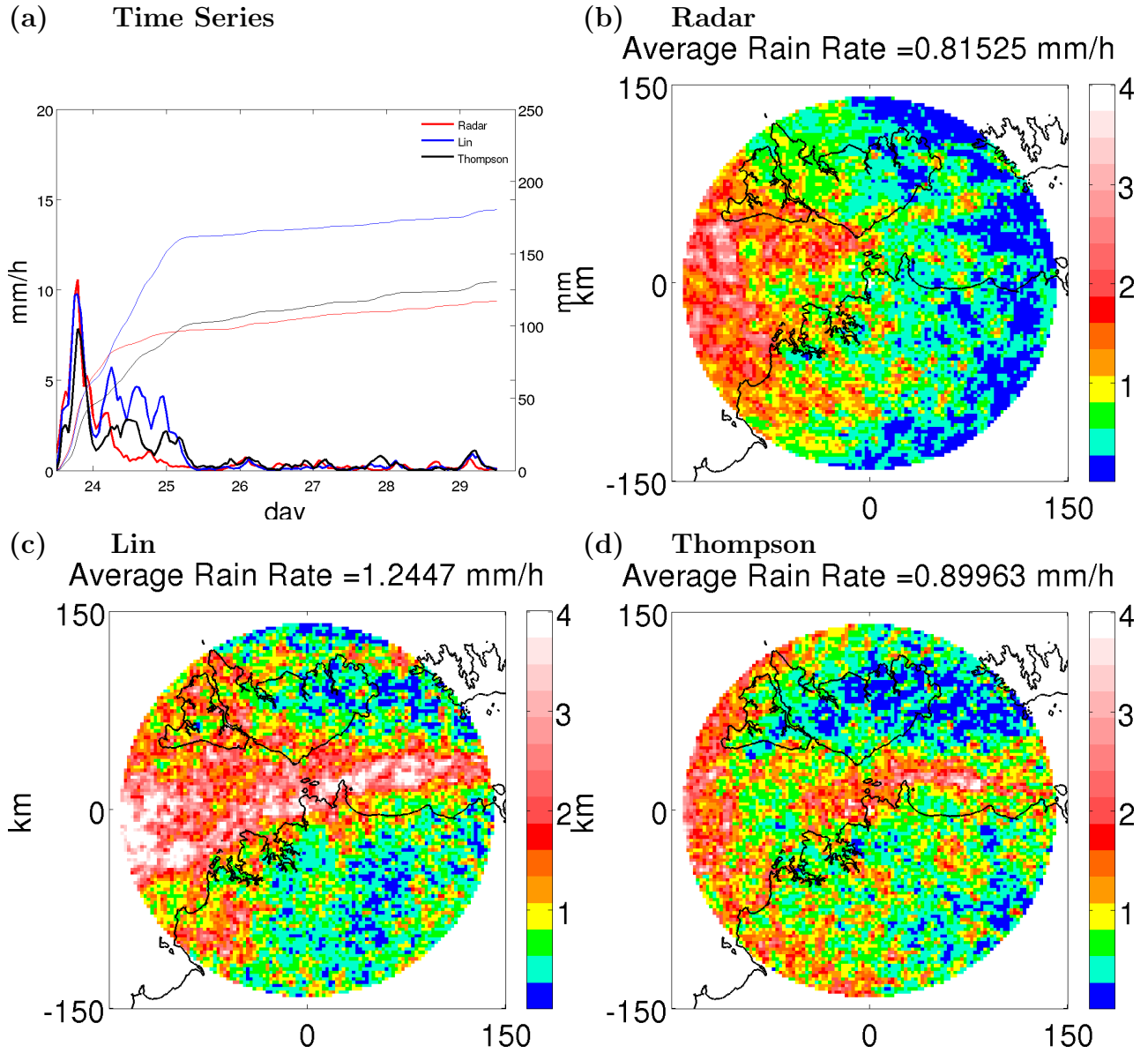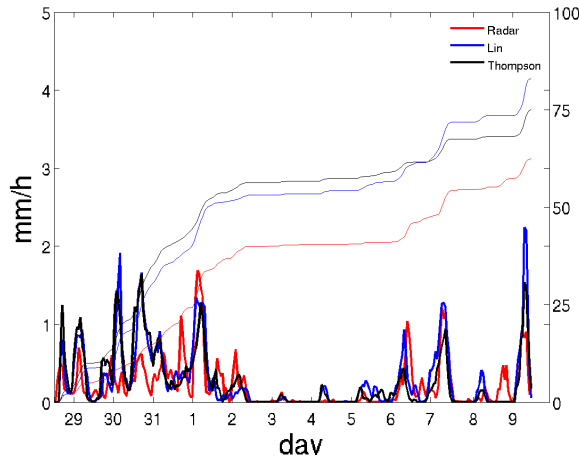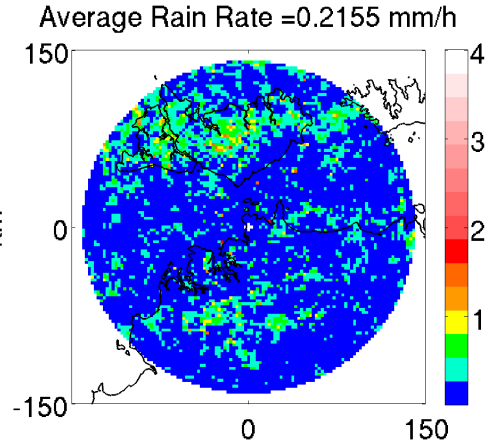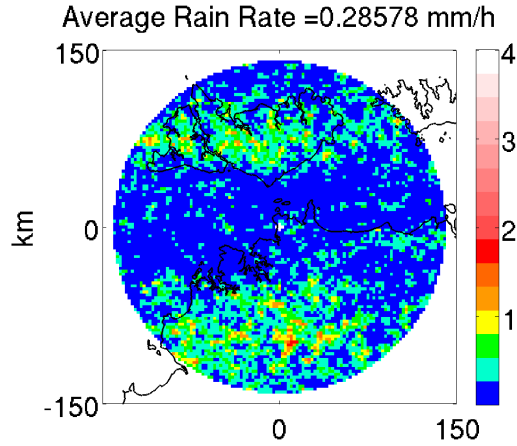


Figure 5.4: **a) Time series of domain-averaged precipitation rates and accumulation over the suppressed period for the radar, Lin and Thompson simulations. b) Spatial distribution of precipitation rates averaged over the suppressed period for the radar c) the Lin simulation and d) the Thompson simulation.**

### 5.2.4 Break Simulation

Figure 5.5 shows the average precipitation rates derived from radar and model simulated reflectivity over the break period. Unlike the previous simulations where the biggest problem with the simulations was an over/under estimation of the precipitation produced within an event, in these simulations there are serious timing issues associated with the precipitation. The models perform well on the 10th, 13th, 14th and 15th of February, when the precipitation is diurnal in nature and occurs during the afternoon, though the Lin simulation does overestimate the precipitation on the 10th. When the observed precipitation occurs during the night both models fail to simulate the events, instead they continue to produce precipitation during the day that is not observed by the radar. On the 11th and 12th the radar does produce a small spike in precipitation associated with the diurnal heating, however very little precipitation is produced.

Figures 5.5 b), c) and d) show that the precipitation in both model simulations occurs almost exclusively over land and both simulations fail capture the oceanic precipitation observed by the radar. Both the model simulations underestimate the domain-averaged precipitation over the break period simulation, which is not surprising considering the simulations fail to capture the correct number of precipitation events. To help explain why the model is performing so poorly during the break period Figure 5.6 shows CAPPI images from the radar during the beginning of each nocturnal precipitation event. Figures 5.6 a), b) and c) show the 2.5 km CAPPI on the 10th of February at 20, 21 and 22 UTC respectively. Figures 5.6 d), e) and f) show the 2.5 km CAPPI on the 11th of February at 18, 19 and 20 UTC and Figures 5.6 g), h) and i) show the 2.5 km CAPPI on the 12th of February at 16, 17 and 18 UTC. From these figures it can be seen that the systems producing the nocturnal precipitation are not initiated in the TWP-ICE domain. On the 10th of February a squall line enters the domain from the south east and on the 11th and

12th of February convective systems enter the domain from the east. The fact that these systems are initiated elsewhere and propagate into the domain explains why the model completely misses the events. If these systems require high resolution to initiate and/or propagate, then the model would be unable to correctly simulate these systems in the outer domains, and consequently they would not enter the cloud resolving domain.

## 5.3    Summary

In this chapter it was shown that the Lin microphysics had a tendency to overestimate precipitation, but was able to adequately capture the spatial distribution of precipitation in most cases. Times in which the model was unable to simulate the spatial distribution of precipitation were attributed to either model spin-up, during the first 48 hours of the Monsoon/suppressed monsoon simulation, or the models inability to initiate convective systems in the outer domains and propagate these systems into the cloud resolving domain.

Despite the underestimation of precipitation coverage found in earlier chapters, the Thompson simulations were found to be remarkably similar to the Lin simulations, with the exception of the last two days in the monsoon period. During the last two days of the monsoon period a MCS generated large amounts of precipitation that was drastically underestimated by the Thompson microphysics. It was hypothesized that the deficiency in the precipitation coverage first seen in Chapter 4 may be associated with poor representation of stratiform precipitation in the Thompson scheme.

Figure 6 of the appendix shows updated figures for the Monsoon period. The conclusions of this chapter remain unchanged when the correct particle size distributions for snow is used in the reflectivity conversion algorithm. The precipitation rates averaged over the domain (Figure 6 a)) are very similar to those obtained when

the exponential particle size distribution was used. Likewise, the spatial distribution of precipitation is also similar to the results obtained when exponential particle size distributions were used. However, the results obtained using the new conversion algorithm show that the variability of rainfall within the domain has increased. Regions of high rainfall rates (above 2 mm/hour) tend to be heavier than previously documented, while regions with low rainfall rates tend to be lighter.
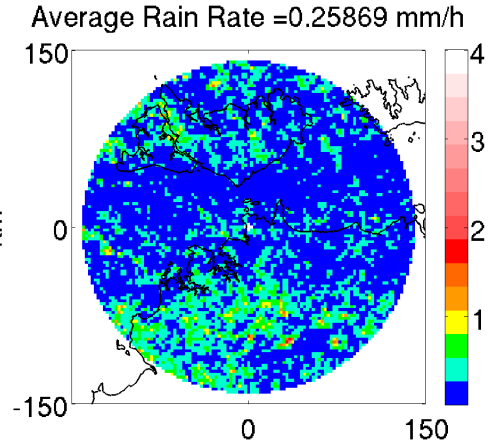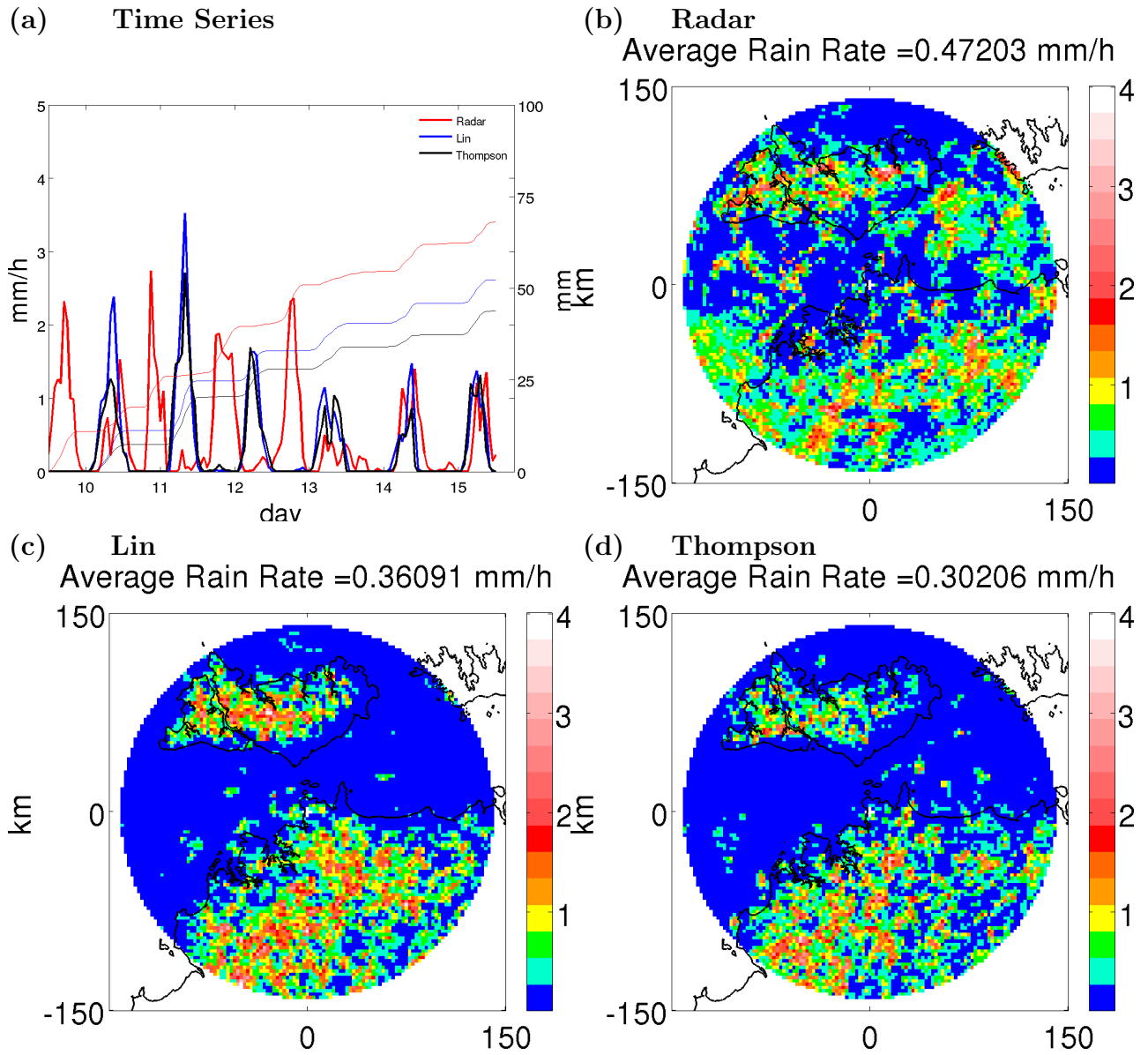
Figure 5.5: **a) Time series of domain-averaged precipitation rates and accumulation over the break period for the radar, Lin and Thompson simulations. b) Spatial distribution of precipitation rates averaged over the break period for the radar c) the Lin simulation and d) the Thompson simulation.**
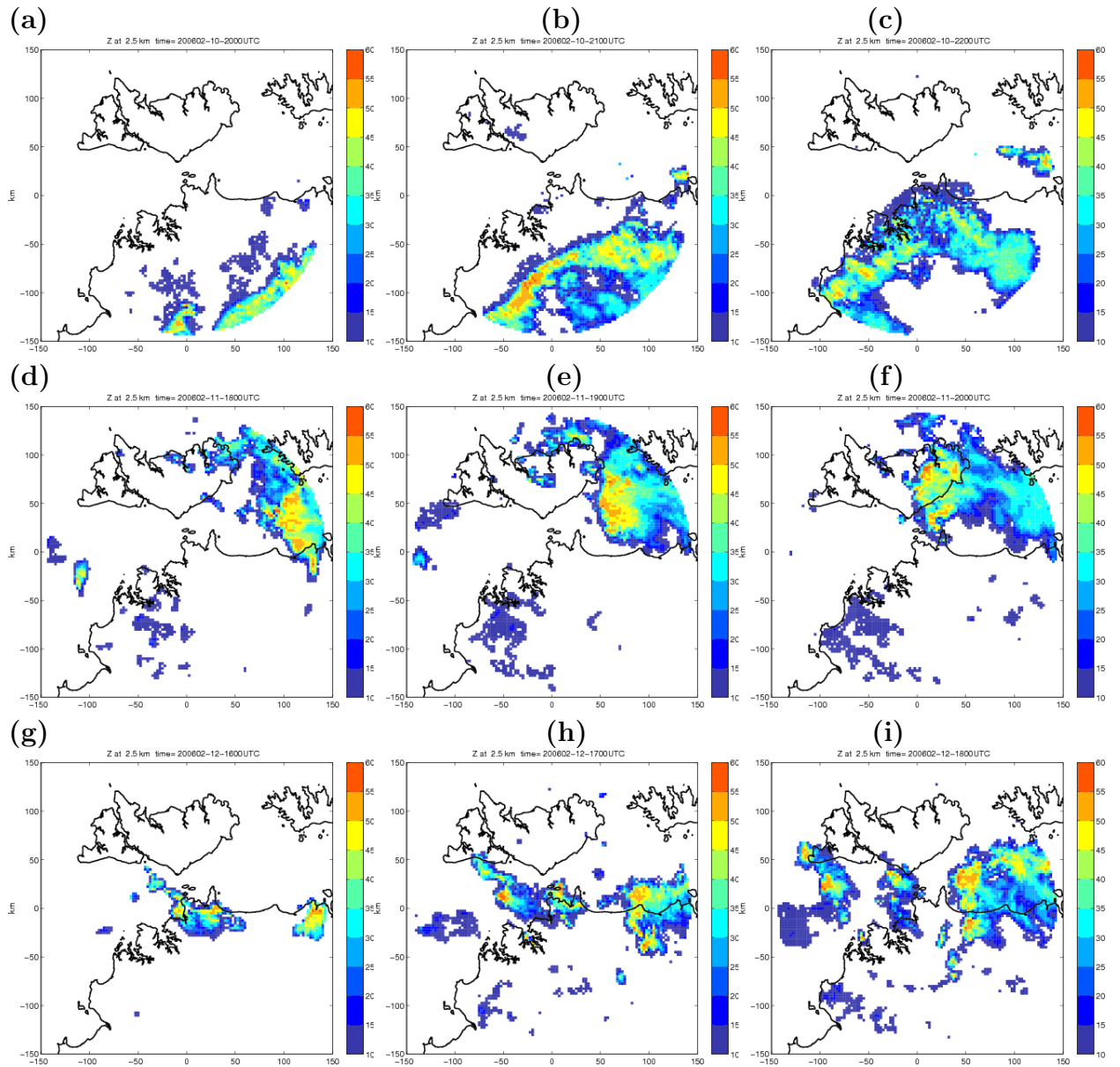
Figure 5.6: **2.5 km CAPPI's during the break period showing three convective systems that propergated into the radar domain. Figures are shown at at 20, 21, 22 UTC on the 10th of February a), b), c) . 18, 19, 20 UTC on the 11th of February d), e), f) and 16, 17, 18 UTC on the 12th of February h), i), j)**

# Chapter 6

# Hydrometeor Classification

In chapter 2 radar data and a clustering algorithm were used to objectively define a number of precipitation regimes that occur over Darwin during the wet season. Chapters 4 and 5 then showed how data from the WRF model could be converted into simulated reflectivity so that radar data could be used to evaluate the WRF model. In the analysis presented so far a major capability of the Gunn Point radar has been overlooked - the polarization information available from a CPOL radar.

The Gunn Point radar uses a polarized beam in both the horizontal and vertical directions (Keenan et al., 1998). The advantage of using a polarized radar is that different hydrometeors (in shape and composition) have different effects on various properties of the polarized beam. The additional information available from a CPOL radar can be used to estimate the dominant species of observed hydrometeors (more detail provided in section 6.1). Knowledge of the observed hydrometeor species offers a unique opportunity for model evaluation because the observed quantities are closely associated with simulated variables in the microphysics schemes. Although the simulated reflectivity data described in chapter 4 is derived from the model microphysics variables (i.e. mixing ratios of rain, graupel and snow), the final result is a combination of the reflectivity factors for each precipitation category. Therefore even if the model simulated reflectivity data closely matches the observed reflectivity data, it is impossible to tell whether the contributions from each precipitation category are in the correct proportions.

Before the radar derived hydrometeor classifications can be used for model evaluation a number of assumptions are required to manipulate the data (both model and

radar) into a form that can be compared directly. The basic problem is that the radar classification algorithm produces ten different microphysical categories and the model microphysics schemes produce five, of which only three are precipitating and of sufficient size to be detected by C-band radar. Furthermore, the model data is in the form of mixing ratios and each model grid point can have positive values for all three precipitation categories. Therefore a method for determining the dominant precipitation category for a model grid point is required. The following section outlines how the classification algorithm works, as well as detailing the assumptions and methods used to manipulate the model and radar data into a form where they can be compared directly.

## 6.1    Assumptions and Method

The algorithm used to estimate the dominant species of the observed hydrometeors uses five polarimetric variables from a CPOL radar to produce ten hydrometer classifications, plus an additional unclassified category. Temperature profiles from a nearby atmospheric sounding are also used by the algorithm to discriminate between hydrometeor types. Below is a brief description of the polarimetric variables used by the classification algorithm, including some examples of how these variables can aid discrimination between hydrometeor species. It should be noted that the examples given are not sufficient in and of themselves to determine the hydrometer species, each polarimetric observation is used in conjunction with the others in the classification process. For a full description of the classification process please refer to Straka et al. (2000) and May and Keenan (2005).

One of the polarimetric variables used by the classification algorithm is the reflectivity factor for horizontally polarized waves ($Z_H$). Particles of the same medium (e.g. liquid water) that have larger horizontal dimensions produce larger values of $Z_H$. Furthermore, the dielectric constant and hence reflectivity factor changes as

a function of ice density, consequently, $Z_H$ can be used (in conjunction with other variables) to distinguish between hydrometeors that have different densities. For example dry low-density graupel typically has $Z_H$ values < 35 dBZ, while wet high-density graupel typically has $Z_H$ values > 45 dBZ (Straka et al., 2000).

The ratio between horizontal and vertical reflectivity factors provides information about the shape of the observed hydrometeors. As spherical particles have similar horizontal and vertical dimensions, the ratio of the horizontal and vertical reflectivity factors will be closer to unity than for non-spherical particles. The polarimetric variable used by the classification algorithm ($Z_{DR}$) is the logarithm of this ratio multiplied by 10. $Z_{DR}$ can be used to discriminate between different types of rain (i.e. drizzle and rain) as raindrops with a diameter less than 1 mm typically have values between 0-0.7, while raindrops with D > 2 mm typically have values greater than 2 (Straka et al., 2000).

The correlation between the signals at the two polarizations ($\rho_{HV}$) can be used to determine when there is a mixture of different types of hydrometeor species. $\rho_{HV}$ is lower when there is a mixture of hydrometeors present rather than just one type. Rain/hail mixtures typically have $\rho_{HV}$ less than 0.95, while rain-only hydrometeors produce $\rho_{HV} > 0.95$ (Straka et al., 2000).

The specific differential phase ($K_{DP}$), which is the rate of change of the phase difference between the vertically and horizontally polarized waves, can be used to differentiate between different types of snow aggregates. Wet snow aggregates have $K_{DP}$ values between 0-0.2 whereas dry snow aggregates have $K_{DP}$ values between 0-0.5. (Straka et al., 2000).

For further details on the procedure of classifying hydrometers from polarimetric values and the definitions of the polarimetric variables the reader is referred to Straka et al. (2000) and May and Keenan (2005) for a comprehensive description. Here it will simply be stated that the rationale behind the classification procedure

comes from the notion that different hydrometer species occupy different parts of the $Z_H, Z_{DR}, \rho_{HV}, K_{DP}, Temperature$ phase space (May and Keenan, 2005).

