Characterizing Wet Season Precipitation in the Central Amazon Using a Mesoscale Convective System Tracking Algorithm

Sheng-Lun Tai[1], Zhe Feng[1], James Marquis[2], and Jerome Fast[3]

[1]Pacific Northwest National Laboratory (DOE)
[2]Unknown
[3]PNNL

March 13, 2024

Abstract

To comprehensively characterize convective precipitation in the central Amazon region, we utilize the Python FLEXible object TRacKeR (PyFLEXTRKR) to track mesoscale convective systems (MCSs) observed through satellite measurements and simulated by the Weather Research and Forecasting (WRF) model at convection-permitting resolution. This study spans a two-month period during the wet seasons of 2014 and 2015. We observe a strong correlation between MCS track density and accumulated precipitation in the Amazon basin. Key factors contributing to precipitation, such as MCS properties (number, size, rainfall intensity, and movement), are thoroughly examined. Our analysis reveals that while the overall model produces fewer MCSs with smaller mean sizes compared to observations, it tends to overpredict total precipitation due to excessive rainfall intensity for heavy rainfall events ([?] 10 mm h-1) and longer traveled distances than observed. These biases in simulated MCS properties vary with the strength of constraints on convective background environment. Moreover, while the wet bias from heavy (convective) rainfall outweighs the dry bias in light (stratiform) rainfall, the latter can be crucial, particularly when MCS cloud cover is significantly underestimated. A relevant case study for April 1, 2014 highlights the influence of environmental conditions on the MCS lifecycle and identifies an unrealistic model representation in convective precipitation features.

Characterizing Wet Season Precipitation in the Central Amazon Using a Mesoscale Convective System Tracking Algorithm

Sheng-Lun Tai, Zhe Feng, James Marquis, and Jerome Fast

Pacific Northwest National Laboratory, Richland, WA.

Corresponding author: Sheng-Lun Tai (sheng-lun.tai@pnnl.gov)

Key Points:

- Simulated and observed MCS clouds and precipitation are tracked during the 2014/15 wet seasons in central Amazon.
- Excessive heavy rain intensity ($\geq$ 10 mm h^{-1}) and relatively long travel distance of simulated MCS lead to overall overprediction of precipitation.
- Dry bias associated with stratiform rainfall may also drive MCS precipitation bias when cloud cover is substantially underpredicted.

Abstract

To comprehensively characterize convective precipitation in the central Amazon region, we utilize the Python FLEXible object TRacKeR (PyFLEXTRKR) to track mesoscale convective systems (MCSs) observed through satellite measurements and simulated by the Weather Research and Forecasting (WRF) model at convection-permitting resolution. This study spans a two-month period during the wet seasons of 2014 and 2015. We observe a strong correlation between MCS track density and accumulated precipitation in the Amazon basin. Key factors contributing to precipitation, such as MCS properties (number, size, rainfall intensity, and movement), are thoroughly examined. Our analysis reveals that while the overall model produces fewer MCSs with smaller mean sizes compared to observations, it tends to overpredict total precipitation due to excessive rainfall intensity for heavy rainfall events ($\geq$ 10 mm h^{-1}) and longer traveled distances than observed. These biases in simulated MCS properties could vary with the strength of constraints on convective background environment. Moreover, while the wet bias from heavy (convective) rainfall outweighs the dry bias in light (stratiform) rainfall, the latter can be crucial, particularly when MCS cloud cover is significantly underestimated. A relevant case study for April 1, 2014 highlights the influence of environmental conditions on the MCS lifecycle and identifies an unrealistic model representation in convective precipitation features.

Plain Language Summary

We tracked large-size rain storms called mesoscale convective systems (MCSs) in the central Amazon during the wet seasons of 2014 and 2015 using an automated feature tracking algorithm. Data generated from MCS tracking helps us understand how MCSs rainfall is produced as a function of the number of storms, as well as their size, rain intensity, and motion, and how those can be better simulated by weather and climate models. We found that generally the model produces less and smaller MCSs than in reality, but the total MCS rainfall amount is often overestimated. This is because simulated MCSs travel longer, and most importantly they produce unrealistically intense heavy rainfall events. On the other hand, light rainfall events are mostly underrepresented by the model. Thus, the model error in total precipitation is determined by how these two compensate each other. Our analysis also suggests accurate model representation in environment is required for simulating realistic MCS properties.

1 Introduction

Mesoscale convective system (MCS) is a deep convective storm with clouds and precipitation organized on spatial scales of 100 km (Houze 2014). In the Amazon, MCSs produce over 60% of the total rainfall during the wet season (March-April-May, MAM), primarily due to their relatively long duration and wide extent than less organized convective storms (Nesbitt et al. 2006; Feng et al. 2021; Schumacher and Rasmussen 2020). In addition to the hydrological impact of MCSs, their extensive anvil clouds have a sizeable impact on the regional radiation budget (Feng et al. 2011). Thus, MCSs play a noteworthy role in regional and global climate.

Most Amazonian MCSs are initiated by near-surface convergence associated with prevailing trade winds, surface friction, and dirunal sea breeze circulations on the Atlantic coast (Sousa et al. 2021). After initiation, many MCSs travel long distances across the Amazon basin with a lifetime over half a day, sustained by environmental conditions that are frequently favorable for deep convection and growing to substantial sizes owing to rich tropical moisture. Hence, the local precipitation diurnal cycle in the Amazon basin is tied closely to westward propagating MCSs (Tai et al. 2021).

Several studies have employed a combination of routine and field campaign observational datasets to characterize MCSs in the Amazon region including weather radars, satellite observations, and surface measurements. Such efforts capture the intricate dynamics of MCS, their spatial distribution, intensity, and lifecycle evolution (e.g., Laurent et al. 2002; Petersen et al. 2002; Cifelli et al. 2002; Rickenbach et al. 2002; Machado et al. 2004). The recent GoAmazon2014/5 field campaign (Martin et al. 2016) collected many valuable observational data sets that have been used to characterize the diurnal variation, morphology, propagation, vertical motion, and precipitation of convective clouds around central Amazon (Burleyson et al. 2016; Rehbein et al. 2019; Giangrande et al. 2020; Wang et al. 2020; Tian et al. 2021; Anselmo et al. 2020). In addition to observations, advanced regional climate and convection-permitting models have been utilized to simulate MCSs over the Amazon under various meteorological conditions. By integrating observations with simulations, researchers have deepened our understanding of the processes governing MCS formation and evolution in the Amazon (Silva Dias et al. 2002; Carvalho et al. 2002; Machado et al. 2004; Tai et al. 2021, Paccini et al. 2023). These studies provide insights into the convective organization within the Amazon basin, shedding light on the complex interactions between atmospheric dynamics, moisture availability, and convective activity.

Nonetheless, remaining model uncertainties in MCSs over Amazon has motivated additional research in the field. Prior studies have shown that simulated MCS precipitation is quite sensitive to model resolution (vertical and horizontal), atmospheric forcing in initial and boundary conditions, soil moisture, and physics parameterizations (land surface, planetary boundary layer, cloud microphysics, and radiation) (e.g., Luo et al. 2015; Stensrud et al. 2000; Feng et al. 2018; Prein et al. 2021; Tai et al 2021; Prein et al. 2022; Na et al. 2022; Rasmos-Valle et al. 2023; Yang et al. 2023). Due to constraints in availability and spatial coverage of observational data in the sparsely populated Amazon, most evaluations of simulated MCS behaviors have been conducted in a confined region and narrow time windows which may be shorter than the MCS lifetime. Thus, additional work is warranted to examine modeled MCSs in realistic Amazonian conditions.

Cataloging MCS frequency, size, precipitation intensity, and movement, are essential for determining precipitation processes that contribute to total accumulative rainfall in the Amazon. However, representation of these MCS characteristics in state-of-the-art atmospheric models lacks rigorous quantitative validation. A pioneering study from Laurent et al. (2002) uses geostationary satellite data with 30-min frequency to enable deep convective cloud tracking, providing a different aspect in assessing modeled storms. With an increasing amount and quality of available satellite data, a number of cloud cluster tracking tools have been developed in recent years to characterize the lifecycle of deep convective clouds (Anselmo et al. 2021; Huang et al. 2018; Feng et al. 2019; 2021, 2023; Rehbein et al. 2018; Galarneau et al. 2023; Prein et al. 2020; Da Silva et al. 2023). One example is the Python FLEXible object TRacKeR (PyFLEXTRKR, Feng et al. 2023) algorithm, which we adopt in this study to facilitate MCS tracking in the central Amazon using satellite observations and a series of convection-permitting (4-km grid spacing) simulations during the 2014/15 wet seasons. The goal of this study is to elucidate the role of key MCS properties in driving the model precipitation errors through an in-depth storm tracking analysis.

The remainder of this paper is organized as following. Section 2 provides the details of the model and experiments as well as the algorithm used for trackings both simulated and observed MCSs. Results of analysis derived from MCS tracking statistics across timescales are demonstrated in Section 3. Finally, summary and conclusion are provided in Section 4.

2 Methods

2.1 Model setup and experiments

We use the WRF model version 3.9.1 (ARW, Skamarock et al. (2008)) to simulate convective clouds over the entire Amazon region, using a general configuration similar to our previous work in this region (Tai et al. 2021). Our study period includes a month in 2014 (March 11 to April 10) and 2015 (March 1 - 31). The model domain encompasses the northern part of the South American continent as well as adjacent oceans (Figure 1). The domain is constructed with a horizontal grid spacing of 4 km and a stretched vertical coordinate of 60 levels. The model top is located at 100 hPa. The physics schemes used for the simulations include: Thompson microphysics parameterization (Thompson et al. 2008), Mellor-Yamada-Nakanishi Niino (MYNN) boundary layer parameterization (Nakanishi and Niino 2009), Mellor-Yamada-Janjic surface layer parameterization (Janjić 2001), Unified Noah land-surface parameterization (Chen and Dudhia 2001), and the RRTMG longwave and shortwave radiation parameterization (Iacono et al. 2008). No cumulus parametrization is used because the model's horizontal grid spacing (4 km) is capable of resolving MCSs (Prein et al 2020; Na et al 2022). We use 6-hourly, 1° × 1° NCEP FNL operational model global tropospheric analysis for model initialization (National Centers for Environmental Prediction 2000).

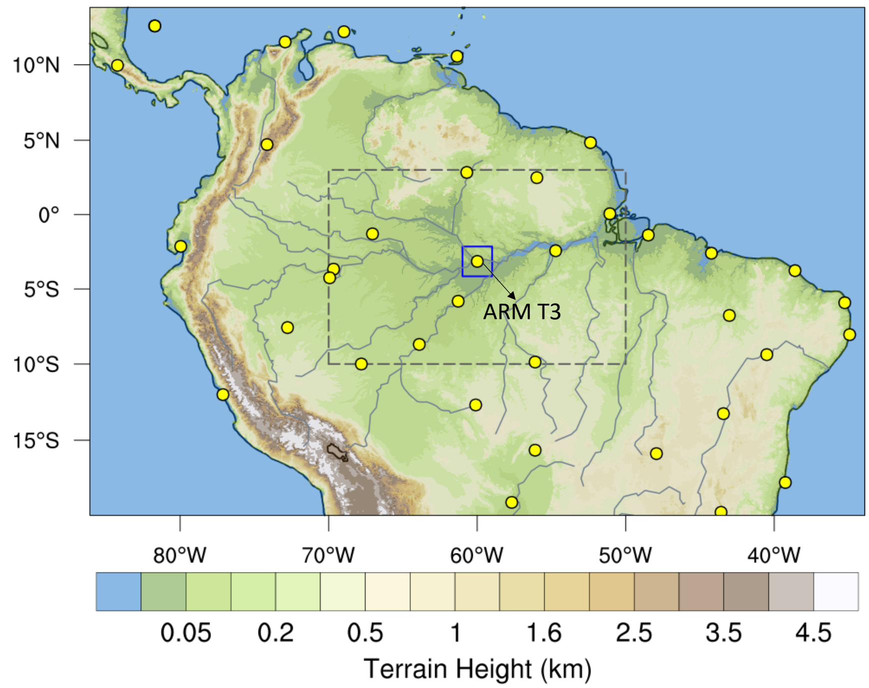

Figure 1 Map shows the configured WRF model domain for the simulations used in this study. Color shading illustrates terrain heights. Yellow dots denote the locations of radiosonde profiles that are assimilated along with the simulations. The location of ARM T3 site during GoAmazon2014/5 is indicated. The dashed rectangle marked by dashed line represents the study area for the MCS tracking analysis. Subdomain denoted by blue box is used for profiles sampling discussed in Section 3.4.

The model is also coupled with a data assimilation (DA) scheme to better constrain the simulation's background meteorological condtions, identical to the approach used in Tai et al. (2021). Conventional observations (e.g., radiosonde profiles, surface meteorology, aircraft, ship and others) and satellite radiances are assimilated by using the three-dimensional variational (3DVar) technique as provided in the package of version 3.6 Community Gridpoint Statistical Interpolation (GSI, Shao et al. 2015). It produces optimized analyses by blending the model data with observations as collected for the NCEP Global Data Assimilation System (GDAS, http://rda.ucar.edu/datasets/ds337.0/) and the radiosonde profiles measured at the T3 site (Figure 1) deployed by the DOE's Atmospheric Radiation Measurement (ARM) Mobile Facility (AMF, Miller et al. 2016) during the GoAmazon2014/5 field campaign (Martin et al. 2016).

Assimilated radiosondes were launched every 6 hours at the ARM T3 (e.g., 00, 06, 12, and18 UTC; 20, 02, 08, and 14 LT) throughout the campaign to measure tropospheric winds, temperature, pressure, and humidity profiles. Over the intensive observational periods (IOPs), one additional radiosonde was launched at 15 UTC (11 LT) to enhance measurement of the diurnal variation of environmental conditions. In addition to ARM's radiosondes, meteorological profiles measured at other sites as archived in the NCEP ADP global upper air and surface weather observations product are also assimilated in our model configuration. Note these observations have lower temporal frequency (up to twice a day at 00 and 12 UTC) and vertical resolution than those performed at the T3 site. The yellow dots denoted in Figure 1 indicate the locations of available radiosonde data at 12 UTC of March 12, 2014. The DA-coupled simulation is initialized at 00 UTC on the first day of each simulated month. The 3DVar data assimilation is performed every 12 hours (at 00 and 12

UTC) throughout the simulation periods. More details regarding the model configuration and DA strategy can be found in Tai et al. (2021).

2.2 Tracking observed MCSs

The MCS tracking is performed using the Python FLEXible object TRacKeR (PyFLEXTRKR, Feng et al. (2021, 2023)), a software package which is designed to track any atmospheric features in 2-D geographic planes using user-prescribed observational data sets or model output. To objectively identify and track the deep convective clouds, the PyFLEXTRKR primarily uses 2-D projections of infrared brightness temperature (Tb) observations commonly measured by geostationary satellites. When tracking observed MCSs with PyFLEXTRKR, we use NASA's Global Merged IR V1 infrared brightness temperature (Janowiak et al. 2017) data set. This dataset comprises multiple operational geostationary satellite data sources and includes viewing angle and parallax corrections. It has a continuous global coverage from 60°S to 60°N with a horizontal grid spacing of ~ 4 km and a temporal resolution of 30 min. We use hourly Tb data to identify and track deep convective clouds associated with MCSs identical to the approach adopted by Feng et al., (2021; 2023). A detailed discussion of the impact of uncertainties in MCS tracking owing to the IR Tb data set are described in Feng et al. (2021). Moreover, PyFLEXTRKR also uses collocated surface precipitation to assist identification of MCS "precipitation features (PF, contiguous area with rain rate > 2 mm h^{-1})" (Feng et al., 2021). Because PyFLEXTRKR tracks all deep convective clouds with a pre-defined minimum area threshold, the tracking data consists of records from early stage (near initiation) of individual deep convective clouds to the decay stage when the area of the cloud system decreases. More details of the PyFLEXTRKR algorithm can be found in Feng et al. (2023).

We use the NASA Integrated Multi-satellitE Retrievals for Global Precipitation Measurement (GPM) (IMERG) V06B precipitation data (Huffman et al. 2019) as a source of observed rainfall data in the MCS tracking algorithm. Precipitation estimates in IMERG are obtained by various precipitation-retrieving satellite passive microwave (PMW) sensors using the Goddard Profiling algorithm (Kummerow et al., 2001, 2015, 2011). Intercalibration is performed using the GPM Combined Radar Radiometer Analysis product. The precipitation product has a grid spacing of 0.1° and is also available every 30 minutes over a large portion of the globe (Huffman et al., 2014; Hou et al., 2014; Tang et al., 2016; Tan et al., 2019). We further averaged the 30 min IMERG data to hourly, and coarsened the 4 km Tb data to match the IMERG grid. Hence, the collocated Tb and IMERG precipitation data at 0.1° and hourly resolution are jointly used for MCS tracking in this study.

2.3 Tracking simulated MCSs

In WRF simulations, the top-of-atmosphere (TOA) outgoing longwave radiation (OLR) is used to infer the Tb. An empirical function is employed in PyFLEXTRKR to convert OLR to Tb following the formula from Yang and Slingo (2001). To avoid observational and simulated data resolution mismatches, the 4 km WRF simulation output is regridded based on the coordinate of observational data in a grid resolution of 0.1°. In this study, the thresholds used to define convective clouds, MCSs, and PFs in terms of cloud top brightness temperature, rain rate, and feature size are listed in Table 1. The sensitivity of MCS tracking due to variations of these thresholds was found to be qualitatively minor based on prior tests (not shown). A recent study comparing six different feature tracking algorithms applied to observed and simulated MCSs over South America found that most of the MCS properties from PyFLEXTRKR are representative of results from other obeject tracking tools (Prein et al. 2023), suggesting our algorithm can produce representative MCS characteristics.

