CONVECTIVE-RESOLVING REGIONAL CLIMATE SIMULATIONS FOR THE AMAZON BASIN: COMPARISON WITH TRMM RAINFALL DATA

A Thesis

by

NICHOLE MARIE KINNEY

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Co-Chairs of Committee,	Craig Epifanio
	Kenneth Bowman
Committee Members,	Ping Chang
Head of Department,	Kenneth Bowman

December 2012

Major Subject: Atmospheric Sciences

ABSTRACT

With increasing computational power, simulations of regional climate are now becoming possible on convective-resolving grids, thus eliminating the need for a convective parameterization. In the present study, a series of seasonal calculations using the Weather Research and Forecasting (WRF) model are computed at 4-km grid spacing, which reasonably resolves most convective systems. Simulations are computed for both the DJF and MAM seasons as averaged over 2005-2008, with a model domain covering the majority of the Amazon Basin and the adjacent South American coastline. Precipitation statistics are computed and compared to satellite rainfall retrieval data from the 13-year Tropical Rainfall Measuring Mission (TRMM) record. For comparison, a set of companion simulations with 12-km grid spacing are also computed, using the Kain-Fritsch convective parameterization.

As compared to the 12-km runs, the 4-km simulations show significant improvement in the overall mean rain rate, the rain rate probability distributions, and the diurnal evolution and timing of precipitation. Both the 4-km and 12-km cases capture the coastal propagating signal and the interior basin-wide diurnal oscillation; however, the 4-km case shows better timing and evolution statistics. Compared to TRMM, the 4-km case rains too infrequently, but is more likely to produce rain events at high rain rates, thus resulting in a similar overall average rain rate.

Overall, the present calculations show significant promise for computing regional rainfall patterns on convective-resolving grids.

ACKNOWLEDGEMENTS

I would like to acknowledge my committee chair and co-chair, Dr. Craig Epifanio and Dr. Ken Bowman, and the other member of my committee, Dr. Ping Chang, for their guidance on this project. Additionally, I would like to thank Justin Joplin for the countless hours of hard work that he put into the TRMM analysis, which he managed to do while finishing up his Bachelor's degree. My personal gratitude also goes to the folks at the Texas A&M Supercomputing Facility for jumping through hoops to meet my rather excessive computation needs, as well as to Dr. Courtney Schumacher for her help in sorting through some of the TRMM data.

I would also like to thank the people at the Tropical Rainfall Measuring Mission and the Goddard Earth Sciences Data and Information Services Center for providing access to the TRMM satellite data.

On a more personal note, I would like to thank my friends and family, and in particular, my significant other, Jonathan, for keeping my sanity somewhat intact during the last two and a half years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

1.1 General Background

The regional climate model (RCM) emerged in the late 1980s as a way to study climate change at regional scales (Wang et al. 2004). In a typical model configuration, a RCM consists of a modified regional prediction model, such as the Weather Research and Forecasting model (WRF), nested inside of a global climate model (GCM). The RCM typically runs on time scales of weeks to months or longer, and with a domain spanning several hundred kilometers to several thousand kilometers. Because of the relatively high computational burden, RCMs have conventionally been run at grid spacings of 10 km or greater, implying that a number of important mesoscale features are not resolved and thus must be parameterized (Leung et al. 2006). One of the important unresolved features is mesoscale convection, which requires a grid resolution of 5 km or less for reasonable representation (Weisman et al. 1997). However, with increases in computing power, RCMs can now be run at convective resolving grid spaces for relatively long periods and large spatial domains, thereby making convective resolving RCM simulations possible.

1.2 Regional Climate Models

As global climate change started to become a major topic among scientists, one area of interest was how this change would impact weather at the regional scale. This interest motivated the development of RCMs in order to address regional climate change

and its consequences (Wang et al. 2004). The first studies using RCMs by Dickinson et al. (1989) and Giorgi and Bates (1989) entailed running a global climate model (GCM) to obtain the synoptic scale pattern to use as the initial and lateral boundary conditions for the Pennsylvania State University (PSU)/National Center for Atmospheric Research (NCAR) regional meteorological model (MM4), one of the first developed RCMs. The goal of this study was to obtain information about the western United States where the complex terrain could not be adequately resolved by the GCM, but where the finer resolution (about 50 km) of the MM4 model could at least partially resolve the mountainous regions (Dickinson et al., 1989, Giorgi and Bates, 1989). The results showed that the RCM nested in the GCM did much better than the GCM by itself in simulating wintertime precipitation patterns in the western United States, showing that RCMs can improve forecasts by better resolving complex features such as mountains and coastlines.

RCMs branched out into a variety of uses as they continued to be developed. As an example, Leung et al. (1996) explored the use of an RCM as a component of an earth system model that simulated hydrology patterns in mountainous terrain. RCMs have also been used to study the impact of small-scale land-atmosphere interactions on regional climate. Schar et al. (1999) and Giorgi et al. (1996) studied the relationship between land moisture and precipitation rates, while Paeth et al. (2009) examined the impacts of land use changes on the northern half of Africa.

Computational resources tend to limit RCMs by forcing tradeoffs between spatial resolution, temporal resolution, size of the domain, and length of the model run. At the

time of the early study of Dickinson et al. (1989), GCM calculations only allowed for 40-50 grid points, amounting to a spatial resolution of about 500 km. The nested RCM had a grid resolution of 60 km and many more points than a GCM, but the RCM ran for a total of only 20 days, considerably less than the GCM. Present day RCMs can run at resolutions of 10-20 km, with run times on the order of weeks or longer. However, in summarizing a workshop on RCMs, Leung et al. (2006) pointed out that RCMs still lack the resolution to explicitly resolve mesoscale events such as convection or turbulence and instead must use parameterizations for these features, resulting in model error. With increasing computational power, one solution entails nesting a smaller domain in the RCM with a resolution fine enough to resolve some mesoscale features such as convection; this requires a resolution of 4-5 km or less (Weisman et al., 1997).

1.3 Convective Resolving Models

Cloud modeling began as one-dimensional models designed to simulate a single cumulus cloud, but the overly simplified physics resulted in high model error except in the most idealized cases (Warner, 1970). Improved technology allowed for the first attempts at three-dimensional cloud models (CMs) in the mid-1970s (Moncrief and Miller, 1976). Modelers initially wanted to study the detailed structure of individual convective clouds, as well as their response to environmental conditions. Klemp and Wilhelmson (1978) developed one of the first three-dimensional CMs, and by introducing a simple shear profile, created a convective storm with split vortices. It should be noted that this model first introduced the idea of time-splitting, in which the

model treats the sound waves separately from the rest of the calculations. Other researchers developed cloud models based on anelastic systems, such as Moncrief and Miller (1976), Clark and Peltier (1977), and Clark (1979). It is worth mentioning that the technology at the time limited their domain size to a few tens of kilometers and their runtimes to a few hours at the spatial and temporal resolutions required to simulate a single convective cell.

In the decade after the introduction of the three-dimensional CM, CM experiments began to focus on mesoscale convective systems such as squall lines. Weisman et al. (1988) and Rotunno et al. (1988) used an evolved version of the Klemp-Wilhelmson model (Klemp and Wilhelmson, 1978) to run idealized simulations of a squall line in which they varied the wind shear environment in the model domain. The squall line showed strong sensitivity to the component of the wind shear perpendicular to the convective system, and the model produced structures ranging from weak cells along the outflow boundary to strong supercells in response to different profiles of that shear component. In a real case application of CMs, Dudhia and Moncrieff (1989) initialized a CM with conditions from the Oklahoma-Kansas Preliminary Regional Experiment for STORM-Central (PRE-STORM) on May 7th, 1985. The model produced a squall typical of late spring in the Great Plains region, including strong right flank supercells. Both of the previously mentioned studies featured grid resolutions of 2 km, coarser than the studies from the late 1970s and not ideal for resolving small scale cloud features, but necessary in order to use a domain large enough to contain a squall line.

Many of the forecast models commonly used today, such as the MM5, the Weather Research and Forecast model (WRF), the Regional Atmospheric Modeling System (RAMS), and the Advanced Regional Prediction System (ARPS), have the ability to explicitly simulate convective processes with a grid spacing of about 4-5 km or less. Modelers typically run a setup consisting of a small convective resolving domain, spanning a few hundred kilometers or less, nested inside a larger domain that utilizes a convective parameterization. Because running this nested grid setup requires considerable computational resources, the runtimes have conventionally not exceed a few days. However, recent improvements in computer power have allowed modelers to expand the convective resolving domain to spatial and temporal scales resembling a RCM. This raises the possibility of regional climate calculations at convective resolving grid scales, thus, in principle, improving the accuracy of computations.

1.4 The Amazon Basin

The Amazon Basin averages roughly 2000 mm of rain per year, with the majority of the rainfall occurring during the December-January-February (DJF) summer season and the March-April-May (MAM) transition season (Salati and Vose, 1984; Villar et al, 2009). In an analysis of the various rain mechanisms in South America, Romatschke and Houze (2010) identify two of the leading precipitation modes over the AB as the diurnal solar heating cycle and the large squall lines that propagate into the region from the northeast coast of South America.

In order to better understand the precipitation patterns over the AB, the Amazon Boundary Layer Experiment (ABLE 2B) was conducted during April and May of 1987. Based on the data from ABLE 2B, Greco et al. (1990) found that the large, propagating squall lines, referred to as "Amazon coastal squall lines" (ACSL), contributed to roughly 40% of the total rainfall during the field campaign. This study also estimated that these ACSL systems average about 1000-2000 km in length, propagate at about 50 km/hr, and last for 24-48 hours. In the first part of the ABLE 2B data analysis, Garstang et al. (1994) examined the dynamical and life cycle compositions of ACSLs. They found that the squall lines can have as many as six possible stages in their lifetime, including a weakening stage followed by a reintensification stage, before the system dissipates, and the study suggested that significant low-level shear in the inflow is responsible for the longevity of the system. In the second part of the ABLE 2B data anaylysis, Greco et al. (1994) determined that within the ACSLs, the trailing stratiform region accounts for more than 50% of the vertical heat transport, with the leading edge convection accounting for the rest. Alcantara et al. (2011) analyzed satellite data to form a climatology of ACSLs between the years 2000 to 2008 and found that these systems reach a peak in frequency between April and June.

Romatschke and Houze (2010) found that the diurnal contribution to the AB precipitation begins in the early afternoon as thunderstorms that form over the eastern portion of the AB and near the Brazilian Highlands due to solar heating, topography, and an ample moisture supply. The convection then spreads to the northwest throughout the afternoon due to solar heating and a possible contribution from the outflow boundary of

an ACSL. The convection reaches a peak in the late afternoon while a broad stratiform region develops and reaches a peak overnight at about 3 A.M. local time (Romatschke and Houze, 2010).

1.5 Project Overview

The present study assessed the potential for convective resolving RCM calculations using the WRF model, with the Amazon Basin (AB) used as a test case and satellite data from the Tropical Rainfall Measuring Mission (TRMM) used as the basis for the comparison. A convective resolving domain roughly the size of the AB is nested inside of a larger, lower resolution domain that uses a parameterized convection scheme. The simulation runs over two 12-week periods (the December-January-February and the March-April-May seasons) for several different years in order to gather climatological precipitation model data for the region. The model is then run again over the same time frames, but with only the lower resolution domain. The main goal of this study is to compare the performance of the convective resolving simulation to that of the lower resolution domain by statistically analyzing the spatial and temporal properties of the rain rates produced by the two model setups. In particular, this study examines the abilities of the model configurations to capture the properties of both the rain mechanisms common to the AB and the seasonal variation between the two seasons for which the models are run for.

Section 2 discusses the model configuration, the details of the TRMM dataset, and the calculation for the three-dimensional rain rate. The results are then laid out in the

next three sections: Section 3 discusses the results of the seasonal surface mean rain rates for both model cases and the TRMM data, Section 4 goes into the details of the results of the diurnal cycle of the seasonal surface mean rain rates for all three products, and Section 5 examines the results of the vertical rain rate distribution comparison between the convective resolving domain and the TRMM data. Lastly, Section 6 summarizes the results and discusses future work for this project.

2. DATA AND METHODS

2.1 Model Configuration

This project uses 3.3.1 of the WRF-ARW model to explore the performance of a convective resolving model calculation when applied at regional climate scales (Skamarock et al. 2008). The two-way nested domain setup consists of a convective resolving domain with a resolution of 4 km nested inside of a domain with 12 km spacing (Figure 1) over the northern portion of South America. Because the ACSL systems are of particular interest, the inner domain was chosen so that, in addition to the AB, it includes the north-eastern coast of the continent where the large squall lines originate. The domains are rotated in order to vastly decrease the number of points that the calculations take place at, thereby significantly reducing the runtime of each model simulation. The 4 km grid dimensions are 696 by 639 mass grid points and the 12 km grid is 413 by 409 mass grid points.

The model runs 12-week simulations twice a year from 2005 to 2008 for a total of eight simulations. The 12 weeks runs span the MAM and DJF seasons in a given year in order to cover the wettest months in the AB. This set of simulations are run twice; once with the 4 km grid on and once with it off, which will be referred to as the high resolution case (HRC) and the low resolution case (LRC), respectively. Due to the computational strain of the HRC, the temporal resolution is two hours, while the LRC outputs data every hour.