### 6.1.1 Radar Assumptions/Simplifications

Applying the hydrometeor classification algorithm to data from a CPOL radar produces the ten (eleven if the unclassified category is included) hydrometeor classifications shown in column two of table 6.1. As the model microphysics schemes simulate only three precipitating categories (rain, graupel and snow) the ten categories derived by the classification algorithm are much too detailed for a straight comparison. Therefore the ten observed categories were grouped into the three broader categories of rain, graupel and snow. The third column of Table 6.1 shows the results of this grouping process. Note that the graupel category defined in column 3 of Table 6.1 is not graupel in the strict definition, but rather encompasses a range of hydrometeors that would likely fit the definition of graupel in the model microphysics schemes (i.e. precipitating ice hydrometeors such as hail or hydrometeors that are a mixture of ice and water).

With the observed hydrometeors grouped into these categories the only remaining radar classification that could not be compared with the model microphysical data was the unclassified category. During the monsoon period unclassified hydrometeors account for between 40-80 percent of the total hydrometeors detected, whereas during the suppressed and break conditions the fraction of unclassified hydrometeors is less, ranging from approximately 20-60 percent. Figure 6.1 shows the percentage of classified hydrometeors (rain, graupel and snow) to the total number of observed hydrometeors for the entire TWP-ICE period (rain, graupel, snow and unclassified). In the analysis below two methodologies were used to deal with the unclassified data, when cross sections were taken through convective cells the unclassified data was kept as a separate category and displayed in each figure. Average

profiles were created, then times that had greater than 60% coverage of unclassified hydrometeors were excluded from the averaging process. During times with less than 60% coverage of unclassified hydrometeors, the unclassified data was estimated by calculating the ratio of each classification (rain, graupel and snow) to the total number of classified hydrometeors at each level. The unclassified data (at each level) was then assumed to be contain hydrometeors in the same proportion as the ratios calculated for the classified species. For example if the domain contained 300 grid points of rain, graupel, snow and unclassified, then the ratio of rain to the total number of classified hydrometeors is one third, consequently one third of the classified data is assumed to be rain.

| # | Observed Category | Final category |
|---|---|---|
| 0 | unclassified | unclassified |
| 1 | drizzle | rain |
| 2 | rain | rain |
| 3 | dry low density snow | snow |
| 4 | dry high density snow | snow |
| 5 | melting snow | graupel |
| 6 | dry graupel | graupel |
| 7 | wet graupel | graupel |
| 8 | small hail | graupel |
| 9 | large hail | graupel |
| 10 | rain hail mix | graupel |

Table 6.1: **The hydrometeor species derived from the classification algorithm (column 2) and the final designations of these categories after each species were grouped into broader categories (column 3).**

## 6.1.2 Model Assumptions/Simplifcations

After applying the assumptions described in section 6.1.1 to the polarimetric radar data an estimate of the bulk hydrometeor species (rain, graupel or snow) for each grid point in the radar domain was obtained. To compare the model microphysical data with these observed quantities a number of assumptions and simplifications
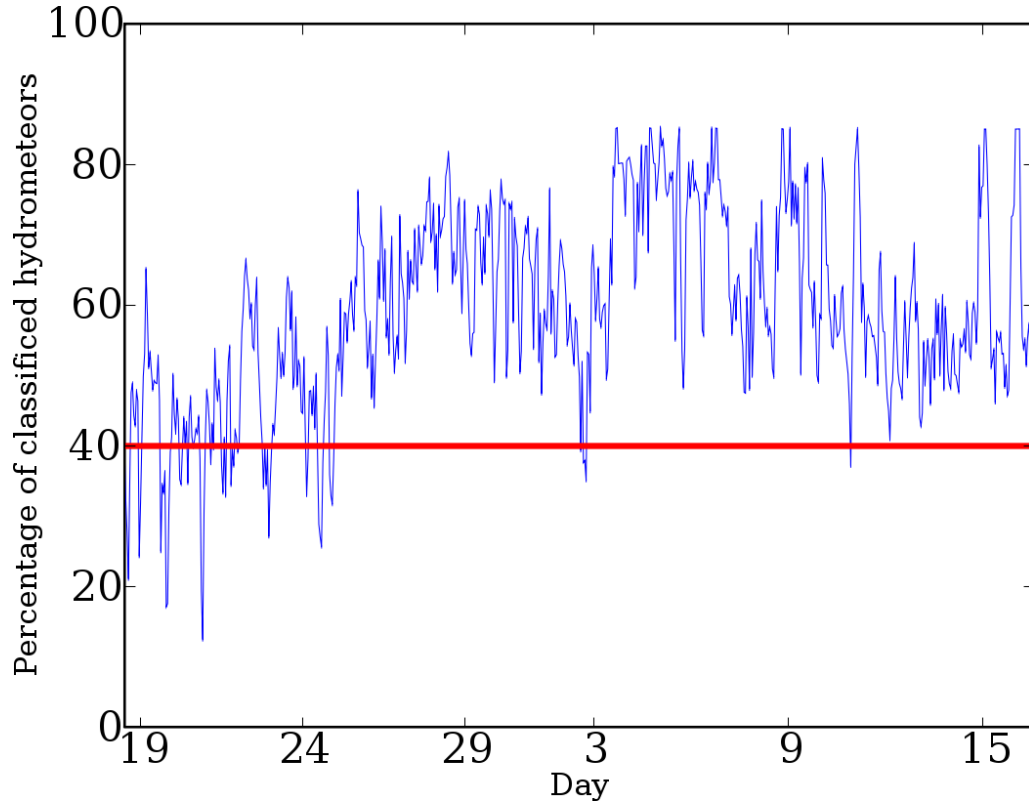
Figure 6.1: **Time series for the TWP-ICE period showing the fraction of classified hydrometeor data.**

were also required. To determine whether or not a model grid point contained a detectable amount of hydrometeors the simulated reflectivity data described in chapter 4 was used. If the simulated reflectivity value for the grid point was below 0 dBZ the grid point was assumed to be clear and if the grid point had a reflectivity value above 0 dBZ the grid point was said to contain a detectable amount of hydrometeors. Once a grid point was defined as containing a detectable amount of hydrometeors the dominant hydrometeor species for the grid point was estimated by determining which of the three precipitation categories contributed most to the reflectivity for a given grid point. The category that had the largest individual reflectivity factor (i.e. equations 4.4, 4.7 and 4.6) was assumed to be the dominant hydrometeor species for the grid point.

## 6.2   Cross sections

Using the above rules both the radar and model data were in a form directly comparable with each other, i.e. each grid point was declared to contain either snow, graupel, rain, or no hydrometeors.

To investigate how the WRF model performs during the different conditions that occurred during the TWP-ICE period, cross sections through convective cells are shown for the radar and WRF simulations during the Monsoon, Break and Suppressed conditions. In all figures displaying cross sections the top 3 panels show results for the radar, the middle panels show the results for the Lin simulation and the bottom panels show the results for the Thompson simulation. From left to right is displayed the 5 km CAPPI showing reflectivity in dBZ, a cross section through the estimated hydrometeor data along the East-West orientated line depicted in the CAPPI, and a cross section through the estimated hydrometeor data along the North-South orientated line depicted in the CAPPI.

Figure 6.2 shows cross sections through convective cells during the monsoon period. As convection in the model simulations sometimes precedes or lags convection observed by the radar, the model cross sections are often taken at slightly different times to the radar cross sections. The radar panels in Figure 6.2 correspond to 1700 UTC on the 23rd of January, while the Lin and Thompson simulations correspond to the 21 UTC and 16 UTC on the 23rd of January respectively.

The cross sections through the radar data show that within a convective cell substantial amounts of graupel are observed up to approximately 14 km, and a small amount of graupel is detected as high as 18 km. Obvious differences between the simulated and the observed distribution of graupel are apparent in Figure 6.2. In general the Thompson simulation underestimates graupel and where graupel does exist the height that it reaches is underestimated by up to 10 km. The Lin simulation on the other hand produces graupel as high as 18 km, and even at cloud top graupel

is defined as the dominant hydrometeor species. In regions of active convection the radar observes graupel as low as 2 km, while both model simulations show a sharp cut-off of graupel at approximately 5 km (the freezing level).

Outside regions of active convection the radar observes graupel uniformly in a layer between approximately 4.5 and 6.5 km (although some of the data is unclassified). The Lin simulation is able to capture the horizontal extent of the graupel well, but performs rather poorly in simulating the vertical distribution graupel. The height that graupel reaches is overestimated everywhere in the Lin simulation and can be seen as high as 10 km in non-convective regions.

The larger dimensions of graupel particles (due to their low density ice core), coupled with a larger dielectric constant (due to a liquid outer skin), mean that graupel particles produce higher reflectivities than snow or water particles of the same mass. Thus the overestimation of maximum reflectivity values above 40 dBZ found in Chapter 4 (Figure 4.1) can be explained as a product of the Lin microphysics overestimating graupel at higher altitudes.

Both the radar and the Lin simulations show wide spread coverage of all three precipitation categories and in general the vertical distribution of the dominant hydrometers (top down) transitions from snow to graupel to rain. The Thompson simulation on the other hand produces very little graupel, and regions exist where snow transitions into rain without going through the graupel phase. Furthermore, there are regions in the Thompson simulation where snow exists above the freezing level with no rain below.

During the break period there were two main type of convection observed by the radar - convection that occurred over land during the afternoon (diurnally forced) and convective systems that propagated into the domain. As discussed in Chapter 5 a number of the systems that propagated into the domain were not captured by the WRF model, leading to serious timing issues in the model simulations. One squall

line that was captured by the simulations occurred at approximately 12 UTC on the 10th of February. Figure 6.3 shows cross sections through the squall line at this time (the model cross sections are taken at 9 UTC on the 10th of February). The radar cross sections show that within the squall line graupel is observed all the way up to cloud top, as this was not observed in the cross section during the monsoon period it is possible that these results are examples of microphysical differences that exist between monsoon and break convection. However, as these figures represent only a single time period no solid conclusions can be made from these figures alone. The Thompson microphysics seems to have picked up on a difference between monsoon and break type convection. Figures 6.3 i) and j) show that significantly more graupel is produced in the squall line than was seen in the monsoon cross sections. Although the Lin simulation produces graupel up to the cloud top, this was also the case during the monsoon period when graupel was not observed at cloud top.

Perhaps the most interesting feature in Figure 6.3 b) is an anvil that can be seen between roughly 50 to 100 km, the majority of which is non-precipitating. Thompson simulation produces a large non-precipitating anvil but for the most part the clouds in the Lin simulation rain uniformly. Note that a small anvil is produced by a convective cell located at approximately -50 km in Figure 6.3 e). This convective cell has a lower cloud top than the other convective cells and graupel reaches approximately 11 km. The fact that the Lin microphysics overestimates graupel may be contributing to the underestimation of snow and anvil outflow. The fall speed of graupel particles is significantly higher than that of snow, therefore if graupel is produced at the expense of snow then the higher fall speeds will mean that the graupel particles will precipitate out too quickly, instead of producing an anvil.

The other type of convection that occurred regularly during the break period was diurnally forced land based convection. Figure 6.4 shows cross sections through

the convection occurring over the Tiwi islands (colloquially referred to as Hector). Similar results are observed during Hector as were seen in the squall line. Again the radar observes graupel up to the cloud top, however here the graupel is patchy and non-continuous. The Lin simulation drastically overestimates graupel during this time - almost everything above the freezing level is classified as graupel leading to unrealistic hydrometeor distributions. In contrast the Thompson simulations produces a more realistic vertical profile of hydrometeors for the most part, however there are some regions where rain and or graupel exist with no snow at higher altitudes.

Figure 6.5 shows cross sections through convective cells during the suppressed period. The patchy nature of the convection can be seen by the thin columns of precipitation inter-spaced with clear regions. Both the Lin and the Thompson simulations appear to capture the patchy nature of the convection well. However, the height of the convection appears to be significantly underestimated, with only a small amount of graupel (snow) produced above the freezing layer in the Lin (Thompson) simulation.

The cross sections presented above are useful for determining how the hydrometeors are distributed within observed and simulated clouds. However, these results represent only a snap shot within the entire simulation. Therefore to determine if these snap shots are representative of the convection that occurs during each period the next section will look at average profiles created for the radar and model simulations.

Figure 6.2: **CAPPI showing the reflectivity values at 5 km for the radar a), the Lin simulation d) and the Thompson simulation h). East-West and North-South cross sections through the estimated hydrometeor data are shown for the radar b) and c), Lin simulation e) and f), and Thompson simulation i) and j) during the Monsoon Period**

Figure 6.3: **CAPPI showing the reflectivity values at 5 km for the radar a), the Lin simulation d) and the Thompson simulation h). East-West and North-South cross sections through the estimated hydrometeor data are shown for the radar b) and c), Lin simulation e) and f), and Thompson simulation i) and j) through a squall line during the Break period**
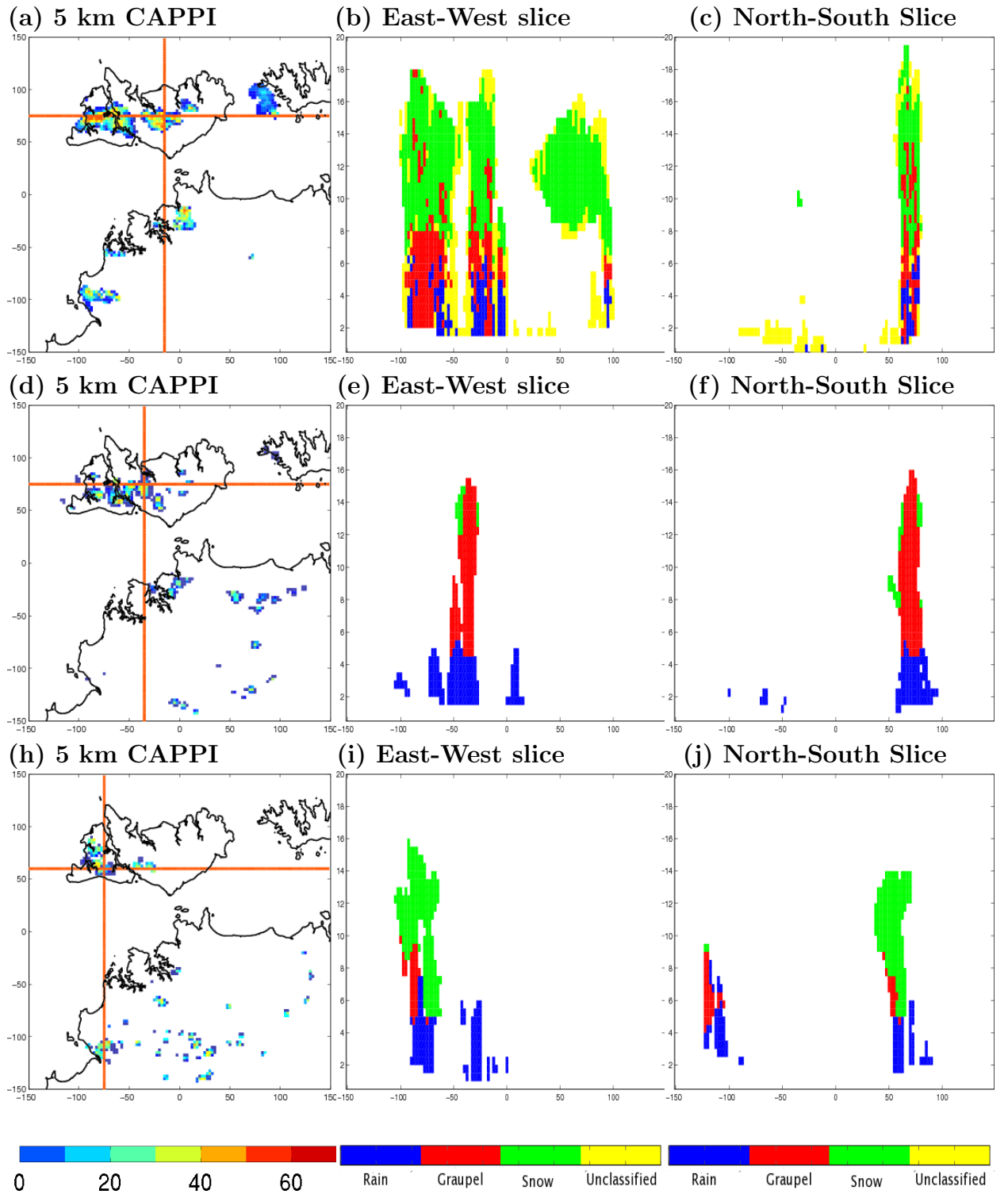
Figure 6.4: **CAPPI showing the reflectivity values at 5 km for the radar a), the Lin simulation d) and the Thompson simulation h). East-West and North-South cross sections through the estimated hydrometeor data are shown for the radar b) and c), Lin simulation e) and f), and Thompson simulation i) and j) through storms over the Tiwi islands during the Break period**
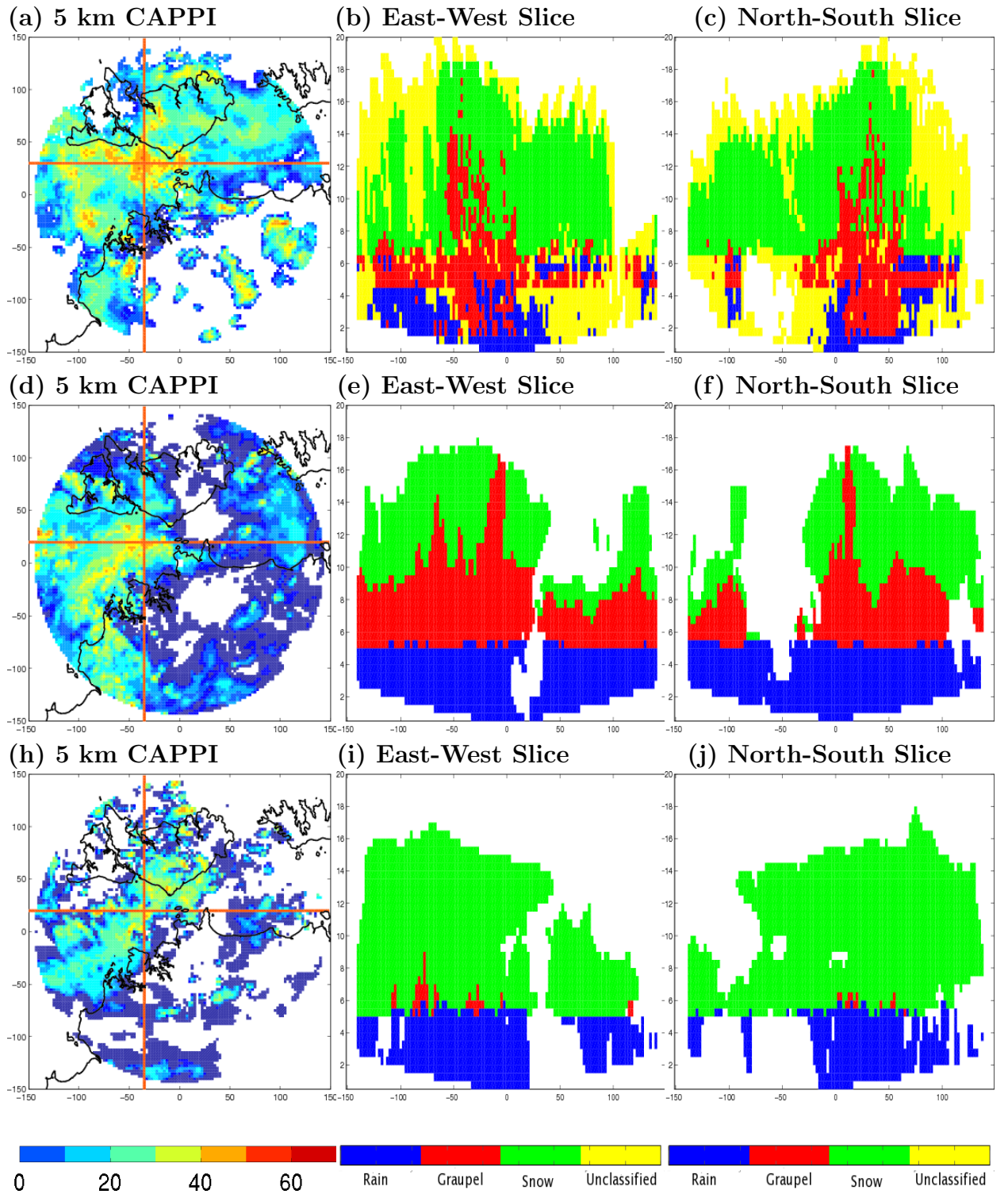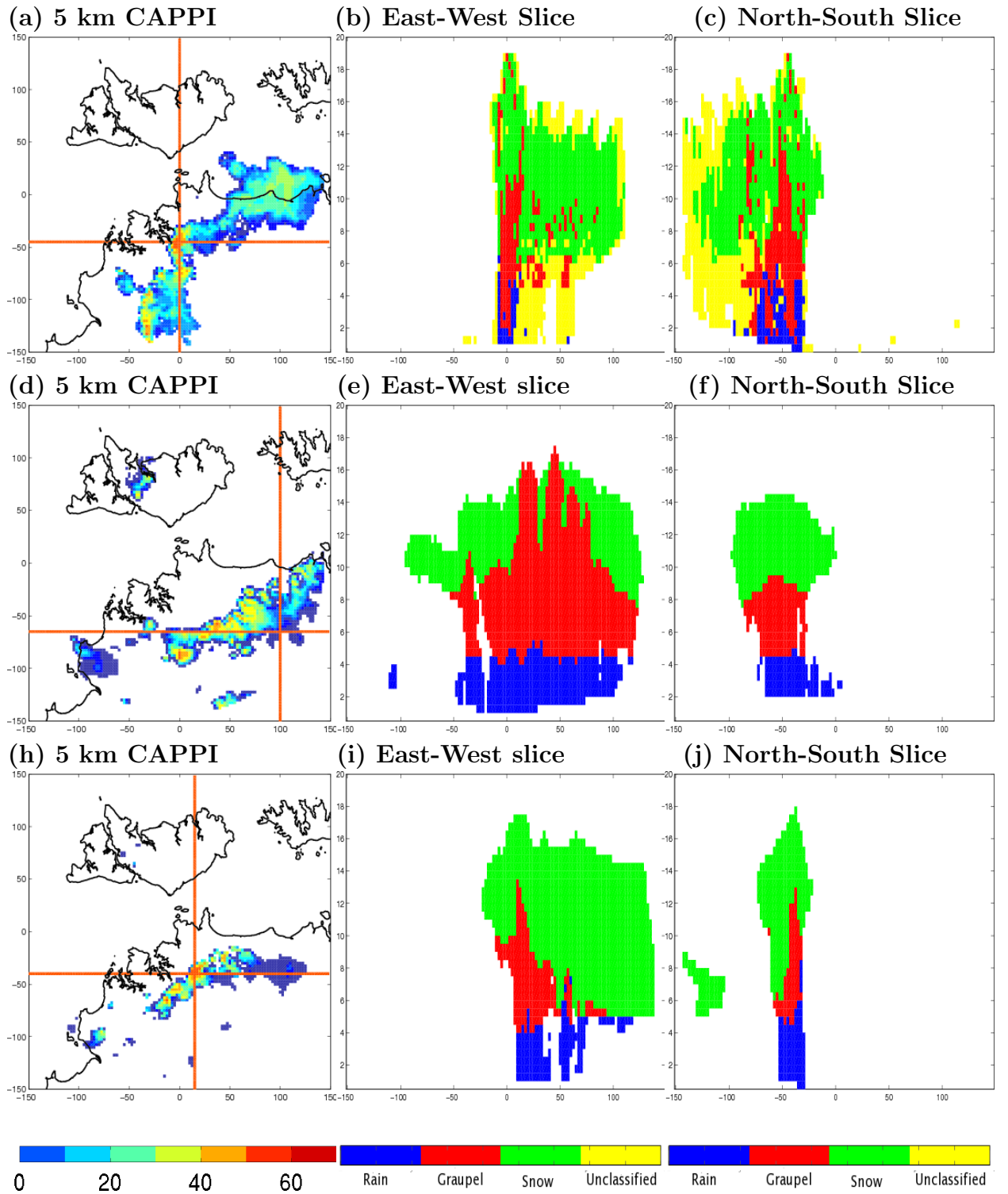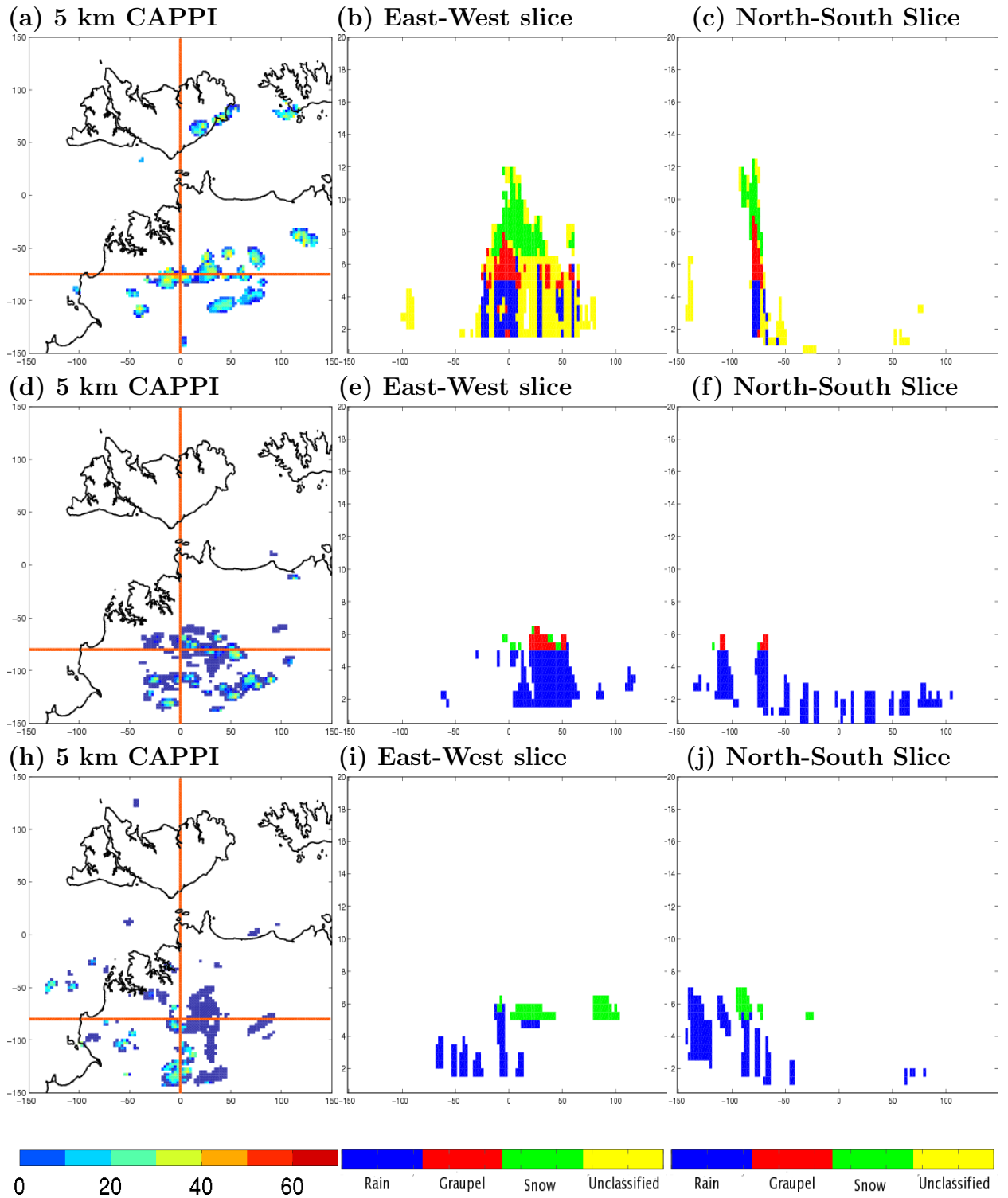
Figure 6.5: **CAPPI showing the reflectivity values at 5 km for the radar a), the Lin simulation d) and the Thompson simulation h). East-West and North-South cross sections through the estimated hydrometeor data are shown for the radar b) and c), Lin simulation e) and f), and Thompson simulation i) and j) during the suppressed period.**