Table 1 Summary of parameters used for the MCS tracking algorithm in the PyFLEXTRKR.

Category	Parameter	Value	Unit	Description
Cloud identification	Warm cloud Tb	261	K	Brightness temperature threshold for identification of "warm anvil"
	Cold cloud Tb	241	K	Brightness temperature threshold for identification of "cold anvil"
	Core cloud Tb	225	K	Brightness temperature threshold for identification of "core cloud"
	Minimum cold core cloud pixels	4	unitless	Mininum number of pixels of cold core cloud in qualification of a "core cloud"
	Minimum area	800	km^2	Minimum area in qualification of a "cloud" object
	Missing data fraction	0.35	unitless	Maximum fraction for missing data
	Minimum area	40000	km^2	Minimum total cloud area in qualification of a MCS
MCS identification	Minimum duration	4	hour	Minimum duration in qualification of a MCS
	Minimum PF rain rate	3	mm h^{-1}	Minimum rain rate in qualification as part of a PF
	Minimum PF link area	648	km^2	Minimum linked area of a PF
PF identification	Minimum PF major axis	100	km	Minimum length for a PF's major axis
	Maximum PF major axis	1800	km	Maximum length for a PF's major axis
	Minimum PF duration	4	hour	Maximum duration for a PF
	Minimum PF rainrate	2	mm h^{-1}	Cut-off rain rate in a PF
	Heavy rain rate threshold	10	mm h^{-1}	Minimum rain rate to be defined as "heavy rain"

2.4 A MCS tracking example

Figure 2 demonstrates an example of MCS detection and tracking at 13 UTC on March 19, 2014, during our study period. In this case, WRF simulates comparable fractions of core and cold anvil clouds. However, due to narrowing between two simulated cloud clusters (Figure 2b), there are two separate MCSs identified rather than a single MCS in observations (Figure 2a). The results of cloud type identification based on the defined thresholds in Table 1 are shown in Figures 2c and 2d. The tracked precipitation features corresponding to this MCS event are shown in Figures 2g and 2h, identified using the rain rates in Figures 2e and 2f. The model produces much higher rain rates than is observed in this time. The simulated maximum rain rate is 49.41 mm h^{-1}; whereas, the maximum satellite-retrieved observational rain rate was less than half of that (21.9 mm h^{-1}). This tendency leads to a relatively large fraction of heavy rainfall area in the simulated MCS than is observed. In addition, the model produces a large area of very light rainfall near the east side of the domain likely associated with a sea breeze circulation (Figure 2f) that is less evident in the satellite observations (Figure 2e).

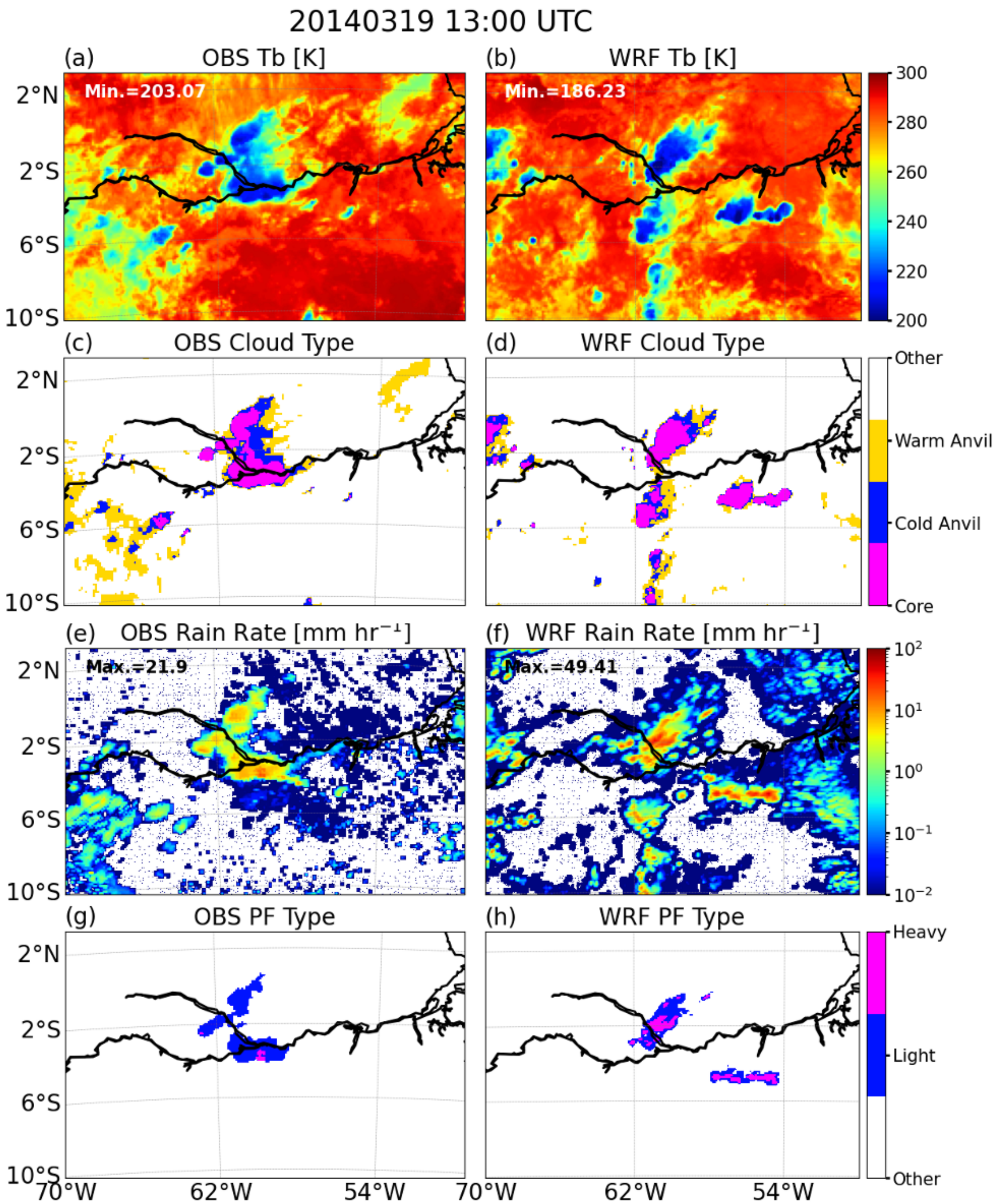

Figure 2 Snapshots of brightness temperature, cloud type, rain rate, and precipitation type as recorded along with the MCS tracking at 13 UTC on March 19, 2014. The results derived from satellite observation and WRF simulation are illustrated in (a), (c), (e), (g) and (b), (d), (f), (h), respectively.

3 Results

3.1 Number and spatial distribution of tracked MCSs

We examine the total number as well as the spatial distribution of all tracked MCSs from both observational and simulated data sets over the study domain (Figure 1). We tracked 125 and 115 observed MCSs occurring during the months of 2014 and 2015, respectively; compared to 120 and 94 MCSs tracked during the same period in the simulations. Therefore, there were slightly fewer simulated MCSs (-8%) in the 2014 period, but a notable simulated deficit of MCSs (-22%) during the 2015 period.

MCS object track density is mapped onto a 1° × 1° grid and illustrated in Figure 4. During the 2014 sampling period, more MCSs are observed near the northwest corner of the domain. The difference map for 2014 (Figure 3c) indicates that while more MCSs are simulated over the northwestern and eastern parts of the domain, fewer MCSs (negative blue patches) occur within grid cells near Amazon river. Interestingly, during the 2015 sampling period, MCSs are also under-predicted along the Amazon river, particularly over the northeastern central Amazon (Figure 3f).

While observed rain maps show relatively a consistent rainfall distribution and amount over the analysis domain (Figures 4a and 4d), simulated rainfall amount is distinctly higher in 2015 than for 2014 (Figures 4b and 4e), despite lower MCS occurrence (Figures 3c and 3f). For instance, during the 2014 period, observed domain-mean precipitation is 160.3 mm, which is slightly lower than is simulated (168.8 mm). Nevertheless, during the 2015 period, simulated and observed domain-mean precipitation are 209.5 and 171.4, respectively. Which suggests the domain-mean precipitation bias dramatically increases from +5% to +22%.

Model bias in MCS occurrence (Figures 3c and 3f) modulates the overall pattern of rainfall bias (Figures 4c and 4f), which confirms that MCS precipitation contributes to a considerably large fraction of total rainfall during these two periods. The dry bias along the Amazon river is analyzed during both years (Figures 4c and 4f) and can be attributed to relatively low MCS occurrence simulated by the model (Figures 3c and 3f) despite potentially higher rain rate (Figure 2). This implies that the current model configuration may be associated with unresolved precipitation processes related to river-atmosphere interactions such as river-breezes that enhance deep convection under easterly trade winds (Burleyson et al. 2016), among other possible factors.

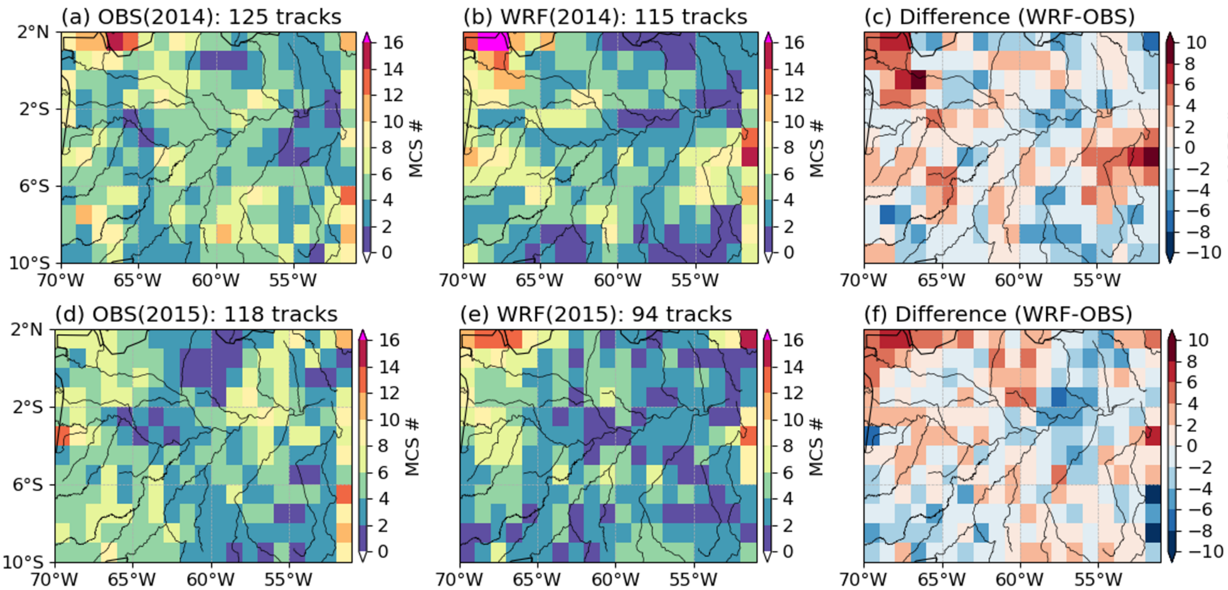

Figure 3 Spatial distribution of gridded MCS track density (i.e., number of MCS objects passing through 1°×1° grid boxes) from observations (a, d) and WRF simulations (b, e) over the study domain. (c) and (f) illustrate the difference between (a), (b) and (d), (e), respectively. The top (bottom) row represents results for the month of 2014 (2015).

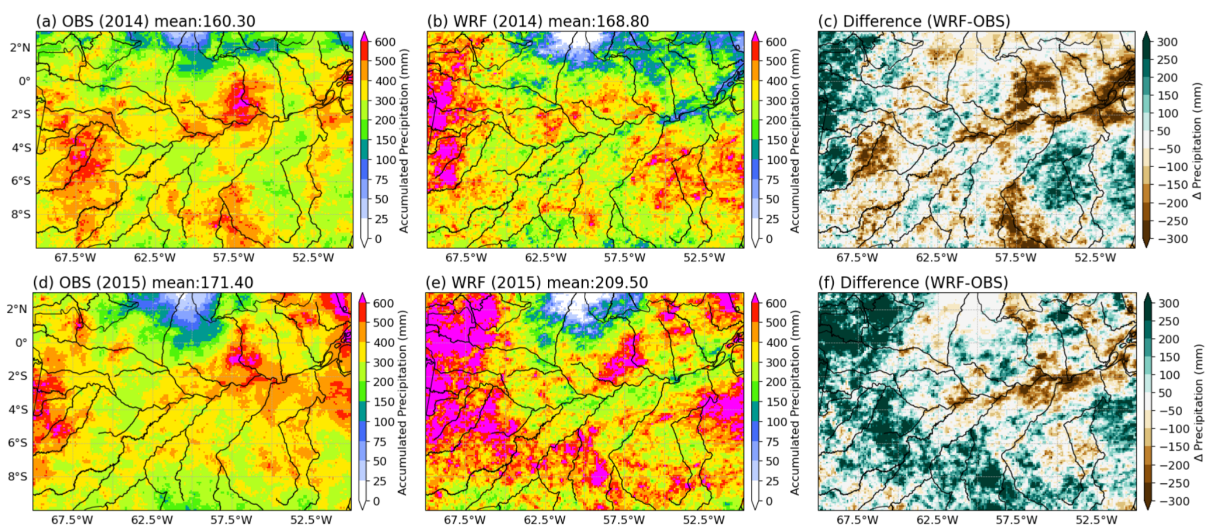

Figure 4 Similar to Figure 3, but for monthly precipitation amount (mm). The domain mean precipitation is given in the title of each panel (a, b, d, and, e)

3.2 MCS properties

The tracking data for all MCSs are used to obtain monthly mean values per MCS in both the 2014/15 periods. Besides the occurrence (number of MCSs) as discussed, the accumulated precipitation of a MCS can be attributed to its size, duration, rainfall intensity, and distance traveled, we analyze properties including MCS cloud area, rain rate, and motion (e.g., duration, speed, and movement distance). This helps elucidate how these factors contribute to overall precipitation amount and what is the fractional bias in each MCS properties. The results of selected 6 MCS properties are given in Figure 5.

We first assess the area covered by the entire MCS cloud shield (Tb < 241 K), and further categorize the cloud cover into convective and stratiform areas (Feng et al. 2021, 2023). The convective (core cloud) area is defined as the continuous area with cloud top temperature below 225 K within a much wider MCS cloudy patch. The remaining area with Tb from 225 to 241 K is attributed to stratiform (cold anvil cloud) type. The statistics show that the total extent of simulated MCS cloud area is generally smaller than observed; while observed mean MCS sizes are larger than 80,000 km^2 in both months, modeled MCSs are approximately 70,000 km^2. Observations suggest the ratios of convective/stratiform cloud cover are 1.13 and 1.05 in 2014 and 2015 period; whereas, simulated ratios increase to 1.36 and 1.43, respectively, most likely due to reduction of stratiform cloud areas.

The results also show simulated MCSs produce robustly higher rainfall intensity than is observed during each analysis period. The MCS precipitation is further partitioned into two types: heavy (≥ 10 mm h^{-1}) and light (< 10 mm h^{-1}). Distinct and consistent positive biases in modeled rain rates are seen in the heavy rainfall regime, driving the simulated total rain rate to be larger than twice of the average observed rate. While total rain rate does not vary much between the two periods (both are 9.1 mm h^{-1}), heavy rain rate substantially increases from 18.3 to 19.2 mm h^{-1}.

Finally, we examine the duration, movement speed, and distance traveled of tracked MCSs. The observations suggest that MCSs had similar durations across both analysis periods. Though simulated MCSs had slightly shorter durations than were observed during 2014 (by ~1 hour), the model significantly under-predicted the MCS lifespan (by > 2 hours) in 2015. Further, simulated MCSs motion was significantly over-predicted by ~3 – 4 m s^{-1} (~80%) during both analysis periods. Hence, the yearly differences in lifespan and propagation speed lead to larger horizontal excursions by the MCSs in both months. The simulated MCSs traveled longer distances than observed (~ +40 km in 2014 and ~ +20 km for 2015 in average).

As a result of the accumulated factors of MCS size, rain intensity, and movement, the modeled total precipitation per MCS is about 500 mm less (-20%) than what is observed in 2014 but about 700 mm more (+35%) in 2015. The much larger simulated rainfall amount in 2015 is most likely due to unrealistically higher rain intensity. The observed ratios between heavy/light precipitation are 0.4 and 0.33 for 2014 and 2015, respectively. However, it becomes nearly opposite in the simulations, as the corresponding ratios are 1.7 and 2 for the two periods. Therefore, besides having biases in total precipitation, the model also poorly represents the fractions of heavy and light precipitation.