Table 1 describes the basic model parameters used in this study. Due to the computational expense of the HRC, the experiment is run using only one model configuration and most of the parameters chosen are ones commonly used in modeling studies. The WRF Single-Moment 6-Class (WSM6) microphysics is chosen over its double-moment counterpart in order to significantly reduce the runtime of the model runs because test runs showed little difference in the rain rates produced by the two schemes.

It was discovered that several unphysical streaks appear in the geographic distributions of the mean rain rates for the LRC output. The left panel of Figure 2 shows the streaks in a sample one week average of the model output for the LRC model setup given in Table 1. The right panel of Figure 2 shows the same sample average, but with the Betts-Miller cumulus parameterization rather than the Kain-Fritsch parameterization. The lack of streaks in the right panel suggests that the Kain-Frisch parameterization is to blame for the unphysical marks that appear in the LRC model setup described in Table 1. It is important to note that while these artifacts are unsightly, they are unlikely to have contaminated the results, so the LRC results will still be used in the analysis, but work will be done in the future to attempt to correct the issue.

2.2 The TRMM Dataset

The Tropical Rainfall Measuring Mission (TRMM) satellite was proposed by Simpson et al. (1988) and launched in November of 1997 (Bowman, 2005). Its goal is to provide thermodynamic information over the tropics, primarily cloud and precipitation

information, where observations are generally limited (Simpson et al., 1988, Bowman, 2005). The instruments used to measure precipitation are the TRMM Microwave Imager (TMI) and the Precipitation Radar (PR), both of which are described by Kummerow et al. (1998). This study makes use of the TRMM products 3B42, 2A25, and 2A23 binned up to a 0.50° by 0.50° grid.

The 3B42 dataset, formally known as the TRMM Multisatellite Precipitation Analysis (TMPA), combines various satellite data with rain gauge data to produce a 3 hourly monthly mean surface rain dataset on a 0.25° by 0.25° grid (Huffman et al. 2007). TMPA relies on two different types of measurements; a set of combined passive microwave data from various low earth orbit satellites, and a set of infrared data from the international geosynchronous earth orbit (GEO) satellites. The products included in the former are the TMI, the Special Sensor Microwave Imager (SSM/I) on Defense Meteorological Satellite Program (DMSP) satellites, Advanced Micro-wave Scanning Radiometer-Earth Observing System (AMSR-E) on the satellite Aqua, and the Advanced Microwave Sounding Unit-B (AMSU-B) on satellites from the National Oceanic and Atmospheric Administration (NOAA) (Huffman et al., 2007). The monthly TMPA is produced by combining these products to produce a rain rate, which is then calibrated using rain gauge data where it is available. (Huffman et al., 2007).

The TRMM 2A25 and 2A23 products derive from satellite data from the TRMM PR instrument (Kozu et al., 2000). A three-dimensional effective reflectivity factor is estimated from the PR data, from which cloud and precipitation information is then calculated. The 2A25 product is the derived three-dimensional rain rate orbital dataset

while the 2A23 product contains orbital properties such as rain type, echo heights, and freezing levels (Kuzo et al., 2000). The 2A25 product uses the information in the 2A23 dataset in its algorithm, although this study also uses the echo heights from the 2A23 data to compare with the model results.

In addition to the TMPA data, the 2A25 can also provide surface rain rates. However, because the 2A25 rain rates are not rain gauge adjusted after processing, it contains a bias relative to the TMPA data. Over the 4 km domain area, the 2A25 surface rain rates are roughly 26% lower than those provided by the TMPA data. The 2A25 data also tends to miss both very low and very high rain rates. The missed low rain rates occur because the PR has a much longer range than a ground radar would, which necessitates a higher noise cutoff; this cutoff for the PR is roughly 17-20 dBZ (Bolen and Chandrasekar, 2000). This cutoff also results in the 2A23 dataset reporting echo top heights that are too low (Schumacher and Houze, 2000). The PR sometimes misses very high rain rates because the attenuation correction algorithm in the PR data can underestimate the PR attenuation in convective rain (Liao and Meneghini, 2009).

2.3 The Three-Dimensional Rain Rate Calculation

Because the HRC model setup receives its precipitation information entirely from the microphysics parameterization while the LRC receives an additional contribution from the cumulus parameterization, the following three-dimensional rain rate calculation is only calculated for the HRC model results. The complexity of the rain

contributions in the LRC model setup prevents an equivalent three-dimensional rain rate calculation for that case.

The WSM6 microphysics scheme originates from the bulk microphysics cloud model described by Lin et al. (1983). The values of any constants are given in Appendix C. The size distributions of each precipitating particle are described by the following exponential equations:

$$n_R(D) = n_{0R} * \exp(-\lambda_R D_R), \quad (1)$$

$$n_S(D) = n_{0S} * \exp(-\lambda_S D_S), \quad (2)$$

$$n_G(D) = n_{0G} * \exp(-\lambda_G D_G), \quad (3)$$

where $n_R(D), n_S(D),$ and $n_G(D)$ denote the distributions for rain, snow, and graupel, respectively. The terms $n_{0R}, n_{0S},$ and n_{0G} are the distribution intercepts, the distribution slopes are given by $\lambda_R, \lambda_S,$ and λ_G, and the particle diameters are $D_R, D_S,$ and D_G. The distribution slopes, which can be found by multiplying (1), (2), and (3) by the particle mass, integrating over all diameters, and setting the result equal to the appropriate water contents, are given by the following equations:

$$\lambda_R = \left(\frac{\pi \rho_R n_{0R}}{\rho q_r}\right)^{\frac{1}{4}}, \quad (4)$$

$$\lambda_S = \left(\frac{\pi \rho_S n_{0S}}{\rho q_S}\right)^{\frac{1}{4}}, \quad (5)$$

$$\lambda_G = \left(\frac{\pi \rho_G n_{0G}}{\rho q_G}\right)^{\frac{1}{4}}, \quad (6)$$

where $\rho_R, \rho_S,$ and ρ_G are the densities of liquid water, snow, and ice. The terms $q_r, q_S,$ and q_G are the mixing ratios for the three precipitation types.

Following the work of Locatelli and Hobbs (1974), Lin et al. (1983), and Hong and Lim (2006), the terminal velocities for rain, snow, and graupel at a single diameter are given by:

$$V_R(D_R) = a_R D_R^{b_R} \left(\frac{\rho_0}{\rho}\right)^{\frac{1}{2}}, \tag{7}$$

$$V_S(D_S) = a_S D_S^{b_S} \left(\frac{\rho_0}{\rho}\right)^{\frac{1}{2}}, \tag{8}$$

$$V_G(D_G) = a_G D_G^{b_G} \left(\frac{\rho_0}{\rho}\right)^{\frac{1}{2}}. \tag{9}$$

Integrating (7), (8), and (9) gives the mass-weighted mean terminal velocity for a given species. The equations for the mean terminal velocity are as follows:

$$\bar{V}_R = \left(\frac{a_R \ (4+b_R)}{6\lambda_R^{b_R}}\right)\left(\frac{\rho_0}{\rho}\right)^{\frac{1}{2}}, \tag{10}$$

$$\bar{V}_S = \left(\frac{a_S \ (4+b_S)}{6\lambda_S^{b_S}}\right)\left(\frac{\rho_0}{\rho}\right)^{\frac{1}{2}}, \tag{11}$$

$$\bar{V}_G = \left(\frac{a_G \ (4+b_G)}{6\lambda_G^{b_G}}\right)\left(\frac{\rho_0}{\rho}\right)^{\frac{1}{2}}. \tag{12}$$

The WSM6 microphysics scheme in the WRF-ARW model uses this bulk microphysics scheme with two modifications. The first change is one of several changes to the ice microphysics detailed in Hong et al. (2004). Based on work from Houze et al. (1979) and Ryan (1996), the size distribution intercept parameter, n_{0S}, is strongly dependent on temperature rather than a constant. The formula for n_{0S} is

$$n_{0S}(T) = (2.0e6) * \exp(0.12(T_0 - T)). \tag{13}$$

The second modification, detailed by Dudhia et al. (2008), changes the scheme so that when there exists a mix of snow and graupel in a given grid box, the particles are treated as lightly rimed particles with regard to the mean terminal velocity. Rather than treating each species as a separate entity, a mass weighted mean terminal velocity for both species can instead be calculated by

$$\bar{V}_{GS} = \left(\frac{q_S\bar{V}_S + q_G\bar{V}_G}{q_S + q_G}\right). \quad (14)$$

Based on the WSM6 microphysics scheme, the liquid water rain rate at a given model level can be calculated as

$$R_l(mm/s) = (\bar{V}_R q_R + \bar{V}_{GS} q_S + \bar{V}_{GS} q_G) * \left(\frac{\rho}{\rho_R}\right) * (1000.0). \quad (15)$$

Because the model levels in the WRF-ARW model are on pressure coordinates whose heights change with time, the rain rate is linearly interpolated in the vertical onto the background geopotential grid where heights remains constant with time. This allows for easy time averaging of the three-dimensional rain rate because the rain rate at a given (i,j,k) point will always be at the same height regardless of time.

2.4 Averages

The average surface rain rates used in this study derive from the HRC and LRC surface rain model results, the TMPA surface rain data, and the TRMM 2A25 surface rain orbital data. The HRC surface rain data comes from the bottom level of the calculated three-dimensional rain rates while the LRC surface rain data is the sum of the surface rain rate contributions from both the microphysics parameterization scheme and

the cumulus parameterization scheme. Seasonal averaging is done over all times in the MAM season, the DJF season, and the six month period across both seasons (DJF-MAM) for every point in the domain. The seasonal geographical surface mean rain rate for each time period is then obtained by averaging up to a 0.50° by 0.50° grid to make it comparable to the TRMM data sets, which are gridded at that resolution. A seasonal domain average rain rate for a particular season is also found by averaging the geographical surface means across all the points in the domain. The diurnal cycles of the previously mentioned averages is found by averaging over each time period rather than all of the time periods at once. For the HRC, the time periods are instantaneous at every two hours, and for the LRC, the time periods are instantaneous and hourly. Both of the model times begin at 0 Z. The TRMM data time periods are every three hours starting at 0 Z, but the rain rates at these times are 10 minute averages rather than instantaneous.

2.5 Histograms

The conditional surface rain histograms used in this study derive from the HRC and LRC model data along with the TRMM 2A25 orbital data. The surface rain rates in all three products are binned in 1 mm/day bins from 0.0 mm/day out to a point where the histogram contains well over 99% of the data. The points in the histograms are plotted at the center of the bins. For the model results, the last bin is centered over 3999.5 mm/day, while for the 2A25 data, the last bin is centered at 1499.5 mm/day. For a fair comparison, the LRC and TRMM histograms are constructed using only the points within the 4 km grid.

The conditional rain rate PDF described by the equation

$$p(r) = \left(\frac{(\sigma(r))}{N}\right) * 100.0, \quad (16)$$

where $p(r)$ is the rain rate PDF as a function of the rain rate r at the center of a bin, $\sigma(r) * \Delta r$ is the number of events in a bin of width Δr, and N is the total number of raining events in the histogram. From the rain rate PDF, a distribution of the net rain is found. The net rain distribution is described by the equation

$$\mu(r) = \left(\frac{(r*\sigma(r))}{\mathcal{R}}\right) * 100.0, \quad (17)$$

where

$$\mathcal{R} = \sum r * \sigma(r) * \Delta r. \quad (18)$$

Using the 2A23 data, an echo top histogram is also calculated at each point by binning all of the echo tops into 1 km bins, and a domain histogram is then found by, for each bin, summing across a given bin at all of the geographic points that lay within the 4 km grid. An equivalent dataset for the HRC rain rates is calculated by locating the highest nonzero rain rate in a column at a given (i,j) point, and then binning the height level of that rain rate into 1 km bins. The domain histogram for the HRC is found in the same manner in which it is for the 2A23 data.

3. SEASONAL MEAN SURFACE RAINFALL

3.1 Seasonal Mean Rain Distributions

Table 2 shows the domain averages for the HRC and the LRC model results, as well as the TMPA dataset. The table also shows the model error for each case relative to the TRMM data. All three products show a higher average rain rate for the MAM season than the DJF season, although this difference is larger for the model mean precipitation rates than for the TMPA mean rate. The HRC domain averages compare favorably to the TMPA data, with a six month error of just under 15% and a MAM error of only around 9.5%. Interestingly, the HRC underestimated the rain rates relative to TRMM in all of the seasons. In contrast to the HRC, the LRC did rather poorly in estimating the domain average rain rates in all three time periods, with a six month error of over 80%. Unlike the HRC, the LRC did better in DJF than MAM by around 13% and it overestimated all of the domain averages relative to the TRMM.