## 6.3   Average Hydrometeor Profiles

At each hour within a simulation the fractional coverage of rain, graupel and snow was calculated at each level for the radar and model data. Average profiles of hydrometeor coverage vs height were then created for each simulation period. The first 24 hours were excluded from the averages to remove any spin-up effect and times in which the radar data had more than 40 % of the observed hydrometeors in the unclassified category were also excluded.

### 6.3.1   Monsoon Simulation

Figure 6.6 shows the hydrometeor profiles averaged over the Monsoon and Monsoon/Supressed period for a) and d) the radar, c) and d) the Lin simulation, and e) and f) the Thompson simulation. The Lin simulation appears to capture the coverage of rain and graupel quite well in both these simulations. However, the overestimation of graupel at higher altitudes previously seen in the cross sections is apparent even in the average profile, especially above 10 km. As discussed earlier the overestimation of graupel is one possible reason that snow is underestimated in the Lin simulation. The greater fall speed of graupel particles means that the residence time for graupel will be much lower than would be the case if the hydrometeors were correctly simulated as snow. Therefore too much mass is precipitating out which leaves a deficiency of water above the freezing level.

The most striking difference between the simulations and the radar profiles is the underestimation of rain below the freezing level in the Thompson simulations. This is most obvious in the monsoon simulation where the maximum coverage of rain approximately 20% compared with 50% in the Lin simulation and observed by the radar. Underestimation of rain below the freezing level in the Thompson simulation comes as no surprise with considering the results in Chapter 4, where precipitation coverage below the freezing level was found to be significantly underestimated in

the monsoon simulation. However, Figure 6.6 e) and f) show that in addition to underestimating the coverage of rain at the lowest levels (greater than a factor of 2 difference in the monsoon simulation) there is a significant underestimation of graupel in the Thompson scheme as well. Figures 6.6 e) and f) show that Thompson simulation actually overestimates the coverage of snow directly above the freezing level. Earlier it was postulated that the Thompson scheme may have a problem converting stratiform precipitation above the freezing level into rain. From these figures it would appear that the problem is also associated with the production of graupel. In section 6.2 both the radar and Lin cross sections showed that snow transitioned through a graupel phase when falling through the melting layer. The underestimation of graupel in the Thompson scheme may indicate that the process of snow turning into rain via the graupel category may be inhibited in the Thompson microphysics. However, when the 10 hydrometeor species were grouped into the three broader categories, melting snow was assigned to the graupel category. This assignment appeared to produce nice correlations between the Lin and observed graupel coverage (at least in the horizontal dimensions). However, if the Thompson microphysics sticks to the strict definition of graupel then melting snow would likely be assigned to the rain category (the assignment of melting snow is rather arbitrary). Changing the assignment of melting snow from graupel to rain should not change any of the conclusions drawn from the above results, however one would have to conclude that the process of snow melting into rain is the problem with the Thompson microphysics.

To further shed light on this problem Figure 6.7 shows mixing ratio profiles for the Lin and Thompson schemes during the passage of the MCS (12 UTC on the 23rd). Comparison of the Lin and Thompson mixing ratios show that the mixing ratios for rain and graupel are substantially smaller in the Thompson simulation, while the maximum mixing ratio for snow is almost four times greater. At approxi-

mately 6 km the graupel mixing ratio reaches its peak value in the Lin simulation, at the exact level where the Lin simulation has large values of graupel mixing ratio the Thompson simulation shows a secondary bulge in the mixing ratio for snow, indicating that snow at this level is not being converted into graupel, or, if graupel in the Thompson simulation represents the strict definition of graupel, then snow is simply not melting into rain. Either way it can be seen that significantly more water mass exists above the freezing level in the Thompson scheme compared with the Lin scheme.
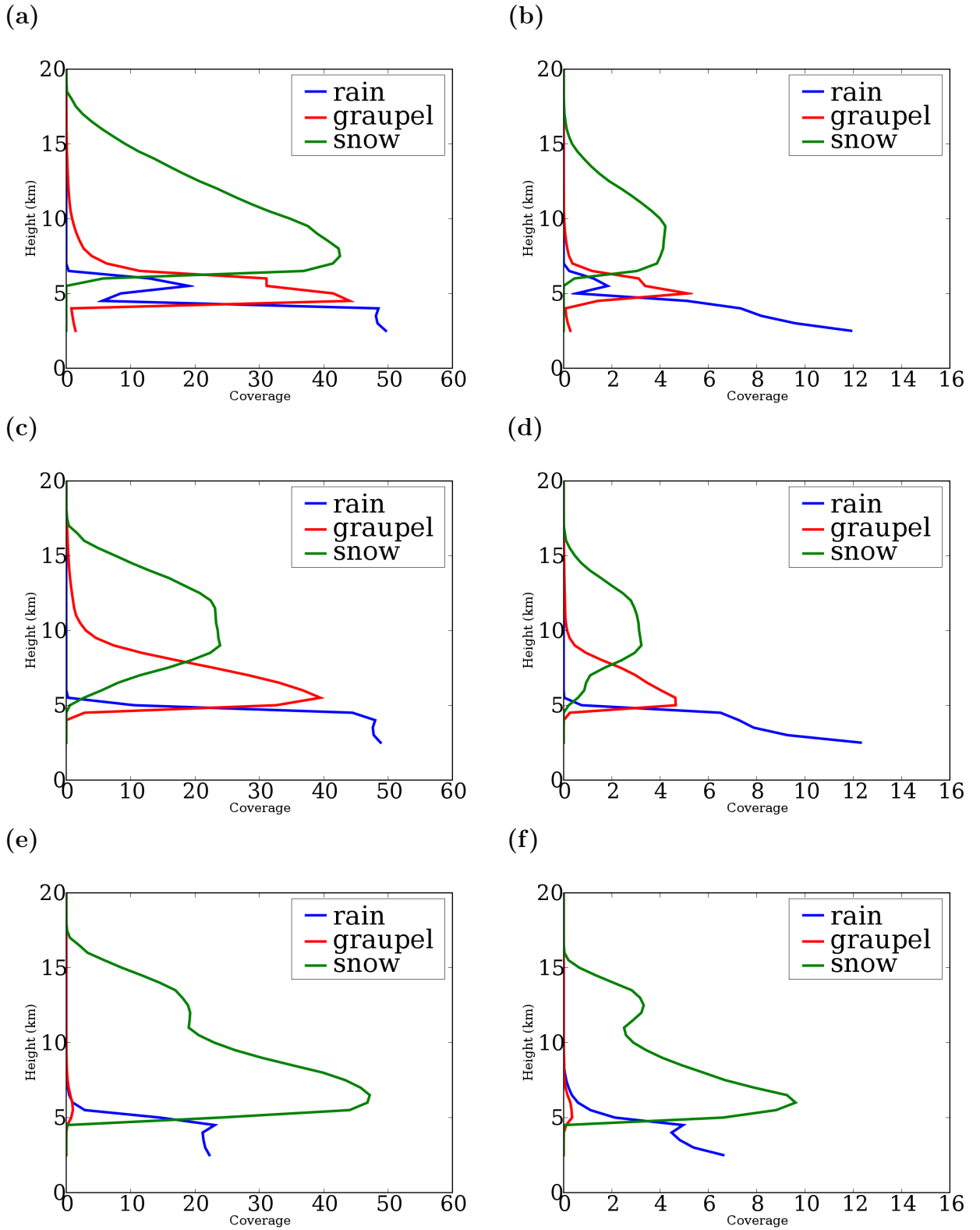
Figure 6.6: **Hydrometeor profiles averaged over the Monsoon period for a) the radar, c) the Lin simulation e) the Thompson simulation. Hydrometeor profiles averaged over the Monsoon/suppressed period for b) the radar, d) the Lin simulation f) the Thompson simulation.**
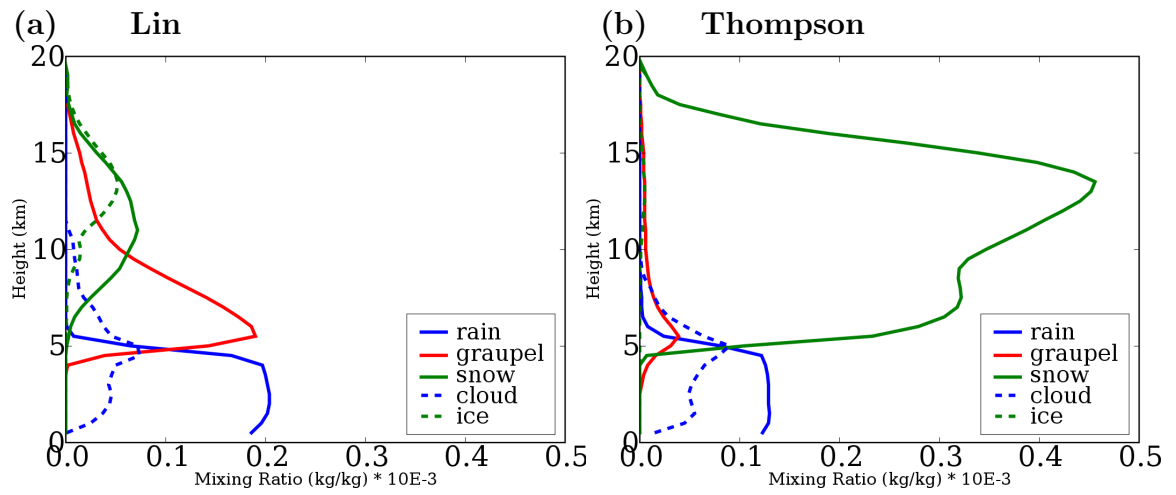
Figure 6.7: **Mixing ratios averaged at 12 UTC on the 23rd of January for a) the Lin simulation and b) the Thompson simulation**

## 6.3.2   The Suppressed Simulations

Figure 6.8 shows the hydrometeor profiles averaged over the suppressed and suppressed/break period for a) and b) the radar, c) and d) the Lin simulation, and e) and f) the Thompson simulation. The hydrometeor profiles during the suppressed conditions show a marked difference to those during the monsoon period, here there is a pronounced decrease in hydrometeor coverage with height. During the monsoon period all three categories had roughly the same peak coverage, whereas during the suppressed conditions the peak coverage of snow is roughly half the peak coverage of rain. While the Lin microphysics does produce less snow coverage (and graupel) than rain, both snow and graupel are greatly underestimated in these simulations. The cross sections in section 6.2 appear to be representative of the convection during the suppressed conditions i.e. coverage of snow and graupel is underestimated at higher altitudes.

Of all the time periods simulated during TWP-ICE, the simulation beginning on the 28th of January appears to be least affected by the lack of rain and graupel in the Thompson scheme. As this simulation was initialised in the middle of the suppressed conditions where deep convection is inhibited and stratiform precipitation is likely to contribute little to the overall rainfall on the ground, a link between stratiform precipitation and the problem below the freezing level appears plausible. However this period is not entirely unaffected, compared with the radar and Lin simulation the coverage of rain is underestimated in the Thompson simulation. Stratiform precipitation may indeed be linked with the problem below the freezing level, it is just that the problem will be more obvious during periods containing large quantities of stratiform precipitation.

### 6.3.3 The Break Period

Figure 6.9 shows the hydrometeor profiles averaged over the break period for a) the radar, b) the Lin simulation and c) the Thompson simulation. In this period the observed peak snow coverage is approximately the same as the peak coverage of rain. The Lin simulation fails to produce the observed snow coverage and instead produces an average profile similar to those produced during the suppressed conditions. Both simulations underestimate the coverage of the three hydrometeor species, however the Thompson simulation is much closer to simulating the correct coverage of snow (the Lin simulation underestimates snow by approximately 50%).

## 6.4 Summary

Polarimetric variables from a CPOL radar can be used to estimate the dominant species of the precipitation particles observed by the radar. In order to use these observations for model evaluation a number of assumptions were required to manipulate both the radar and model data into a form that would allow a direct comparison. The assumption which is perhaps the weakest in the analysis presented here is the assignment of melting snow in the graupel category. In future work it would be useful to investigate how sensitive results are to the categories outlined in column 3 of table 6.1. Despite the ambiguity of where to place melting snow, the observed hydrometeors provided valuable insight into how the model microphysics packages performed in terms of hydrometeor coverage. In general the Lin simulation was found to overestimate the coverage of graupel and underestimate the coverage of snow. It was hypothesized that by incorrectly simulating too much graupel, water mass precipitates much too quickly, thereby removing water mass from the upper levels of the atmosphere. Conversely, the Thompson simulation underestimated graupel and was found to produce way too much snow.

Figure 7 and 8 of the appendix show updated figures for the Thompson simulation when the correct particle size distribution for snow is used within the reflectivity conversion algorithm. The conclusions in this chapter for the Thompson simulations are somewhat different to the conclusions drawn when exponential particle size distributions were used. The more sophisticated treatment of graupel and melting snow within the reflectivity conversion algorithm appears to have produced the most notable difference. Within a convective cell graupel is now more frequently defined as the dominant hydrometeor species and can now be seen up to the cloud top. The averaged profiles also show the coverage of graupel has increased and extends much higher than observed previously. However, the averaged profiles (Figure 8) show that even with the more sophisticated treatment of graupel, compared with both the Radar and Lin simulations graupel is still significantly underestimated in the Thompson simulation. The averaged profiles show a maximum coverage of approximately 5% compared with approximately 40% observed by the radar. Using the correct particle size distribution for snow within the conversion algorithm leads to a more realistic profile of snow coverage above 10 km. The rapid and unrealistic decrease in snow coverage previously seen between approximately 13 and 10 km is now a smooth transition. The coverage of rain below the freezing level is still significantly under-estimated and consistent with the updated Figure in Chapter 4.
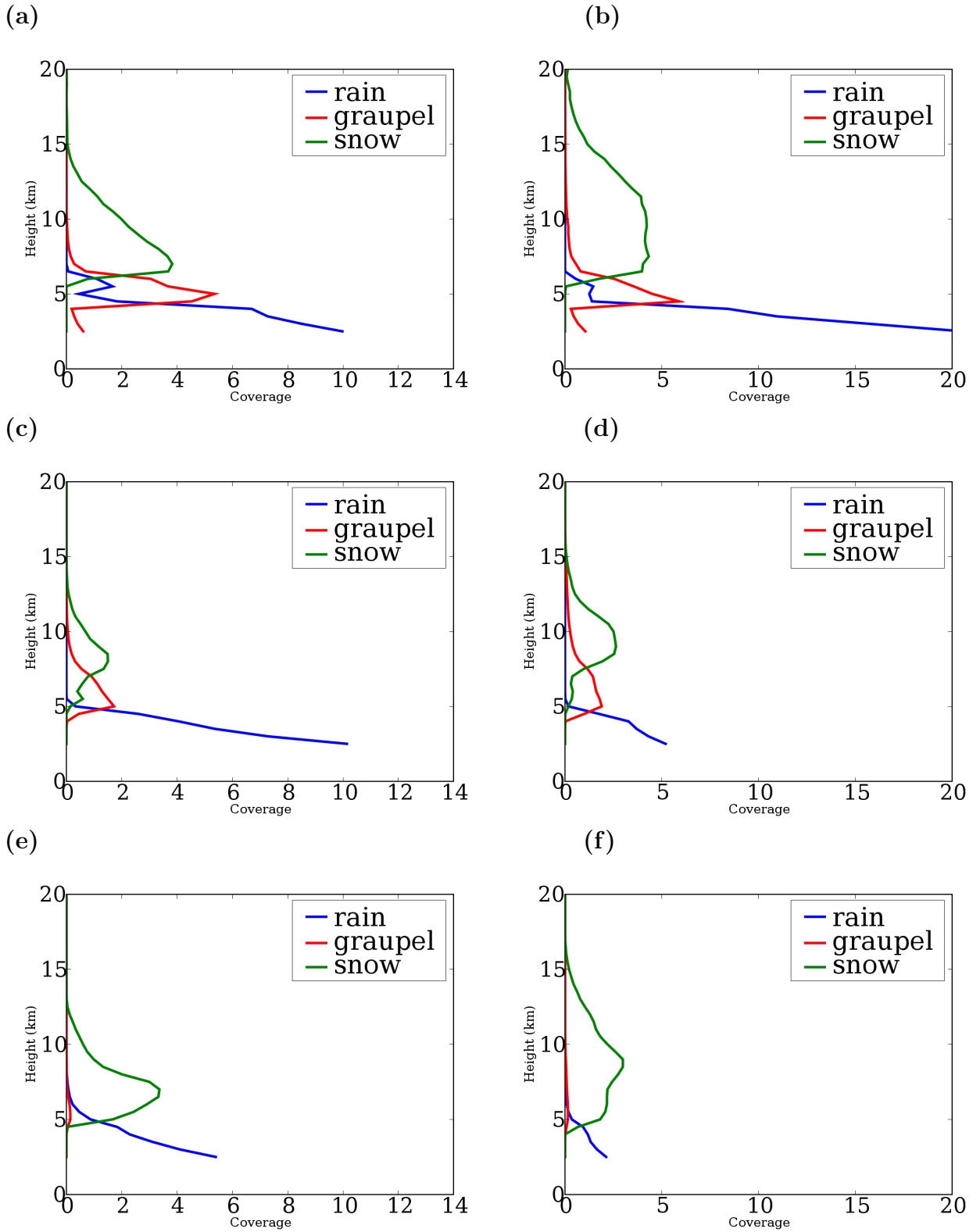
Figure 6.8: **Hydrometeor profiles averaged over the suppressed period for a) the radar, c) the Lin simulation e) the Thompson simulation. Hydrometeor profiles averaged over the suppressed/break period for b) the radar, d) the Lin simulation f) the Thompson simulation.**
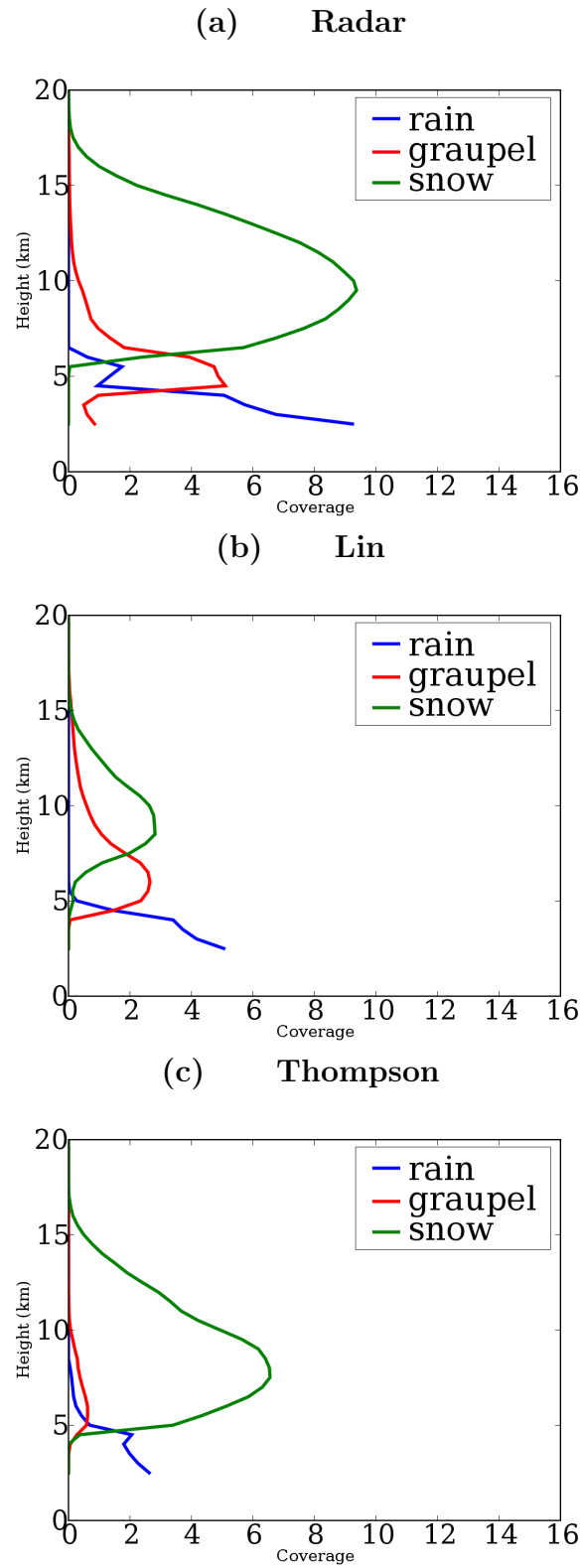
Figure 6.9: **Hydrometeor profiles averaged over the Break period for a) the radar, b) the Lin simulation c) the Thompson simulation.**

# Chapter 7

# Precipitation Regimes

This thesis has explored how data from the WRF model can be converted to simulated radar reflectivity and compared with radar data to assess the model's performance in a tropical environment. In this chapter an attempt is made to continue the evaluation process by using the precipitation regimes defined in Chapter 2. Each of the four precipitation regimes represent distinct and reoccurring states of the atmosphere, based on the statistical distribution of precipitation produced by convective systems that occur over Darwin. Together the four regimes encompass the wide range of precipitation structures produced during a typical wet season, and each have been shown to have physical links with the tropical atmosphere. It is hoped that conditions under which the model performs poorly (or well) will become evident by determining how well the WRF model is able to simulate each precipitation regime.