The dramatic increase of errors in MCS occurrence (Section 3.1) as well as total precipitation from 2014 to 2015 periods catches attention because model skill in rainfall prediction usually does not significantly change from year to year. One possible source of this discrepancy in our simulations comes from biases in meterological conditions resulting from our data assimilation scheme. The quantitiy of radiosonde observations for during 2015 is approixmatley half of what was available to assimilate during 2014 (Figure 6). Thus, a much weaker constraint in simulated environmental conditions during 2015 most likely leads to enlarged biases in convective background conditions, which is expected to affect the examined MCS properties. Moreover, results of partitioned precipitation quantity (heavy vs. light) indicate the excessive MCS total precipitation in the model for the 2015 period is primarily due to a much larger fraction of heavy rainfall. The overpredicated amount of heavy rainfall reaches more than three times of the observed value. Conversely, the simulated light rainfall amount is much less than what was observed (by ~ 50%), which partially compensates the positive bias in heavy rain.

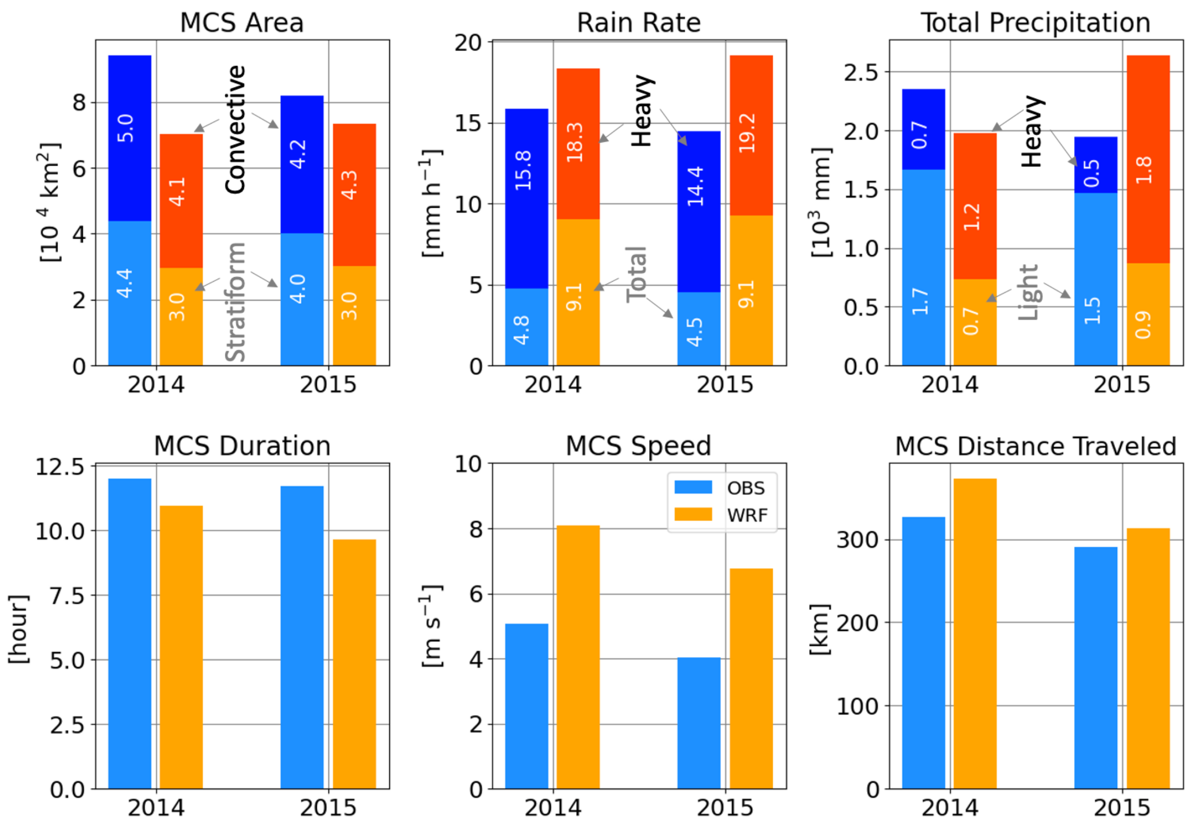

Figure 5 Comparison of monthly-mean values for MCS properties, including: MCS area, rain rate, total MCS precipitation, duration, movement speed, and distance traveled. The statistics obtained from observations and simulations are represented by blue and orange bars. In the plots for MCS area, the fractions of stratiform- and convective-type clouds are indicated by light blue (orange) and dark blue (red) bars, respectively. For rain rate comparison, the total and heavy ($\geq$ 10 mm h^{-1}) rain rates are denoted by light blue (orange) and light blue + dark blue (orange + red) bars. Total precipitation is partitioned into light (light blue and orange bars) and heavy (dark blue and red bars) rain types.

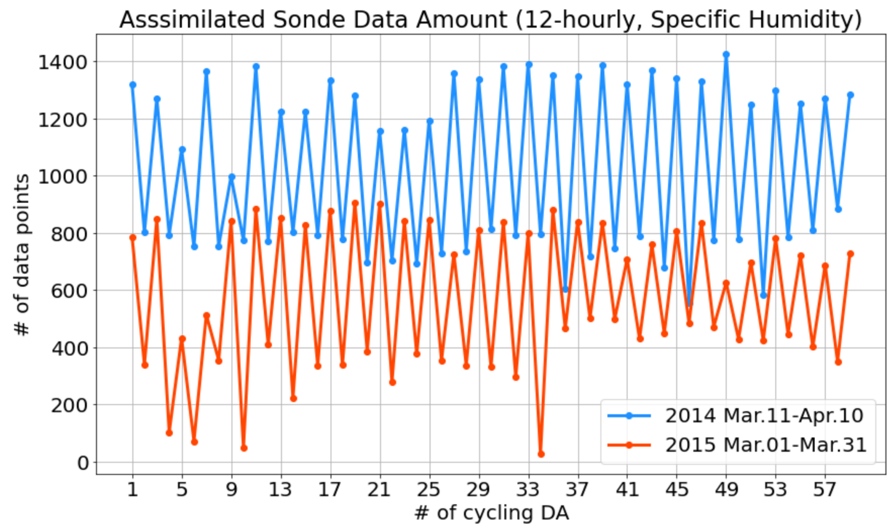

Figure 6 The number of assimilated data points from radiosonde specific humidity observations over the model domain (Figure 1). The blue and red curves represent timeseries results for 2014 and 2015 analysis periods. Y-axis denotes the number of assimilation cycles.

3.3 Daily variability

The tracked MCS properties, including the number, area, movement distance, and rain rate are further broken down to facilitate model validation on daily timescales. Approximately four MCSs were observed daily, on average, during the 2014 period, with only two days in which no MCSs were identified (Figure 7a). Although the model reproduced only slightly fewer MCSs in terms of the monthly mean value, it does not fully capture the daily variations, particularly during the first half of the month. Overall, the MCS cloud area is underpredicted by the model. Over the entire month, an average MCS size of ~9.4×10^4 km^2 is observed in satellite data; whereas, the model yields ~7×10^4 km^2. The model has more difficulty in simulating large MCSs (area > 10×10^4 km^2), such as the ones observed on 3/13, 3/22 and 4/1. While the correlation between the observed and simulated MCS movement distance is much lower than for other properties, simulated MCSs more frequently travel farther than observed ones (20 out of 31 days). Lastly, the simulated mean MCS rain rate is much higher than in observations (approximately +90%) every day.

Duringthe 2015 period (Figure 7b), simulated MCS number is notably underpredicted after 3/22 despite qualitative agreement in the trend. A noticeable contrast on the first day of simulations (3/1) may be due to model spin-up. There are no days during the 2015 period with observed MCS size larger than 15×10^4 km^2; thus, the MCS size is generally smaller than during the 2014 period. As a result, although the simulated mean MCS area is still smaller than observed, the deficit is not as large as it is during 2014. Similar to results obtained for 2014 period, there are 19 out of 31 days that MCSs traveled farther than observed in 2015. Despite variability in model biases of MCS properties, positive biases in MCS rain rate is robustly observed and are largest during the2015 period.

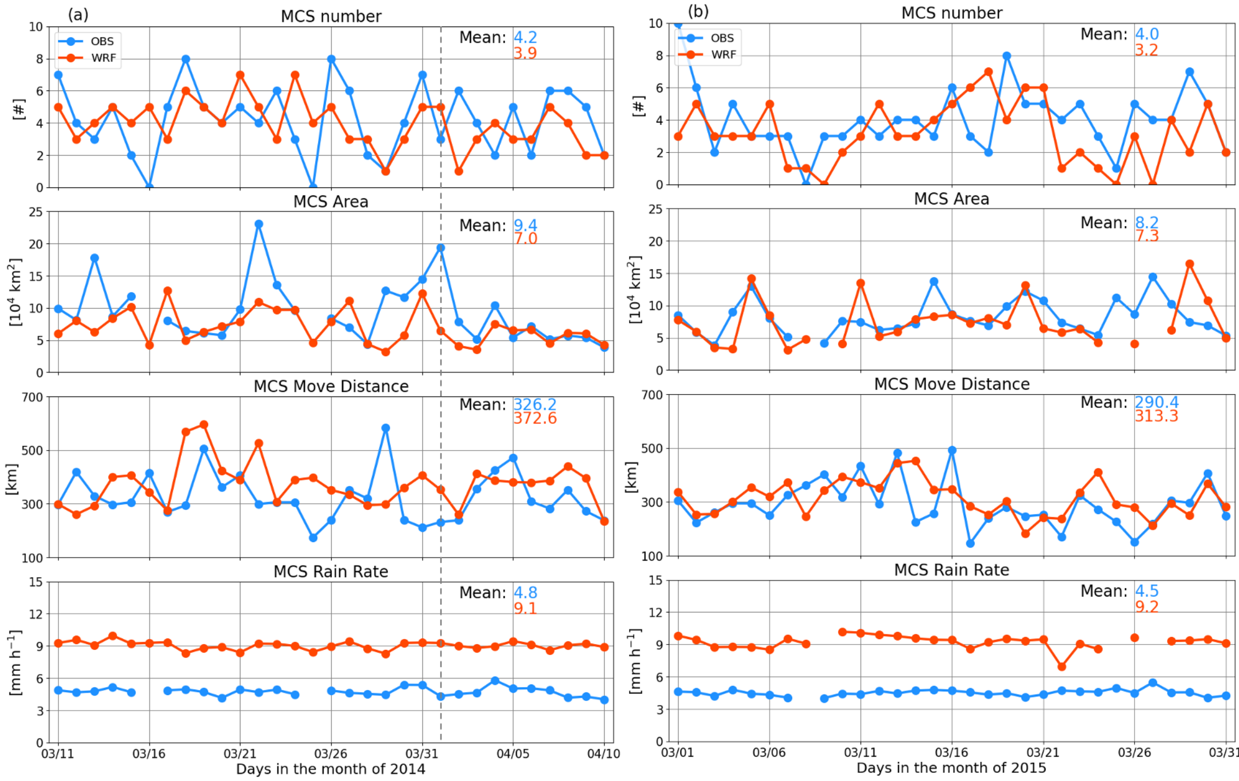

Figure 7 Daily variations of MCS tracking statistics including number, area, movement distance, and mean rain rate for the analysis period in (a) 2014 and (b) 2015. The dashed line denoted in (a) identifies the date (April 1, 2014) selected for case study (Section 3.4).

We next quantify heavy versus light rainfall-regime dependent model biases relative to biases in rain rate and cloud area (Figure 8). During most of the 2014 analysis period (Figure 8a), MCS precipitation is overpredicted on fewer than half of the days (13 out of 31) and underpredicted for the remaining 18 days. Moreover, on 3/13, 3/22, and 4/1, the negative biases in light rain are much more distinct than otherwise typical positive biases in heavy rainfall. This happens when MCS cloud areas in both convective and stratiform types are significantly under-predicted. During the 2015 analysis period, there are only two days such as 3/4 and 3/15 when negative biases of light rainfall are lower than -2000 mm and thus compensate or even lead to a negative total MCS precipitation bias. Other than those days, heavy rainfall bias dominates the total precipitation bias. Rain rate biases in heavy precipitation are noticeably higher in 2015 than 2014. Given relative minor model-observation differences in MCS area, excessive simulated rain rate is responsible for large positive biases in total MCS precipitation for 2015 analysis period as shown in Figure 6. With varied scenarios as observed during the two sampling periods, it suggests model validation using only accumulated precipitation amount will most likely ignore crucial model errors that may be revealed from examination of many MCS properties.

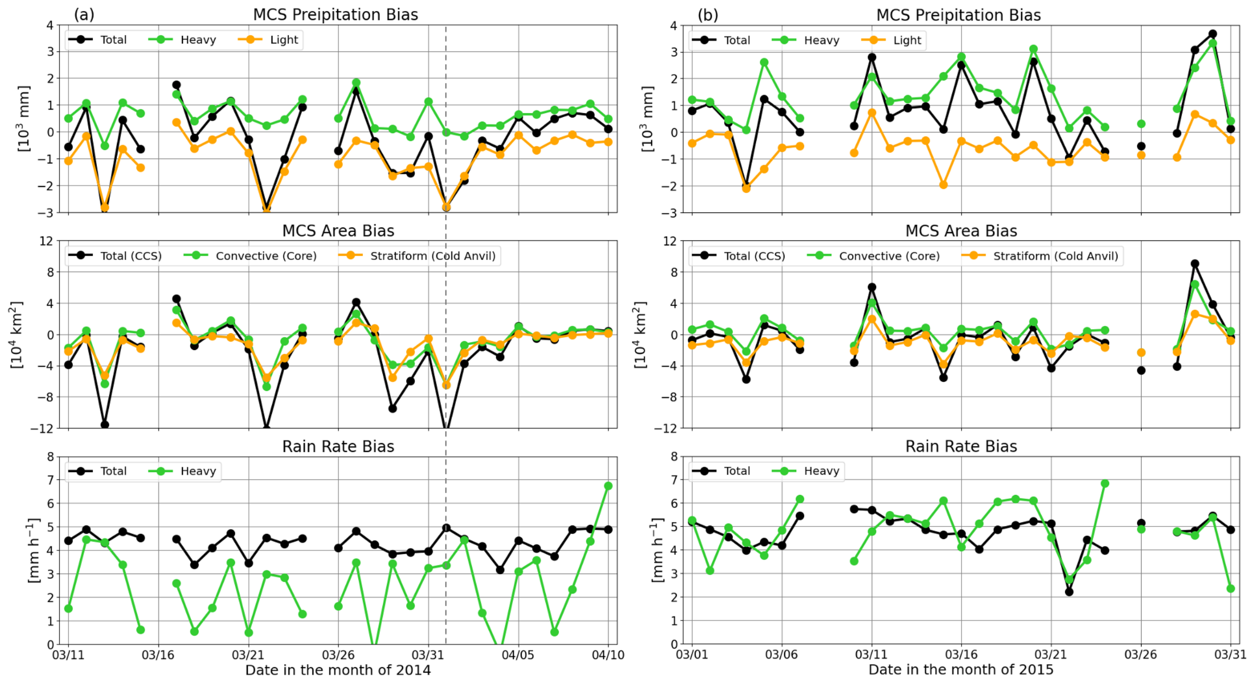

Figure 8 Similar to Figure 7 but for biases in MCS precipitation amount, rain rate, and cloud area. The black lines denote results for the total biases. Green lines illustrate fractional results for the heavy rainfall (≥ 10 mm h^{-1}) and convective-type (core) clouds. Orange lines represent results for the part with light rain (< 10 mm h^{-1}) and stratiform-type (cold anvil) clouds. The dashed line denoted in (a) identifies the date (April 1, 2014) selected for case study (Section 3.4).

3.4 Case study: April 1, 2014

Following the discussion in the previous section, we further investigate the contrasts between observed and simulated MCS clouds and precipitation by highlighting a case study for April 1, 2014. On this day, a MCS (denoted by white arrows at first row in Figure 9) initiated over northeastern corner of the analysis domain and experienced upscale growth through decay as it propagated westward (Figure 9). In general, the model reasonably reproduces the MCS's lifecycle with relatively accurate placement of clouds in time, allowing us to examine the evolution of clouds (Figure 9) and corresponding precipitation processes (Figure 10) with confidence. In this case, the model tends to simulate smaller cloud cover regardless of cloud types. The simulated stratiform clouds dissipated faster than observed after 12 UTC, particularly on the southeastern flank of the MCS, as the brightness temperature of cloud top significantly increases. Nevertheless, more isolated convective initiation remains active on the leading edge of the propagating MCS. Therefore, only a relatively narrow core cloud band is sustained in simulations. In reality, satellite observations suggest a much wider MCS cloud patch propagating toward the southwest through the domain.

The contrast between observed and simulated heavy/light rain distribution is shown in Figure 10. Though observations suggest heavy rainfall patches are scattered and mostly located in the center of a much wider light rain area, simulated heavy rainfall patches tend to have a much larger fractional area than observed and appear on the leading edge of relatively narrow cloud bands. This structural difference in the precipitation features implies that, most likely there are model deficiencies in representing the dynamics and/or microphysics within the simulated MCS causing the issue and warrants in-depth investigation in a future study to provide further insight.

The diurnal variations of cloud and precipitation type relative area fractions are shown in Figure 11. The evolution of simulated core cloud fraction is overall aligned with the observed trend until 12 UTC. Similar fractions (~20 – 30%) are obtained for cold anvil clouds. While the observed cold anvil cloud fraction is consistently around 30%, simulated cold anvil cloud fraction dropped to ~20% at the end of the day. Cold anvil clouds started dissipating at 16 UTC, which is about 3–4 hours later than the dissipation of core clouds.