Figure 3 shows the geographic distribution of the surface mean rate rates for each product and for all three time periods. Again, the HRC means compare fairly well with the TMPA data overall, although the HRC performs better with some time periods and geographical areas than others. The HRC does particularly well with the geographic distribution of the mean rain rates in MAM, although it produces too much rain over the mouth of the Amazon River and too little rain over the interior of the AB. Additionally, the HRC places the area of maximum rain rate a little farther north than the TMPA data does. The HRC is somewhat less successful in DJF than in MAM. The TMPA mean rain

rate distribution shows more widespread rain and higher precipitation rates in the interior of the AB than the HRC, while the HRC places an area of maximum precipitation along the coast near the northern edge of the domain that is not seen in the TMPA mean distribution.

Compared to the HRC, the LRC mean rain rate geographical distribution produced significantly higher rain rates than the TMPA rain data across almost all of the 4 km domain area. Despite the amplified rain rates, the LRC performs well with the geographic distribution of the rain across the domain in that it gets the general shape of the distribution correct. In particular, the mean rain rate geographical distribution for the LRC line up a little better with the TMPA data than the HRC, which is most noticeable in the DJF season.

3.2 Seasonal Surface Rain Histograms

Figure 4a shows an immediate discrepancy between the model rain rates and the TRMM rain rates; the models show a large number of very small rain rates while the TRMM shows very few values below 8 to 10 mm/day. The difference is likely due to a combination of the high noise threshold in the PR algorithm and the models overestimating the frequency of light rain rates. It would be ideal to know exactly what percentage of that discrepancy can be attributed to the TRMM PR instrument and what percentage can be attributed to the models so as to better compare the histograms in Figure 4a; however, that is not possible. For the sake of simplicity and to produce the best possible comparison between the histograms, the assumption is made that the

TRMM PR instrument missed all of the events below 10 mm/day, and in particular, that the 2A25 data has the same percentage of its events below 10 mm/day as the HRC. Following that assumption, an adjusted TRMM rain rate PDF is included in Figure 4a, in which the normalization factor N from the original TRMM histogram is algebraically scaled so that the integrated frequency of the first 10 bins are equal for both the TRMM data and the HRC. It should be stressed that these assumptions and the subsequent adjustment are done strictly for the sake of the comparison and are not meant to suggest that frequencies reported by the models below 10 mm/day are all real.

The HRC and the adjusted TRMM rain rate PDFs show a good match past the point from about 10 mm/day out to about 300 mm/day, indicating that both products have a similar probability of producing rain rates in that range (Figure 4a). However, with higher rain rates than 300 mm/day, the HRC is increasingly more likely to rain at those rates than the TRMM. In comparison to the HRC, the LRC rain rate PDF shows a relatively poor match to the TRMM data beyond 10 mm/day. The LRC rain rate PDF an unusual feature in which there exists an equal probability of the LRC producing rain rates from about 10 mm/day out to about 100 mm/day before the distribution falls off sharply. The value of the PDF over that range of bins is also much higher in the LRC than either the HRC or the TRMM data.

Figure 4b shows the net rain distributions, which can be thought of as the relative importance of each bin to the domain average rain rate. The very low rain rate bins where the models and the TRMM disagree only contribute to about 1.6% of the domain average rain rate for either model setup, making those bins relatively unimportant.

Figure 4b also shows that the HRC and the TRMM do not agree nearly as much as they have prior to this point. While these two products show good agreement with regards to the domain average rain rate, the HRC sees less of a contribution 10 mm/day to about 400 mm/day and more of a contribution from 400 mm/day onward than the TRMM data. However, it was noted in Section 2.2 that the 2A25 data has a 26% negative bias relative to the TMPA dataset, which is likely more representative of the true rain rates in the AB. This bias impacts the TRMM distribution in Figure 4b because rain rate appears in the equation for $\mathcal{R}$, so an adjusted TRMM distribution is calculated by dividing $\mathcal{R}$ for the original TRMM distribution by (1.0-0.26) to account for the bias in the two TRMM datasets. The bias adjustment is also shown on Figure 4b, and while the adjusted curve agrees with the HRC curve better than the original TRMM curve, all of the previous observations about the comparison of the original TRMM curve and the HRC curve still stand. Despite the differences between the HRC and the TRMM in Figure 4b, the HRC agreed with the TRMM data much more than the LRC. The LRC curve shows that the cumulus parameterization produces far too much rain between 10 mm/day and 300 mm/day.

Table 3 shows integrated ranges of the rain rate PDF (Figure 4a) and the net rain contribution (Figure 4b) for each product. Low is the integration over the first ten bins, medium covers from 10 mm/day to 400 mm/day, and high ranges from 400 mm/day onward. The integrations for the hypothetical TRMM curves are also included in parenthesis. For the purposes for the integration of the adjusted TRMM curve shown in

Figure 4b, it is assumed that the bias between the 2A25 and the TMPA datasets is contained in the highest bins.

In all of the products, the low range only contributed a percent or two of the domain average rain rate, highlighting its relative unimportance to the average rain rate (Table 3). The vast majority of the rain rate frequencies for both the LRC and the adjusted TRMM curves are contained in the medium range; this range also constitutes the largest contribution to the domain average rain rate for both of these products. In contrast, the HRC only has about 5% of its rain events in the high range, but those events account for over 57% of the average rain rate. The large contribution of medium rain rates and low contribution of high rain rates to the TRMM domain average rain rate likely reflects the tendency of the PR attenuation algorithm to miss some high rain rates. However, even after adjusting for the bias in the 2A25 and TMPA datasets, the HRC receives about 20% more of its rain from the highest rain rates than the TRMM data does, indicating that the model is producing too much rain in that range.

Table 4 shows the percentage of total events (rain and non-rain) that consist of rain events. An adjustment to the percent raining for both of the models setups is included to account for the difference between the models and the 2A25 data in the lower bins, which treats all of the events for 8 mm/day or less in the models as non-raining events. The adjustment cuts the frequency of raining events for the HRC by a little over 50%, while the frequency of raining events for the LRC is reduced by about a third. The HRC, the adjusted HRC, and the TRMM dataset all rain for roughly the same amount, although the adjustment changes the HRC from raining slightly more than the

TRMM data to raining slightly less. While the HRC and the TRMM rain for about the same amount of time, Table 3 shows that the domain average rain rate in the HRC receives a large contribution from a few deep convective events that produce copious amounts of rain per event, while the domain average rain rate for the TRMM dataset usually consists of events with moderate rain rates. The domain average rain rate for the LRC is roughly twice as high as the other two products (Table 2), which can be attributed to the fact that the LRC model setup rains significantly more than either of the other two products (Table 4) while producing a high frequency of moderate to high rain rates (Table 3) that contribute to the vast majority of the domain average rain rate (Table 3).

4. SEASONAL DIURNAL CYCLE

4.1 Mean Diurnal Evolution

Figure 5 shows the diurnal cycle for the MAM season for all three products, with frames displayed every six hours. This figure contains a prominent signature from the ACSL in the form of a concentrated line of larger rain rate values that form along the coast and propagate inward. All three products show similar timing for the evolution of the ACSL, although the LRC shows much larger rain rates for the system while propagating it farther than either the HRC or the TMPA data. While the HRC rain rates for the ACSL show better agreement with the TRMM rain rates than the LRC, the HRC also shows too much rain for the propagating systems while keeping them intact for too long. In addition to the ACSL, Figure 5 also exhibits a diurnal signal over the interior of the AB that features an increase in convective activity in the afternoon before the activity dies down the next morning. The LRC greatly overestimates the magnitude of the interior diurnal cycle relative to the TRMM while the HRC appears to slightly under-represent the cycle.

Figure 6 shows the diurnal cycle for the DJF season. The coastal propagating signal, while still existent, is considerable weaker than in the MAM season, while the interior diurnal signal features more prominently than in MAM. The convection in the interior of the AB has also shifted southward from its location in Figure 5. As with the MAM season, all three products show a similar time evolution of both the interior diurnal cycle and the coastal propagation. Once again, the magnitude of the rain rates in

the LRC is considerably higher across the entire domain at all times compared to the rain rates in the other two products. The LRC also produces coastal signals that are too strong and long-lived. The HRC underestimates the rain rates for the convection in the interior of the AB at 0 Z and 6 Z (Figure 6a, 6b, 6d, and 6e). By 12 Z, the HRC switches from underestimating the interior mean to overestimating it (Figure 6g and 6h). The somewhat poor representation of the rain in the interior of the AB indicates that the HRC may not be producing enough of the nocturnal stratiform rain found in the TRMM data and mentioned by Romatschke and Houze (2010), but then keeps the stratiform rain that it did develop around for too long.

4.2 Diurnal Extremes

Figure 7 shows the diurnal cycle of the domain average rain rate, which are simply the domain averages of the diurnal geographical mean rain rate distributions from Section 4.1. The first thing that stands out is the similarity of the curves between both seasons (and by extension, the six month mean), which is surprising given the obvious differences between the geographical mean rain rate distributions for the MAM and DJF seasons. Figure 7 also highlights differences in the timing of the maximum and minimum rain rate between all three products. Both model cases hit a local minimum in the rain rate sooner than the TMPA data. The HRC minimum is only an hour earlier than the TMPA data, but the LRC minimum rain rate is four hours earlier than the TMPA. The model performance is worse for the timing of the maximum rain rate; the HRC

maximum occurs two to three hours too early relative to the TMPA maximum, while the LRC maximum is a full six hours too early.

It is also worth noting that the range between the maximum and minimum rain rates is much larger in the LRC than in either the TMPA dataset or the HRC; that range in almost twice as large in the LRC data as in the TMPA dataset and almost four times larger than the HRC (Figure 7). Figures 8 and 9 show the geographic distribution of the mean rain rate at the time of the maximum domain average rain rate for each product (Figure 8) and at the time of the minimum domain average rain rate for each product (Figure 9). Regardless of the how the HRC represents the ACSL, in both seasons it underestimates the rain rates in the interior of the AB at the time of the maximum domain average rain rate and overestimates them at the time of the minimum domain average rain rate; this leads to the possibility that the relatively small range between the maximum and minimum domain average rain rates for the HRC could be primarily due to its somewhat poor representation of the rain rates in the interior of the AB.

A Fourier analysis is done for all three diurnal mean rain products to better understand the diurnal structure of each rain mechanism. The time evolution of the mean rain rate at every point in the domain is fit to a sinusoidal function. Figures 10 and 11 show the geographic distributions of the phase (Figure 10) and the amplitude (Figure 11) of the diurnal harmonic, respectively. The coastal propagation pattern is clearly seen in all of the rain products and appears to be relatively uniform across all of the time periods for a given product. The coastal propagation signatures for the HRC and the TMPA match well, while the LRC starts the coastal propagation a little too early, but otherwise

also matches well with the TMPA data. While the models perform well with the timing of the coastal propagation, both setups develop the convection in the interior of the AB too early, and in the case of the LRC, several hours too early.

The geographic distribution in Figure 11 resemble the patterns of the seasonal mean rain distribution from Figure 3, indicating that the strongest diurnal signal occurs in the areas with the largest average rain rates. The HRC shows better agreement with the TRMM data than the LRC does, although the HRC overemphasizes the coastal signature in MAM and underestimates the interior signature in DJF. The LRC overestimates both the coastal and the interior amplitudes relative to the TRMM, likely because it overestimates the average rain rate everywhere in general (Figure 3).

5. THE VERTICAL RAIN STRUCTURE

Figure 12 shows a vertical cross section, taken from both the 2A25 data and the HRC output, which starts from a point on the edge of the 4 km grid near the Andes Mountains and runs perpendicular to the domain boarder across the AB and through the mouth of the Amazon River to the other side of the domain. This line is chosen so that it extends through the area with the highest rain rates. The HRC rain rates in this cross section are several mm/day higher than in the 2A25 data, with the rain rates extending over twice as high vertically as they do in the 2A25 data. Given that the HRC setup is over twice as likely to produce very high rain rates as the 2A25 data and receives the majority of its rainfall from those rain rates (Table 3), the possibility exists that the HRC is producing convection that is too deep. However, the 2A25 data shows an unusually sharp drop off in rain rates at about 5-6 km, potentially due to problems with the TRMM ice retrievals. Figure 13 shows another cross section taken farther north so that the magnitudes of the rain rates are more comparable between the two products. Figures 13c and 13d in this case show a close match in the rain rates between the two products at lower levels, but again the 2A25 shows an artificial cutoff at around 5-6 km.

Figure 14 shows the storm histograms for both products as described in Section 2.5. The total domain histogram for MAM and DJF are shown in Figures 14a and 14d, while the diurnal cycle of the domain histograms is shown in the other four panels of Figure 14. For each product, the diurnal cycle is shown every other hour, with the late evening and overnight (0 Z through 10 Z) designated as “night” and the morning and

afternoon hours (12 Z through 22 Z) designated as “day”. Figure 14d shows a large maximum in the histogram at around 5 km before a sharp drop off between 5km and 10 km, which possibly corresponds to the 2A25 rain rate drop at that height seen in Figures 12 and 13. This maximum is much larger in DJF than in MAM, as opposed to the HRC domain histogram, which shows very similar curves in both seasons (Figure 14a). In contrast, the HRC shows a maximum near the 3 km bin and a significantly higher frequency in the highest bins.