One of the methods often used to compare model simulations with regimes created using the KMEANS algorithm is to create histograms from model data in the same format as the observational data, then assign the model histograms to one of the regimes (Hume and Jakob, 2007; Gordon et al., 2005). However as the 2006 wet season was not used in the definition of the precipitation regimes, it was first necessary to assign histograms of reflectivity vs height from the TWP-ICE period to one the precipitation regimes (see Chapter 2 for how the histograms are created). Assigning a histogram to a precipitation regime involves calculating the euclidian distance between the histogram and each of the four precipitation regimes, then associating the histogram with the regime that had the most similar reflectivity

structure (defined by the minimum euclidian distance). Note that this is how histograms are assigned to centroids in the clustering algorithm used in Chapter 2, the only difference is that in the clustering process after all histograms have been assigned to a centroid, new centroids are calculated and the algorithm is iterated until the clusters become stable. Assigning histograms to the precipitation regimes this fashion has the benefit of leaving the regimes unaltered by the additional data, while also being computationally efficient.

After the radar histograms from TWP-ICE were assigned to the precipitation regimes a similar process was used to assign model data to the precipitation regimes. Conceptually there is little difference between the model simulated reflectivity data and data obtained from the Gunn Point radar, only the methods that lead to the data generation differ. Therefore the model simulated reflectivity data described in chapter 4 were used to create histograms of simulated reflectivity vs height, which were assigned to the precipitation regimes using the same method described above.

The following section uses the precipitation regime assignments for both model and radar derived histograms to determine how well the WRF model performs during different conditions.

# 7.1 Time Series of Model and Radar Regimes

## 7.1.1 Monsoon Period:- Lin microphysics

Figure 7.1 a) shows the regimes that the radar (black), Lin (red) and Thompson (green) histograms were assigned to during the monsoon period. Domain-averaged vertical velocity profiles derived from the Xie-Klein forcing dataset have been reproduced in Figure 7.1 b), c) and d) to aid discussion. The following text is dedicated to the results from Lin simulation only, any reference to the "model histograms" refer to the Lin histograms. A separate section analyzing the Thompson time series is provided in section 7.1.2.

In chapter 2 the convective intensity of each regime was inferred from the reflectivity structure of the regime centroids by considering the maximum echo top height and maximum reflectivity in the lowest 5 km for each regime. There it was found that regime 1 contained weak/medium convective periods, regime 2 contained strong convective periods, regime 3 contained weak convective periods with large stratiform regions and regime 4 contained strong convective periods with large stratiform regions.

Comparing the time-series of radar regime assignments (black) with the domain-averaged upward motion derived from the Xie-Klein dataset (Figure 7.1 d) there appears to be a nice correlation between the domain-averaged vertical ascent and the assignment of the radar histograms to regimes 2 and 4. Just before 0 UTC on the 20th of January the radar derived histograms change their assignment from regime 1 to regime 2 (indicating a change from weak to strong convective updrafts) at the same time that an increase in the domain-averaged upward motion is observed in figure 7.1 d). While it may appear obvious that regimes 2 and 4 should be linked with stronger domain-averaged vertical ascent, vertical motion was never used in the definition or analysis of the regimes in chapter 2.

In addition to providing information about convective updraft strengths, the precipitation regime assignments also offer information about the amount and type of precipitation generated. For example the change from regime 2 to regime 4 (just after 0 UTC on the 20th of January) in the radar assignments indicates that significant amounts of stratiform precipitation were produced in the convective outflow. Furthermore, by the continued assignment of the radar histograms to regime 3, it can be concluded that even after the convective updrafts have weakened (see Figure 7.1 d)) the stratiform precipitation previously generated is still present. A similar pattern occurs at around 0 UTC on the 21st of January, the radar histograms are first assigned to the strong convective regime (regime 2), followed by a period of strong convection and stratiform precipitation (regime 4), followed by a sustained assignment of regime 3 when the convective event has weakened.

From the time series of regime assignments for the model (red), it can be seen that the Lin histograms are assigned to regime 1 during the first three days of the simulation despite the frequent occurrence of the higher precipitation regimes observed by the radar. From this incorrect assignment it can be concluded that the convective updraft strengths (and precipitation) are underestimated by the model during this time, a finding that is consistent with the results found in chapter 3 and 5.

At approximately 6 UTC on the 21st of January the Lin derived histograms leave regime 1 for the first time. While the model is correctly assigned to a stronger convective regime, the timing of this convection appears to be delayed, the radar histograms are assigned to the strong convective regimes approximately 4-6 hours earlier. Judging from the regime assignments alone, it can be concluded that the model underestimates the precipitation during this period. The model histograms are assigned to regimes 2 and 3 for only a few hours before returning to regime 1, whereas the radar histograms are continually assigned to regime 3.

Comparing the time series of the model and radar histograms to figures 7.1 b) and d) it can be seen that whenever strong domain-averaged vertical ascent occurs, the histograms are assigned to either regime 2 or 4. Using the precipitation regimes instead of the vertical velocity profiles has the added benefit of providing information about the precipitation produced in the convective events, as well as providing information about the persistence of stratiform precipitation after the convective event has finished. For example the model fails to capture the decrease in precipitation observed by the radar on the 23rd of January, and is then assigned to regime 4 earlier than the radar. From these incorrect regime assignments it would be expected that the model has overestimated the precipitation during the last two days of the simulation.

From the regime assignments it can be concluded that the Lin simulation has two main problems during the monsoon period. Firstly the convective updraft strength and precipitation is underestimated during the first half of the simulation, as indicated by the incorrect assignment of the model histograms to regime 1. Secondly the model overestimates precipitation during the last 2 days of the simulation, as indicated by the assignment of the model histograms to regimes 3 and 4 more often than observed by the radar. While this analysis is qualitative in nature and does not provide any additional information than has already been discovered in previous chapters, the results are just the first step in using the precipitation regimes to evaluate the models performance. Section 7.2 will provide a more quantitate and useful method for evaluating the models performance, however before showing these results section 7.1.2 will continue to investigate if assigning the model derived histograms to the precipitation regimes can provide useful information about the models performance.
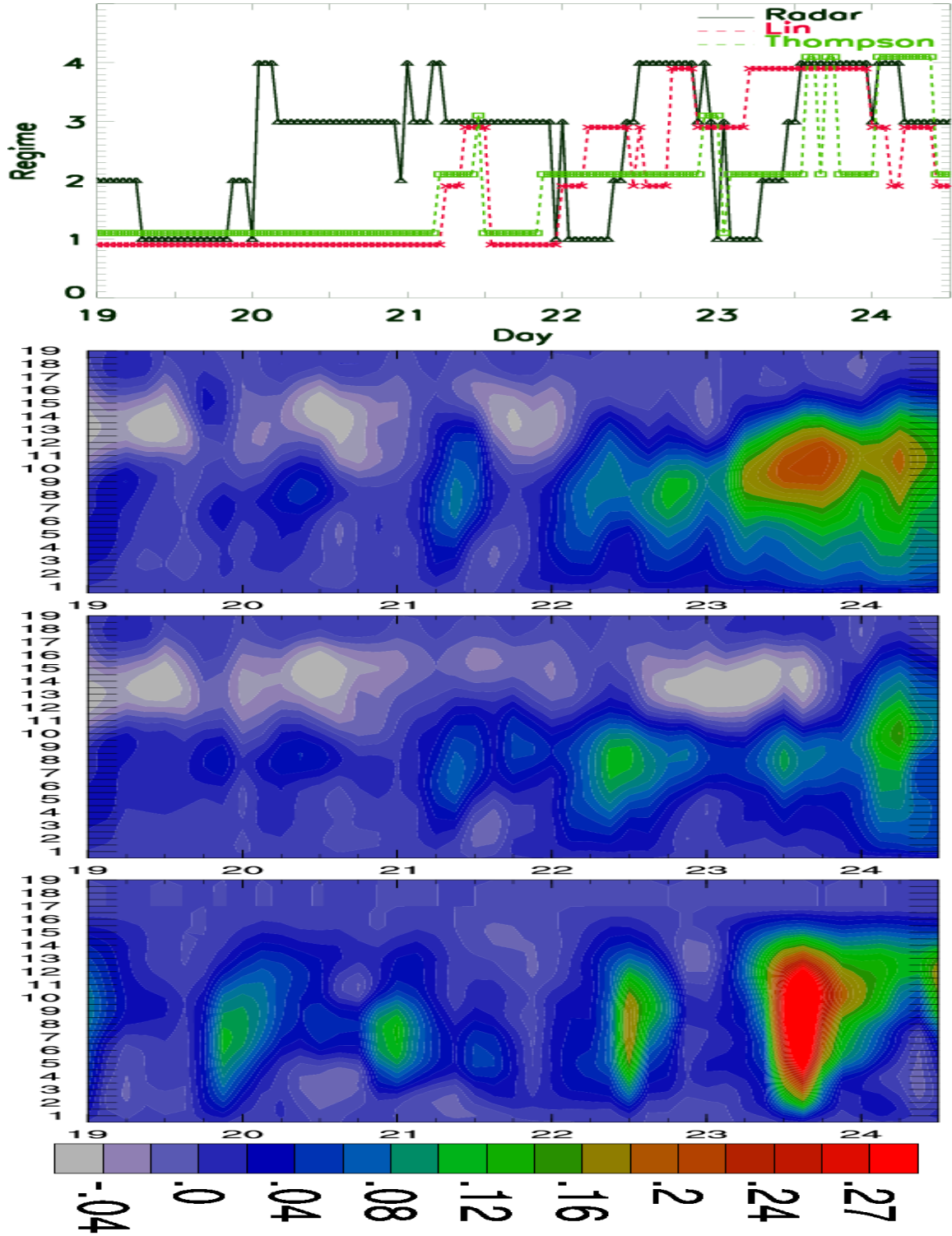
Figure 7.1: **a) Time series of radar and model assigned precipitation regimes for the monsoon period. b) Domain-averaged vertical velocity profiles for the Lin simulation. c) Domain-averaged vertical velocity profiles for the Thompson simulation. d) Observed vertical velocity profiles from the Xie-Klein forcing dataset.**

## 7.1.2 Monsoon Period:- Thompson microphysics

Looking at the time series of regime assignments for the Thompson derived histograms it can be seen that the occurrence of the regimes with strong convection (2 and 4) again correlate well with the domain-averaged vertical velocity profiles in figure 7.1 c). One of the major differences between the Lin and the Thompson histograms for this period appears to be the infrequent assignment of the Thompson histograms to regime 3 and 4. The Thompson histograms are rarely assigned to regime 3, and the rapid change between regime 2 and 4 during the last 12 hours on the 23rd gives some indication that the assignment of the Thompson histograms to a given regime is unstable.

Chapters 4 and 6 have shown that the Thompson simulations have a significant problem below the freezing level. To investigate if this problem may be linked with infrequent assignment of the Thompson histograms to regime 3 and the unstable assignment of the Thompson histograms to regimes 2 and 4, Figure 7.2 a) shows the average of all histograms during the monsoon period for the Thompson simulation (the average Lin histogram and precipitation regimes 3 and 4 from chapter 2 are shown for completeness).

The average Thompson histogram shows that the problem below the freezing level occurs primarily between 5 and 30 dbz. Comparing figures 7.2 a) and c) it can be seen that regime 3 has a high frequency of particles with reflectivity values in the same range that the Thompson simulation is deficient. This would explain why regime 3 appears to be under represented when the Thompson histograms are assigned to the precipitation regimes. It should be noted that regime 4 also has a high frequency of particles with reflectivity values in the range that the Thompson physics has a problem simulating, however regime 4 also has a high frequency of particles above the freezing layer. If the Thompson simulation correctly produces large amounts of precipitation above the freezing level, then the Thompson histograms

may be assigned to regime 4 if the correct simulation above the freezing layer makes up for the lack between 5 and 30 dbz below the freezing layer (however such an assignment would likely be unstable, as was observed).

Earlier it was postulated that the problem with the Thompson simulation below the freezing level was linked with stratiform precipitation. As stratiform precipitation usually falls within the 5-30 dBz range (Houze, 1997; Steiner et al., 1995), figure 7.2 a) provides the best evidence so far that stratiform precipitation is contributing to the underestimation of precipitation in the Thompson simulations. While it is possible that some of the convective precipitation is affected by the problem below the freezing level, the most intense convective precipitation (35 dBz and greater) appears to be relatively unaffected. Therefore only during periods in which stratiform precipitation contributes significantly to the overall rain rate will the problem with the Thompson microphysics be obvious.
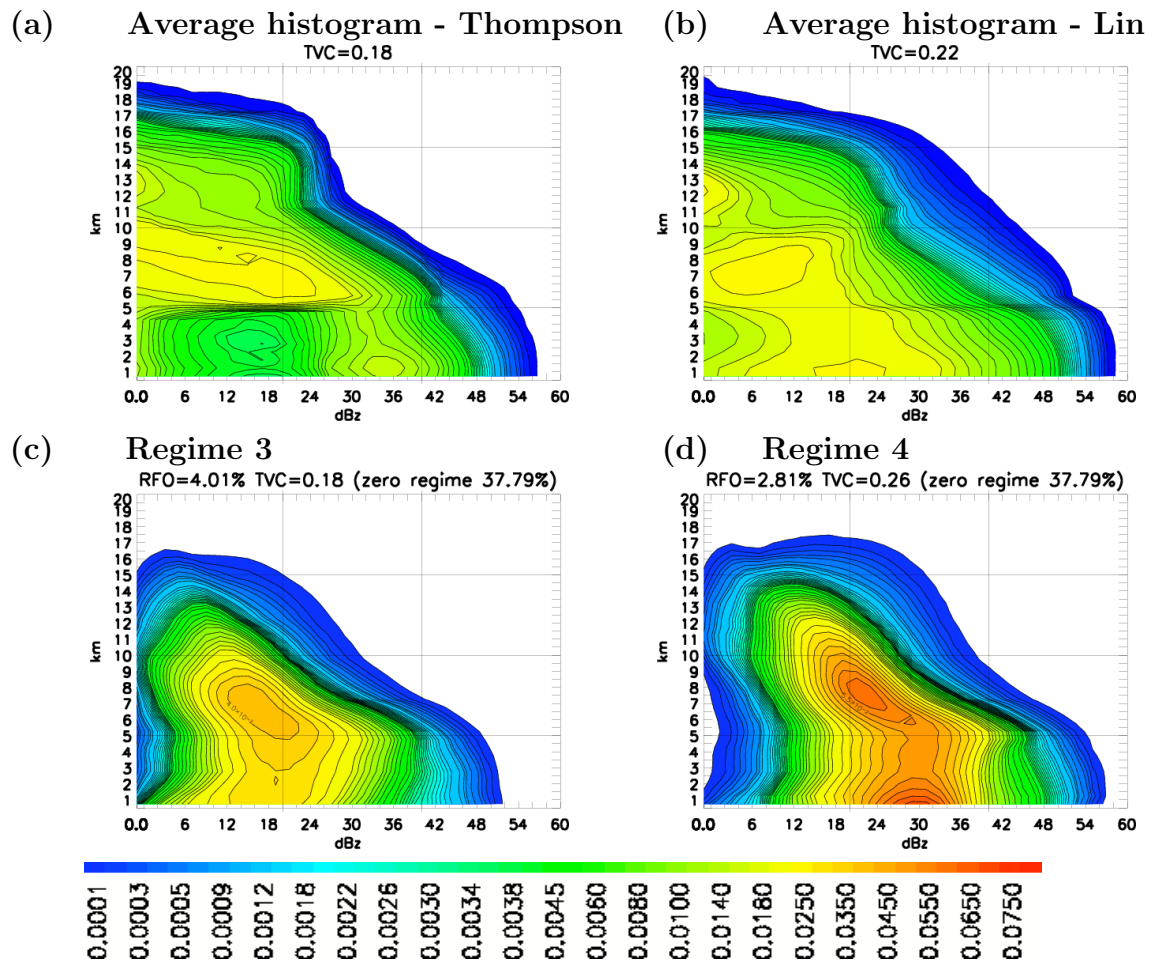
Figure 7.2: **a) Average histogram for the Thompson simulation over the monsoon period. b) Average histogram for the Lin simulation over the monsoon period. c) Precipitation regime 3. d) Precipitation regime 4.**

### 7.1.3   Regime assignments for TWP-ICE

Figure 7.3 shows the time series of regime assignments for the Lin and Thompson simulation during the entire TWP-ICE period. Data from all five simulations described in section 3.1.3 have been used to create the time series. During suppressed conditions convection is generally patchy, relatively weak, and produces relatively small amounts of precipitation. It would be therefore be expected that the suppressed period would be dominated by the occurrence of regime 1. Indeed both model simulations appear to be performing well during the suppressed conditions, the model and the radar histograms are assigned to regime 1 for the majority of the time, and both model simulations capture the transition from monsoon to suppressed conditions with only a slight delay.

During the break period the WRF model does not appear to be performing well, both the Lin and Thompson simulations rarely leave regime 1 despite the frequent assignment of the radar histograms to regime 2 and 3. The Thompson simulation is assigned more frequently to the zero precipitation regime than the Lin simulation, but strangely is also assigned more frequently to regime 2. Insight into why the Thompson simulation has a higher occurrence of regime 2 during the break period, even though it was found to have weaker vertical velocity profiles during this time can be gained by recalling Figure 6.9) in chapter 6. There it was found that the Thompson simulation performed poorly in simulating the coverage of rain below the freezing level, but the coverage of snow above the freezing level was much closer to the observed values than the Lin simulation. Thus it would appear plausible that even with the poor representation of rain below the freezing level, a more accurate representation of snow above the freezing level makes up for this deficiency, and the histograms are assigned to regime 2. Indeed this is exactly what has occurred, figure 7.4 shows the difference between the model histogram and the first two precipitation regimes on the 11 of February at 11 UTC. In the top of each figure is shown the

total difference between the histogram and the precipitation regime. Recall that the histogram is assigned to the regime with the minimum euclidian distance. Here we can see why the Lin histogram is assigned to regime 1 for this time period (it has a difference of 4.14 compared to 5.3 for regime 2), whereas the Thompson histogram is assigned to regime 2 with a difference of 6.07 compared with 7.45 for regime 1. Although there is a larger difference below the freezing level for medium reflectivity values (20-40 dbz) in the Thompson histogram, it more closely matches reflectivity values above the freezing level (5-10 km). This figure also shows that the Lin histogram is actually closer to regime 2 than the Thompson histogram (5.29 compared with 6.07), however the lack of reflectivity values above the freezing level means that the histogram is assigned to regime 1.

Judging purely from how the model derived histograms are assigned to the four precipitation regimes it would seem that the Lin simulation performs well in the Monsoon period and not so well in the break period, and the Thompson simulation performs poorly in the Monsoon period but "better" in the break period. Quotations are used because the above analysis highlights a potential problem with assigning the model histograms to a precipitation regime. On first glance it would appear the Thompson simulation performs better than the Lin simulation during the break period because it is assigned to the stronger convective regime more often. However looking at the differences it could be seen that the Lin histogram is actually closer to regime 2 than the Thompson histogram (it has a difference of 5.29 compared with 6.07). Therefore simply assigning a histogram to the regime with the smallest euclidian distance has a problem, it gives potentially misleading information about how realistic the simulated histograms are. No matter how un-realistic the histogram it will always be assigned to one of the four precipitation regimes. Even a fictitious histogram with 100 % coverage of 30 dbz reflectivity values at all levels would still be assigned to a regime (regime 4).

Therefore is must be concluded that while the assignment histograms to the precipitation regimes are useful as a first indication of how well a simulation performs, a separate analysis is required to determine the validity of the simulated histograms in terms of a realistic precipitation structure. This is the focus on the next section of this chapter.