The area fraction of heavy rain is consistent between observations and simulations and does not exceed 6% over the course of the day. However, we see a notable contrast in the fraction of light rain area. Overall, the observed light rain area covers more than twice of what it does in simulations. Hourly rain rate comparison indicates light rain rates in both observations and simulations are never greater than 5 mm hr^{-1}; whereas simulated heavy rain rate is mostly near 20 mm hr^{-1}, approximately 4 mm hr^{-1} more than observed, on average. Given much larger negative bias in anvil cloud cover and relatively small heavy rain rate bias during this event, the total MCS precipitation is thus primarily driven by the significant under-prediction of stratiform cloud cover. However, it is worth noting that validations of IMERG data against ground based observations (either radar or rain gauges) reveal that IMERG tends to significantly overestimate the frequency of weak precipitation (1-2 mm h^{-1}) while underestimating intense precipitation, particularly over land (Cui et al. 2020; Zhang et al. 2021; Ayat et al. 2021). Moreover, the actual resolution of IMERG is significantly coarser than its grid spacing (Guilloteau & Foufoula-Georgiou, 2020). Thus, associated rainrate model biases themselves may be overestimated.

To elucidate how environment conditions may influence differences in this MCS's lifecycle, we examine the pre-storm environment (before 12 UTC) as observed by radiosonde profiles at the AMF T3 site and simulated profiles within a 2° by 2° box centered at the AMF site (Figure 1). Vertical interpolation with an interval of 0.1 km was carried out for both radiosonde and model profiles. To exclude profiles affected by convective clouds, only the model profiles with column maximum reflectivity less than 0 dBZ are sampled. Resulting wind profiles valid at 00, 06, and 12 UTC are illustrated in Figure 12.

There are two jets evident in the observed wind speed profiles; one peaking at z = 2–3 km, and another above mid-troposphere (z = 7–9 km) (Figure 12a). Owing to small vertical heterogeneity in the meridional wind, these jets are primarily a result of variations in the zonal wind (Figure 12 b-c). Wind conditions do not vary significantly in the pre-storm environment. Although the model simulated wind conditions over the period are qualitatively similar to observations, simulated maximum wind speed of the lower jet is consistently smaller (~2–3 m s^{-1}

less than radiosonde). Furthermore, at 12 UTC, the lower jet descends to height below 1 km in the model. This may imply a shallower and weaker low-level jet being simulated, leading to much weaker moisture transport and convergence in the lower troposphere. Results of convective available potential energy (CAPE) and convective inhibition (CIN) further demonstrate while the available energy for convective growth increases from 00 to 12 UTC in reality, the model simulates a completely opposite trend (Figure 13), where CAPE dropped from nearly 2000 to 500 J kg^{-1} over the 12-hour period. Despite good agreement in CIN values, the simulated environment does not favor convective growth as observed. This evidence may at least partially explain why the simulated MCS quickly dissipates after 12 UTC and thus has a much smaller area cover by stratiform clouds.

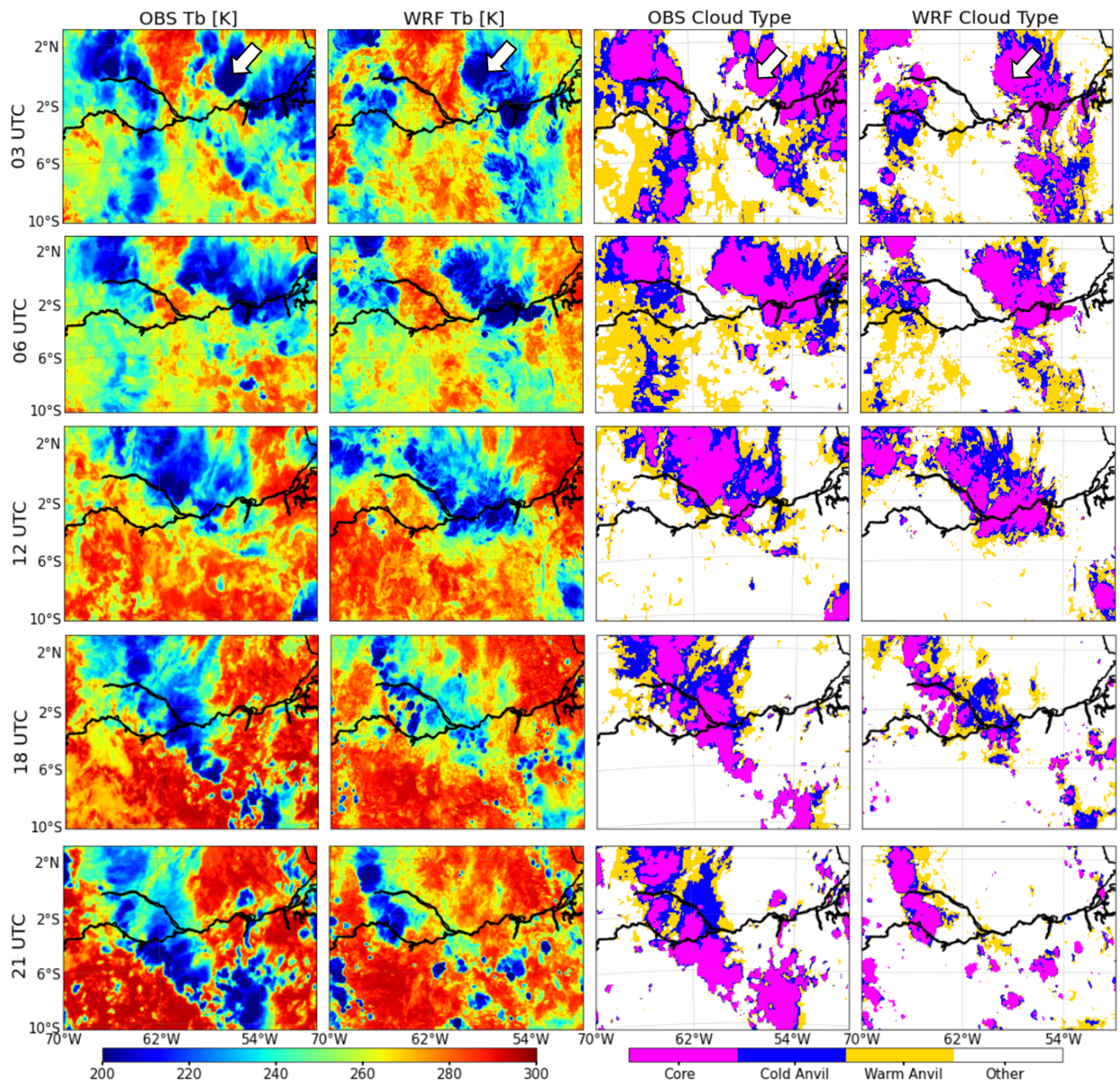

Figure 9 Similar to Figure 2a – 2d, but for 03 to 21 UTC on April 1, 2014. White arrows on the panels for 03 UTC point to the initiating MCS of interest.

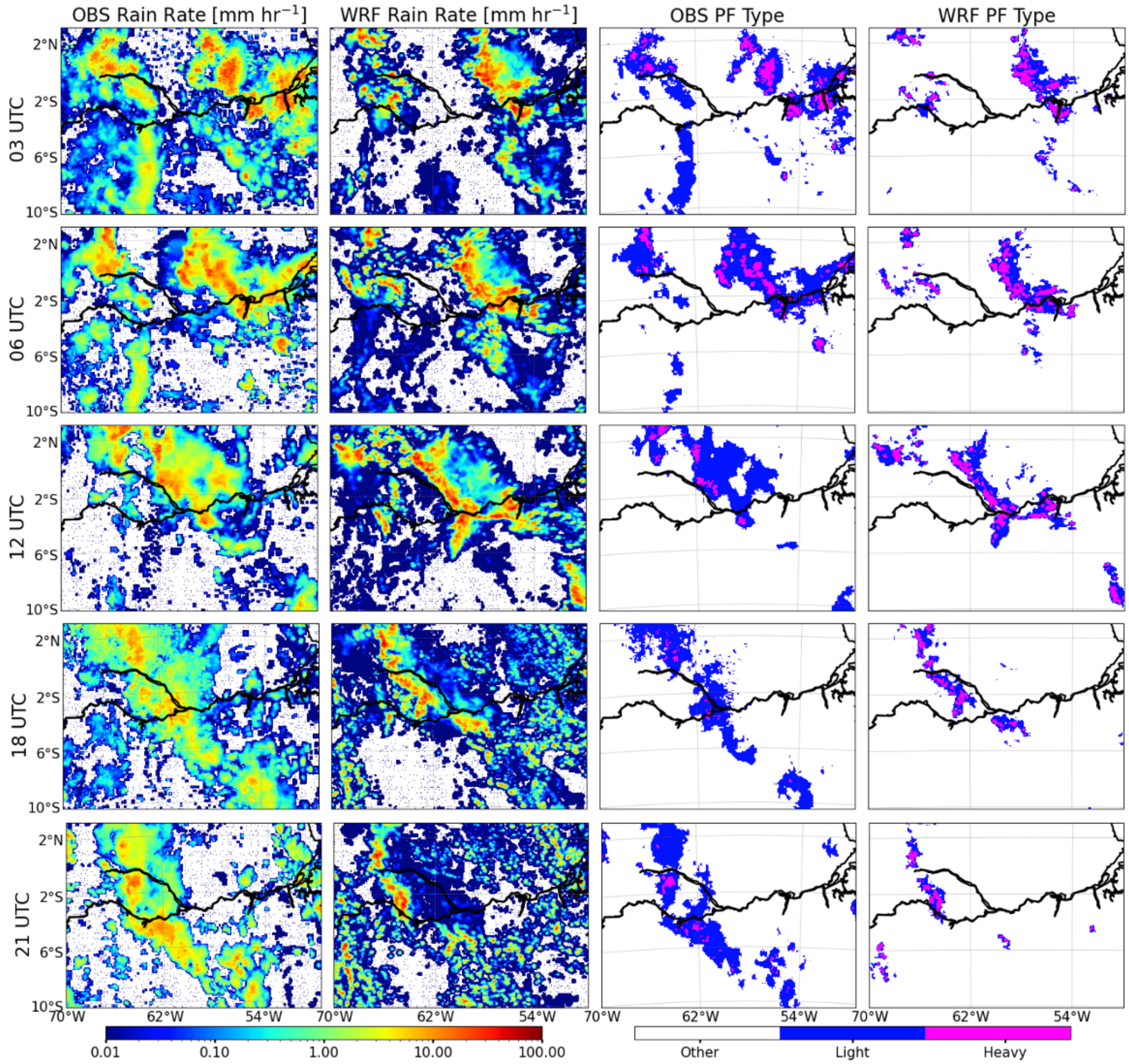

Figure 10 Similar to Figure 2e – 2h, but for 03 to 21 UTC on April 1, 2014.

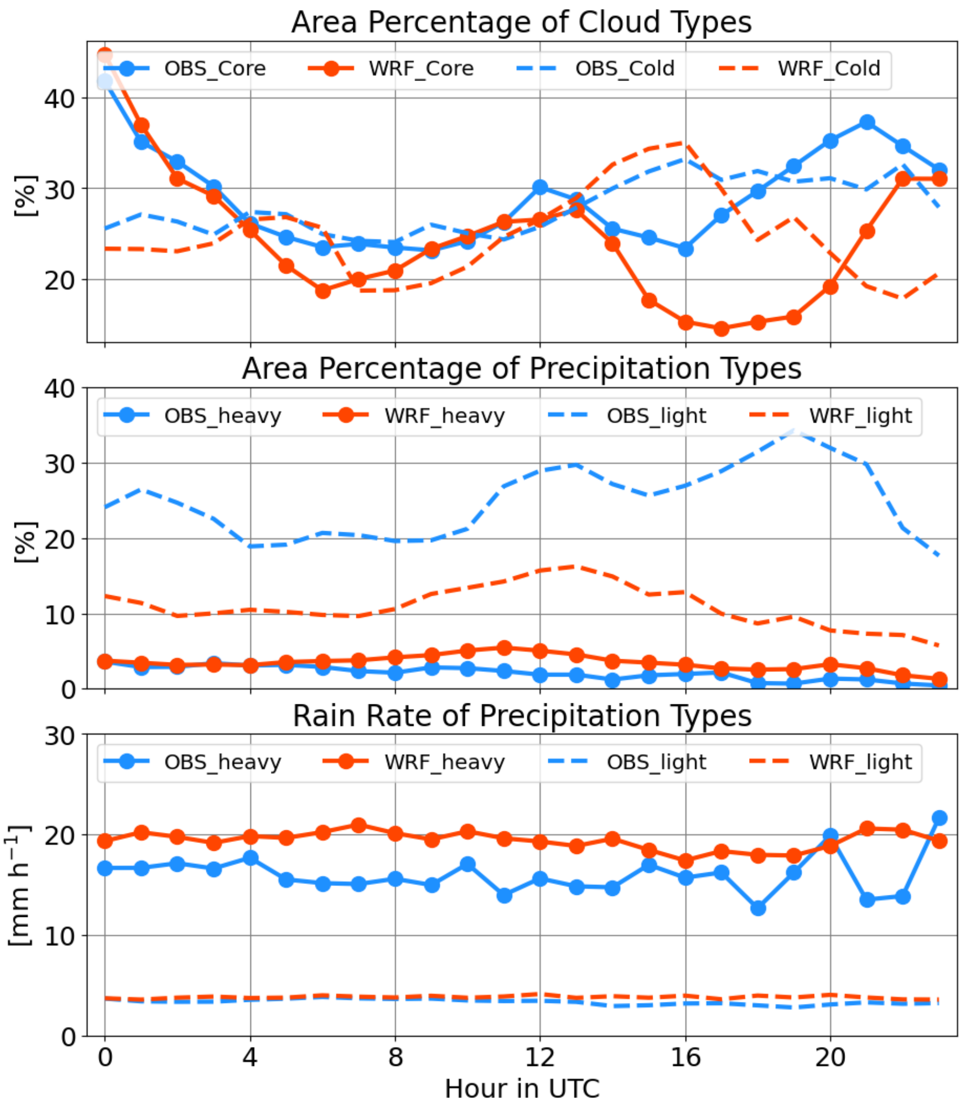

Figure 11 Comparison in diurnal varations of area percentages of cloud (top panel) and precipitation (middle panel) types as computed over the analysis domain on April 1, 2014. Correspondng rain rate comparison in dependency of precipitation type is displayed in the bottom panel.

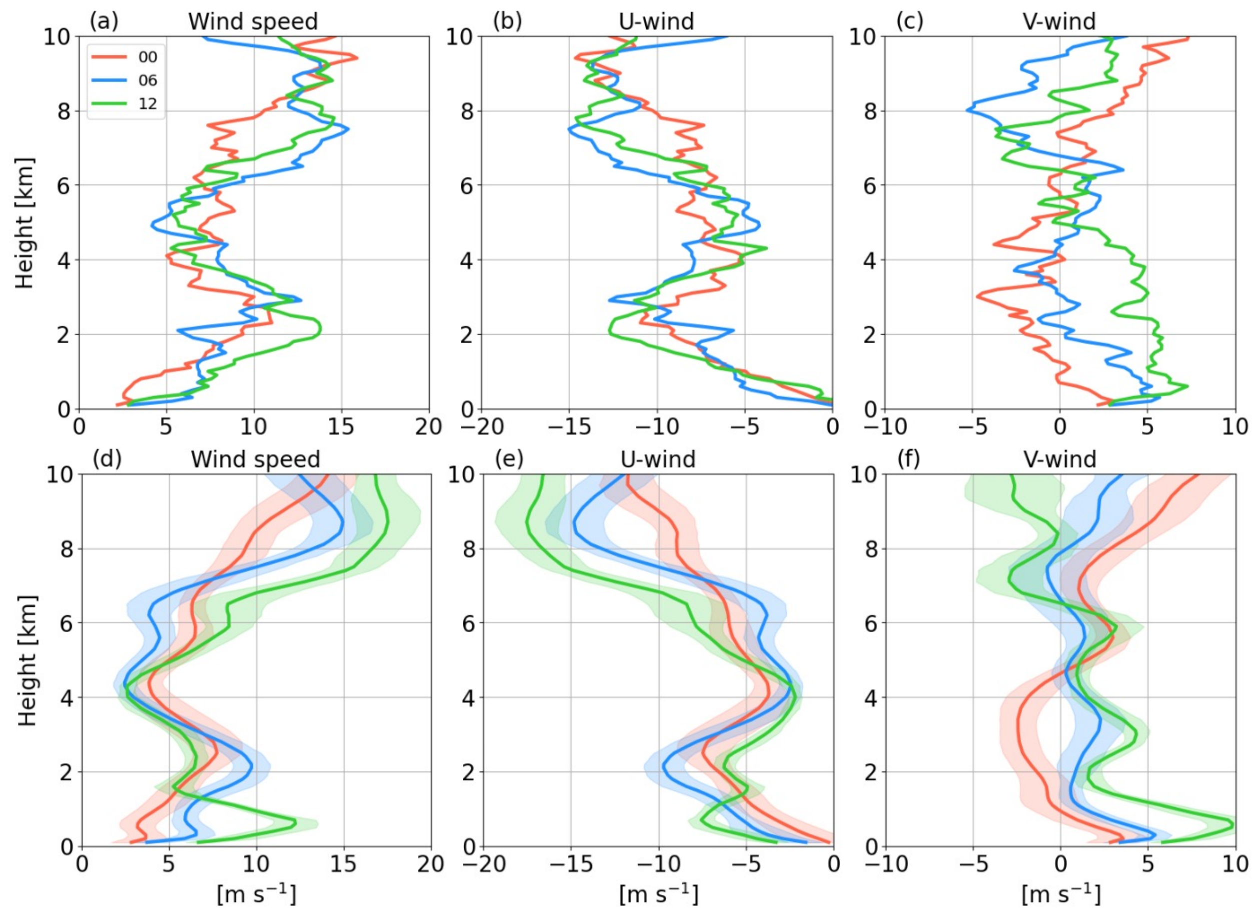

Figure 12 Wind speed, zonal (U-) wind, and meridional (V-) wind profiles as observed by radiosondes launched (a, b, and c) and simulated by WRF (d, e, and f) at the location of ARM T3 site. Colors denoted in legend indicate results for 00, 06, and 12 UTC of April 1, 2014. Swath of each line in d, e, and f represents the range within $\pm$ 1 standard deviation among the samples.