The HRC domain histogram shows very little in the way of changes across the time periods at night, but there is a noticeable decrease with time in the frequency of rain rates in the 3-5 km range and an increase with time in the frequency of events in the 5-7 km range (Figure 14b). The 2A23 data shows even less variability within the histogram than the HRC at night (Figure 14e). However, both products display a clear pattern during the day hours (Figures 14c and 14f). The frequency of lowest bins, up to about 2 or 3 km, hits a maximum at 16 Z, while the maximum of the next few bins, up to about 5 km for the HRC and 6 km for the TRMM data, occurs at 18 Z. The HRC then displays a frequency maximum for the 9km bin to the 16 km bin at 20 Z and 22 Z while the 2A23 data shows the afternoon maximum for bins 8 km to 11 km. This daytime pattern reveals the afternoon cycle of convection that initiates near noon and develops into deep convection (“deep” relative to the product) by the late afternoon.

Figure 15 shows the geographical distribution of the HRC storm top histograms for each season. The height are rebinned into three catagories: shallow tops (ground to 5 km), medium tops (5 km to 10 km), and high tops (10 km and higher). The frequency is

found by normalizing the number of events in each bin by the six-month total number of events as averaged over the domain, which makes all of the panels comparable because they are all normalized by the same number. Figures 15a and 15d reveal high occurrences of shallow tops along and off of the coastline north of the Amazon River mouth and moderate occurrences in the interior of the AB, much of which can be attributed to showers and the initiation of convective systems. The medium tops show a frequency maximum over the Amazon River mouth in MAM (Figure 15b) and over the interior of the AB in DJF (Figure 15e), with moderate occurrences over the AB in MAM and along the coast in DJF. Figures 15c and 15f show that the moderate to high occurrences in high tops for each season coincide with their counterparts in the medium tops panels, suggesting that they represent the cloud structures in the ACSL along the coast and the interior diurnal cycle in the AB.

6. SUMMARY AND CONCLUSIONS

Results show that the HRC model setup equals or outperforms the LRC in nearly all of the comparisons done is this in this study. The HRC does particularly well with the domain averaged mean surface rain rates as compared to the LRC, with six month average rain rate errors relative to the TMPA data of -15% and 81% respectively. The two model setups are much closer in their success in reproducing the geographical distribution of the surface mean rain rate. Both model setups do fairly well in reproducing the general pattern of mean rain rates in MAM, DJF, and DJF-MAM, although there are a couple of exceptions to that generalization. While the LRC mean precipitation pattern matches the TMPA well, the magnitudes of the rain rates are as high as twice as large as the rain rates the TMPA data shows. The HRC tends to overestimate the rain rates along the coast in MAM and underestimate them in the interior of the AB in both seasons. In addition, the HRC geographical mean rain rate distribution along the coast in DJF did not match the TMPA pattern particularly well.

The HRC performs fairly well in the rain rate PDF when compared to the 2A25 TRMM dataset, with the exception of very low and very high rain rates. The disagreement at low rain rates is primarily due to the high noise cutoff of the TRMM PR instrument, which limits the observations to values over 10 mm/day. At high rain rates, the model is found to greatly overestimate the number of high rain events as compared to the observations, which is partially due to attenuation issues in the TRMM PR instrument. Even after adjusting for the attenuation bias, the HRC still produced

significantly more rain at high rain rates than the TRMM data, indicating that the HRC overestimates the frequency of high rain events by roughly 20%. However, despite the error in the HRC at high rain rates, it still performs much better than the LRC; the LRC produces far too much rain at medium rain rates and the rain rate PDF generally does not resemble that of the 2A25.

The one area in which both model setups both perform poorly is the domain mean diurnal cycle. The magnitude of the diurnal mean rain cycle for the LRC is far too high relative to the TMPA data, while the peak and valley of that cycle both occur four to six hours too early. The HRC does much better with the timing of that cycle; it only misses the extremes by an hour or so, but the magnitude of the diurnal mean rain cycle is only about half as much as the TMPA cycle. The HRC tend to allow the nocturnal stratiform rain to persist for too long, which in turn could reduce the solar heating that is needed to power the diurnal rain cycle, thereby dampening it. The overall conclusion with regards to the diurnal mean cycle is that the HRC does much better with the timing of the cycle relative to the TMPA data, but underestimates the magnitude of the cycle almost as much as the LRC overestimates it. The Fourier analysis reinforces the idea that the HRC setup better represents the timing of the rain mechanisms than the LRC setup; while the HRC representation of the interior diurnal rain cycle is somewhat mediocre, it does remarkably well in its representation of the coastal propagating system, versus the LRC, which initiates both of these rain mechanisms too early.

The storm top histograms and the vertical cross sections for the 2A23 and the HRC datasets possibly reflect some of the issues of their respective datasets. The 2A23

contains a large frequency of medium storm top cases, which could be inflated because the TRMM PR instrument sometimes misses information related to deep convective events. This idea is reinforced by the abrupt cutoff of rain rates at the freezing level seen in all four vertical cross sections for the 2A25 data. In contrast, because the HRC is suspected of producing too much rain at high rain rates, the frequency of the high storm tops bin is likely too high. Regardless of these issues, the afternoon convective cycle is still apparent in both storm top histograms.

Overall in this study, the convective resolving model setup performs significantly better than the model setup that parameterizes convection. However, computational resources limited the study in several ways, suggesting that future work is needed to confirm the results. First, computational resources limited the model runs to eight seasons over four years, which is a fairly small sample as far as the number of years goes, raising the possibility that a bias exists in the model data due to under-sampling. Also due to computational limitations, only one combination of model parameters was chosen, so there is likely room for improvement in fine-tuning the model. Lastly, while this likely does not impact the model statistics, it should be kept in mind that the LRC model runs using the Kain-Fritsch cumulus parameterization produced unphysical artifacts. Future work on this project should take all of these points into account, while also testing other regions in the world to see if convectively resolving domains can consistently outperform domains that parameterize the convection.

REFERENCES

Alcantara, C. R., M. A. F. S. Dias, E. P. Souza, J. C. P. Cohen (2011), Verification of the role of the low level jets in Amazon squall lines, *Atmos. Res.*, *100*, 36-44.

Bolen, S. M. and V. Chandrasekar (2000), Quantitative cross validation of space-based and ground-base radar observations, *J. Appl. Meteor.*, *39*, 2071-2079

Bowman, K. P. (2005), Comparison of TRMM precipitation retrievals with rain gauge data from ocean buoys, *J. Climate*, *18*, 178-190

Clark, T. L. (1979), Numerical simulations with a three-dimensional cloud model: Lateral boundary condition experiments and multicellular severe storm simulations. *J. Atmos. Sci.*, *36*, 2191-2215.

Clark, T. L. and W. R. Peltier (1977), On the evolution and stability of finite-amplitude mountain waves. *J. Atmos. Sci.*, *34*, 1715–1730.

Dickinson, R. E., R. M. Errico, F. Giorgi, and G. T. Bates (1989), A regional climate model for the western United States. *Climatic Change*, *15*, 383-422.

Dudhia, J. and M. W. Moncrieff (1989), A three-dimensional numerical study of an Oklahoma squall line containing right-flank supercells. *J. Atmos. Sci.*, *46*, 3363-3391.

Dudhia, J., S.-Y. Hong, and K.-S. Lim (2008), A new method for representing mixed-phase particle fall speeds in bulk microphysics parameterizations, *J. Meteor. Soc. Japan*, *86A*, 33-44.

Garstang, M., S. Ulanski, S. Greco, J. Scala, R. Swap, D. Fitzjarrald, D. Martin, E. Browell, M. Shipman, V. Connors, R. Harriss, and R. Talbot (1990), The Amazon boundary-layer experiment (ABLE 2B): A meteorological perspective, *Bull. Amer. Meteor. Soc.*, *71*, 19-32.

Garstang, M., H. L. Massie, Jr., J. Halverson, S. Greco, and J. Scala (1994), Amazon coastal squall lines. Part I: Structure and kinematics, *Mon. Wea. Rev.*, *122*, 608-622

Giorgi, F. and G. T. Bates (1989), The climatological skill of a regional model over complex terrain. *Mon. Wea. Rev.*, *117*, 2325-2347.

Giorgi, F., L. O. Mearns, C. Shields, and L. Mayer (1996), A regional model study of the importance of local versus remote controls of the 1988 drought and the 1993 flood over the central United States. *J. Climate*, *9*, 1150–1162.

Greco, S., R. Swap, M. Garstang, S. Ulanski, M. Shipham, R. C. Harriss, R. Talbot, M. O. Andreae, and P. Artaxo (1990), Rainfall and surface kinematic conditions over central Amazonia during ABLE 2B, *J. Geophys. Res.*, *95*, 17001-17014.

Greco, S., J. Scala, J. Halverson, H. L. Massie, Jr., W.-K. Tao, and M. Garstang (1994), Amazon coastal squall lines. Part II: Heat and moisture transports, *Mon. Wea. Rev.*, *122*, 623-635.

Hong, S.-Y. and J.-O. J. Lim (2006), The WRF single-moment 6-class microphysics scheme (WSM6), *J. Korean Meteor. Soc.*, *42*, 129-151.

Hong, S.-Y., J. Dudhia, and S.-H. Chen (2004), A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation, *Mon. Wea. Rev.*, *132*, 103-120.

Huffman, G. J., R. F. Adler, D. T. Bolvin, G. Gu, E. J. Nelkin, K. P. Bowman, Y. Hong, E. F. Stocker, and D. B. Wolff (2007), The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales, *J. Hydrometeor.*, *8*, 38-55.

Iguchi, T., T. Kozu, R. Meneghini, J. Awaka, and K.'I. Okamoto (2000), Rain-profiling algorithm for the TRMM precipitation radar, *J. Appl. Meteor.*, *39*, 2038-2052.

Klemp, J. B. and R. B. Wilhelmson (1978), The simulation of three-dimensional convective storm dynamics. *J. Atmos. Sci.*, *35*, 1070-1096.

Kummerow, C., W. Barnes, T. Kozu, J. Shue, and J. Simpson (1998), The tropical rainfall measuring mission (TRMM) sensor package, *J. Atmos. Ocean. Tech.*, *15*, 809-817.

Moncrieff, M. W. and M.J. Miller (1976), The dynamics and simulation of tropical cumulonimbus and squall lines. *Q.J.R. Meteor. Soc.*, *102*, 373–394. doi: 10.1002/qj.49710243208

Leung, L. R., M. S. Wigmosta, S. J. Ghan, D. J. Epstein, and L. W. Vail (1996), Application of a subgrid orographic precipitation/surface hydrology scheme to a mountain watershed, *J. Geophys. Res.*, *101*, 12,803–12,817, doi:10.1029/96JD00441.

Leung, L. R., Y.H. Kuo, and J. Tribbia (2006), Research needs and directions of regional climate modeling using WRF and CCSM. *Bull. Amer. Meteor. Soc.*, *87*, 1747-1751, doi:10.1175/BAMS-87-12-1747.

Liao, L. and R. Meneghini (2009), Validation of TRMM precipitation radar through comparison of its multiyear measurements with ground-based radar, *J. Appl. Meteor. Climatol.*, *48*, 804-817.

Lin, Y.-L., R. D. Farley, and H. D. Orville (1983), Bulk parameterization of the snow field in a cloud model, *J. Climate Appl. Meteor.*, *22*, 1065-1092.

Paeth, H., K. Born, R. Girmes, R. Podzun, and D. Jacob (2009), Regional climate change in tropical and northern Africa due to greenhouse forcing and land use changes. *J. Climate*, *22*, 114–132.

Romatschke, U. and R. A. Houze, Jr. (2010), Extreme summer convection in South America, *J. Climate*, *23*, 3761-3791

Rotunno, R., J. B. Klemp, and M. L. Weisman (1988), A theory for strong, long-lived squall lines. *J. Atmos. Sci.*, *45*, 463–485.

Salati, E. and P. B. Vose (1984), Amazon basin: A system of equilibriums, *Science*, *225*, 129-138.

Schär, C., D. Lüthi, U. Beyerle, and E. Heise (1999), The soil–precipitation feedback: A process study with a regional climate model. *J. Climate*, *12*, 722–741.

Schumacher, C. and R. A. Houze, Jr. (2000), Comparison of radar data from the TRMM satellite and Kwajalein oceanic validation site, *J. Appl. Meteor.*, *39*, 2151-2164.

Simpson, J., R. F. Adler, and G. R. North (1988), A proposed tropical rainfall measuring

mission (TRMM) satellite, *Bull. Amer. Meteor. Soc.*, *69*, 278-295.

Skamarock, W. C., and Coauthors, cited 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN–4751STR, 113 pp. [Available online at http:// www.mmm.ucar.edu/wrf/users/docs/arw_v3.pdf.]

Villar, J. C. E., J. Ronchail, J. L. Guyot, G. Cochonneau, F. Naziano, W. Lavado, E. D. Oliveria, R. Pombosa, and P. Vauchel (2009), Spatio-temporal rainfall variability in the Amazon basin countries (Brazil, Peru, Bolivia, Columbia, and Equador), *Int. J. Climatol.*, *29*, 1574-1594.

Wang, Y., L. R. Leung, J. L. McGregor, D. K. Lee, W. C. Wang, Y. Ding, and F. Kimura (2004), Regional climate modeling: Progress, challenges, and prospects. *J. Meteor. Soc. Japan*, *82*, 1599-1628.