**(a)  Time Series of Precipitation Regimes:- All TWPICE**
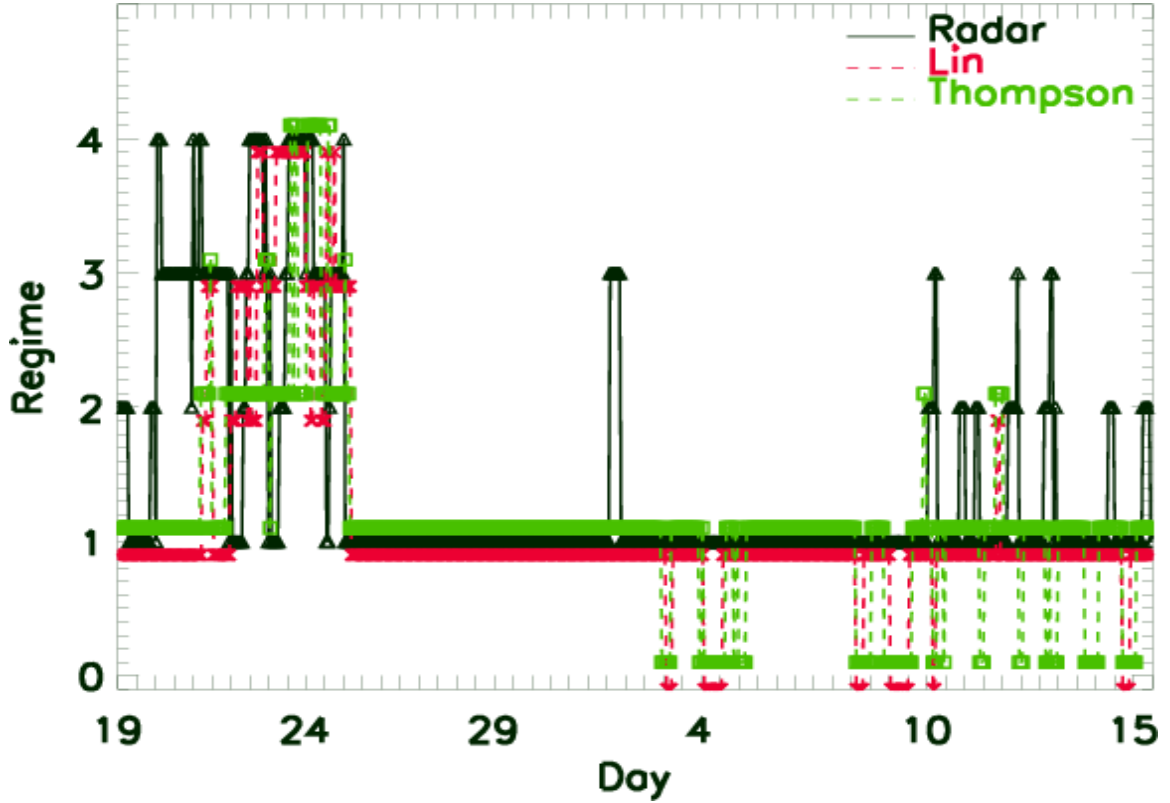


Figure 7.3: **Time series of radar and model assigned precipitation regimes for the TWP-ICE period.**
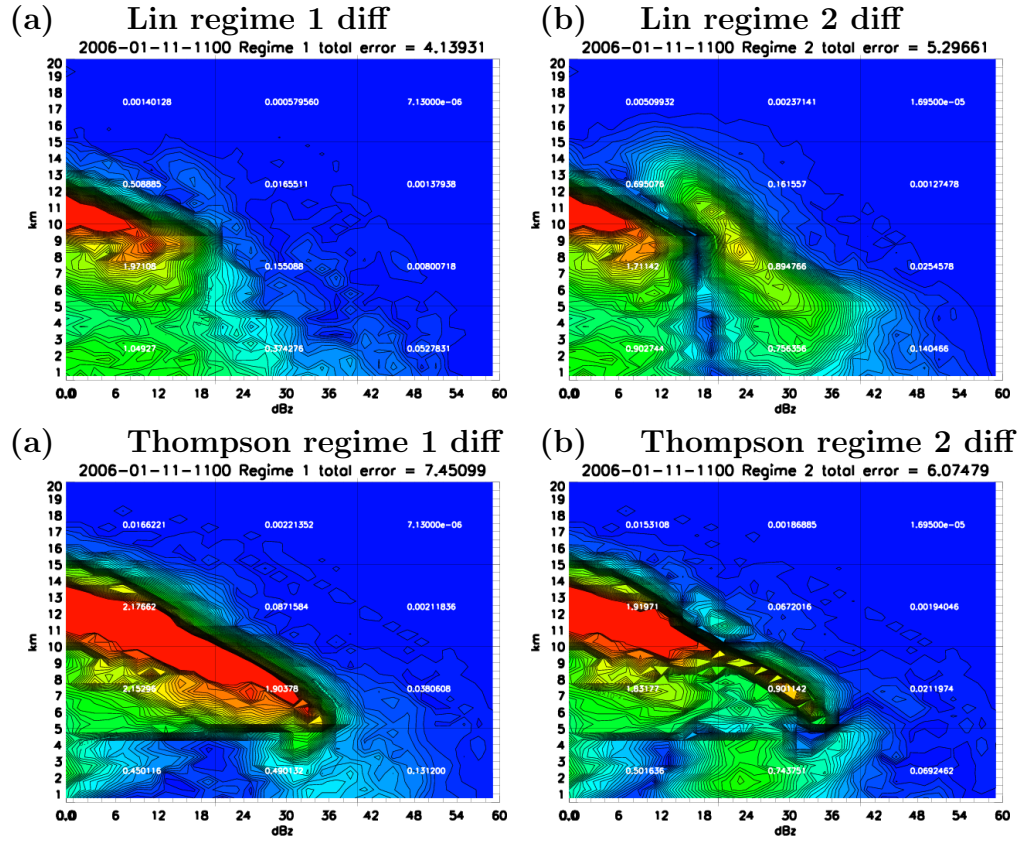
Figure 7.4: **Difference between the Lin histogram and regime 1 on the 23rd of January 11 UTC. b) Difference between the Lin histogram and regime 2 on the 23rd of January 11 UTC. c) Difference between the Thompson histogram and regime 1 on the 23rd of January 11 UTC. d) Difference between the Thompson histogram and regime 2 on the 23rd of January 11 UTC.**

## 7.2 Cluster Analysis with Model and Radar Histograms

The previous section has shown that the histograms derived from model simulated reflectivity data can be assigned to each of the four precipitation regimes at some point over the TWP-ICE period. It has also been shown that even if a histogram is assigned to a given precipitation regime this does not necessarily mean the histogram is physically sensible. In an attempt to overcome this problem and derive a more useful method for evaluating a models performance, the clustering technique used to define the four precipitation regimes in chapter 2 is employed a second time. Unlike previous studies where model derived data was clustered separately (Williams et al., 2005; Williams and Tselioudis, 2007), here it was decided that both radar and model histograms for the entire TWP-ICE period should be used as inputs into the KMEANS algorithm. The rationale behind the decision came from the notion that the clustering algorithm does not differentiate between model and radar derived histograms, it assigns each histogram to a cluster based purely on how similar the reflectivity structures in each histogram are to the cluster centroids.

It is hoped that information about how well (or poorly) the model performs in different conditions can be gained by looking at how the clustering algorithm assigns the model and radar histograms to the cluster centroids. For example, if the model performs well under certain conditions then it would be expected that the clustering algorithm would assign both model and radar histograms to the same cluster centroid, indicating that the model is able to produce precipitation structures that are similar to those observed by the radar. If however the model performs poorly under certain conditions (is unable to simulate certain types of convection), then it would be expected the clustering algorithm would "see" the model and radar histograms as different and assign each to seperate clusters, one that represents

reality and is comprised purely of radar histograms, and the other representing a distorted model version of the same conditions. Finally, if the model is unable to simulate certain meteorological conditions completely, then it would be expected that a radar-only cluster would be produced that contained no model histograms and has no model-equivalent.

Using both model and radar histograms as input in the clustering algorithm has the potential to allow conditions in which the model performs poorly to be ranked in terms of their importance, or their difference from reality. Recall that the KMEANS algorithm requires the user to choose the number of defined clusters. This is often seen as a disadvantage because the act of choosing the number of clusters the algorithm searches for includes a subjective step in an otherwise objective process. However in this case it can be turned into an advantage. By initially searching for a low number of clusters, model and radar histograms will separate first during the conditions in which the model performs the "worst", i.e. the histograms with the most unrealistic precipitation structures compared with observations will be the first to separate. The number of clusters the algorithm will search for can then be increased, and if the model and radar histograms again separate, then conditions in which the model "second worst" have been found. This process can be continued for as long as the user believes benefit is gained.

Figure 7.5 a) shows the results from the clustering algorithm when 5 clusters were sought using histograms derived from the radar and Lin simulations as input into the algorithm (similar results were obtained for 4 clusters). It appears that the clustering process has indeed sorted the model and radar histograms in a way that provides useful information about the model's performance. Looking at the histograms that make up each cluster it can be seen that cluster 2 is comprised solely of radar derived histograms. This type of separation is exactly what would be expected of a model unable to simulate conditions/convection with statistical

properties similar to those observed. Insight into the type of convection associated with cluster 2 can be gained by looking at the centroid of this cluster. Although it is not necessary to have performed prior analysis using only radar histograms as input to the clustering algorithm, the regimes defined in chapter 2 represent the four dominant reflectivity structures that occur over a wet season and should prove useful when analyzing the current cluster centroids.

Figure 7.5 b)-f) shows the centroids produced when 5 clusters are required from the KMEANS algorithm. The reflectivity profile associated with cluster 2 (Figure 7.5 c)) appears similar in structure and coverage to what would be expected of weak/medium intensity convection. Comparing this centroid to the regimes defined in chapter 2 it can be seen that this cluster is most similar to precipitation regime 1. Likewise, the centroid for cluster 1 (Figure 7.5 b)) also has a similar reflectivity structure to regime 1. Thus it appears that clusters 1 and 2 represent the same physical conditions, cluster 2 is comprised solely of the observed reflectivity structures associated with weak/medium convection, while centroid 1 contains the model version of weak/medium convection. It should be noted that some of the time the model is able to produce reflectivity structures that are similar enough to the radar histograms to allow concurrent assignment to cluster 1. However, the fact that clusters 1 and 2 were the first clusters to naturally separate shows that the biggest problem with the WRF model is its ability to simulate weak convection (or regime 1).

Looking at the histograms that make up the other three clusters it can be seen that both model and radar derived histograms are assigned to clusters 3, 4 and 5 at some stage of the simulation. Therefore it can be concluded that the model is able to create histograms with statistical properties similar to the radar histograms to allow concurrent assignment within these clusters. Of course searching for additional clusters are may lead to further separation and highlight additional deficiencies in

the models performance. However before searching for additional clusters a similar analysis will be performed using input from the radar and Thompson derived histograms.

Figure 7.6 a) shows the results from the clustering algorithm when 5 clusters were sought using histograms derived from the radar and the Thompson simulations as input into the algorithm. As was the case with the Lin simulation there are no model histograms assigned to cluster 2, indicating that the WRF model has a problem simulating weak convective periods (or regime 1) independent of the microphysics scheme used. It is possible that the grid spacing used in these simulations is not sufficient to accurately resolve this type of convection. Models that use grid spacings of approximately 1 km are often referred to as cloud resolving models, however the patchy convection associated with regime 1 may be too small to be resolved by grid spacings of this size. It would be interesting to see if increasing the resolution model has an effect on how the model and radar histograms are assigned to centroids by the clustering algorithm, however such an analysis is beyond the scope of the current investigation.

Previous chapters have shown that in the tropics the Thompson microphysics has a problem simulating precipitation coverage below the freezing level. In section 7.1.2 this deficiency was seen as a decrease in reflectivity values between 5 and 30 dBZ in the average histogram for the Thompson simulation (Figure 7.2 a)). The clustering algorithm appears to have picked up on differences in the reflectivity structure between the Thompson and radar histograms and created a second radar-only cluster (cluster 4) that was not created when Lin histograms were used instead of the Thompson histograms. Figure 7.6 b)-f) shows the centroids for the 5 clusters. Cluster 3 shows the characteristic "forbidden" zone in the reflectivity values first discovered in section 7.1.2. While some of the radar histograms have been assigned to this cluster (likely because their reflectivity structure is somewhere between cluster

2 and cluster 4), the unrealistic structure in the model histograms can clearly be seen to dominate this cluster.

Using the clustering algorithm to sort the radar and model histograms into clusters it can be concluded that the biggest problem with the model simulations (both Lin and Thompson) is the unrealistic representation of weak convection. Furthermore, the Lin microphysics scheme produces more realistic histograms than the Thompson microphysics scheme, as evidenced by the creation of a second radar-only cluster when the radar and Thompson histograms are used as input into the KMEANS algorithm.

The fact that the histograms separate into radar-only and model-equivalent clusters allows information about the models performance to be easily obtained, and provides a nice way of determining under which circumstances the model is performing poorly. The algorithm was able to pick out the problem below the freezing level in the Thompson histograms, as seen by the creation second radar-only cluster. The "forbidden" range of reflectivity values between 6 and 30 dBZ is easily seen in clusters 1 and 3 of figure 7.6, whereas the Lin simulation appears overestimate precipitation at low levels during periods of weak convection.

There is nothing special about 5 clusters, if one wanted to highlight further problems with the WRF model the algorithm could be set to find more clusters and run again. Figure 7.7 shows the time series of clusters assignments when 6 clusters were sought using histograms derived from the radar and a) the Lin simulations, and b) the Thompson simulations. By searching for an additional cluster the algorithm now produces two radar-only clusters when the Lin histograms are used (cluster 2 and cluster 6), and three radar-only clusters when the Thompson histograms are used (cluster 2, cluster 4 and cluster 6). Figures 7.8 shows the cluster centroids produced when the radar and Lin histograms were used as input into the clustering algorithm and Figure 7.9 shows cluster centroids produced when the radar and

Thompson histograms were used as input into the algorithm. From these figures it can be seen that the cluster centroids for clusters 1 and 2 are practically the same as those found when 5 clusters where sought. Again cluster 1 is a shared radar/model cluster and cluster 2 is a radar only cluster representing convection typical of that associated with regime 1. Cluster 6 appears to be an extremely strong convective regime with very large coverage of precipitation, reflectivity values of 60 dbz are observed below 3 km and reflectivity values up to 30 dbz are observed all the way up to 20 km. In chapter 3 it was found that the model simulated vertical velocity was always underestimated, therefore it is not unexpected that both microphysics schemes are unable unable to produce histograms with this converge and intensity. The Lin simulation is however able to produce strong convection with large regions of stratiform precipitation; cluster 5 is a shared model/radar cluster and has reflectivity values of 60 dbz below 3 km, an echo top height of 19.5 km and a TVC of 0.42.

Depending how well or poorly a model performs will dictate how useful it will be to continue searching for additional clusters, a possible application of this type of analysis involves determining how successful a change to the model setup will be. For example if one were to change the resultion of the model and weak convective periods were assigned to the same cluster as the radar histograms (without detrimentally effecting the other clusters) then one could state rather objectively that progress had been made. The process of evaluating a new microphysics scheme (or changes to a microphysics scheme) may involve trying to get all k clusters to be shared radar/model centroids. One could then investigate the differences or spread of the histograms within these clusters to determine if the model contains a systemic bias, or one could chose to search for additional clusters until radar or model only clusters are produced.

Figure 9 in the appendix shows the average histogram over the monsoon period derived from a) simulated reflectivity data generated using an exponential functions

to represent the particle size distribution for snow and b) simulated reflectivity data generated using gamma functions to representing the particle size distribution of snow. Figure 9 indicates that although the reflectivity structure has changed, the unrealistic nature of the precipitation within the Thompson simulation is still present. The under-representation of precipitation can still be seen in this figure, although the reflectivity range most effected has shifted slightly (now between 12 and 40 dBZ rather than 5 and 35 dBZ seen previously).