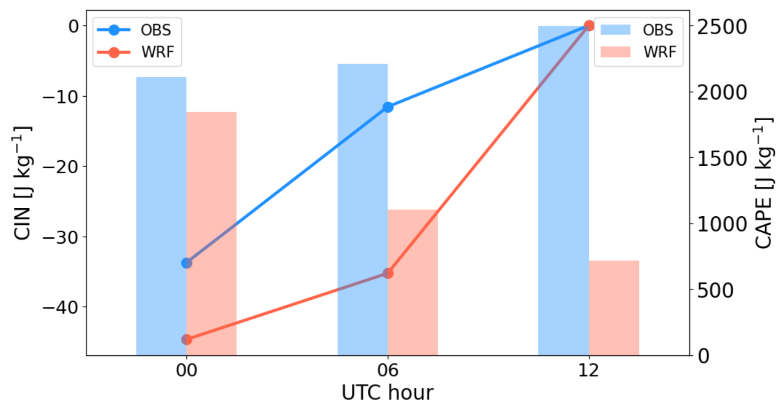

Figure 13 CIN (line with dots; left y-axis) and CAPE (bar; righy y-axis) values computed using ARM T3 radiosonde and corresponding WRF-simulated profiles at 00, 06, and 12 UTC of April 1, 2014.

4 Summary and conclusions

Mesoscale convective systems (MCSs) are responsible for a large fraction of the total precipitation in the Amazon. However, various uncertainties in state-of-the-art atmospheric models hinder them from reproducing a realistic lifecycle and morphology of MCSs within this region. To facilitate comprehensive characterization of MCS precipitation in the central Amazon, the Python FLEXible object TRacKeR (PyFLEXTRKR) is employed to track individual MCSs that are simulated by convection permitting (4 km grid) Weather Research and Forecasting (WRF) simulations. The WRF simulations are performed over two separate months during the 2014 and 2015 Amazon wet seasons. A 3DVar data assimilation scheme is used to constrain environmental conditions throughout simulations. These results are then compared to observed satellite analogs to examine possible mechanisms of MCS model biases.

First, we examined the MCS occurrence and its relation to accumulated precipitation. Overall, the model tended to produce fewer MCSs than were observed within the study area. While only 8% fewer MCSs are reproduced in 2014 period, we observe a difference of -22% for the month in 2015. The heterogenous precipitation bias distribution is closely tied with how well the MCS track density was reproduced. A distinct feature of dry biases along the Amazon river is identified and found to be well explained by model error in reproducing realistic MCS occurrence near the river.

Analysis of monthly means of tracked MCS characteristics further reveal the contrasts between observed and modeled MCS properties in general. Although simulated MCSs are generally smaller than observed ones, they produced far more rain and propagated farther than observed. Moreover, we find the model-observation discrepencies in various MCS properties must be considered when accounting for the sources of MCS total precipitation bias. For example, in 2014 period, MCS total precipitation is underestimated by the model due in part to relatively large negative bias in MCS size and minor positive rain rate bias. Whereas in 2015, while model bias in MCS size is relatively small, substantial positive bias in rain rate results in severe overpredication of MCS total precipitation. Aside from biases in total precipitation, we also show that the model has difficulty in reproducing realistic fraction of heavy/light precipitation.

The model errors in MCS number, rain rate, and precipitation (MCS and domain-mean) notably increase from 2014 to 2015 (Figures 3, 4, and 5). We find the degraded model performance is most likely driven by the availability of observational data for assimilation. The amount of assimilated radiosonde moisture data in 2015 dropped to only half of what was available in 2014 despite consistent assimilation of radiosonde profiles collected at the ARM T3 site in central Amazon. This reinforces the importance of additional observation sites to constrain simulated synoptic environments over the continent.

We further break down the statistics by each day and demonstrate that model skill in reproducing MCS properties, including number, size, and distance traveled, could vary significantly from day to day. On many days, the bias in total precipitation can be attributed to the wet bias in heavy rainfall, which result in overall overpredicted precipitation. However, light rainfall may occasionally drive the total precipitation error. Such events happened when both convective and stratiform cloud cover are under-predicted, hence both contribute to dry biases in precipitation. This suggests that it is critical to validate simulated precipitation by considering its dependency per rainfall regime because the biases sourced in different regimes may imply unrealistic model representations of various dynamical and/or microphysics processes. Analysis of daily bias provides more details in terms of model biases in MCS characteristics.

Finally, an analysis of an MCS on April 1, 2014 is provided to illustrate how differently the clouds and precipitation are resolved in both observational and model data. We showed that in this particular event, while relatively small wet bias in heavy rainfall is analyzed, the large dry bias in light rain controls the total precipitation bias. This is mainly caused by significant under-prediction in area cover of light rain. Examination of the pre-storm environment suggests the jet in lower troposphere is relatively shallow and weak in the simulations compared to observations. This could lead to insufficient moisture transport and hence weaker convergence that are essential for convective growth and sustainability. Moreover, weaker simulated CAPE also indicated unfavorable conditions for convective growth. Given the evidence, we conclude the environmental conditions may be causing the early dissipation of MCSs and significant negative bias in stratiform cloud cover.

In addition to environmental conditions as discussed in Section 3.4, potential sources of model uncertainties in reproducing observed MCS clouds/precipitation may also relate to 1) model resolution, which directly influences how MCS's dynamic structure (e.g. vertical motion) may be resolved and thus alters the secondary circulation accordingly (Varble et al. 2020); and 2) paramterization of microphysical processes. For instance, the magnitude of simulated stratiform precipitation is found to be associated with ice particle mass fluxes as predicted by the employed microphysics schemes (Han et al. 2019). Heating profiles could be changed drastically by replacing one microphysics scheme by another (Feng et al. 2018).

Compared to a mesoscale model, climate models tend to simulate even more unrealistic representations of tropical precipitation features (e.g., Tai et al. 2021), due in part to coarse grid spacing and much more simplified physics parameterizations. Given the substantial increase in computational power, climate models are now more frequently run at cloud-resolving scales (e.g., Tang et al. 2021; Liu et al. 2023). Despite promise as seen in selected case studies (Liu et al. 2023), a high-resolution configuration does not always lead to distinct improvements in general precipitation features (e.g., diurnal cycle) and associated meteorological conditions. We note climate models should use mesoscale model (e.g., WRF) simulations as benchmarks when assessing their performance. In this way, the behaviors of state-of-the-art climate models can be constrained by both success and failure of relatively well-developed mesoscale models.

Acknowledgments

This research was supported by the Atmospheric Science Research (ASR) program as part of the DOE Office of Biological and Environmental Research. Pacific Northwest National Laboratory is operated by DOE by the Battelle Memorial Institute under contract DE-A06-76RLO 1830. The modeling computation was performed on PNNL's Research Computing clusters.

Data Availability Statement

The WRF community model is available from the National Ceter for Atmospheric Research (NCAR) at http://www2.mmm.ucar.edu/wrf/users/. The NCEP FNL operational model Global Tropospheric Analysis is accessible at https://rda.ucar.edu/datasets/ds083.2/. The NCEP ADP global upper air and surface weather obserbations (prepbufr format) can be downloaded from https://rda.ucar.edu/datasets/ds337.0/. The dataset of National Aeronautics and Space Administration (NASA) Global Merged IR V1 infrared brightness temperature can be accessed at https://disc.gsfc.nasa.gov/datasets/GPM_MERGIR_1/summary/. The IMERG Final-Run V06B precipitation products used in this study were acquired from http://pmm.nasa.gov/data-access/downloads/gpm/. The GoAmazon2014/15 data used in this manuscript are freely available from the ARM data archive (https://www.arm.gov/data). The WRF model outputs generated by the simulations in this study are saved on a long-term storage system at PNNL (rc-support@pnnl.gov)

References

Nesbitt, S. W., R. Cifelli, and S. A. Rutledge, 2006: Storm Morphology and Rainfall Characteristics of TRMM Precipitation Features. *Mon. Wea. Rev.*, **134**, 2702–2721, https://doi.org/10.1175/MWR3200.1.

Feng, Z., Leung, L. R., Liu, N., Wang, J., Houze, R. A., Li, J., et al. (2021). A global high-resolution mesoscale convective system database using satellite-derived cloud tops, surface precipitation, and tracking. *Journal of Geophysical Research: Atmospheres*, 126, e2020JD034202. https://doi.org/10.1029/2020JD034202

Sousa, A. C., L. A. Candido, and P. Satyamurty, 2021: Convective Cloud Clusters and Squall Lines along the Coastal Amazon. *Mon. Wea. Rev.*, **149**, 3589–3608, https://doi.org/10.1175/MWR-D-21-0045.1.

Schumacher, R.S., Rasmussen, K.L. The formation, character and changing nature of mesoscale convective systems. *Nat Rev Earth Environ* **1**, 300–314 (2020). https://doi.org/10.1038/s43017-020-0057-7

Feng, Z., Dong, X., Xi, B., Schumacher, C., Minnis, P., and Khaiyer, M. (2011), Top-of-atmosphere radiation budget of convective core/stratiform rain and anvil clouds from deep convective systems, *J. Geophys. Res.*, 116, D23202, doi:10.1029/2011JD016451.

Laurent, H., Machado, L. A. T., Morales, C. A., and Durieux, L., Characteristics of the Amazonian mesoscale convective systems observed from satellite and radar during the WETAMC/LBA experiment, *J. Geophys. Res.*, 107(D20), 8054, doi:10.1029/2001JD000337, 2002.

Petersen, W. A., S. W. Nesbitt, R. J. Blakeslee, R. Cifelli, P. Hein, and S. A. Rutledge, 2002: TRMM Observations of Intraseasonal Variability in Convective Regimes over the Amazon. *J. Climate*, **15**, 1278–1294, https://doi.org/10.1175/1520-0442(2002)015<1278:TOOIVI>2.0.CO;2.

Cifelli, R., Petersen, W. A., Carey, L. D., Rutledge, S. A., and da Silva Dias, M. A. F., Radar observations of the kinematic, microphysical, and precipitation characteristics of two MCSs in TRMM LBA, J. Geophys. Res., 107(D20), 8077, doi: ref="info:doi/doi:10.1029/2000JD000264">doi:10.1029/2000JD000264 2002.

Da Silva, N. A., & Haerter, J. O. (2023). The precipitation characteristics of mesoscale convective systems over Europe. Journal of Geophysical Research: Atmospheres, 128, e2023JD039045. https://doi.org/10.1029/2023JD039045.

Rickenbach, T. M., Ferreira, R. N., Halverson, J. B., Herdies, D. L., and Silva Dias, M. A. F., Modulation of convection in the southwestern Amazon basin by extratropical stationary fronts, J. Geophys. Res., 107(D20), doi:10.1029/2000JD000263, 2002.

Machado, L. A. T., Morales, C., & Avissar, R. (2004). The Large-Scale Biosphere-Atmosphere Experiment in Amazonia (LBA): Insights and future research needs. Journal of Geophysical Research: Atmospheres, 109(D20).

Burleyson, C. D., Z. Feng, S. M. Hagos, J. Fast, L. A. T. Machado, and S. T. Martin, 2016: Spatial Variability of the Background Diurnal Cycle of Deep Convection around the GoAmazon2014/5 Field Campaign Sites. *J. Appl. Meteor. Climatol.*, **55**, 1579–1598, https://doi.org/10.1175/JAMC-D-15-0229.1.

Rehbein, A, Ambrizzi, T, Mechoso, CR, Espinosa, SAI, Myers, TA. Mesoscale convective systems over the Amazon basin: The GoAmazon2014/5 program. *Int J Climatol*. 2019; 39: 5599–5618. https://doi.org/10.1002/joc.6173

Giangrande, S. E., Wang, D., and Mechem, D. B.: Cloud regimes over the Amazon Basin: perspectives from the GoAmazon2014/5 campaign, Atmos. Chem. Phys., 20, 7489–7507, https://doi.org/10.5194/acp-20-7489-2020, 2020.

Wang, D., Giangrande, S. E., Feng, Z., Hardin, J. C., & Prein, A. F. (2020). Updraft and downdraft core size and intensity as revealed by radar wind profilers: MCS observations and idealized model comparisons. *Journal of Geophysical Research: Atmospheres*, 125, e2019JD031774. https://doi.org/10.1029/2019JD031774

Tian, Y., Zhang, Y., Klein, S. A., & Schumacher, C. (2021). Interpreting the diurnal cycle of clouds and precipitation in the ARM GoAmazon observations: Shallow to deep convection transition. *Journal of Geophysical Research: Atmospheres*, 126, e2020JD033766. https://doi.org/10.1029/2020JD033766

Anselmo, E. M., C. Schumacher, and L. A. T. Machado, 2020: The Amazonian Low-Level Jet and Its Connection to Convective Cloud Propagation and Evolution. *Mon. Wea. Rev.*, **148**, 4083–4099, https://doi.org/10.1175/MWR-D-19-0414.1.

Carvalho, L. M. V., Silva Dias, M. A. F., & Fitzjarrald, D. R. (2002). Atmospheric conditions for deep convective clouds in the Amazon. Journal of Geophysical Research: Atmospheres, 107(D20).

Machado, L. A. T., Laurent, H., Gherboudj, I., Lima, A. P., Morales, C., Blakeslee, R., ... & Cifelli, R. (2004). The convective, turbulent, and microphysical characteristics of the Amazon boundary layer. Monthly Weather Review, 132(11), 2919-2931.

Prein, A. F., Ge, M., Valle, A. R., Wang, D., & Giangrande, S. E. (2022). Towards a unified setup to simulate mid-latitude and tropical mesoscale convective systems at kilometer-scales. Earth and Space Science, 9, e2022EA002295. https://doi.org/10.1029/2022EA002295

Paccini, L., & Stevens, B. (2023). Assessing precipitation over the Amazon basin as simulated by a storm-resolving model. *Journal of Geophysical Research: Atmospheres*, 128, e2022JD037436. https://doi.org/10.1029/2022JD037436

Silva Dias, M. A. F., et al., A case study of convective organization into precipitating lines in the Southwest Amazon during the WETAMC and TRMM-LBA, *J. Geophys. Res.*, 107(D20), 8078, doi:10.1029/2001JD000375, 2002.

Tai, S.-L., Feng, Z., Ma, P.-L., Schumacher, C., & Fast, J. D. (2021). Representations of precipitation diurnal cycle in the amazon as simulated by observationally constrained cloud-system resolving and global climate models. Journal of Advances in Modeling Earth Systems, 13, e2021MS002586. https://doi.org/10.1029/2021MS002586

Feng, Z., Leung, L. R., Houze, R. A., Hagos, S., Hardin, J., Yang, Q., et al. (2018). Structure and evolution of mesoscale convective systems: Sensitivity to cloud microphysics in convection-permitting simulations over the United States. *Journal of Advances in Modeling Earth Systems*, 10, 1470–1494. https://doi.org/10.1029/2018MS001305

Prein, A. F., Ge, M., Valle, A. R., Wang, D., & Giangrande, S. E. (2022). Towards a unified setup to simulate mid-latitude and tropical mesoscale convective systems at kilometer-scales. *Earth and Space Science*, 9, e2022EA002295. https://doi.org/10.1029/2022EA002295

Prein A. F., Rasmussen R. M., Wang D. and Giangrande S. E. (2021). Sensitivity of organized convective storms to model grid spacing in current and future climates. Phil. Trans. R. Soc. A. 379:20190546.

Ramos-Valle, A. N., Prein, A. F., Ge, M., Wang, D., & Giangrande, S. E. (2023). Grid spacing sensitivities of simulated mid-latitude and tropical mesoscale convective systems in the convective gray zone. *Journal of Geophysical Research: Atmospheres*, 128, e2022JD037043. https://doi.org/10.1029/2022JD037043

Yang, Z., Varble, A., Berg, L. K., Qian, Y., Tai, S.-L., Chen, J., et al. (2023). Sensitivity of precipitation displacement of a simulated MCS to changes in land Surface conditions. *Journal of Geophysical Research: Atmospheres*, 128, e2022JD037642. https://doi.org/10.1029/2022JD037642

Laurent, H., Machado, L. A. T., Morales, C. A., and Durieux, L., Characteristics of the Amazonian mesoscale convective systems observed from satellite and radar during the WETAMC/LBA experiment, *J. Geophys. Res.*, 107(D20), 8054, doi:10.1029/2001JD000337, 2002.