Warner, J. (1970), On steady-state one-dimensional models of cumulus convection. *J. Atmos. Sci.*, *27*, 1035-1040.

Weisman, M. L., J. B. Klemp, and R. Rotunno (1988), Structure and evolution of numerically simulated squall lines. *J. Atmos. Sci.*, *45*, 1990-2013.

Weisman, M. L., W. C. Skamarock, and J. B. Klemp (1997), The resolution dependence of explicitly modeled convective systems. *Mon. Wea. Rev.*, *125*, 527-548.

APPENDIX A

Table 1. The WRF-ARW model parameters.

Model Parameter	12 km	4 km
Horizontal Grid Spacing	12 km	4 km
Vertical Levels	50	50
Model Top	50 mb	50 mb
Time Step	50 s	50/3 s
Initial Conditions	NCEP FNL Analysis	NCEP FNL Analysis
Boundary Conditions	NCEP FNL Analysis (6 h)	NCEP FNL Analysis (6 h)
Cumulus Convection	Kain-Fritsch	Explicit
Boundary Layer	YSU Scheme	YSU Scheme
Surface Layer	Monin-Obukhov	Monin-Obukhov
Microphysics	WSM-6	WSM-6
Land Surface	Noah	Noah
Turbulance	Horizontal Smagorinksy	Horizontal Smagorinksy
Shortwave Radiation	Goddard	Goddard
Longwave Radiation	RRTM	RRTM

Table 2. The domain average rain rates for the TMPA dataset, the HRC (4 km), and the LRC (12 km) in mm/day. The percent error relative to the TMPA data is also given in parenthesis.

	TMPA	4 km	12 km
MAM	7.81	7.08 (-9.46%)	14.69 (87.91%)
DJF	7.80	6.27 (-19.61%)	13.65 (74.98%)
DJF-MAM	7.81	6.67 (-14.54%)	14.15 (81.22%)

Table 3. The integration of the rain rate PDF (rain events) and the net rain contribution (net rain) expressed as a percent over the rain rate bins (mm/day) in column 1. This integration is done for the TRMM 2A25 data, the HRC (4 km), and the LRC (12 km). The numbers in parenthesis are the integration of the adjusted TRMM curves in Figure 4a and 4b.

	TRMM 2A25		4 km		12 km	
Rain Bins	Rain Events	Net Rain	Rain Events	Net Rain	Rain Events	Net Rain
0-10	12.51 (53.33)	1.33 (0.98)	53.46	2.12	29.96	1.54
10-400	85.57 (45.64)	82.73 (61.22)	41.19	40.68	69.83	95.67
400+	1.93 (1.03)	15.94 (37.79)	5.36	57.31	0.22	2.89

Table 4. The percentage of total events that consists of rain events for the TRMM 2A25 data, the HRC (4 km), and the LRC (12 km). This is done for the MAM season, the DJF season, and the six month period over the two seasons.

	TRMM 2A25	4 km	12 km
MAM	6.94	9.32 (4.67)	24.74 (17.43)
DJF	7.95	8.78 (4.39)	24.97 (17.99)
DJF-MAM	7.44	9.05 (4.53)	24.85 (17.71)

APPENDIX B

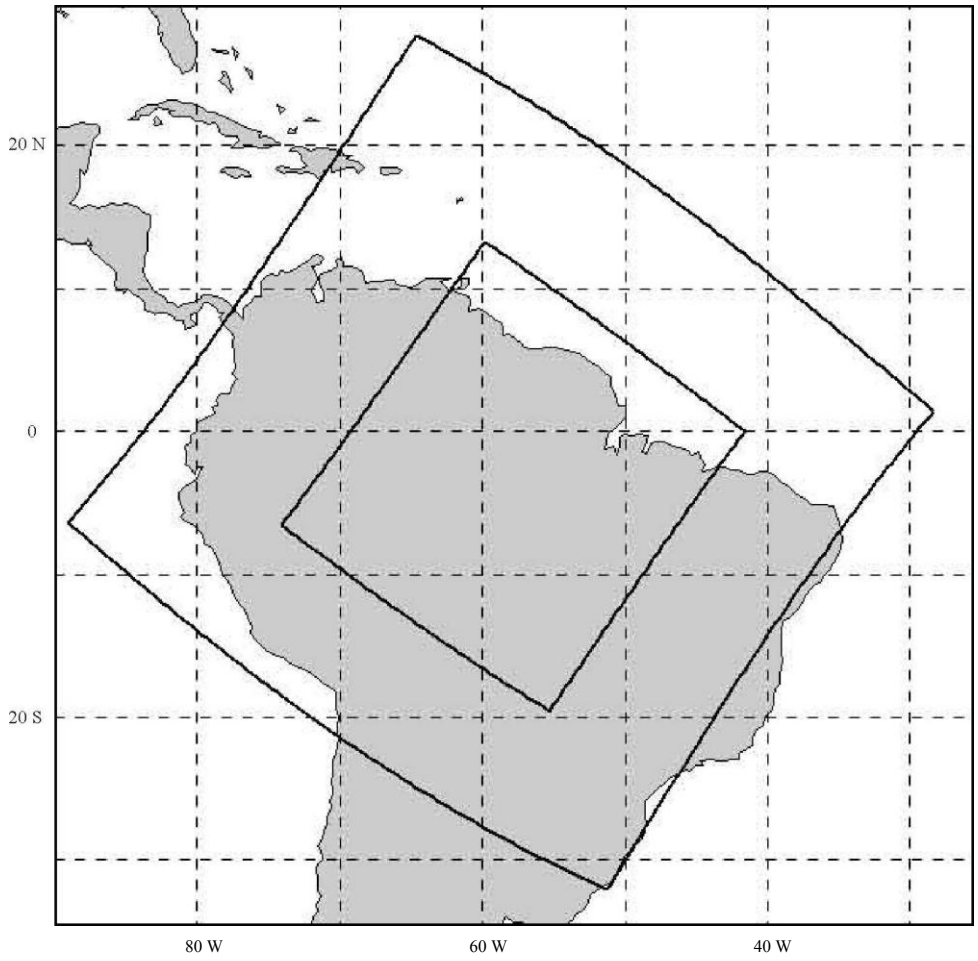

Figure 1. The location of the 12 km domain (outer) and the 4 km domain (inner).

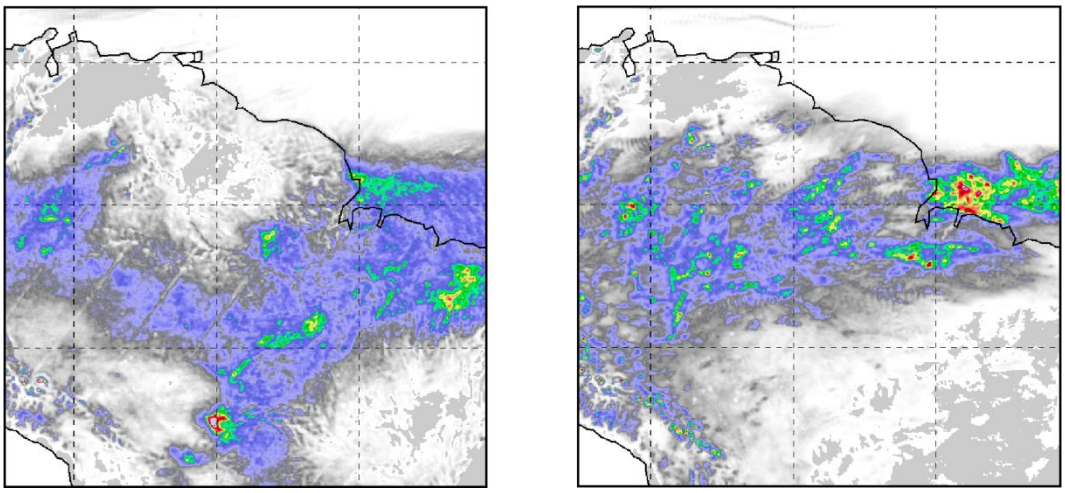

Figure 2. The geographic distribution of a one week average of the model data for the Kain-Fritsch cumulus parameterization (left) and the Betts-Miller cumulus parameterization (right).

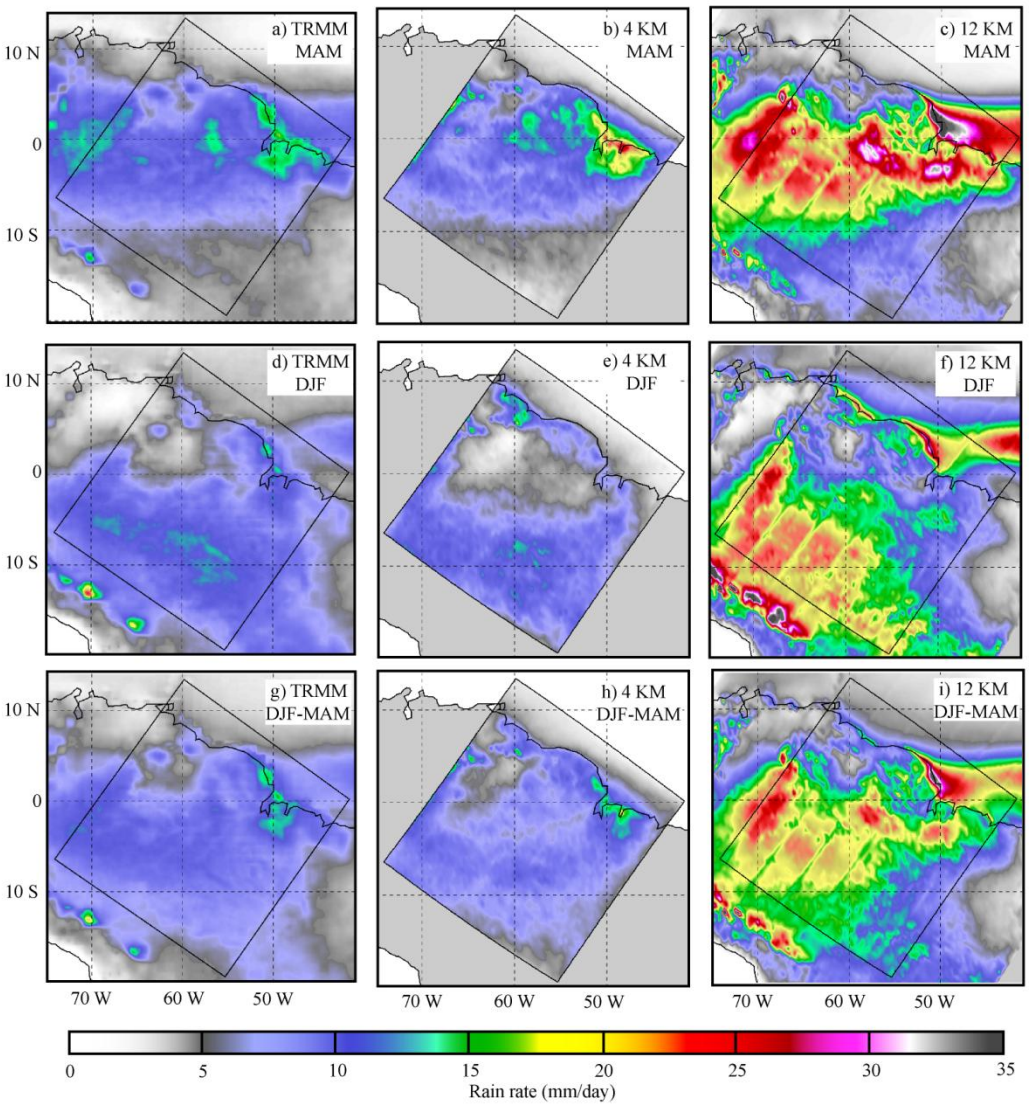

Figure 3. The geographic distribution of mean rain rate for the TMPA data, the HRC (4 km), and the LRC (12 km). The means are taken over the MAM season, the DJF season, and the six month period.