Given that reflectivity structure in the Thompson derived FAD's change slightly when the correct gamma functions are used in the simulated reflectivity algorithm, it is expected that the clusters created in section 7.2 would also be slighted modified were this data used as input the KMEANS algorithm. However, as this chapter focuses primarily on the methodology used for model evaluation rather than specifically focusing on the performance of the Thompson microphysics scheme, the general conclusions remain the same. The FADs derived from the simulated reflectivity data generated using the incorrect particle size distribution for snow can be thought of as providing simplified test case for a bad simulation.
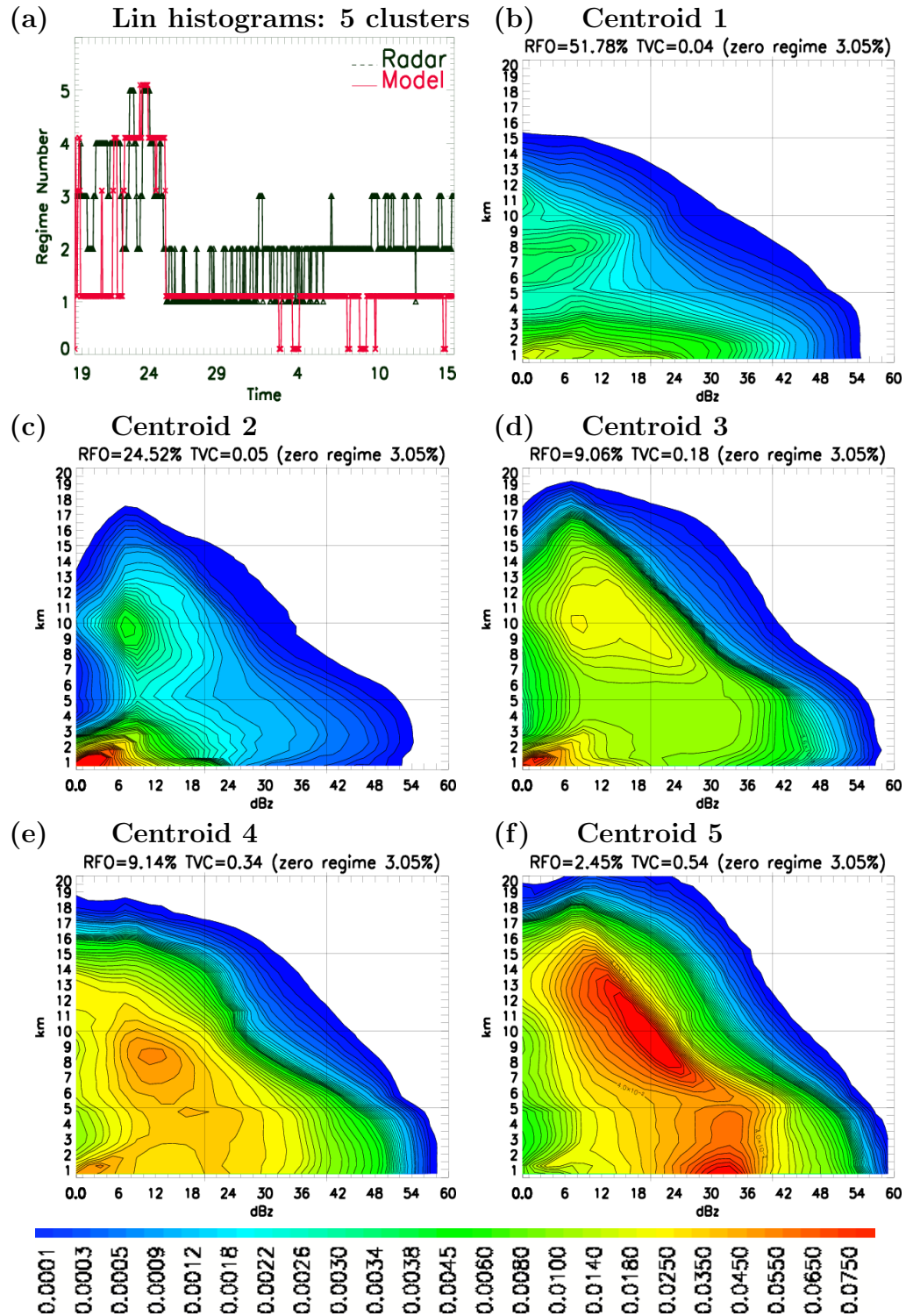
Figure 7.5: **a) Time series and centroids for the Lin and radar cluster analysis with 5 centroids. b)-f) the five centroids from the clustering algorithm**
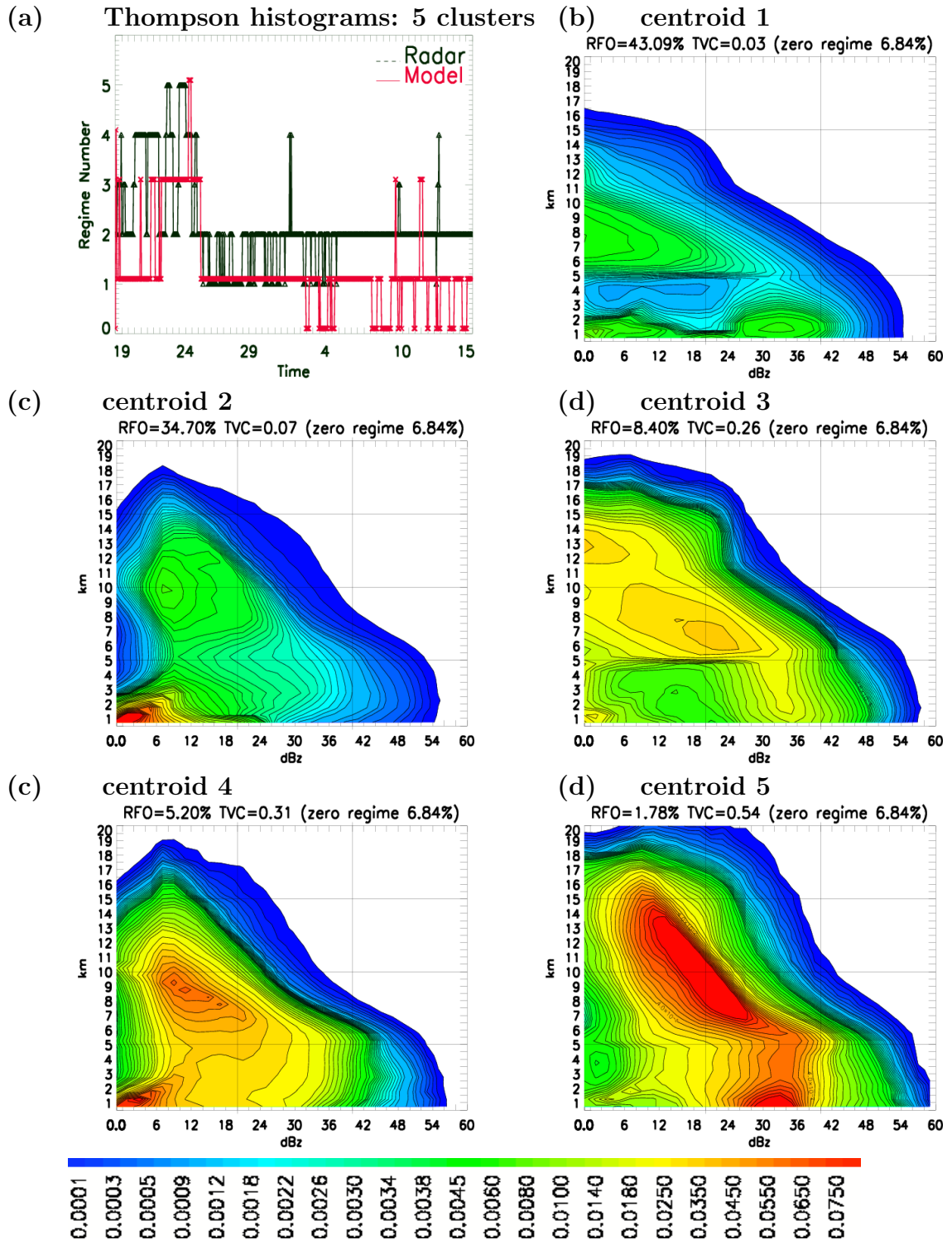
Figure 7.6: **a) Time series and centroids for the Thompson and radar cluster analysis with 5 centroids. b)-f) the five centroids from the clustering algorithm**
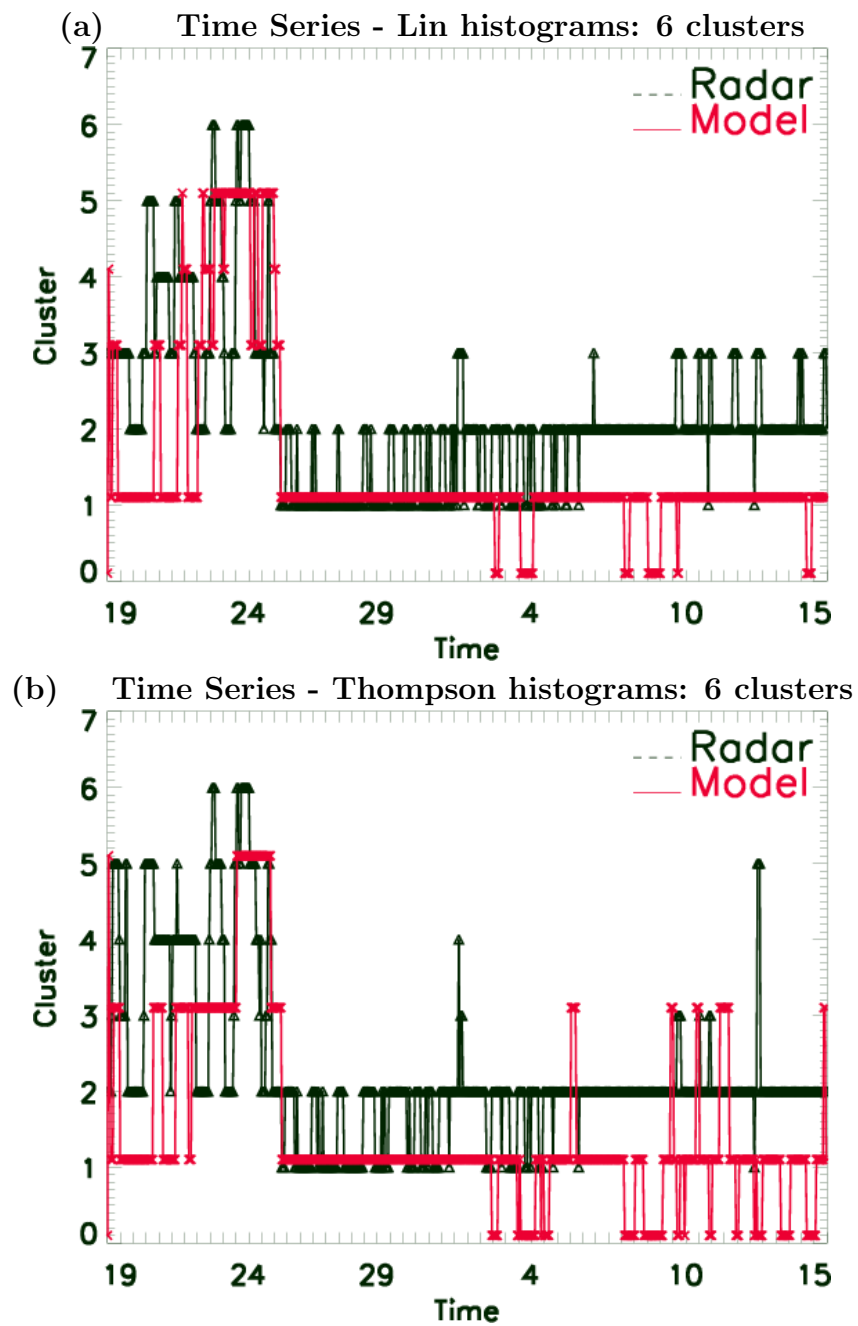
Figure 7.7: **a) Time series for the Lin and radar cluster analysis with 6 centroids. b) Time series for the Thompson and radar cluster analysis with 6 centroids.**
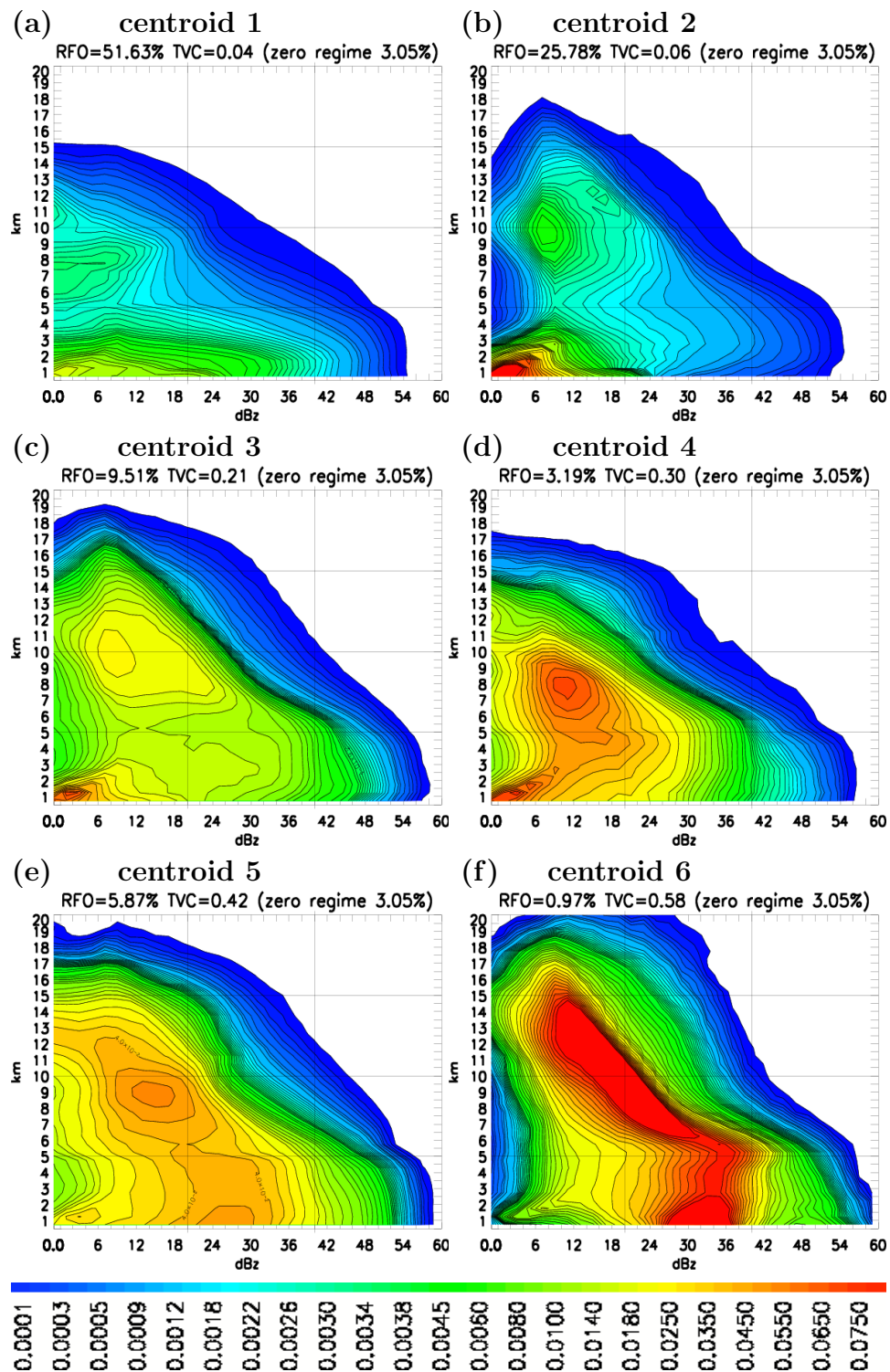
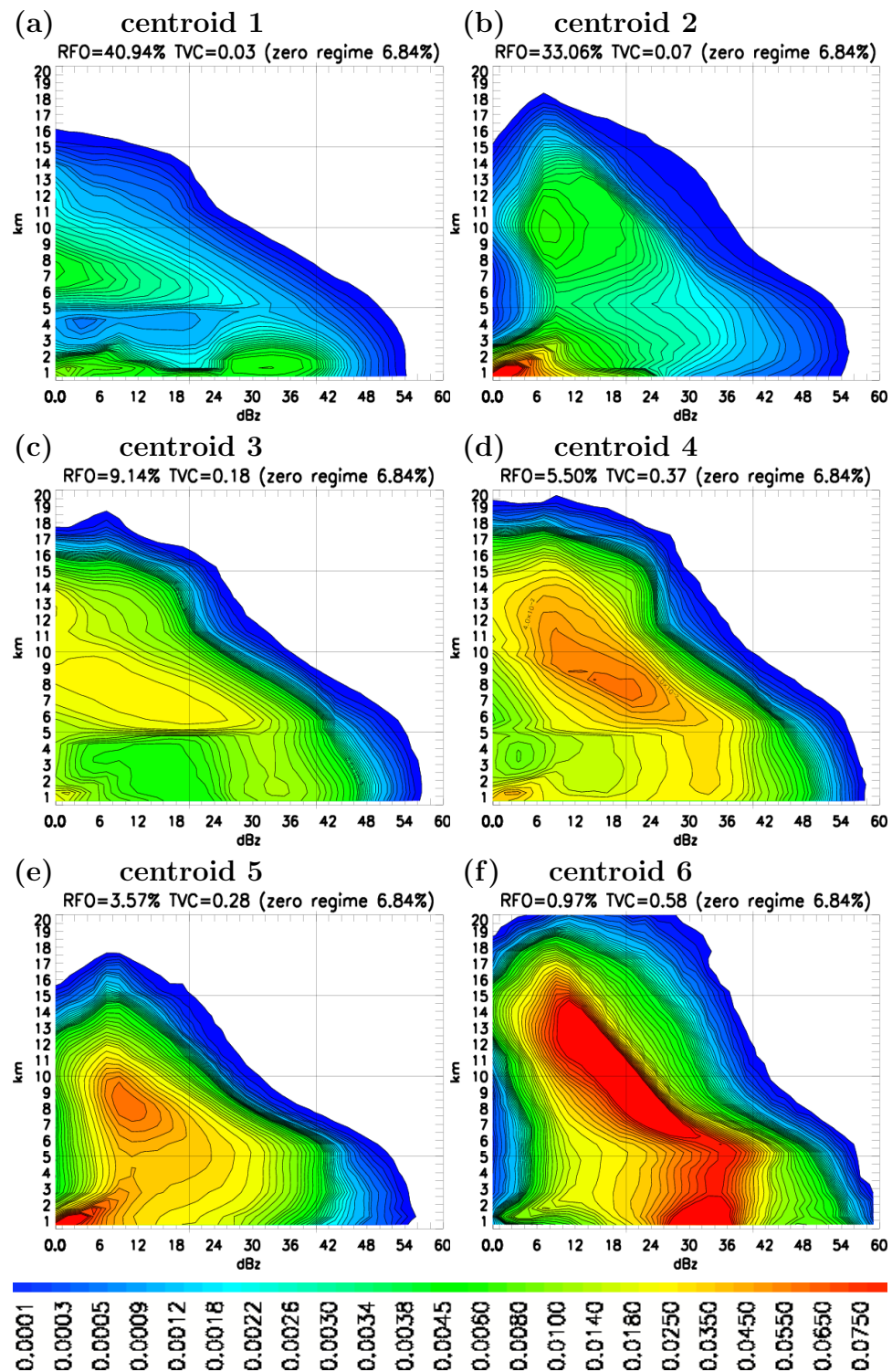Figure 7.8: **a)-f) Centroids for the Lin and radar cluster analysis with 6 centroids**

Figure 7.9: **a)-f) Centroids for the Thomson and radar cluster analysis with 6 centroids**

# Chapter 8

# Conclusions

Tropical convection is known to have a significant impact on the environment from the local to the global scale. The complex nature and wide variety of clouds that occur in the tropics makes accurately simulating tropical convection in numerical models is a notoriously difficult task. This thesis has investigated the regime nature of tropical convection in an attempt to simplify the large range of cloud types that occur over a tropical region into a discrete and manageable number.

Chapter 2 has shown that by using a simple clustering algorithm on multiple seasons of hourly radar data it was possible to objectively define four precipitation regimes that had physical connections to the known tropical atmosphere. Regimes 1 and 2 were found to be predominately convective in nature and diurnally driven, showing a distinct peak in the afternoon and occurring most frequently during break conditions. Regimes 3 and 4 contained precipitating clouds with a larger fraction of stratiform precipitation. Furthermore, regimes 3 and 4 had links with monsoon conditions and tended to peak in the late night or early morning, typical of convection occurring over tropical oceans. By linking the precipitation regimes with the tropical cloud regimes previously identified from ISCCP data it was possible to estimate the precipitation rates associated with each of the ISCCP regimes.

The remainder of the thesis was dedicated to evaluating the WRF model's performance in tropical north Australia, specifically, evaluating the simulated precipitation with data from a CPOL radar. Five consecutive simulations were run during the TWP-ICE period with the aim of capturing the large range of meteorological conditions present during a typical wet season over Darwin. The sequence of simulations

was run twice using both Lin and Thompson microphysics schemes.

A major source of inaccuracies in the model simulations was found to come not from the model itself, but rather the data used to force the model. In chapter 3 it was found that the GFS data used to force the WRF model contained significant biases in the temperature fields. Comparison of the GFS forcing data with observations taken from TWP-ICE revealed a warm bias at 17 km (stratosphere) reaching as high as 5 degrees, and a second bias of approximately 2 degrees at 15 km. Evidence of the detrimental effect these biases was found in the resulting analysis of the WRF model's performance. vertical velocity profiles indicated convection was capped by the warm bias and precipitation coverage was underestimated at higher altitudes.

The choice of microphysics scheme was found to have a large impact on the quality of the simulations. The Thompson scheme was shown to have a significant problem simulating the precipitation coverage below the freezing level, as demonstrated in chapter 4 where an unrealistic drop-off in precipitation coverage below 5 km was seen in the domain-averaged results. Chapter 6 then showed that the problem extended to both the coverage of rain and graupel production. Finally, in chapter 7 the average Thompson FAD showed that the reduction in precipitation coverage below the freezing level occurred primarily in values below 30 dbz. Times in which quantities of stratiform precipitation contributed significantly to the domain-averaged rain rates, such as during the passage of the MCS during the monsoon period, were most strongly affected by the problem in the Thompson microphysics.

Using the polarimetric variables available from a CPOL radar to estimate the observed hydrometeor species provided valuable information that could not be gained by using radar reflectivity alone. In chapter 6 it was found that the Lin microphysics significantly overestimated graupel, which was then hypothesized to be a cause of the underestimation of snow in the Lin simulations (the higher fall speeds of graupel particles allowed water mass to precipitate too quickly).

One of the main goals of the thesis was to determine whether the precipitation regimes defined in chapter 2 could be used to evaluate the model's performance. In chapter 7 each of the radar and model simulated FADs were assigned to one of the precipitation regimes and the time series of these regime assignments allowed for a quick and easy analysis of the models performance. Information about the strength of convection, presence of stratiform precipitation and duration of the stratiform precipitation associated with a convective event was evident in the time series. Assigning model and radar FADs to the precipitation regimes was a useful first check of the models performance. However, a limitation of this technique is that it gives no indication of how realistic the model simulated precipitation structure is when compared with observations of real clouds. In an attempt to gain such information the clustering algorithm used to initially derive the precipitation regimes was employed again, using both model and radar FADs as inputs to the clustering algorithm. This technique allowed model and radar FADs to naturally separate into a different clusters due to differences in their precipitation structure. As the clustering algorithm is objective in the way it assigns FADs to each centroid, this technique provides an objective method for determining where efforts should be focused when improving model performance. By initially searching for a low number of clusters then incrementally increasing this number, conditions in which the model performs poorly can be seen by the order in which the FADs separate into different clusters. One possible outcome of this technique is that two clusters representing the same physical conditions are produced, one containing radar FADs and the other containing model FADs. This occurred when 5 clusters were sought using input from the Lin and radar FADs. Cluster 1 contained all the model FADs associated with low/medium convective intensity and low precipitation coverage (although some radar FADs were also assigned to this cluster), while cluster 2 was purely a radar cluster. As these were the first clusters to naturally separate, one can draw the conclusion that the biggest

problem with the WRF model's performance in simulating the tropical atmosphere over Darwin (initialised with GFS data) is the representation of weak convective time periods (or regime 1). A second possible outcome of this technique is that a radar-only cluster is produced that has no model-equivalent. This occurred when 6 clusters were sought using both Lin and radar FADs as input in the clustering algorithm, cluster 6 was a radar-only cluster that had very high precipitation coverage and contained extremely strong convection. As there was no model-equivalent of cluster 6 one can objectively state that the second biggest problem with the WRF model is its inability to simulate periods strong deep convection with large coverage of stratiform precipitation.

While it is possible to continue searching for additional clusters and further highlighting inconsistencies in the WRF model, searching for additional problems with the model in the current setup was deemed an inefficient use of resources. As the temperature bias in the GFS data used to force the model simulations was found to have a large effect on the simulations, the logical next step in any research would be to run the WRF model initialised with a different source. The clustering technique described in chapter 7 could then be used to determine if any improvements were gained by using the additional data source, as well as investigating any deficiencies the model produces in the new model setup.

In summary it would seem that the use of the precipitation regimes by themselves are useful, but of limited value for model evaluation. By far the greatest potential is to the use precipitation regimes in conjunction with the clustering algorithm applied to both model and radar derived FADs. Such a process appears to have potential to be used as an objective method in determining where efforts are needed to improve model performance. Furthermore, this technique has a much wider applicability than was demonstrated in this thesis, any variable or quantity that can be both observed and simulated by a numerical model can be used to create histograms and

used as input into the clustering algorithm.

# Bibliography

Adler, R. F., et al., 2003: The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979Present). *Journal of Hydrometeorology*, **4**, 11471167.

Anderberg, M. R., 1973: *Cluster Analysis for Applications*. Academic Press, 359 pp.

Arakawa, A., 2004: The Cumulus Parameterization Problem: Past, Present, and Future. *Journal of Climate*, **17**, 2493–2525.

Boccippio, D. J., W. A. Petersen, and D. J. Cecil, 2005: The Tropical Convective Spectrum. Part I: Archetypal Vertical Structures. *Journal of Climate*, **18**, 2744–2769.

Brandes, E. A., G. Zhang, and J. Vivekanandan, 2002: The Distribution with size of aggregate snowflakes. . *Journal of Applied Meteorology*, **41**, 674–685.

Bringi, V. N., G.-J. Huang, V. Chandrasekar, and T. D. Keenan, 2001: An Areal Rainfall Estimator Using Differential Propagation Phase: Evaluation Using a C-Band Radar and a Dense Gauge Network in the Tropics. *Journal of Atmospheric and Oceanic Technology*, **18**, 1810–1818.

Bringi, V. N., T. Tang, and V. Chandrasekar, 2004: Evaluation of a new polarimetrically-based Z-R relation. *Journal of Atmospheric and Oceanic Technology*, **21**, 612–622.

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**, 2687–2710.

Cetrone, J. and R. A. Houze, 2006: Characteristics of Tropical Convection over the Ocean near Kwajalein. *Monthly Weather Review*, **134**, 834–853.

Chen, S.-H. and W.-Y. Sun, 2002: A One-dimensional Time Dependent Cloud Model. *Journal of the Meteorological Society of Japan*, **80**, 99–118.

Cheng, W. and W. Steenburgh, 2005: Evaluation of Surface Sensible Weather Forecasts by the WRF and the Eta Models over the Western United States . *Weather and Forecasting*, **20**, 812–821.

Chou, C. and J. D. Neelin, 1999: Cirrus Detrainment-Temperature Feedback. *Geophysical Research Letters*, **26**, 1295–1298.

Crosson, W. L., C. E. Duchon, R. Raghavan, and S. J. Goodman, 1996: Assessment of Rainfall Estimates Using a Standard Z-R Relationship and the Probability Matching Method Applied to Composite Radar Data in Central Florida . *Journal of Applied Meteorology*, **35**, 1203–1219.

Demott, C. A. and S. A. Rutledge, 1998: The Vertical Structure of TOGA COARE Convection. Part I: Radar Echo Distributions. *Journal of the Atmospheric Sciences*, **55**, 2730–2747.

Done, J., C. A. Davis, and M. Weisman, 2004: The next generation of NWP: explicit forecasts of convection using the weather research and forecasting (WRF) model. *Atmospheric Science Letters*, **5**, 110–117.

Drosdowsky, W., 1996: Variability of the Australian Summer Monsoon at Darwin: 1957-1992. *Journal of Climate*, **9**, 85–96.

Fujiyoshi, Y., T. Endoh, T. Yamada, K. Tsuboki, Y. Tachibana, and G. Wakahama, 1990: Determination of a Z-R relationship for snowfall using a radar and high sensitivity snow guages. . *Journal of Applied Meteorology*, **29**, 147–152.

Gordon, N. D., J. R. Norris, C. P. Weaver, and S. A. Klein, 2005: Cluster analysis of cloud regimes and characteristic dynamics of midlatitude synoptic systems in observations and a model. *Journal of Geophysical Research.*, **110**, D15S17, doi:10.1029/2004JD005 027.

Gray, W. M. and R. W. Jacobson, 1977: Diurnal Variation of Deep Cumulus Convection. *Monthly Weather Review*, **105**, 1171–1188.

Gunn, K. L. S. and J. S. Marshall, 1958: The Distribution with size of aggregate snowflakes. . *Journal of Meteorology*, **15**, 452–461.

Hamilton, K., R. A. Vincent, and P. T. May, 2004: Darwin Area Wave Experiment (DAWEX) field campaign to study gravity wave generation and propagation. *Journal of Geophysical Research.*, **109**, D20S01, doi:10.1029/2003JD004 393.

Höglund, S., 2005: Clouds in Darwin and their relation to large-scale conditions. M.S. thesis, Applied Physics and Mechanical Engineering, Lulea University of Technology. http://epubl.luth.se/1402-1617/2005/169/LTU-EX-05169-SE.pdf.

Holland, G. J., J. L. McBride, R. K. Smith, D. Jasper, and T. D. Keenan, 1986: The BMRC Australian Monsoon Experiment. *Bulletin of the American Meteorological Society*, **67**, 1466–1472.

Houze, R. A., 1997: Stratiform Precipitation in Regions of Convection: A meteorological Paradox? *Bulletin of the American Meteorological Society*, **78**, 2179–2196.

Houze, R. A., C.-P. Cheng, C. A. Leary, and J. F. Gamache, 1980: Diagnosis of Cloud Mass and Heat Fluxes from Radar and Synoptic Data. *Journal of the Atmospheric Sciences*, **37**, 754–773.

Houze, R. A. and D. D. Churchill, 1984: Microphysical Structure of Winter Monsoon Cloud Clusters. *Journal of the Atmospheric Sciences*, **41**, 3405–3411.

Houze, R. A. and D. D. Churchill, 1987: Mesoscale Organization and Cloud Microphysics in a Bay of Bengal Depression. *Journal of the Atmospheric Sciences*, **44**, 1846–1867.

Houze, R. A., S. G. Geotis, F. D. Marks, D. D. Churchill, and P. H. Herzegh, 1981: Comparison of Airborne and Land-Based Radar Measurements of Precipitation during Winter MONEX. *Journal of Applied Meteorology.*, **20**, 772–783.

Hume, T. and C. Jakob, 2007: Ensemble single column model validation in the tropical western pacific. *Journal of Geophysical Research.*, **112**, D10 206,doi:10.1029/2006JD008 018.

Jakob, C. and G. Tselioudis, 2003: Objective identification of cloud regimes in the Tropical Western Pacific. *Geophysical Research Letters*, **30**, 2082, doi:10.1029/2003GL018 367.

Jakob, C., G. Tselioudis, and T. Hume, 2005: The Radiative, Cloud, and Thermodynamic Properties of the Major Tropical Western Pacific Cloud Regimes. *Journal of Climate*, **18**, 1203–1215.

Jankov, I., W. A. Gallus, M. Segal, B. Shaw, and S. E. Koch, 2005: The Impact of Different WRF Model Physical Parameterizations and Their Interactions on Warm Season MCS Rainfall. *Weather and Forecasting*, **20**, 1048–1060.

Kawashima, M., Y. Fujiyoshi, M. Ohi, T. Honda, T. Kozu, T. Shimomat, and H. Hashiguchi, 2006: Overview of Doppler Radar Observations of Precipitating Cloud Systems in Sumatera Island During the First CPEA Campaign. *Journal of the Meteorological Society of Japan*, **84A**, 33–56.

Keenan, T., K. Glasson, F. Cummings, T. S. Bird, J. Keeler, and J. Lutz, 1998: The BMRC/NCAR C-Band Polarimetric (C-POL) Radar System. *Journal of Atmospheric and Oceanic Technology*, **15**, 871–886.

Keenan, T., et al., 2000: The Maritime Continent Thunderstorm Experiment (MC-TEX): Overview and Some Results. *Bulletin of the American Meteorological Society*, **81**, 2433–2455.

Keenan, T. D. and R. E. Carbone, 1992: A preliminary morphology of precipitation systems in tropical northern Australia. *Quarterly Journal of the Royal Meteorological Society*, **118**, 283–326.