Anselmo, EM, Machado, LAT, Schumacher, C, Kiladis, GN. Amazonian mesoscale convective systems: Life cycle and propagation characteristics. Int J Climatol. 2021; 41: 3968–3981. https://doi.org/10.1002/joc.7053

Huang, X., Hu, C., Huang, X. *et al.* A long-term tropical mesoscale convective systems dataset based on a novel objective automatic tracking algorithm. *Clim Dyn* **51**, 3145–3159 (2018). https://doi.org/10.1007/s00382-018-4071-0

Feng Z., L. Leung, N. Liu, J. Wang, R.A. Houze, J. Li, and J.C. Hardin, et al. 2021. "A Global High-resolution Mesoscale Convective System Database using Satellite-derived Cloud Tops, Surface Precipitation, and Tracking." *Journal of Geophysical Research: Atmospheres* 126, no. 8:e2020JD034202. PNNL-SA-157800. doi:10.1029/2020JD034202

Feng Z., J.C. Hardin, H.C. Barnes, J. Li, L. Leung, A.C. Varble, and Z. Zhang. 2023. "PyFLEXTRKR: a Flexible Feature Tracking Python Software for Convective Cloud Analysis." *Geoscientific Model Development* 16, no. 10:2753-2776. PNNL-SA-179219. doi:10.5194/gmd-16-2753-2023

Rehbein, A., Ambrizzi, T. and Mechoso, C.R. (2018), Mesoscale convective systems over the Amazon basin. Part I: climatological aspects. Int. J. Climatol, 38: 215-229. https://doi.org/10.1002/joc.5171

Galarneau, T. J., Zeng, X., Dixon, R. D., Ouyed, A., Su, H., & Cui, W. (2023). Tropical mesoscale convective system formation environments. *Atmospheric Science Letters*, 24(5), e1152. https://doi.org/10.1002/asl.1152

Prein, A.F., Liu, C., Ikeda, K. *et al.* Simulating North American mesoscale convective systems with a convection-permitting climate model. *Clim Dyn* **55**, 95–110 (2020). https://doi.org/10.1007/s00382-017-3993-2

Emanuel, K., 2023: On the Physics of High CAPE. *J. Atmos. Sci.*, **80**, 2669–2683, https://doi.org/10.1175/JAS-D-23-0060.1.

Martin, S. T., Artaxo, P., Machado, L. A. T., Manzi, A. O., Souza, R. A. F., Schumacher, C., Wang, J., Andreae, M. O., Barbosa, H. M. J., Fan, J., Fisch, G., Goldstein, A. H., Guenther, A., Jimenez, J. L., Pöschl, U., Silva Dias, M. A., Smith, J. N., and Wendisch, M.: Introduction: Observations and Modeling of the Green Ocean Amazon (GoAmazon2014/5), Atmos. Chem. Phys., 16, 4785–4797, https://doi.org/10.5194/acp-16-4785-2016, 2016.

Na, Y., Fu, Q., Leung, L. R., Kodama, C., & Lu, R. (2022). Mesoscale convective systems simulated by a high-resolution global nonhydrostatic model over the United States and China. *Journal of Geophysical Research: Atmospheres*, 127, e2021JD035916. https://doi.org/10.1029/2021JD035916

Cheng, W. Y. Y., and W. R. Cotton, 2004: Sensitivity of a Cloud-Resolving Simulation of the Genesis of a Mesoscale Convective System to Horizontal Heterogeneities in Soil Moisture Initialization. *J. Hydrometeor.*, **5**, 934–958, https://doi.org/10.1175/1525-7541(2004)005<0934:SOACSO>2.0.CO;2.

National Centers for Environmental Prediction/National Weather Service/NOAA/U.S. Department of Commerce. (2000). NCEP FNL Operational Model Global Tropospheric Analyses, continuing from July 1999 (Updated daily) [Dataset]. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. https://doi.org/10.5065/D6M043C6. Accessed† 01 Apr 2020.

Andreas Franz Prein, Zhe Feng, Thomas fiolleau, et al. Km-Scale Simulations of Mesoscale Convective Systems (MCSs) Over South America - A Feature Tracker Intercomparison. *ESS Open Archive* . October 27, 2023. DOI: 10.22541/essoar.169841723.36785590/v1

Da Silva, N. A., & Haerter, J. O. (2023). The precipitation characteristics of mesoscale convective systems over Europe. *Journal of Geophysical Research: Atmospheres*, **128**, e2023JD039045. https://doi.org/10.1029/2023JD039045

Liu, W., Ullrich, P. A., Li, J., Zarzycki, C., Caldwell, P. M., Leung, L. R., & Qian, Y. (2023). The June 2012 North American derecho: A testbed for evaluating regional and global climate modeling systems at cloud-resolving scales. *Journal of Advances in Modeling Earth Systems*, **15**, e2022MS003358. https://doi.org/10.1029/2022MS003358

Tang, Q., Prather, M. J., Hsu, J., Ruiz, D. J., Cameron-Smith, P. J., Xie, S., and Golaz, J.-C. (2021). Evaluation of the interactive stratospheric ozone (O3v2) module in the E3SM version 1 Earth system model, *Geosci. Model Dev.*, **14**, 1219–1236, https://doi.org/10.5194/gmd-14-1219-2021

Han, B., Fan, J., Varble, A., Morrison, H., Williams, C. R., Chen, B., et al. (2019). Cloud-resolving model intercomparison of an MC3E squall line case: Part II. Stratiform precipitation properties. *Journal of Geophysical Research: Atmospheres*, **124**, 1090–1117. https://doi.org/10.1029/2018JD029596

Varble, A., H. Morrison, and E. Zipser, 2020: Effects of Under-Resolved Convective Dynamics on the Evolution of a Squall Line. *Mon. Wea. Rev.*, **148**, 289–311, https://doi.org/10.1175/MWR-D-19-0187.1.

Houze, R. A. (2014). Cloud Dynamics (2nd ed.): Elsevier/Academic Press, Oxford.

Cui, W., X. Dong, B. Xi, Z. Feng, and J. Fan, 2020: Can the GPM IMERG Final Product Accurately Represent MCSs' Precipitation Characteristics over the Central and Eastern United States?. J. Hydrometeor., 21, 39–57, https://doi.org/10.1175/JHM-D-19-0123.1.

Zhang, Z., A. Varble, Z. Feng, J. Hardin, and E. Zipser, 2021: Growth of Mesoscale Convective Systems in Observations and a Seasonal Convection-Permitting Simulation over Argentina. Mon. Wea. Rev., 149, 3469–3490, https://doi.org/10.1175/MWR-D-20-0411.1.

Ayat, H., J. P. Evans, S. Sherwood, and A. Behrangi, 2020: Are Storm Characteristics the Same When Viewed Using Merged Surface Radars or a Merged Satellite Product?. J. Hydrometeor., 22, 43–62, https://doi.org/10.1175/JHM-D-20-0187.1.

Guilloteau, C., Foufoula-Georgiou, E. (2020). Multiscale Evaluation of Satellite Precipitation Products: Effective Resolution of IMERG. In: Levizzani, V., Kidd, C., Kirschbaum, D., Kummerow, C., Nakamura, K., Turk, F. (eds) Satellite Precipitation Measurement. Advances in Global Change Research, vol 69. Springer, Cham. https://doi.org/10.1007/978-3-030-35798-6_5

Characterizing Wet Season Precipitation in the Central Amazon Using a Mesoscale Convective System Tracking Algorithm

Sheng-Lun Tai, Zhe Feng, James Marquis, and Jerome Fast

Pacific Northwest National Laboratory, Richland, WA.

Corresponding author: Sheng-Lun Tai (sheng-lun.tai@pnnl.gov)

Key Points:

- Simulated and observed MCS clouds and precipitation are tracked during the 2014/15 wet seasons in central Amazon.
- Excessive heavy rain intensity ($\geq$ 10 mm h^{-1}) and relatively long travel distance of simulated MCS lead to overall overprediction of precipitation.
- Dry bias associated with stratiform rainfall may also drive MCS precipitation bias when cloud cover is substantially underpredicted.

Abstract

To comprehensively characterize convective precipitation in the central Amazon region, we utilize the Python FLEXible object TRacKeR (PyFLEXTRKR) to track mesoscale convective systems (MCSs) observed through satellite measurements and simulated by the Weather Research and Forecasting (WRF) model at convection-permitting resolution. This study spans a two-month period during the wet seasons of 2014 and 2015. We observe a strong correlation between MCS track density and accumulated precipitation in the Amazon basin. Key factors contributing to precipitation, such as MCS properties (number, size, rainfall intensity, and movement), are thoroughly examined. Our analysis reveals that while the overall model produces fewer MCSs with smaller mean sizes compared to observations, it tends to overpredict total precipitation due to excessive rainfall intensity for heavy rainfall events (≥ 10 mm h^{-1}) and longer traveled distances than observed. These biases in simulated MCS properties could vary with the strength of constraints on convective background environment. Moreover, while the wet bias from heavy (convective) rainfall outweighs the dry bias in light (stratiform) rainfall, the latter can be crucial, particularly when MCS cloud cover is significantly underestimated. A relevant case study for April 1, 2014 highlights the influence of environmental conditions on the MCS lifecycle and identifies an unrealistic model representation in convective precipitation features.

Plain Language Summary

We tracked large-size rain storms called mesoscale convective systems (MCSs) in the central Amazon during the wet seasons of 2014 and 2015 using an automated feature tracking algorithm. Data generated from MCS tracking helps us understand how MCSs rainfall is produced as a function of the number of storms, as well as their size, rain intensity, and motion, and how those can be better simulated by weather and climate models. We found that generally the model produces less and smaller MCSs than in reality, but the total MCS rainfall amount is often overestimated. This is because simulated MCSs travel longer, and most importantly they produce unrealistically intense heavy rainfall events. On the other hand, light rainfall events are mostly underrepresented by the model. Thus, the model error in total precipitation is determined by how these two compensate each other. Our analysis also suggests accurate model representation in environment is required for simulating realistic MCS properties.

1 Introduction

Mesoscale convective system (MCS) is a deep convective storm with clouds and precipitation organized on spatial scales of 100 km (Houze 2014). In the Amazon, MCSs produce over 60% of the total rainfall during the wet season (March-April-May, MAM), primarily due to their relatively long duration and wide extent than less organized convective storms (Nesbitt et al. 2006; Feng et al. 2021; Schumacher and Rasmussen 2020). In addition to the hydrological impact of MCSs, their extensive anvil clouds have a sizeable impact on the regional radiation budget (Feng et al. 2011). Thus, MCSs play a noteworthy role in regional and global climate.

Most Amazonian MCSs are initiated by near-surface convergence associated with prevailing trade winds, surface friction, and dirunal sea breeze circulations on the Atlantic coast (Sousa et al. 2021). After initiation, many MCSs travel long distances across the Amazon basin with a lifetime over half a day, sustained by environmental conditions that are frequently favorable for deep convection and growing to substantial sizes owing to rich tropical moisture. Hence, the local precipitation diurnal cycle in the Amazon basin is tied closely to westward propagating MCSs (Tai et al. 2021).

Several studies have employed a combination of routine and field campaign observational datasets to characterize MCSs in the Amazon region including weather radars, satellite observations, and surface measurements. Such efforts capture the intricate dynamics of MCS, their spatial distribution, intensity, and lifecycle evolution (e.g., Laurent et al. 2002; Petersen et al. 2002; Cifelli et al. 2002; Rickenbach et al. 2002; Machado et al. 2004). The recent GoAmazon2014/5 field campaign (Martin et al. 2016) collected many valuable observational data sets that have been used to characterize the diurnal variation, morphology, propagation, vertical motion, and precipitation of convective clouds around central Amazon (Burleyson et al. 2016; Rehbein et al. 2019; Giangrande et al. 2020; Wang et al. 2020; Tian et al. 2021; Anselmo et al. 2020). In addition to observations, advanced regional climate and convection-permitting models have been utilized to simulate MCSs over the Amazon under various meteorological conditions. By integrating observations with simulations, researchers have deepened our understanding of the processes governing MCS formation and evolution in the Amazon (Silva Dias et al. 2002; Carvalho et al. 2002; Machado et al. 2004; Tai et al. 2021, Paccini et al. 2023). These studies provide insights into the convective organization within the Amazon basin, shedding light on the complex interactions between atmospheric dynamics, moisture availability, and convective activity.

Nonetheless, remaining model uncertainties in MCSs over Amazon has motivated additional research in the field. Prior studies have shown that simulated MCS precipitation is quite sensitive to model resolution (vertical and horizontal), atmospheric forcing in initial and boundary conditions, soil moisture, and physics parameterizations (land surface, planetary boundary layer, cloud microphysics, and radiation) (e.g., Luo et al. 2015; Stensrud et al. 2000; Feng et al. 2018; Prein et al. 2021; Tai et al 2021; Prein et al. 2022; Na et al. 2022; Rasmos-Valle et al. 2023; Yang et al. 2023). Due to constraints in availability and spatial coverage of observational data in the sparsely populated Amazon, most evaluations of simulated MCS behaviors have been conducted in a confined region and narrow time windows which may be shorter than the MCS lifetime. Thus, additional work is warranted to examine modeled MCSs in realistic Amazonian conditions.

Cataloging MCS frequency, size, precipitation intensity, and movement, are essential for determining precipitation processes that contribute to total accumulative rainfall in the Amazon. However, representation of these MCS characteristics in state-of-the-art atmospheric models lacks rigorous quantitative validation. A pioneering study from Laurent et al. (2002) uses geostationary satellite data with 30-min frequency to enable deep convective cloud tracking, providing a different aspect in assessing modeled storms. With an increasing amount and quality of available satellite data, a number of cloud cluster tracking tools have been developed in recent years to characterize the lifecycle of deep convective clouds (Anselmo et al. 2021; Huang et al. 2018; Feng et al. 2019; 2021, 2023; Rehbein et al. 2018; Galarneau et al. 2023; Prein et al. 2020; Da Silva et al. 2023). One example is the Python FLEXible object TRacKeR (PyFLEXTRKR, Feng et al. 2023) algorithm, which we adopt in this study to facilitate MCS tracking in the central Amazon using satellite observations and a series of convection-permitting (4-km grid spacing) simulations during the 2014/15 wet seasons. The goal of this study is to elucidate the role of key MCS properties in driving the model precipitation errors through an in-depth storm tracking analysis.

The remainder of this paper is organized as following. Section 2 provides the details of the model and experiments as well as the algorithm used for trackings both simulated and observed MCSs. Results of analysis derived from MCS tracking statistics across timescales are demonstrated in Section 3. Finally, summary and conclusion are provided in Section 4.

2 Methods

2.1 Model setup and experiments

We use the WRF model version 3.9.1 (ARW, Skamarock et al. (2008)) to simulate convective clouds over the entire Amazon region, using a general configuration similar to our previous work in this region (Tai et al. 2021). Our study period includes a month in 2014 (March 11 to April 10) and 2015 (March 1 - 31). The model domain encompasses the northern part of the South American continent as well as adjacent oceans (Figure 1). The domain is constructed with a horizontal grid spacing of 4 km and a stretched vertical coordinate of 60 levels. The model top is located at 100 hPa. The physics schemes used for the simulations include: Thompson microphysics parameterization (Thompson et al. 2008), Mellor-Yamada-Nakanishi Niino (MYNN) boundary layer parameterization (Nakanishi and Niino 2009), Mellor-Yamada-Janjic surface layer parameterization (Janjić 2001), Unified Noah land-surface parameterization (Chen and Dudhia 2001), and the RRTMG longwave and shortwave radiation parameterization (Iacono et al. 2008). No cumulus parametrization is used because the model's horizontal grid spacing (4 km) is capable of resolving MCSs (Prein et al 2020; Na et al 2022). We use 6-hourly, $1° \times 1°$ NCEP FNL operational model global tropospheric analysis for model initialization (National Centers for Environmental Prediction 2000).

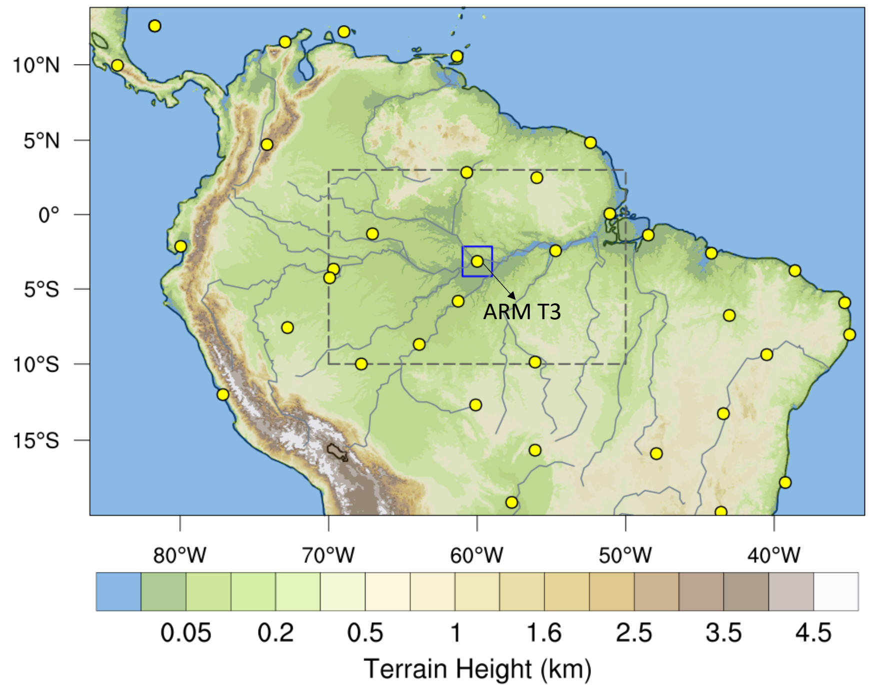

Figure 1 Map shows the configured WRF model domain for the simulations used in this study. Color shading illustrates terrain heights. Yellow dots denote the locations of radiosonde profiles that are assimilated along with the simulations. The location of ARM T3 site during GoAmazon2014/5 is indicated. The dashed rectangle marked by dashed line represents the study area for the MCS tracking analysis. Subdomain denoted by blue box is used for profiles sampling discussed in Section 3.4.