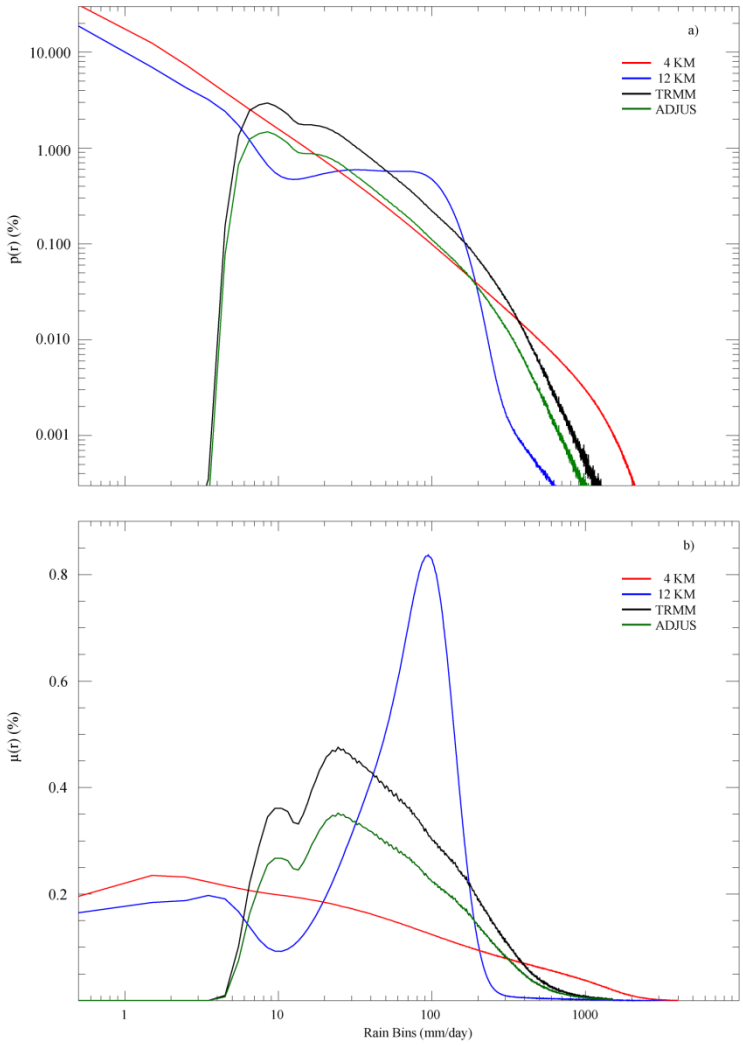

Figure 4. The distributions for the HRC (4 km), the LRC (12 km), and the TRMM 2A25 data of a) the rain rate PDF and b) the net rain rate contribution to the domain average mean rain rate. The line labeled "ADJUS" in a) is the adjustment to the TRMM data to account for possible missed rain events below 8 mm/day, and in b) is the adjustment to the TRMM data to account for missed high rain rate events that significantly contribute to the net average rain. The rain bins in a) and b) and the distribution in a) are plotted on a logarithmic scale.

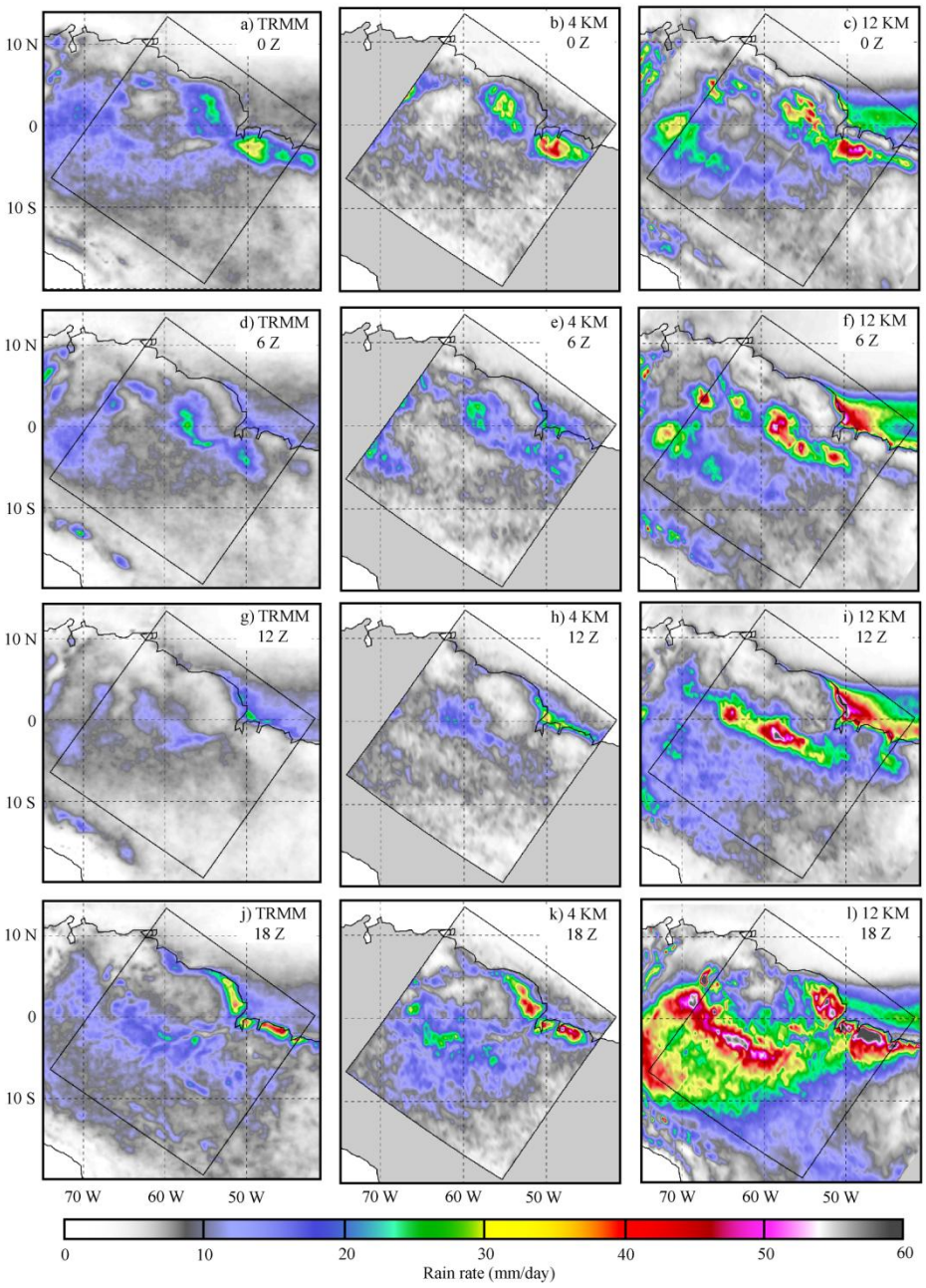

Figure 5. The diurnal cycle for the geographic distribution of the MAM mean rain rate for the TMPA data, the HRC (4 km), and the LRC (12 km). The cycle is shown every six hours (0 Z, 6 Z, 12 Z, and 18 Z).

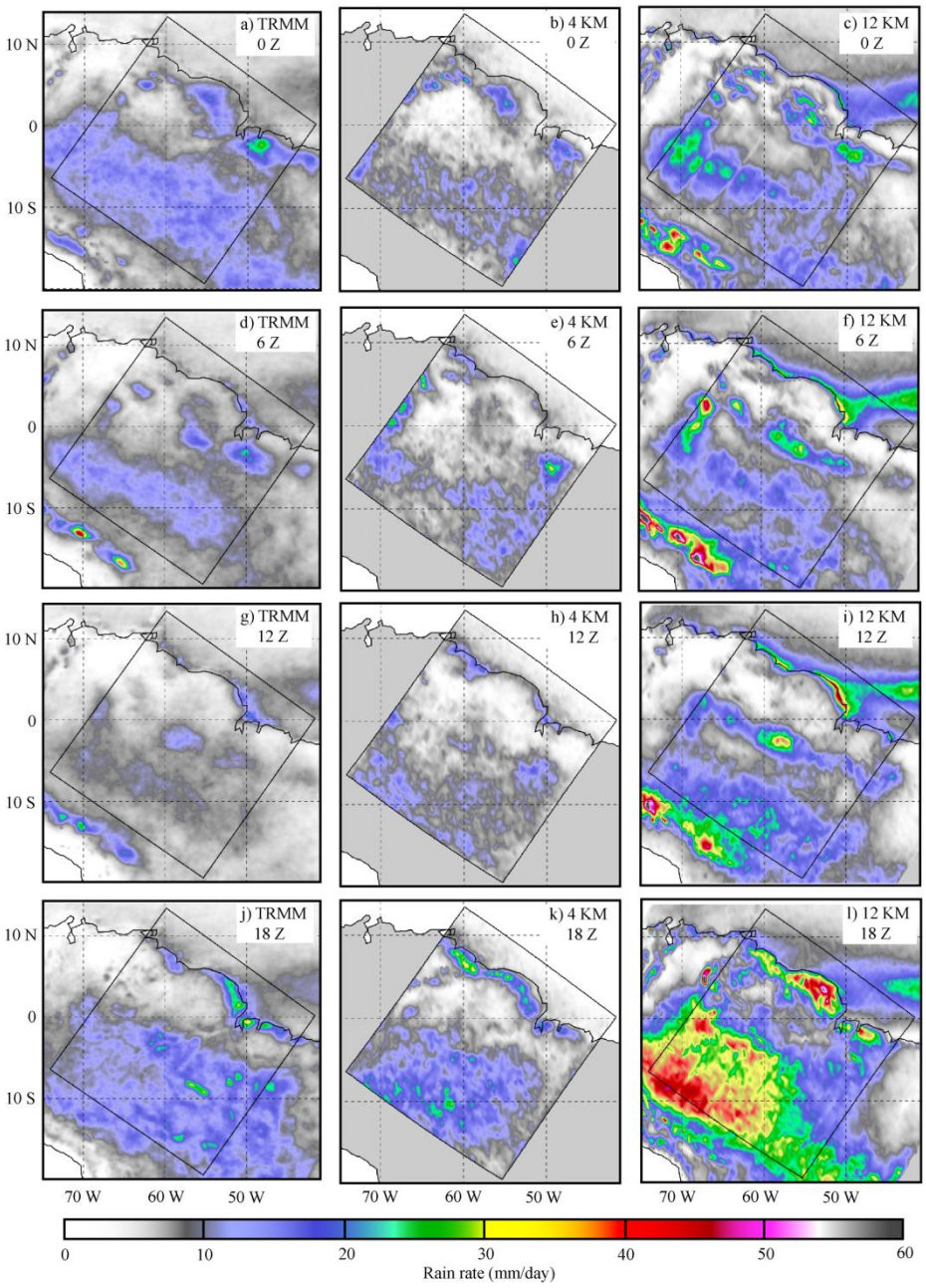

Figure 6. The same as for Figure 5, but for the DJF season.

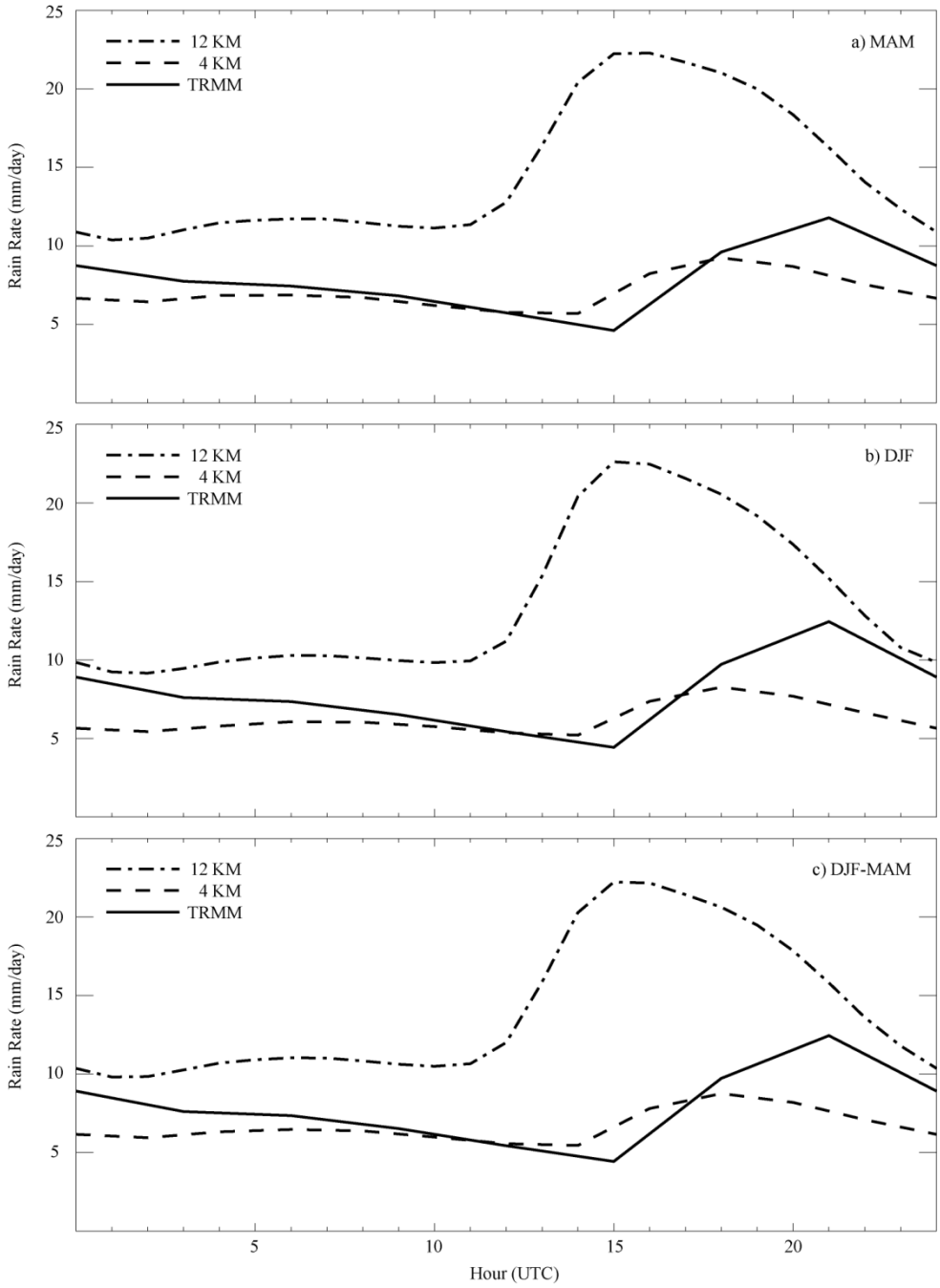

Figure 7. The diurnal cycle of the domain mean rain rate for seasons a) MAM and b) DJF, as well as c) the six month period of DJF-MAM. The comparison is between the HRC (4 km), the LRC (12 km), and the TMPA data.