Lin, J., B. Mapes, M. Zhang, and M. Newman, 2004: Stratiform Precipitation, Vertical Heating Profiles, and the Madden-Julian Oscillation . *Journal of the Atmospheric Sciences*, **61**, 296–309.

Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk Parameterization of the Snow Field in a Cloud Model. *Journal of Climate and Applied Meteorology*, **22**, 1065–1092.

Mapes, B. E., 1993: Gregarious Tropical Convection. *Journal of the Atmospheric Sciences*, **50**, 2026–2037.

Mapes, B. E. and R. A. Houze, 1993: Cloud Clusters and Superclusters over the Oceanic Warm Pool. *Monthly Weather Review*, **121**, 1398–1415.

Mapes, B. E. and J. Lin, 2005: Doppler Radar Observations of Mesoscale Wind Divergence in Regions of Tropical Convection . *Monthly Weather Review*, **133**, 1808–1824.

Mapes, B. E., S. Tulich, J. Lin, and P. Zuidema, 2006: The mesoscale convection life cycle: building block or prototype for large-scale tropical waves . *Dynamics of Atmospheres and Oceans*, **42**, 3–29.

May, P. T. and T. D. Keenan, 2005: Evaluation of Microphysical Retrievals from Polarimetric Radar with Wind Profiler Data. *Journal of Applied Meteorology*, **44**, 827–838.

May, P. T., T. D. Keenan, D. S. Zrnić, L. D. Carey, and S. A. Rutledge, 1999: Polarimetric Radar Measurements of Tropical Rain at a 5-cm Wavelength. *Journal of Applied Meteorology*, **38**, 750–765.

May, P. T., J. H. Mather, G. Vaughan, C. Jakob, G. M. McFaquhar, K. Bower, and G. G. Mace, 2008: The Tropical Warm Pool International Cloud Experiment. *Bulletin of the American Meteorological Society*, 629–645.

Neiman, P. J., G. A. Wick, F. M. Ralph, B. E. Martner, A. B. White, and D. E. Kingsmill, 2005: Wintertime Nonbrightband Rain in California and Oregon during CALJET and PACJET: Geographic, Interannual, and Synoptic Variability. *Monthly Weather Review*, **133**, 1199–1223.

Nesbitt, S. and E. Zipser, 2003: The Diurnal Cycle of Rainfall and Convective Intensity according to Three Years of TRMM Measurements. *Journal of Climate*, **16**, 1456–1475.

Nesbitt, S. W., R. Cifelli, and S. A. Rutledge, 2006: Storm Morphology and Rainfall Characteristics of TRMM Precipitation Features. *Monthly Weather Review*, **134**, 2702–2721.

Nesbitt, S. W., E. J. Zipser, and D. J. Cecil, 2000: A Census of Precipitation Features in the Tropics Using TRMM: Radar, Ice Scattering, and Lightning Observations . *Journal of Climate*, **13**, 4087–4106.

Orlanski, I., 1975: A Rational Subdivision of Scales for Atmospheric Processes. *Bulletin of the American Meteorological Society*, **56**, 527–530.

Pattanayak, S. and U. C. Mohanty, 2006: Evaluation of WRF-ARW High-Resolution Tropical Storm Forecasts in 2005 Season. *current science*, **95**, 923–936.

Pope, M., C. Jakob, and M. J. Reeder, 2009: Regimes of the North Australian wet season. *Journal of Climate*, **accepted**.

Rasmussen, R. M., I. Geresdi, G. Thompson, K. Manning, and E. Karplus, 2002: Freezing Drizzle Formation in Stably Stratified Layer Clouds: The Role of Radiative Cooling of Droplets, Cloud Condensation Nuclei, and Ice Initiation. *Journal of the Atmospheric Sciences*, **59**, 837–860.

Rickenbach, T. M. and S. Rutledge, 1998: Convection in TOGA COARE: Horizontal Scale, Morphology, and Rainfall Production. *Journal of the Atmospheric Sciences*, **55**, 2715–2729.

Rogers, R., M. Black, S. Chen, and R. Black, 2007: An Evaluation of Microphysics Fields from Mesoscale Model Simulations of Tropical Cyclones. Part I: Comparisons with Observations. *Journal of the Atmospheric Sciences*, **64**, 1811–1834.

Rossow, W. B., G. Tselioudis, A. Polak, and C. Jakob., 2005: Tropical climate described as a distribution of weather states indicated by distinct mesoscale cloud properties. *Geophysical Research Letters*, **32**, L21 812,doi:10.1029/2005GL024 584.

Schiffer, R. A. and W. B. Rossow, 1983: The International Satellite Cloud Climatology Project (ISCCP): The First Project of the World Climate Research Programme. *Bulletin of the American Meteorological Society*, **64**, 779–784.

Schumacher, C., R. A. Houze, Jr., and I. Kraucunas, 2004: The Tropical Dynamical Response to Latent Heating Estimates Derived from the TRMM Precipitation Radar . *Journal of the Atmospheric Sciences*, **61**, 1341–1358.

Skamarock, W., J. Klemp, J. Dudhia, D. Gill, D. Baker, W. Wang, and J. Powers, 2005: A Description of the Advanced Research WRF Version 2. *NCAR/TN-468+STR*, 88 pp.

Steiner, M., R. A. Houze, Jr., and S. E. Yuter, 1995: Climatological Characterization of Three-Dimensional Storm Structure from Operational Radar and Rain Gauge Data. *Journal of Applied Meteorology*, **34**, 1978–2007.

Straka, J. M., D. S. Zrnic, and A. V. Ryzhkov, 2000: Bulk Hydrometeor Classification and Quantification Using Polarimetric Radar Data: Synthesis of Relations. *Journal of Applied Meteorology*, **39**, 1341–1372.

Tao, W.-K., et al., 2006: Retrieval of latent heating from TRMM measurements. *Bulletin of the American Meteorological Society*, 1555–1572.

Thompson, G., R. Rasmussen, and K. Manning, 2004: Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part I: Description and Sensitivity Analysis. *Monthly Weather Review*, **132**, 519–542.

Toracinta, E. R., D. J. Cecil, E. J. Zipser, and S. W. Nesbitt, 2002: Radar, Passive Microwave, and Lightning Characteristics of Precipitating Systems in the Tropics. *Monthly Weather Review*, **130**, 802–824.

Williams, K., C. A. Senior, A. Slingo, and J. F. B. Mitchell, 2005: Towards evaluating cloud response to climate change using clustering technique identification of cloud regimes. *Climate Dynamics*, **24**, 701–719.

Williams, K. and G. Tselioudis, 2007: GCM intercomparison of global cloud regimes: present-day evaluation and climate change response. *Climate Dynamics*, **29**, 231–250.

Xie, S., T. Hume, C. Jakob, S. A. Klein, R. McCoy, and M. Zhang, 2009: Observed Large-Scale Structures and Diabatic Heating and Drying Profiles during TWP-ICE . *Journal of Climate*, **accepted**.

Yang, G.-Y. and J. Slingo, 2001: The Diurnal Cycle in the Tropics. *Monthly Weather Review*, **129**, 784–801.

Yuter, S. E. and R. A. Houze, Jr., 1995: Three-Dimensional Kinematic and Microphysical Evolution of Florida Cumulonimbus. Part II: Frequency Distributions

of Vertical Velocity, Reflectivity, and Differential Reflectivity. *Monthly Weather Review*, **123**, 1941–1963.

Zhang, M. H. and J. L. Lin, 1997: Constrained Variational Analysis of Sounding Data Based on Column-Integrated Budgets of Mass, Heat, Moisture, and Momentum: Approach and Application to ARM Measurements . *Journal of the Atmospheric Sciences*, **54**, 1503–1524.

Zhang, M. H., J. L. Lin, R. T. Cederwall, J. J. Yio, and S. C. Xie, 2001: Objective Analysis of ARM IOP Data: Method and Sensitivity. *Monthly Weather Review*, **129**, 295–311.

# Appendix

$$Z_e = \int D^6 N_0 e^{-\lambda D} \, \mathrm{dD} \tag{A1}$$

Setting $U = D^6$ and $\frac{dV}{dD} = N_0 e^{-\lambda D} \Rightarrow V = \frac{-N_0 e^{-\lambda D}}{\lambda}$ and integrating equation A1 by parts we have

$$\int D^6 N_0 e^{-\lambda D} \, \mathrm{dD} = \frac{-D^6 N_0 e^{-\lambda D}}{\lambda} + \int \frac{6 D^5 N_0 e^{-\lambda D}}{\lambda} \, \mathrm{dD} \tag{A2}$$

where we integrate by parts again with $U = 6D^5$ and $\frac{dV}{dD} = \frac{N_0 e^{-\lambda D}}{\lambda} \Rightarrow V = \frac{-N_0 e^{-\lambda D}}{\lambda^2}$ thus

$$\int \frac{6 D^5 N_0 e^{-\lambda D}}{\lambda} \mathrm{dD} = \frac{-6 D^5 N_0 e^{-\lambda D}}{\lambda^2} + \int \frac{30 D^4 N_0 e^{-\lambda D}}{\lambda^2} \, \mathrm{dD} \tag{A3}$$

where we integrate by parts again with $U = 30D^4$ and $\frac{dV}{dD} = \frac{N_0 e^{-\lambda D}}{\lambda^2} \Rightarrow V = \frac{-N_0 e^{-\lambda D}}{\lambda^3}$ thus

$$\int \frac{30 D^4 N_0 e^{-\lambda D}}{\lambda^2} \mathrm{dD} = \frac{-30 D^4 N_0 e^{-\lambda D}}{\lambda^3} + \int \frac{120 D^3 N_0 e^{-\lambda D}}{\lambda^3} \mathrm{dD} \tag{A4}$$

where we integrate by parts again with $U = 120D^3$ and $\frac{dV}{dx} = \frac{N_0 e^{-\lambda D}}{\lambda^3} \Rightarrow V = \frac{-N_0 e^{-\lambda D}}{\lambda^4}$ thus

$$\int \frac{120 D^3 N_0 e^{-\lambda D}}{\lambda^3} \, \mathrm{dD} = \frac{-N_0 120 D^3 e^{-\lambda D}}{\lambda^4} + \int \frac{360 D^2 N_0 e^{-\lambda D}}{\lambda^4} \mathrm{dD} \tag{A5}$$

where we integrate by parts again

with $U = 360D^2$ and $\dfrac{dV}{dD} = \dfrac{N_0 e^{-\lambda D}}{\lambda^4} \Rightarrow V = \dfrac{-N_0 e^{-\lambda D}}{\lambda^5}$ thus

$$\int \frac{360D^2 N_0 e^{-\lambda D}}{\lambda^4} dD = \frac{-360D^2 N_0 e^{-\lambda D}}{\lambda^5} + \int \frac{720D N_0 e^{-\lambda D}}{\lambda^5} dD \quad (A6)$$

where we integrate by parts again with $U = 720D$ and $\dfrac{dV}{dD} = \dfrac{N_0 e^{-\lambda D}}{\lambda^5} \Rightarrow V = \dfrac{-N_0 e^{-\lambda D}}{\lambda^6}$ thus

$$\int \frac{720D N_0 e^{-\lambda D}}{\lambda^6} dD = \frac{-720D N_0 e^{-\lambda D}}{\lambda^6} + \int \frac{720 N_0 e^{-\lambda D}}{\lambda^6} dD \quad (A7)$$

where

$$\int \frac{720 N_0 e^{-\lambda D}}{\lambda^6} dD = \frac{-720 N_0 e^{-\lambda D}}{\lambda^7} \quad (A8)$$

Combining equations A8 - A2 gives

$$Z_e = \int D^6 N_0 e^{-\lambda D} \, dD =$$

$$\frac{-D^6 N_0 e^{-\lambda D}}{\lambda} \frac{-6D^5 N_0 e^{-\lambda D}}{\lambda^2} \frac{-30D^4 N_0 e^{-\lambda D}}{\lambda^3} \frac{-N_0 120D^3 e^{-\lambda D}}{\lambda^4}$$

$$\frac{-360D^2 N_0 e^{-\lambda D}}{\lambda^5} \frac{-720D N_0 e^{-\lambda D}}{\lambda^6} \frac{-720 N_0 e^{-\lambda D}}{\lambda^7}$$

$$= \frac{-N_0 e^{-\lambda D}}{\lambda^7} (D^6 \lambda^6 + 6D^5 \lambda^5 + 30D^4 \lambda^4 + 120D^3 \lambda^3 + 360D^2 \lambda^2 + 720D\lambda + 720)$$

with $D << 0$ we have

$$Z_e = \frac{-720N_0}{\lambda^7}$$
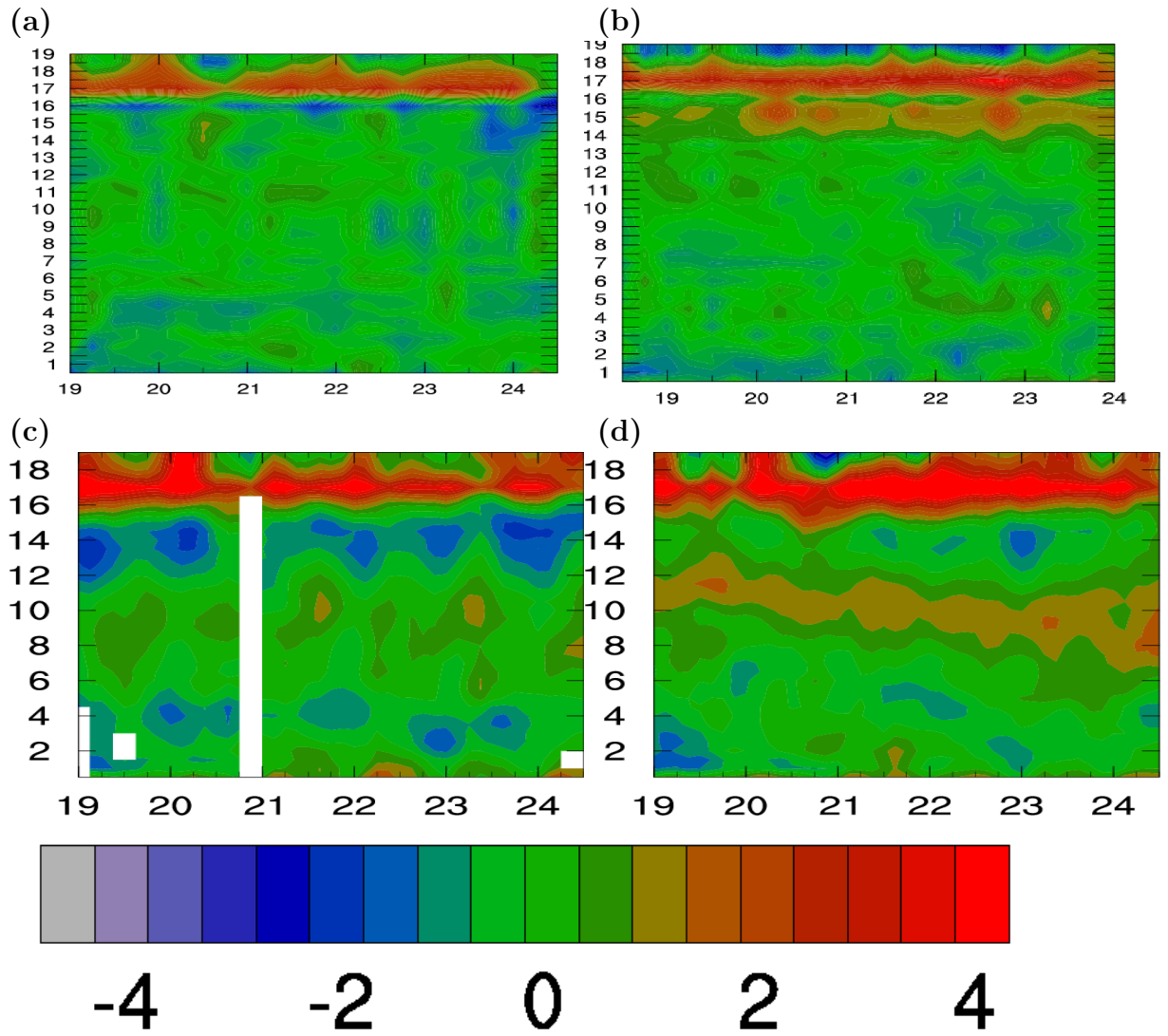
Figure 1: **Difference between the Xie-Klein forcing dataset and a) EWMWF input data. b) GFS input data. c) WRF model using ECMWF data. d) WRF model using GFS data.**
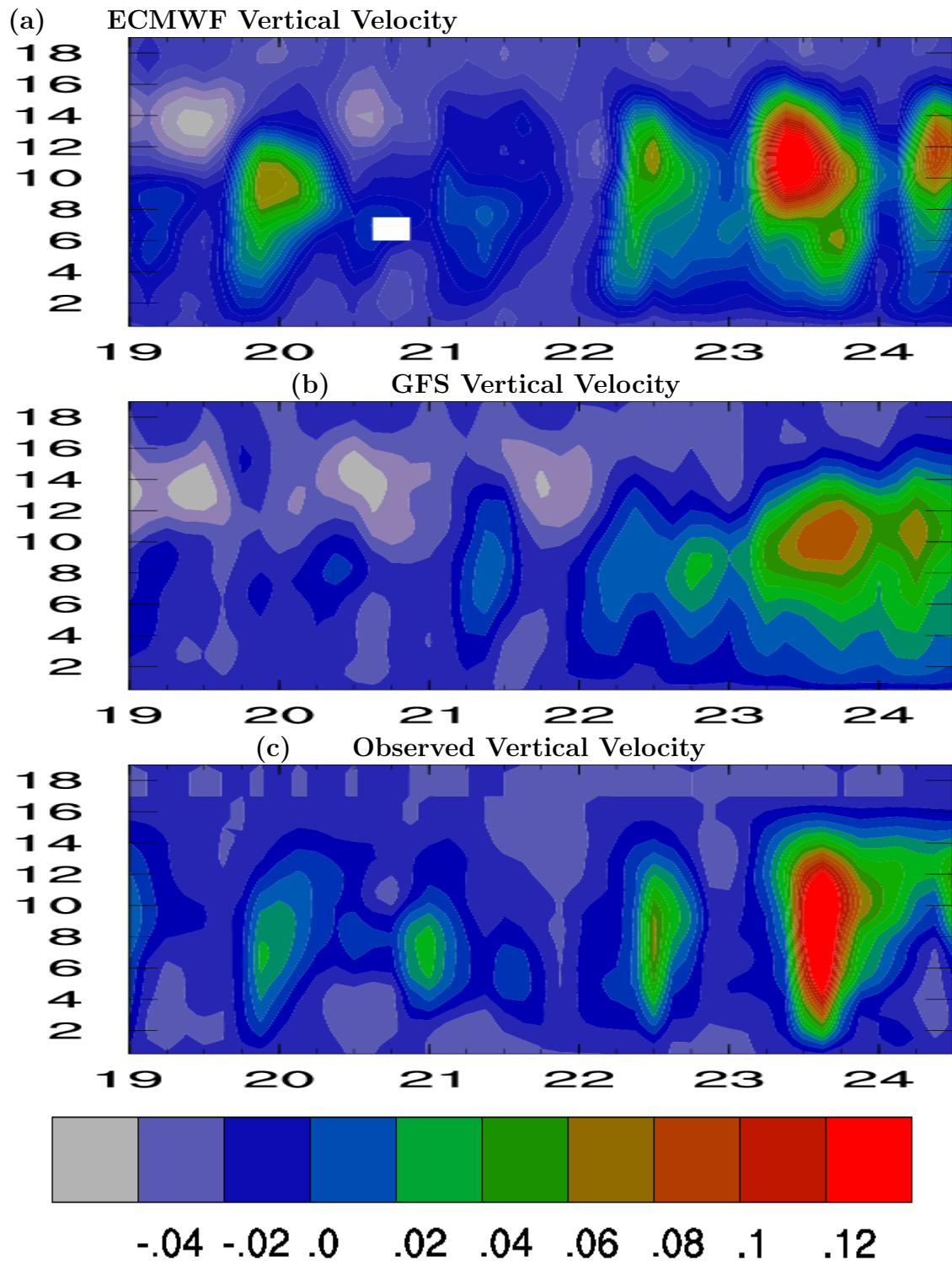
Figure 2: **Domain-averaged vertical velocity profiles for a) The WRF model forced with ECMWF data. b) the WRF model forced with GFS data. c) the Xie-Klein forcing dataset**
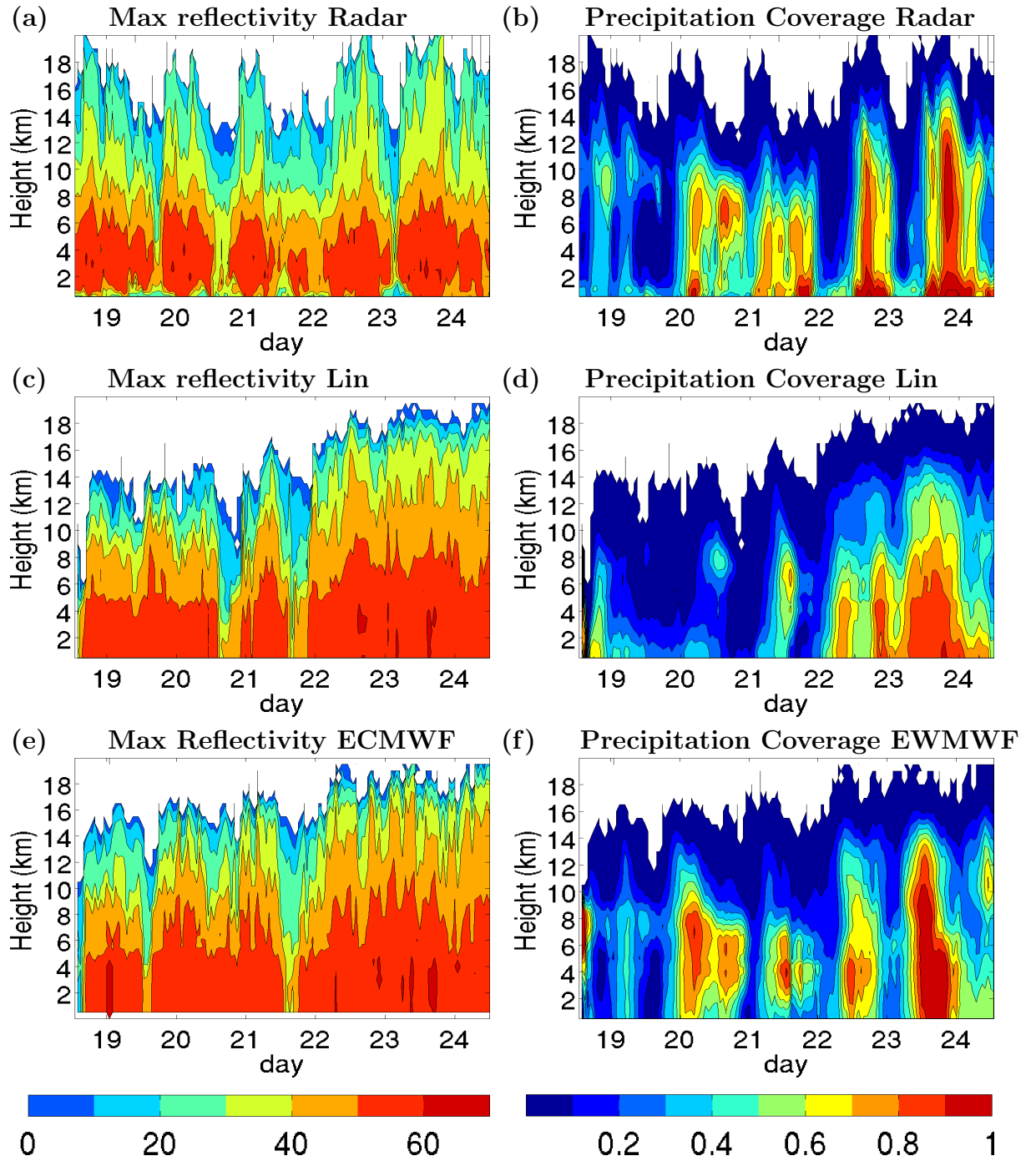
Figure 3: **Maximum reflectivity and precipitation coverage plots for a,b) the Lin simulation forced with GFS data. c),d) the radar and e) f) the Lin simulation forced with ECMWF data.**
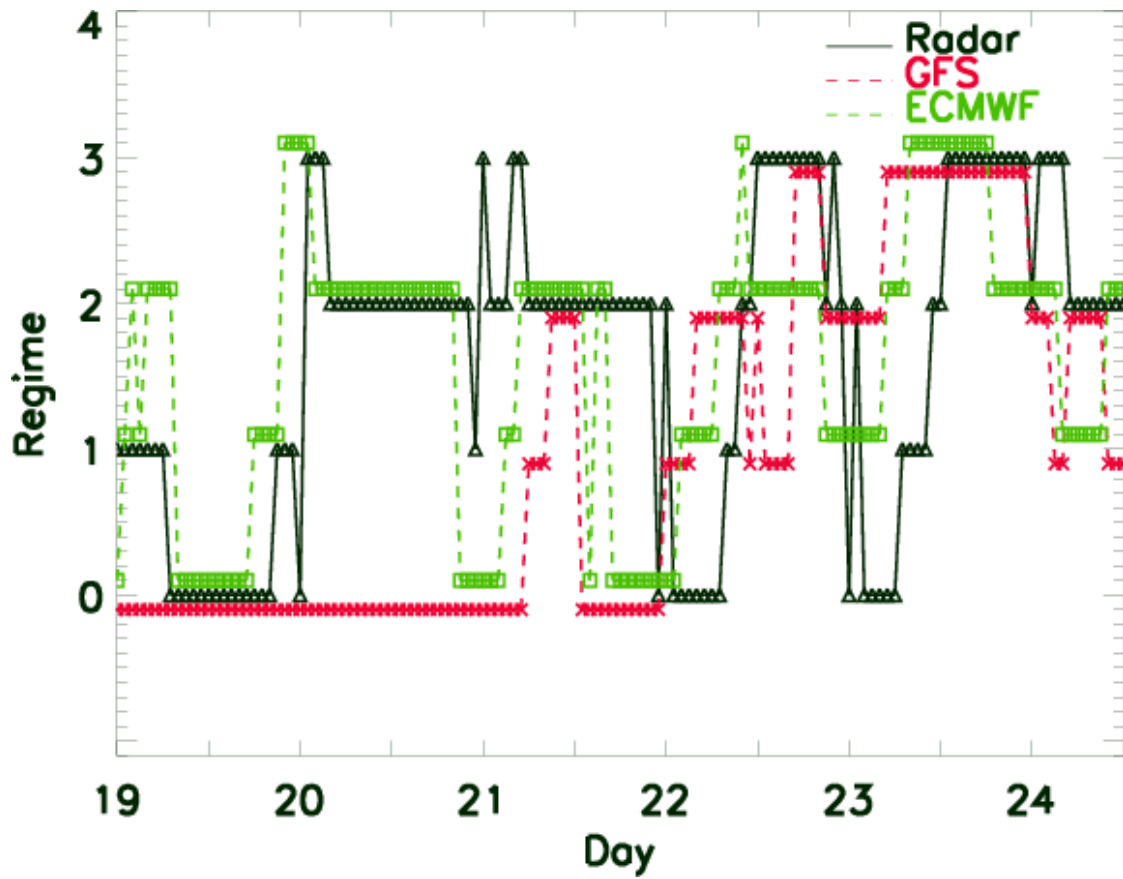
Figure 4: **Time series of radar and model assigned precipitation regimes for the monsoon period. Black line shows the radar assignments. Red line shows the GFS forced WRF assignments and the green line shows the ECMWRF forced WRF assignments**
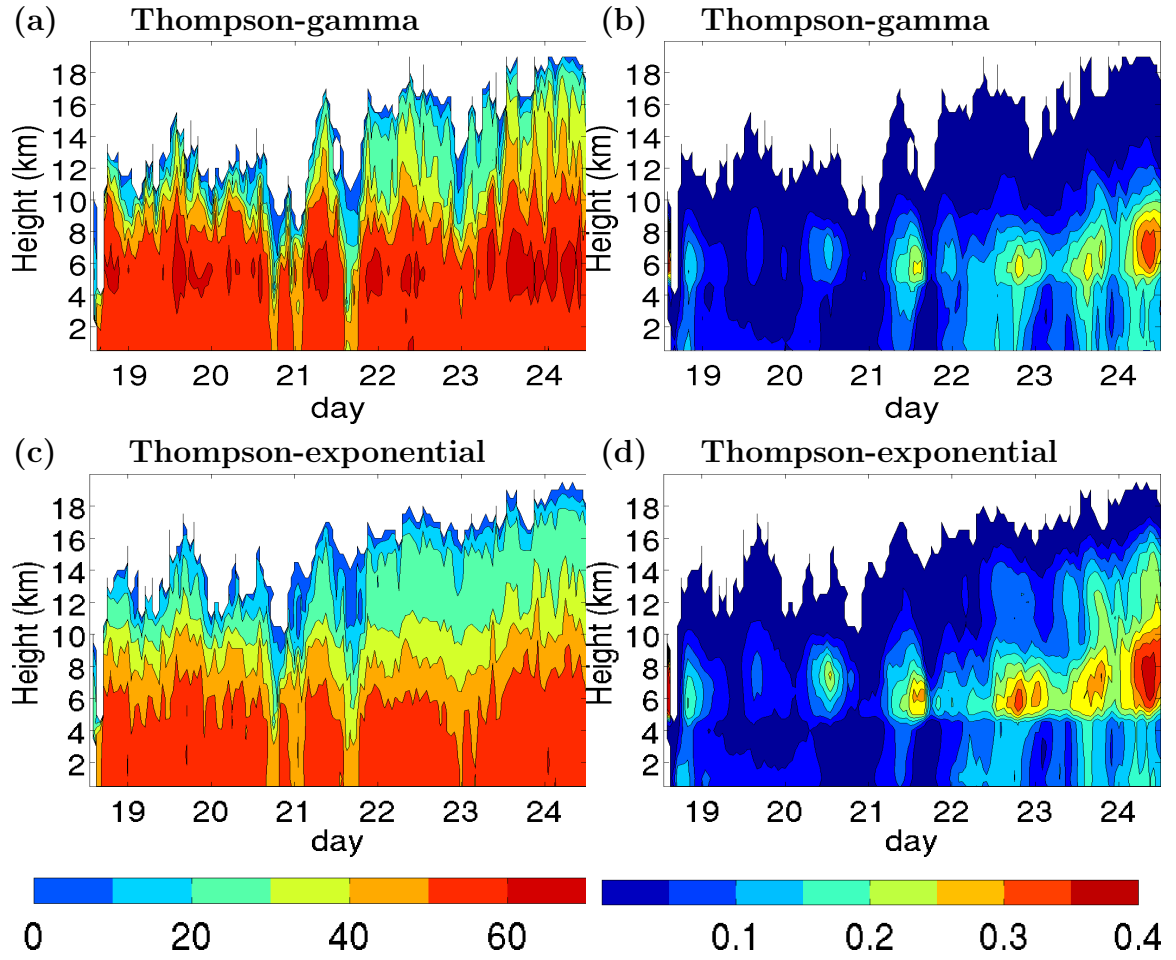
Figure 5: **Maximum Reflectivity and Precipitation Coverage profiles for the Monsoon period (Thompson simulation) using gamma equations for the particle size distributions a) and b), and using exponential equations for the particle size distributions c) and d)**