The model is also coupled with a data assimilation (DA) scheme to better constrain the simulation's background meteorological condtions, identical to the approach used in Tai et al. (2021). Conventional observations (e.g., radiosonde profiles, surface meteorology, aircraft, ship and others) and satellite radiances are assimilated by using the three-dimensional variational (3DVar) technique as provided in the package of version 3.6 Community Gridpoint Statistical Interpolation (GSI, Shao et al. 2015). It produces optimized analyses by blending the model data with observations as collected for the NCEP Global Data Assimilation System (GDAS, http://rda.ucar.edu/datasets/ds337.0/) and the radiosonde profiles measured at the T3 site (Figure 1) deployed by the DOE's Atmospheric Radiation Measurement (ARM) Mobile Facility (AMF, Miller et al. 2016) during the GoAmazon2014/5 field campaign (Martin et al. 2016).

Assimilated radiosondes were launched every 6 hours at the ARM T3 (e.g., 00, 06, 12, and18 UTC; 20, 02, 08, and 14 LT) throughout the campaign to measure tropospheric winds, temperature, pressure, and humidity profiles. Over the intensive observational periods (IOPs), one additional radiosonde was launched at 15 UTC (11 LT) to enhance measurement of the diurnal variation of environmental conditions. In addition to ARM's radiosondes, meteorological profiles measured at other sites as archived in the NCEP ADP global upper air and surface weather observations product are also assimilated in our model configuration. Note these observations have lower temporal frequency (up to twice a day at 00 and 12 UTC) and vertical resolution than those performed at the T3 site. The yellow dots denoted in Figure 1 indicate the locations of available radiosonde data at 12 UTC of March 12, 2014. The DA-coupled simulation is initialized at 00 UTC on the first day of each simulated month. The 3DVar data assimilation is performed every 12 hours (at 00 and 12

UTC) throughout the simulation periods. More details regarding the model configuration and DA strategy can be found in Tai et al. (2021).

2.2 Tracking observed MCSs

The MCS tracking is performed using the Python FLEXible object TRacKeR (PyFLEXTRKR, Feng et al. (2021, 2023)), a software package which is designed to track any atmospheric features in 2-D geographic planes using user-prescribed observational data sets or model output. To objectively identify and track the deep convective clouds, the PyFLEXTRKR primarily uses 2-D projections of infrared brightness temperature (Tb) observations commonly measured by geostationary satellites. When tracking observed MCSs with PyFLEXTRKR, we use NASA's Global Merged IR V1 infrared brightness temperature (Janowiak et al. 2017) data set. This dataset comprises multiple operational geostationary satellite data sources and includes viewing angle and parallax corrections. It has a continuous global coverage from 60°S to 60°N with a horizontal grid spacing of ~ 4 km and a temporal resolution of 30 min. We use hourly Tb data to identify and track deep convective clouds associated with MCSs identical to the approach adopted by Feng et al., (2021; 2023). A detailed discussion of the impact of uncertainties in MCS tracking owing to the IR Tb data set are described in Feng et al. (2021). Moreover, PyFLEXTRKR also uses collocated surface precipitation to assist identification of MCS "precipitation features (PF, contiguous area with rain rate > 2 mm h^{-1})" (Feng et al., 2021). Because PyFLEXTRKR tracks all deep convective clouds with a pre-defined minimum area threshold, the tracking data consists of records from early stage (near initiation) of individual deep convective clouds to the decay stage when the area of the cloud system decreases. More details of the PyFLEXTRKR algorithm can be found in Feng et al. (2023).

We use the NASA Integrated Multi-satellitE Retrievals for Global Precipitation Measurement (GPM) (IMERG) V06B precipitation data (Huffman et al. 2019) as a source of observed rainfall data in the MCS tracking algorithm. Precipitation estimates in IMERG are obtained by various precipitation-retrieving satellite passive microwave (PMW) sensors using the Goddard Profiling algorithm (Kummerow et al., 2001, 2015, 2011). Intercalibration is performed using the GPM Combined Radar Radiometer Analysis product. The precipitation product has a grid spacing of 0.1° and is also available every 30 minutes over a large portion of the globe (Huffman et al., 2014; Hou et al., 2014; Tang et al., 2016; Tan et al., 2019). We further averaged the 30 min IMERG data to hourly, and coarsened the 4 km Tb data to match the IMERG grid. Hence, the collocated Tb and IMERG precipitation data at 0.1° and hourly resolution are jointly used for MCS tracking in this study.

2.3 Tracking simulated MCSs

In WRF simulations, the top-of-atmosphere (TOA) outgoing longwave radiation (OLR) is used to infer the Tb. An empirical function is employed in PyFLEXTRKR to convert OLR to Tb following the formula from Yang and Slingo (2001). To avoid observational and simulated data resolution mismatches, the 4 km WRF simulation output is regridded based on the coordinate of observational data in a grid resolution of 0.1°. In this study, the thresholds used to define convective clouds, MCSs, and PFs in terms of cloud top brightness temperature, rain rate, and feature size are listed in Table 1. The sensitivity of MCS tracking due to variations of these thresholds was found to be qualitatively minor based on prior tests (not shown). A recent study comparing six different feature tracking algorithms applied to observed and simulated MCSs over South America found that most of the MCS properties from PyFLEXTRKR are representative of results from other obeject tracking tools (Prein et al. 2023), suggesting our algorithm can produce representative MCS characteristics.

Table 1 Summary of parameters used for the MCS tracking algorithm in the PyFLEXTRKR.

Category	Parameter	Value	Unit	Description
Cloud identification	Warm cloud Tb	261	K	Brightness temperature threshold for identification of “warm anvil”
	Cold cloud Tb	241	K	Brightness temperature threshold for identification of “cold anvil”
	Core cloud Tb	225	K	Brightness temperature threshold for identification of “core cloud”
	Minimum cold core cloud pixels	4	unitless	Mininum number of pixels of cold core cloud in qualification of a “core cloud”
	Minimum area	800	km^2	Minimum area in qualification of a “cloud” object
	Missing data fraction	0.35	unitless	Maximum fraction for missing data
	Minimum area	40000	km^2	Minimum total cloud area in qualification of a MCS
MCS identification	Minimum duration	4	hour	Minimum duration in qualification of a MCS
	Minimum PF rain rate	3	$mm\ h^{-1}$	Minimum rain rate in qualification as part of a PF
	Minimum PF link area	648	km^2	Minimum linked area of a PF
PF identification	Minimum PF major axis	100	km	Minimum length for a PF’s major axis
	Maximum PF major axis	1800	km	Maximum length for a PF’s major axis
	Minimum PF duration	4	hour	Maximum duration for a PF
	Minimum PF rainrate	2	$mm\ h^{-1}$	Cut-off rain rate in a PF
	Heavy rain rate threshold	10	$mm\ h^{-1}$	Minimum rain rate to be defined as “heavy rain”

2.4 A MCS tracking example

Figure 2 demonstrates an example of MCS detection and tracking at 13 UTC on March 19, 2014, during our study period. In this case, WRF simulates comparable fractions of core and cold anvil clouds. However, due to narrowing between two simulated cloud clusters (Figure 2b), there are two separate MCSs identified rather than a single MCS in observations (Figure 2a). The results of cloud type identification based on the defined thresholds in Table 1 are shown in Figures 2c and 2d. The tracked precipitation features corresponding to this MCS event are shown in Figures 2g and 2h, identified using the rain rates in Figures 2e and 2f. The model produces much higher rain rates than is observed in this time. The simulated maximum rain rate is 49.41 mm h^{-1}; whereas, the maximum satellite-retrieved observational rain rate was less than half of that (21.9 mm h^{-1}). This tendency leads to a relatively large fraction of heavy rainfall area in the simulated MCS than is observed. In addition, the model produces a large area of very light rainfall near the east side of the domain likely associated with a sea breeze circulation (Figure 2f) that is less evident in the satellite observations (Figure 2e).

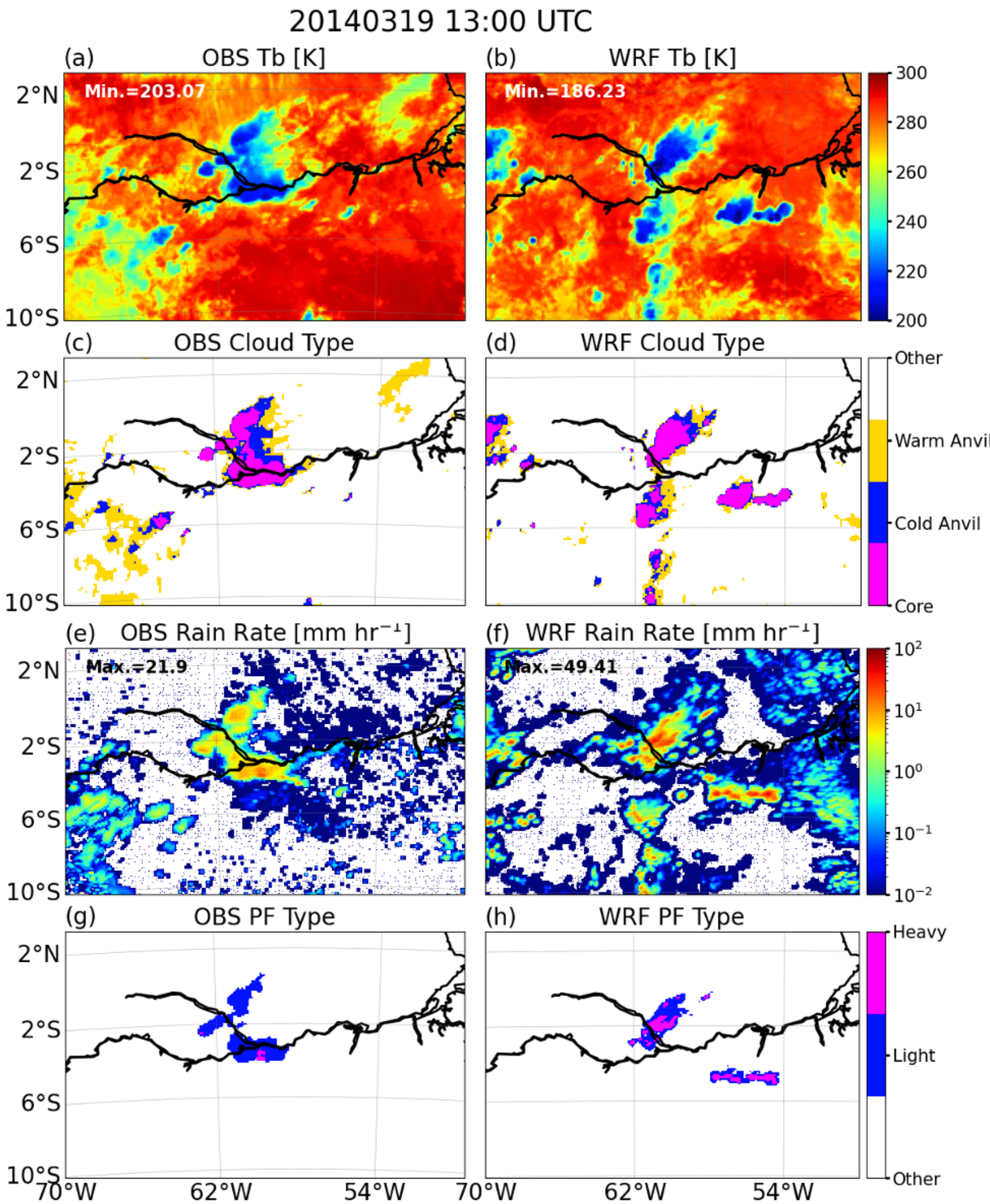

Figure 2 Snapshots of brightness temperature, cloud type, rain rate, and precipitation type as recorded along with the MCS tracking at 13 UTC on March 19, 2014. The results derived from satellite observation and WRF simulation are illustrated in (a), (c), (e), (g) and (b), (d), (f), (h), respectively.

3 Results

3.1 Number and spatial distribution of tracked MCSs

We examine the total number as well as the spatial distribution of all tracked MCSs from both observational and simulated data sets over the study domain (Figure 1). We tracked 125 and 115 observed MCSs occurring during the months of 2014 and 2015, respectively; compared to 120 and 94 MCSs tracked during the same period in the simulations. Therefore, there were slightly fewer simulated MCSs (-8%) in the 2014 period, but a notable simulated deficit of MCSs (-22%) during the 2015 period.

MCS object track density is mapped onto a 1° × 1° grid and illustrated in Figure 4. During the 2014 sampling period, more MCSs are observed near the northwest corner of the domain. The difference map for 2014 (Figure 3c) indicates that while more MCSs are simulated over the northwestern and eastern parts of the domain, fewer MCSs (negative blue patches) occur within grid cells near Amazon river. Interestingly, during the 2015 sampling period, MCSs are also under-predicted along the Amazon river, particularly over the northeastern central Amazon (Figure 3f).

While observed rain maps show relatively a consistent rainfall distribution and amount over the analysis domain (Figures 4a and 4d), simulated rainfall amount is distinctly higher in 2015 than for 2014 (Figures 4b and 4e), despite lower MCS occurrence (Figures 3c and 3f). For instance, during the 2014 period, observed domain-mean precipitation is 160.3 mm, which is slightly lower than is simulated (168.8 mm). Nevertheless, during the 2015 period, simulated and observed domain-mean precipitation are 209.5 and 171.4, respectively. Which suggests the domain-mean precipitation bias dramatically increases from +5% to +22%.

Model bias in MCS occurrence (Figures 3c and 3f) modulates the overall pattern of rainfall bias (Figures 4c and 4f), which confirms that MCS precipitation contributes to a considerably large fraction of total rainfall during these two periods. The dry bias along the Amazon river is analyzed during both years (Figures 4c and 4f) and can be attributed to relatively low MCS occurrence simulated by the model (Figures 3c and 3f) despite potentially higher rain rate (Figure 2). This implies that the current model configuration may be associated with unresolved precipitation processes related to river-atmosphere interactions such as river-breezes that enhance deep convection under easterly trade winds (Burleyson et al. 2016), among other possible factors.

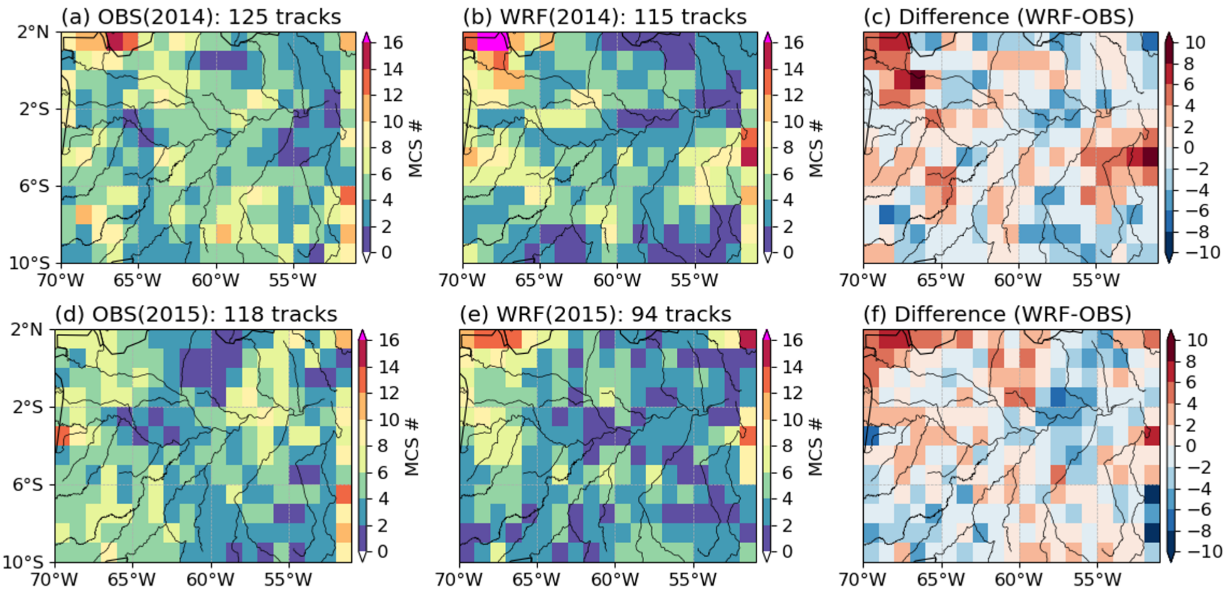

Figure 3 Spatial distribution of gridded MCS track density (i.e., number of MCS objects passing through 1°×1° grid boxes) from observations (a, d) and WRF simulations (b, e) over the study domain. (c) and (f) illustrate the difference between (a), (b) and (d), (e), respectively. The top (bottom) row represents results for the month of 2014 (2015).