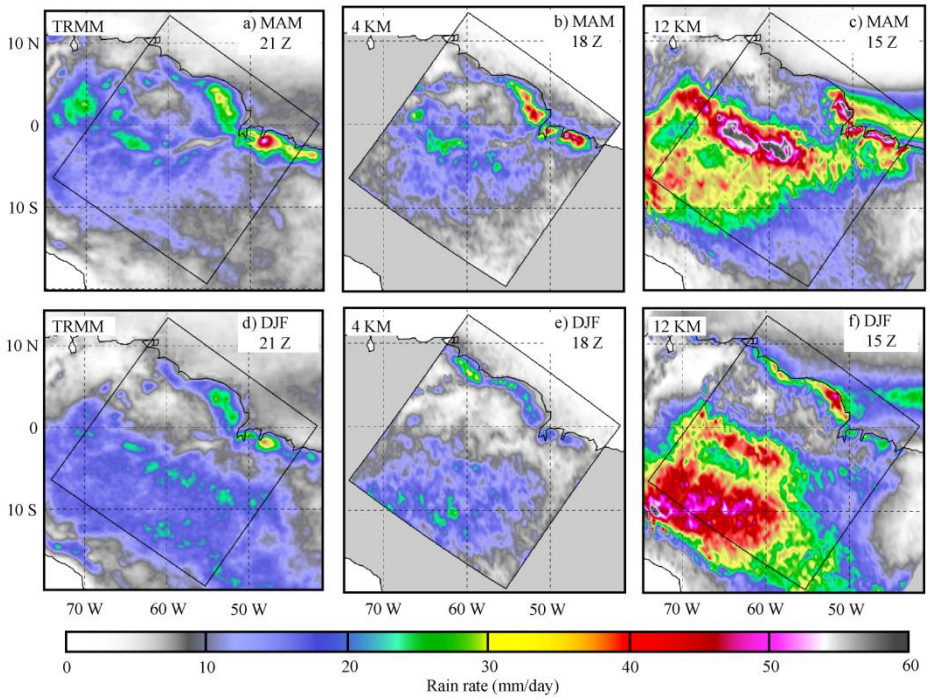

Figure 8. The geographic distribution for the hour of the maximum domain average rain rate for the HRC (4 km), the LRC (12 km), and the TMPA data for the MAM and DJF seasons. This maximum occurs at 21 Z for the TRMM data, 18 Z for the HRC, and 15 Z for the LRC.

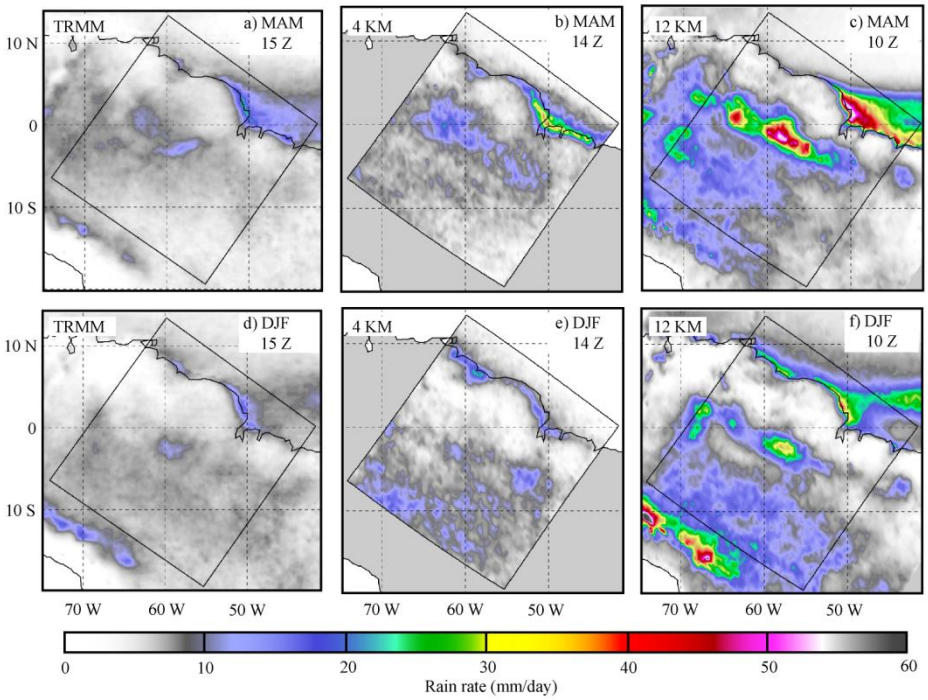

Figure 9. The geographic distribution for the hour of the minimum domain average rain rate for the HRC (4 km), the LRC (12 km), and the TMPA data for the MAM and DJF seasons. This minimum occurs at 15 Z for the TRMM data, 14 Z for the HRC, and 10 Z for the LRC.

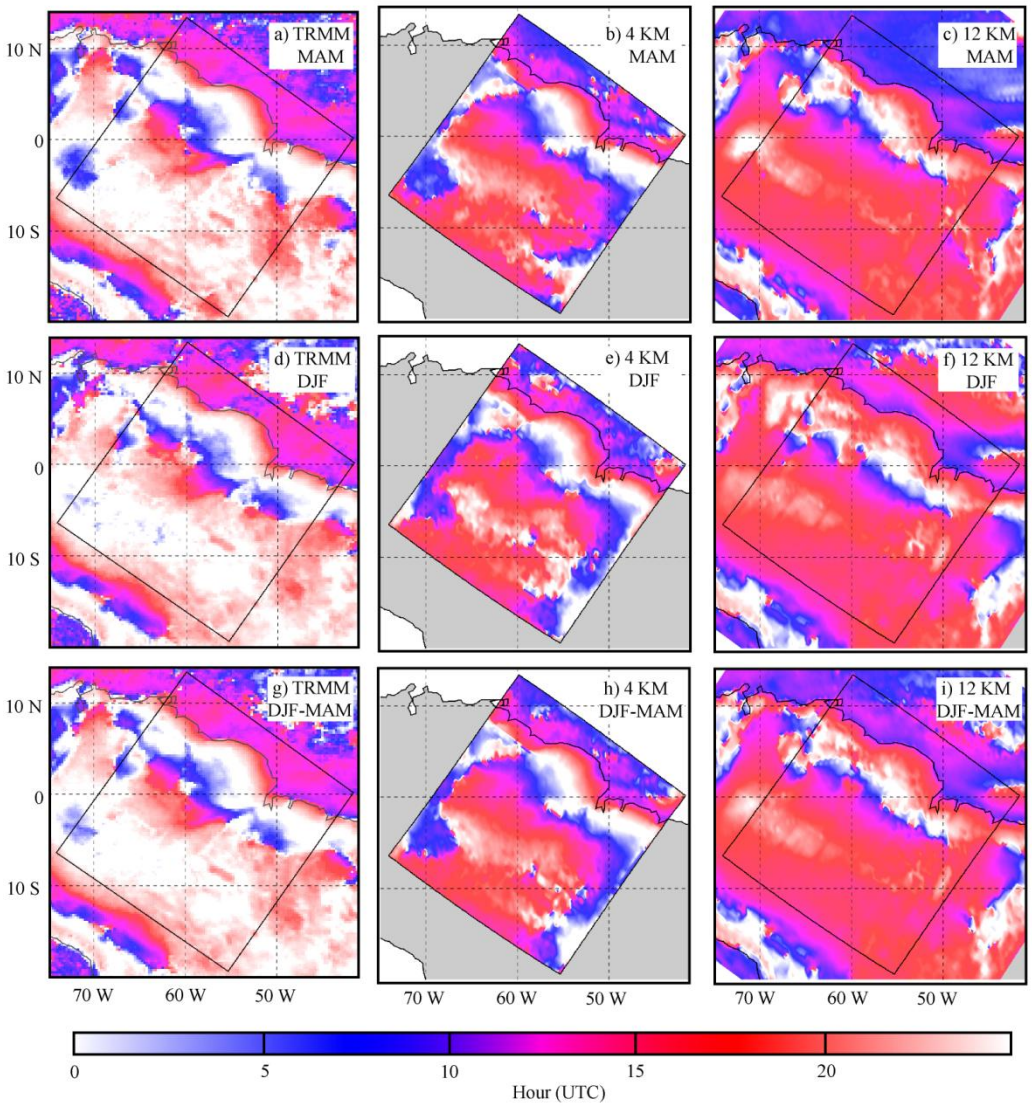

Figure 10. The phase of the Fourier diurnal harmonic mode of the geographic mean rain rate for the TMPA data, the HRC (4 km) rain rates, and the LRC (12 km) rain rates, and for the MAM, DJF, and DJF-MAM time periods.

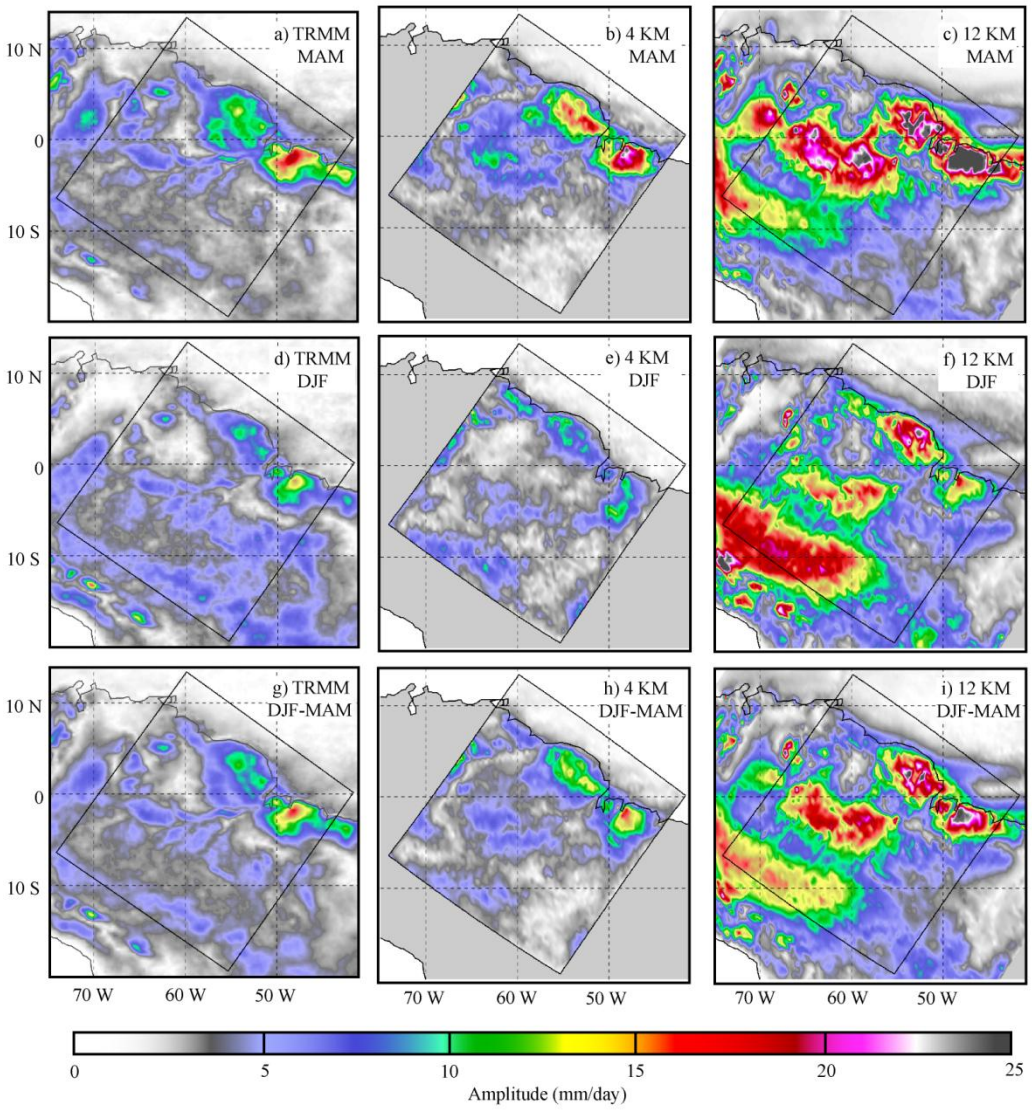

Figure 11. The amplitude of the Fourier diurnal harmonic mode of the geographic mean rain rate for the TMPA data, the HRC (4 km) rain rates, and the LRC (12 km) rain rates, and for the MAM, DJF, and DJF-MAM time periods.