**(a)** **Time Series**
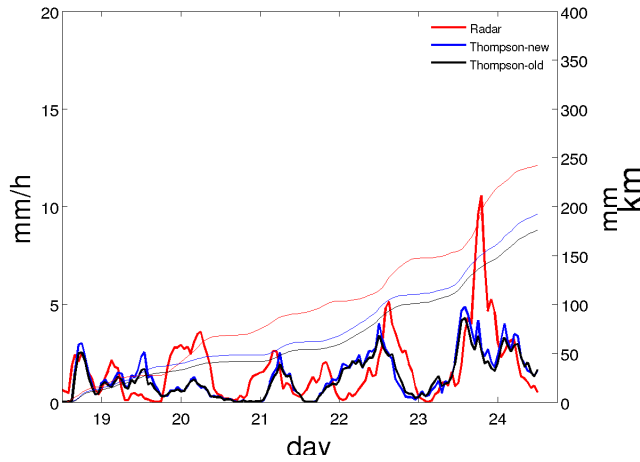
**(b)** **Radar**
**Average Rain Rate =1.6728 mm/h**
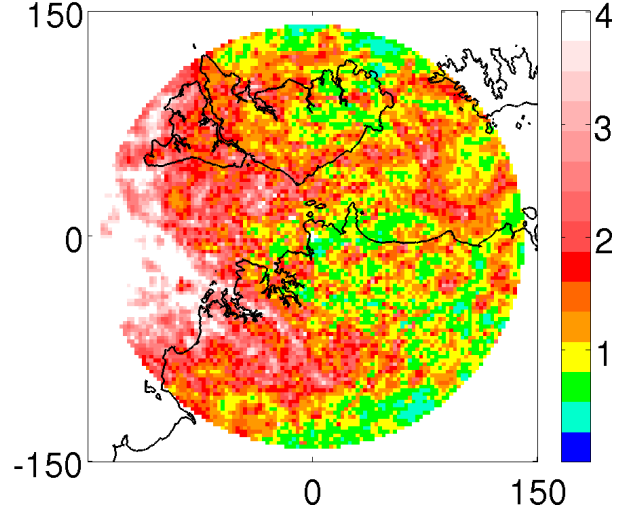
**(c)** **Thompson-gamma**
**Average Rain Rate =1.3237 mm/h**
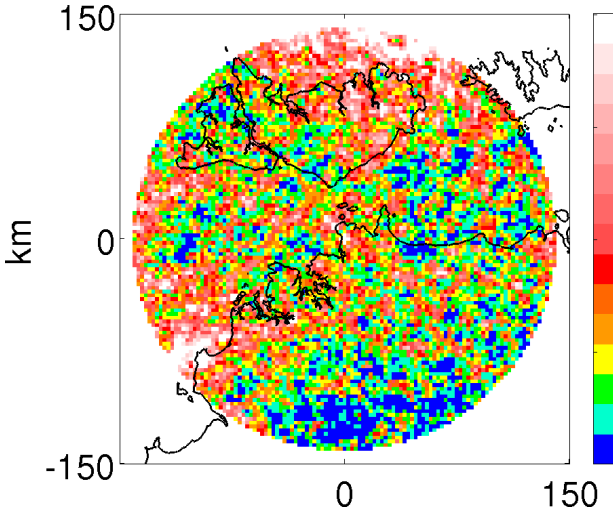
**(d)** **Thompson-exponential**
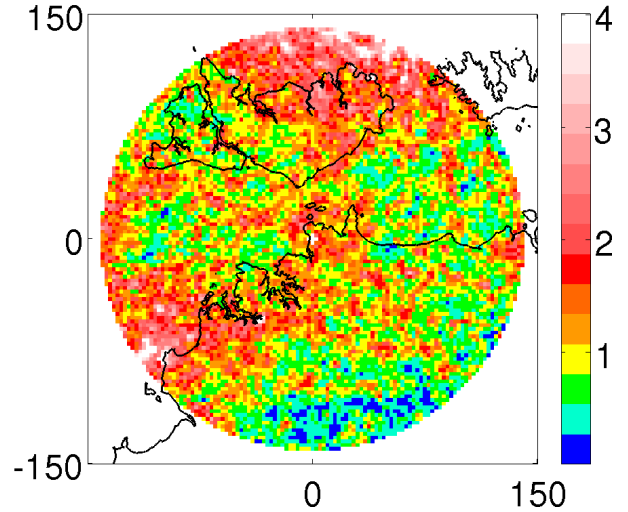**Average Rain Rate =1.2144 mm/h**

Figure 6: **a) Time series of domain-averaged precipitation rates and accumulation over the monsoon period for the radar and Thompson simulations using two versions of the reflectivity conversion algorithm. b) Spatial distribution of precipitation rates averaged over the monsoon period for the radar, c) the Thompson simulation using gamma equations to represent the particle size distributions and d) the Thompson simulation using exponential equations to represent the particle size distributions**
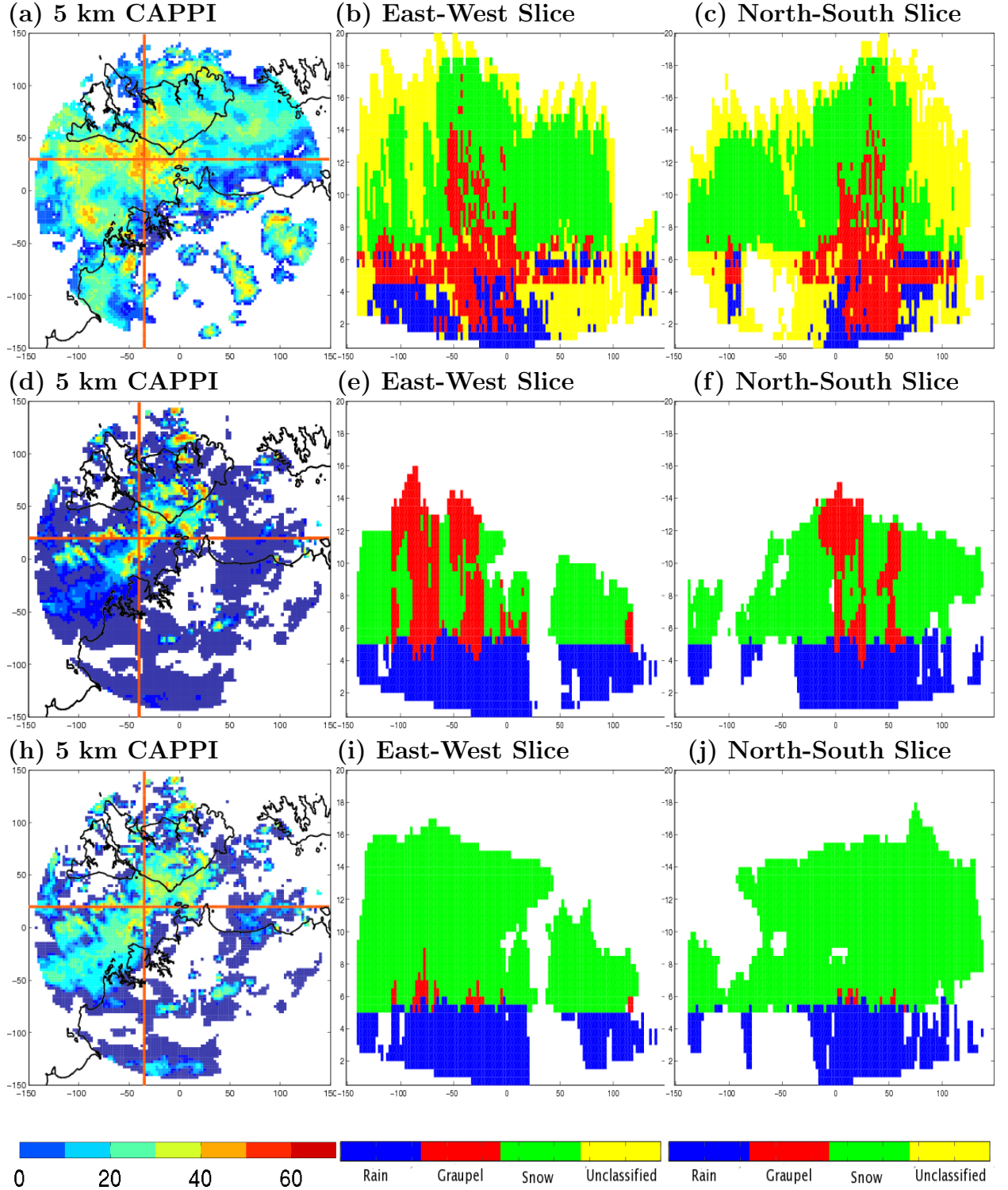
Figure 7: **CAPPI showing the reflectivity values at 5 km for the radar a), the Thompson gamma simulation d) and the Thompson exponential simulation h). East-West and North-South cross sections through the estimated hydrometeor data are shown for the radar b) and c), the Thompson gamma simulation e) and f), and the Thompson exponential simulation i) and j) during the Monsoon Period**
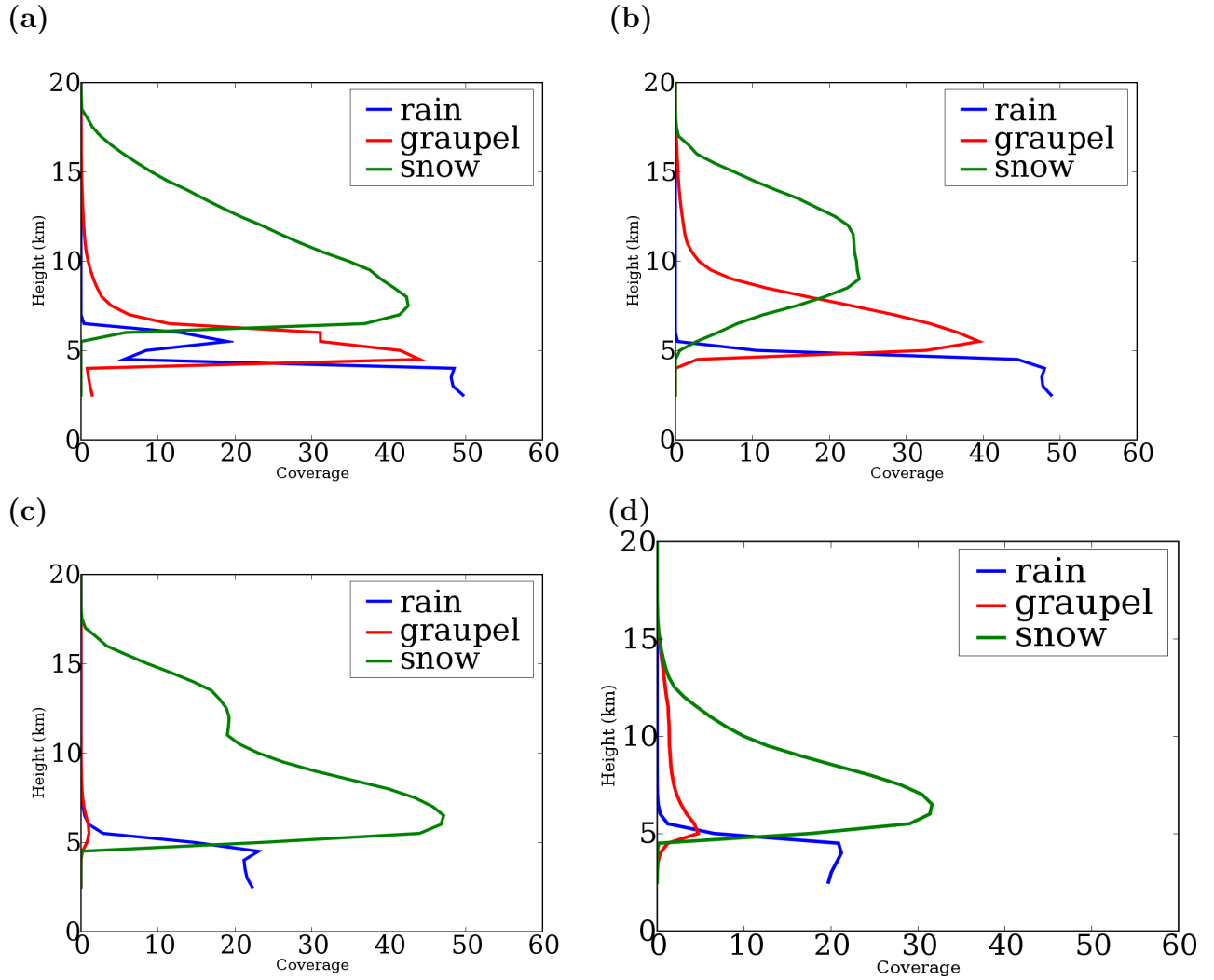
Figure 8: **Hydrometeor profiles averaged over the Monsoon period for a) the radar, b) the Lin simulation, c) the Thompson exponential simulation and d) the Thompson gamma simulation.**
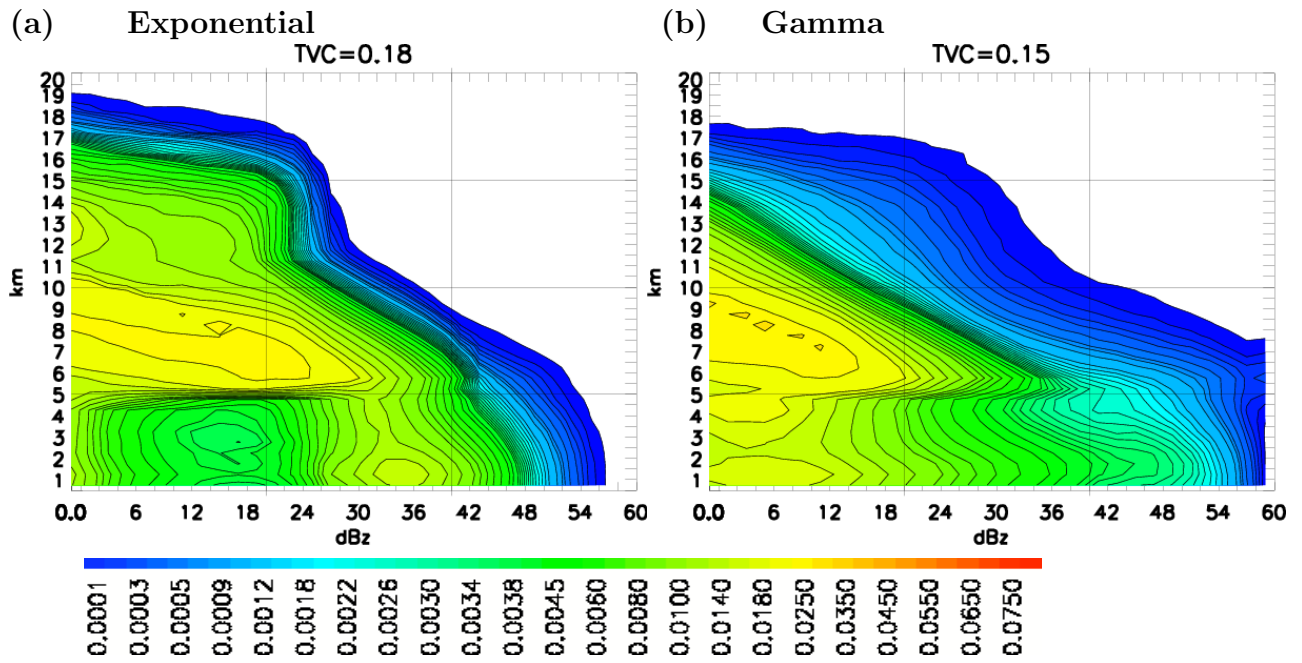
Figure 9: **a) Average histogram for the Thompson simulation over the monsoon period with exponential functions modeling the particle size distributions, b) average histogram for the Thompson simulation over the monsoon period with gamma functions modeling the particle size distributions.**