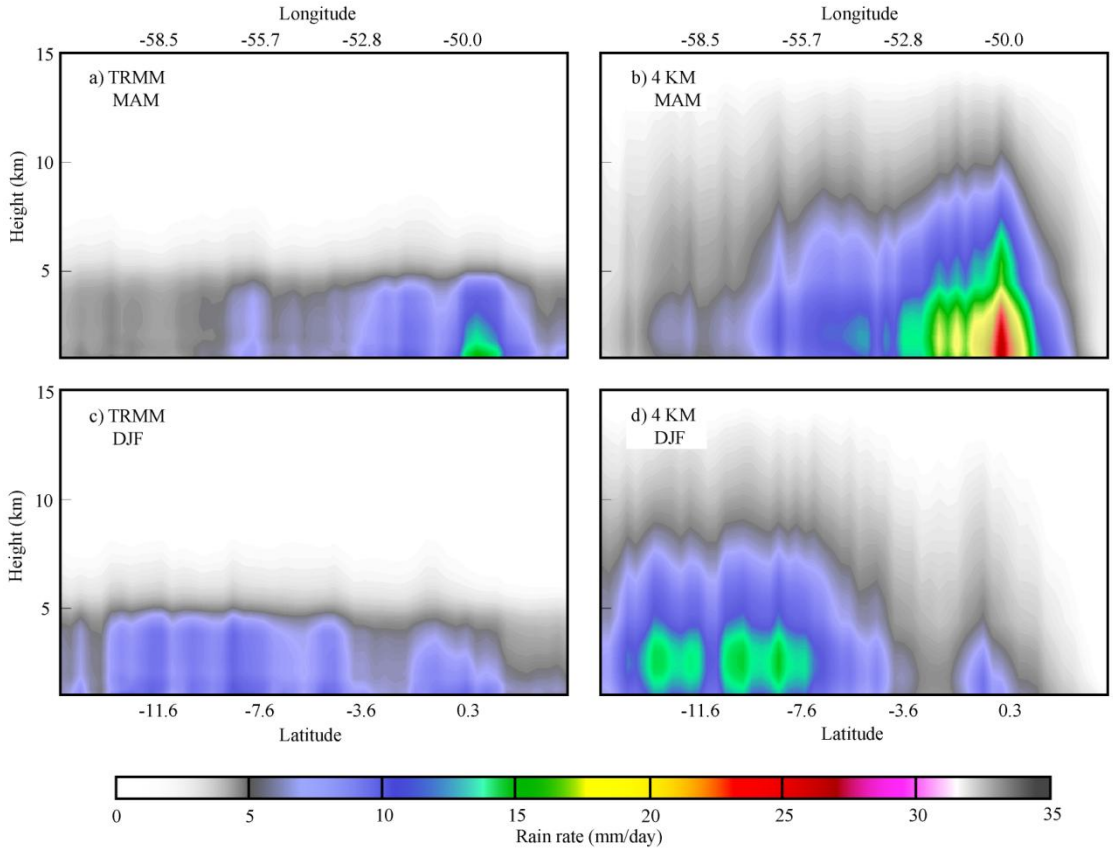

Figure 12. The vertical cross section through the mouth of the Amazon River, where some of the highest average rain rates are located. This cross section is done for the HRC (4 km) rain rates and the TRMM 2A25 rain rates, and for DJF and MAM seasons.

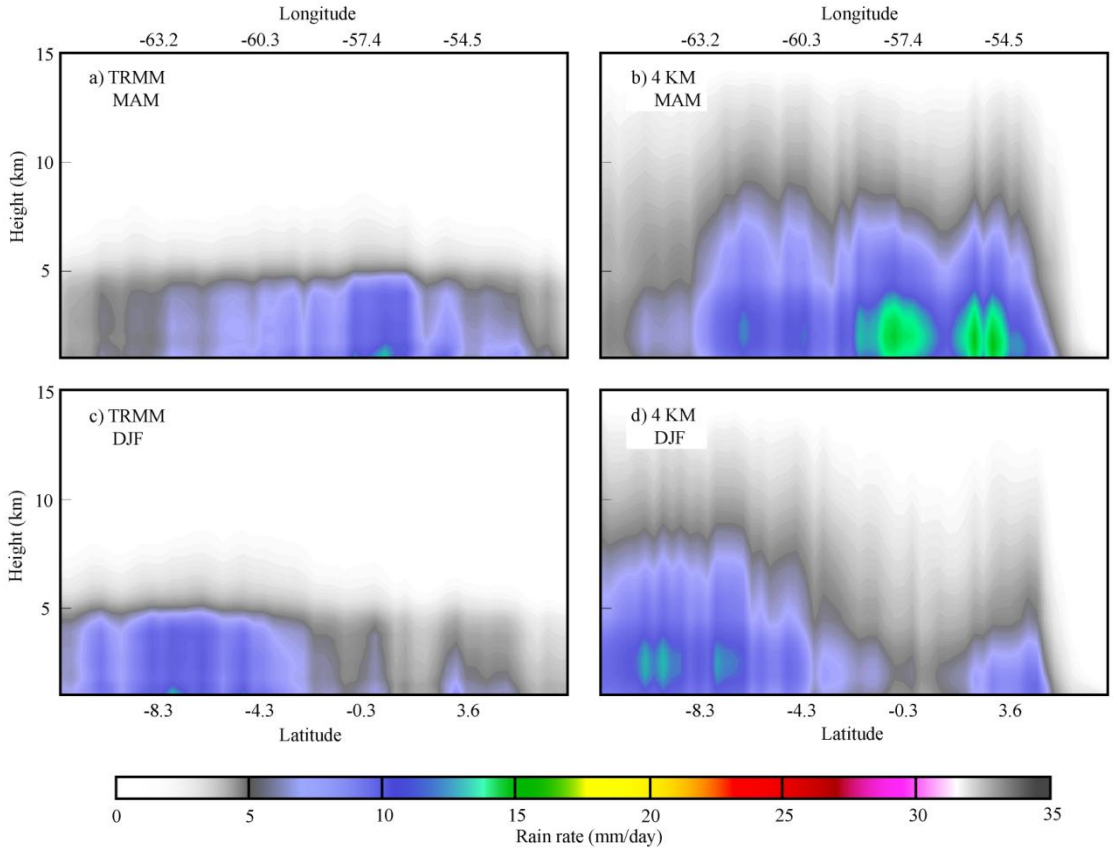

Figure 13. Same as for Figure 12, but for a parallel cross section farther north and away from the higher rain rates.

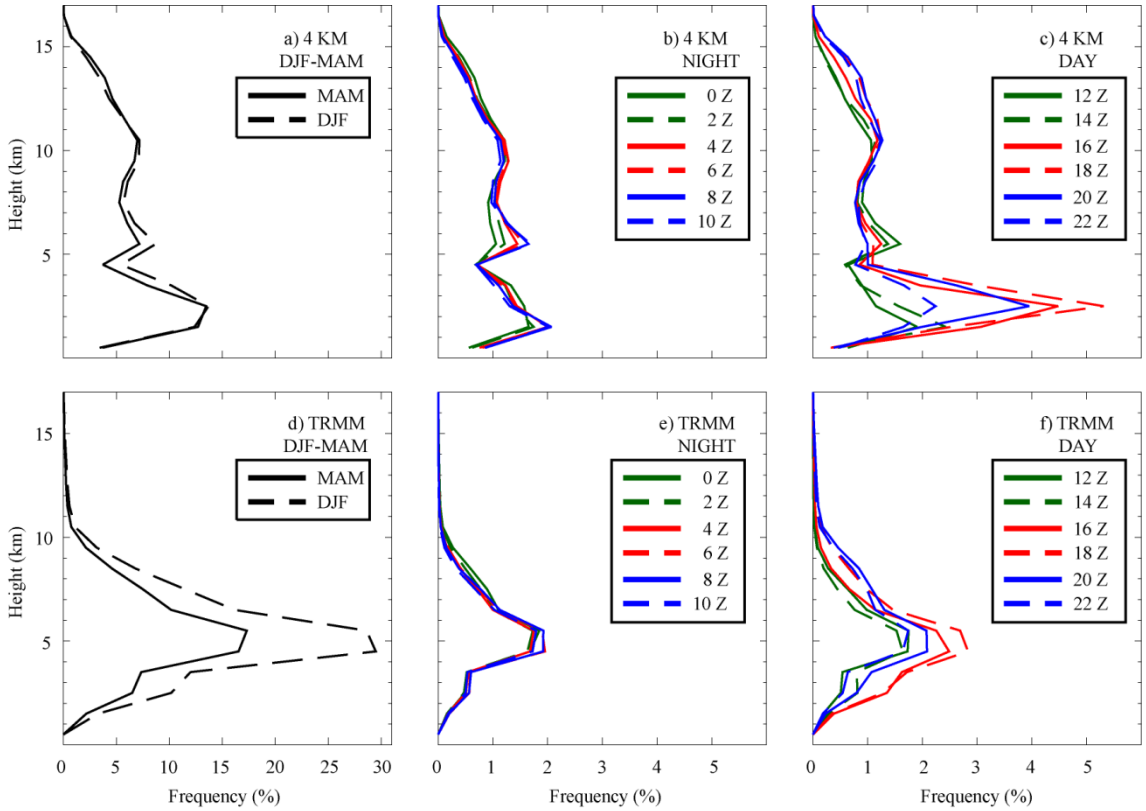

Figure 14. The histograms for the storm top heights of the TRMM 2A23 data and the HRC (4 km) rain rates. Panels a) and d) show a seasonal breakdown of the histogram for MAM and DJF for each product. The diurnal breakdown of the six month histogram is split into b) and e) the late evening and night times hours and c) and f) the early morning and afternoon hours. The histograms are all normalized by the total number of rain events in the six month period DJF-MAM.

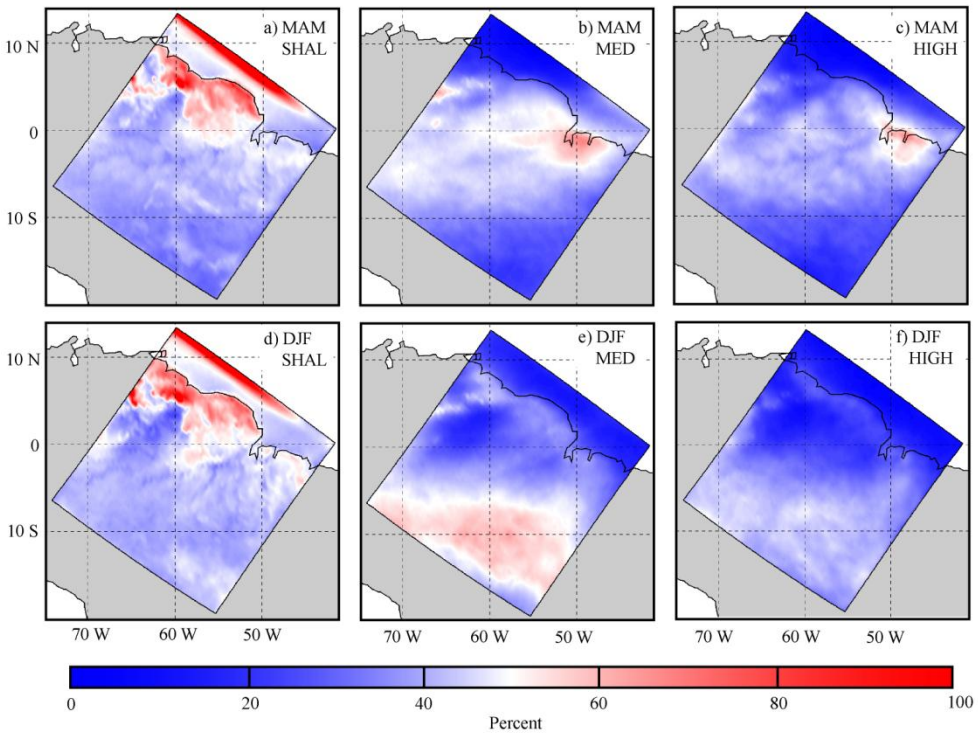

Figure 15. The geographic distribution of the storm top histograms for the HRC (4 km) rain rate output for MAM and DJF. The histogram bins are rebinned into a shallow storm top bin (0 km to 5 km), a medium storm top bin (5 km to 10 km), and a high storm top bin (10 km and higher). The number of event in each bin is normalized by the six month average number of total events at a given (i,j) point and the distribution is shown as a percent.

APPENDIX C

a_G	330
a_R	841.9
a_S	11.72
b_G	0.8
b_R	0.8
b_S	0.41
n_{0G}	4e6 m^{-4}
n_{0R}	8e6 m^{-4}
n_{0S}	2e6 * exp[0.12*(T-T_0)] m^{-4}
T_0	273.16 K
ρ_G	500 kg m^{-3}
ρ_R	$1e^3$ kg m^{-3}
ρ_S	100 kg m^{-3}
ρ_0	1.28 kg m^{-3}