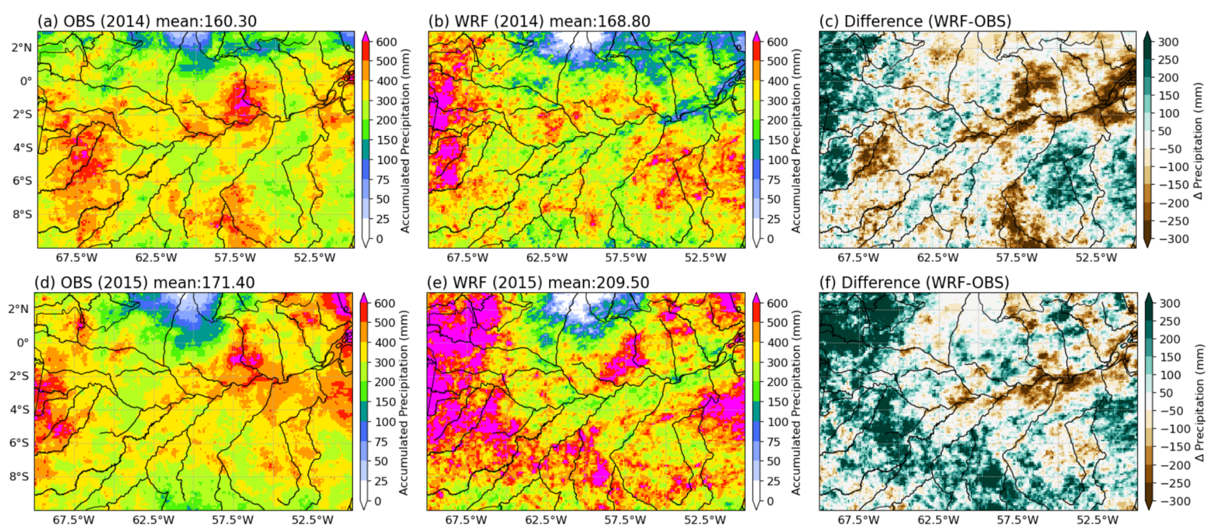

Figure 4 Similar to Figure 3, but for monthly precipitation amount (mm). The domain mean precipitation is given in the title of each panel (a, b, d, and, e)

3.2 MCS properties

The tracking data for all MCSs are used to obtain monthly mean values per MCS in both the 2014/15 periods. Besides the occurrence (number of MCSs) as discussed, the accumulated precipitation of a MCS can be attributed to its size, duration, rainfall intensity, and distance traveled, we analyze properties including MCS cloud area, rain rate, and motion (e.g., duration, speed, and movement distance). This helps elucidate how these factors contribute to overall precipitation amount and what is the fractional bias in each MCS properties. The results of selected 6 MCS properties are given in Figure 5.

We first assess the area covered by the entire MCS cloud shield (Tb < 241 K), and further categorize the cloud cover into convective and stratiform areas (Feng et al. 2021, 2023). The convective (core cloud) area is defined as the continuous area with cloud top temperature below 225 K within a much wider MCS cloudy patch. The remaining area with Tb from 225 to 241 K is attributed to stratiform (cold anvil cloud) type. The statistics show that the total extent of simulated MCS cloud area is generally smaller than observed; while observed mean MCS sizes are larger than 80,000 km^2 in both months, modeled MCSs are approximately 70,000 km^2. Observations suggest the ratios of convective/stratiform cloud cover are 1.13 and 1.05 in 2014 and 2015 period; whereas, simulated ratios increase to 1.36 and 1.43, respectively, most likely due to reduction of stratiform cloud areas.

The results also show simulated MCSs produce robustly higher rainfall intensity than is observed during each analysis period. The MCS precipitation is further partitioned into two types: heavy (≥ 10 mm h^{-1}) and light (< 10 mm h^{-1}). Distinct and consistent positive biases in modeled rain rates are seen in the heavy rainfall regime, driving the simulated total rain rate to be larger than twice of the average observed rate. While total rain rate does not vary much between the two periods (both are 9.1 mm h^{-1}), heavy rain rate substantially increases from 18.3 to 19.2 mm h^{-1}.

Finally, we examine the duration, movement speed, and distance traveled of tracked MCSs. The observations suggest that MCSs had similar durations across both analysis periods. Though simulated MCSs had slightly shorter durations than were observed during 2014 (by ~1 hour), the model significantly under-predicted the MCS lifespan (by > 2 hours) in 2015. Further, simulated MCSs motion was significantly over-predicted by ~3 – 4 m s^{-1} (~80%) during both analysis periods. Hence, the yearly differences in lifespan and propagation speed lead to larger horizontal excursions by the MCSs in both months. The simulated MCSs traveled longer distances than observed (~ +40 km in 2014 and ~ +20 km for 2015 in average).

As a result of the accumulated factors of MCS size, rain intensity, and movement, the modeled total precipitation per MCS is about 500 mm less (-20%) than what is observed in 2014 but about 700 mm more (+35%) in 2015. The much larger simulated rainfall amount in 2015 is most likely due to unrealistically higher rain intensity. The observed ratios between heavy/light precipitation are 0.4 and 0.33 for 2014 and 2015, respectively. However, it becomes nearly opposite in the simulations, as the corresponding ratios are 1.7 and 2 for the two periods. Therefore, besides having biases in total precipitation, the model also poorly represents the fractions of heavy and light precipitation.

The dramatic increase of errors in MCS occurrence (Section 3.1) as well as total precipitation from 2014 to 2015 periods catches attention because model skill in rainfall prediction usually does not significantly change from year to year. One possible source of this discrepancy in our simulations comes from biases in meterological conditions resulting from our data assimilation scheme. The quantitiy of radiosonde observations for during 2015 is approixmatley half of what was available to assimilate during 2014 (Figure 6). Thus, a much weaker constraint in simulated environmental conditions during 2015 most likely leads to enlarged biases in convective background conditions, which is expected to affect the examined MCS properties. Moreover, results of partitioned precipitation quantity (heavy vs. light) indicate the excessive MCS total precipitation in the model for the 2015 period is primarily due to a much larger fraction of heavy rainfall. The overpredicated amount of heavy rainfall reaches more than three times of the observed value. Conversely, the simulated light rainfall amount is much less than what was observed (by ~ 50%), which partially compensates the positive bias in heavy rain.

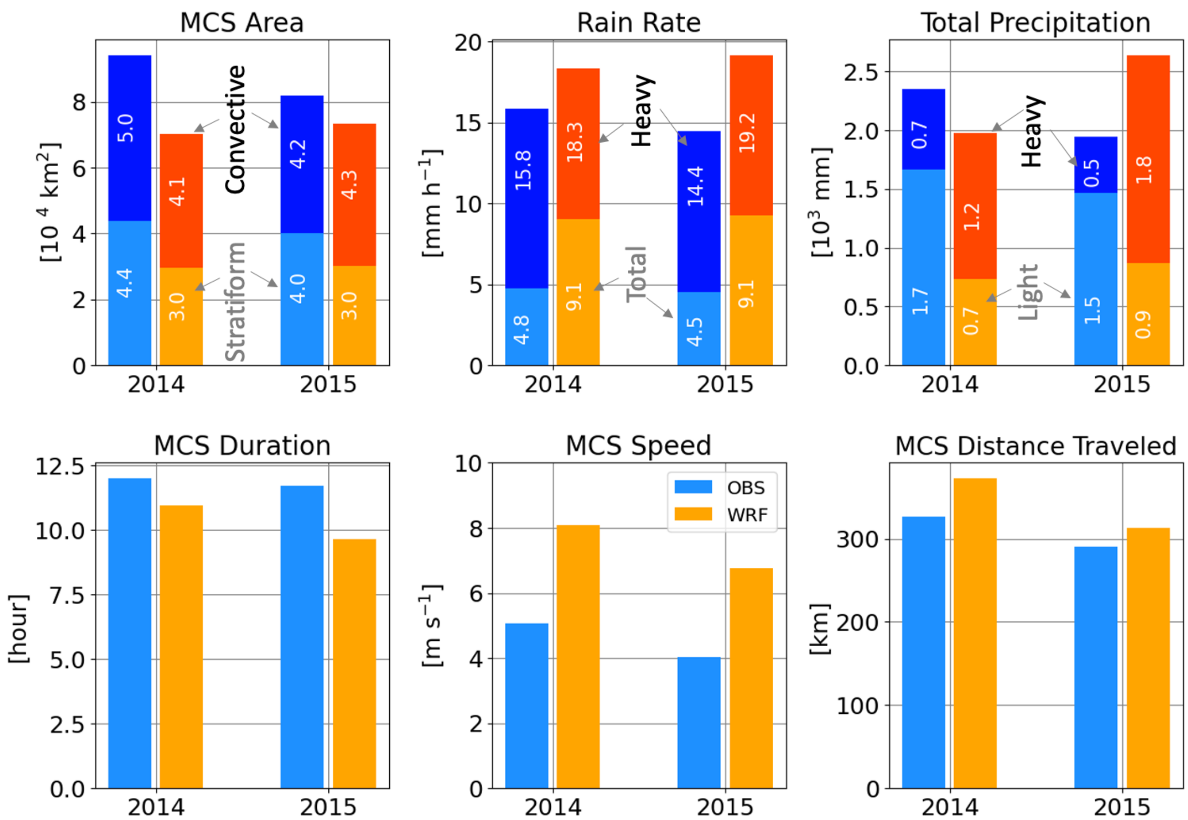

Figure 5 Comparison of monthly-mean values for MCS properties, including: MCS area, rain rate, total MCS precipitation, duration, movement speed, and distance traveled. The statistics obtained from observations and simulations are represented by blue and orange bars. In the plots for MCS area, the fractions of stratiform- and convective-type clouds are indicated by light blue (orange) and dark blue (red) bars, respectively. For rain rate comparison, the total and heavy (≥ 10 mm h^{-1}) rain rates are denoted by light blue (orange) and light blue + dark blue (orange + red) bars. Total precipitation is partitioned into light (light blue and orange bars) and heavy (dark blue and red bars) rain types.

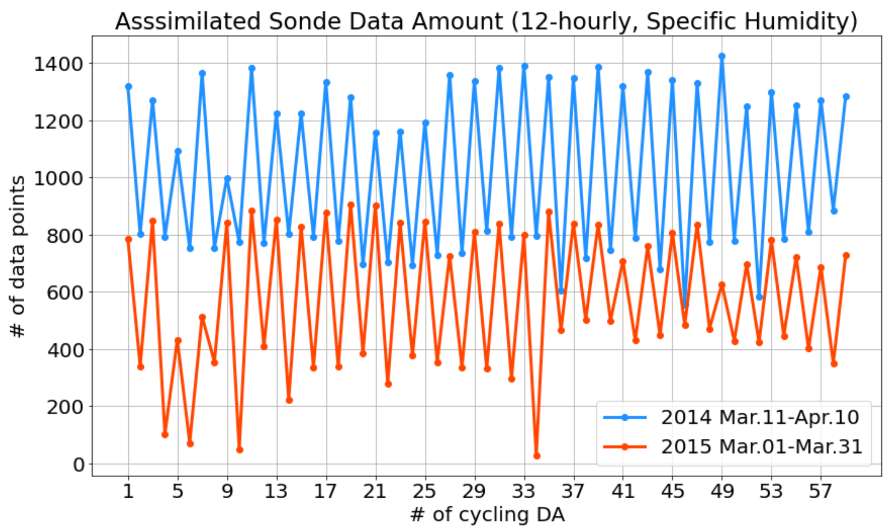

Figure 6 The number of assimilated data points from radiosonde specific humidity observations over the model domain (Figure 1). The blue and red curves represent timeseries results for 2014 and 2015 analysis periods. Y-axis denotes the number of assimilation cycles.

3.3 Daily variability

The tracked MCS properties, including the number, area, movement distance, and rain rate are further broken down to facilitate model validation on daily timescales. Approximately four MCSs were observed daily, on average, during the 2014 period, with only two days in which no MCSs were identified (Figure 7a). Although the model reproduced only slightly fewer MCSs in terms of the monthly mean value, it does not fully capture the daily variations, particularly during the first half of the month. Overall, the MCS cloud area is underpredicted by the model. Over the entire month, an average MCS size of ~9.4×10^4 km^2 is observed in satellite data; whereas, the model yields ~7×10^4 km^2. The model has more difficulty in simulating large MCSs (area > 10×10^4 km^2), such as the ones observed on 3/13, 3/22 and 4/1. While the correlation between the observed and simulated MCS movement distance is much lower than for other properties, simulated MCSs more frequently travel farther than observed ones (20 out of 31 days). Lastly, the simulated mean MCS rain rate is much higher than in observations (approximately +90%) every day.

Duringthe 2015 period (Figure 7b), simulated MCS number is notably underpredicted after 3/22 despite qualitative agreement in the trend. A noticeable contrast on the first day of simulations (3/1) may be due to model spin-up. There are no days during the 2015 period with observed MCS size larger than 15×10^4 km^2; thus, the MCS size is generally smaller than during the 2014 period. As a result, although the simulated mean MCS area is still smaller than observed, the deficit is not as large as it is during 2014. Similar to results obtained for 2014 period, there are 19 out of 31 days that MCSs traveled farther than observed in 2015. Despite variability in model biases of MCS properties, positive biases in MCS rain rate is robustly observed and are largest during the2015 period.

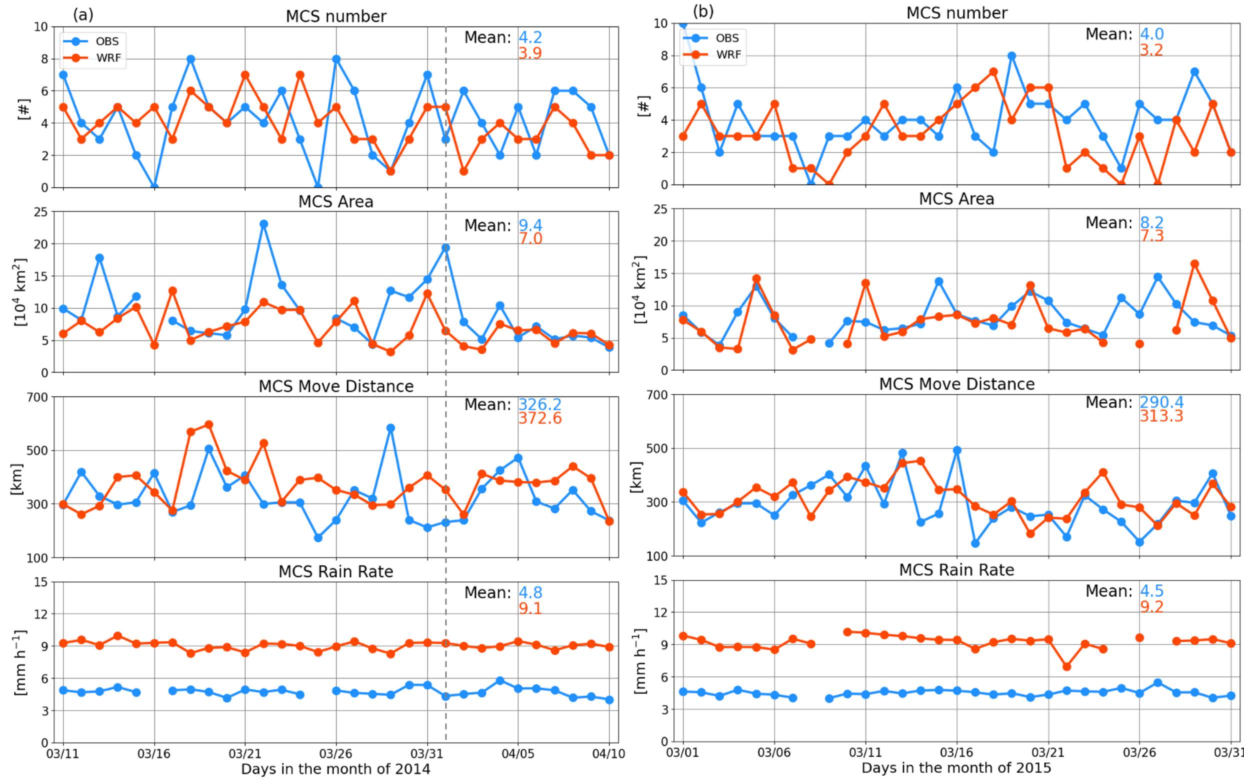

Figure 7 Daily variations of MCS tracking statistics including number, area, movement distance, and mean rain rate for the analysis period in (a) 2014 and (b) 2015. The dashed line denoted in (a) identifies the date (April 1, 2014) selected for case study (Section 3.4).

We next quantify heavy versus light rainfall-regime dependent model biases relative to biases in rain rate and cloud area (Figure 8). During most of the 2014 analysis period (Figure 8a), MCS precipitation is overpredicted on fewer than half of the days (13 out of 31) and underpredicted for the remaining 18 days. Moreover, on 3/13, 3/22, and 4/1, the negative biases in light rain are much more distinct than otherwise typical positive biases in heavy rainfall. This happens when MCS cloud areas in both convective and stratiform types are significantly under-predicted. During the 2015 analysis period, there are only two days such as 3/4 and 3/15 when negative biases of light rainfall are lower than -2000 mm and thus compensate or even lead to a negative total MCS precipitation bias. Other than those days, heavy rainfall bias dominates the total precipitation bias. Rain rate biases in heavy precipitation are noticeably higher in 2015 than 2014. Given relative minor model-observation differences in MCS area, excessive simulated rain rate is responsible for large positive biases in total MCS precipitation for 2015 analysis period as shown in Figure 6. With varied scenarios as observed during the two sampling periods, it suggests model validation using only accumulated precipitation amount will most likely ignore crucial model errors that may be revealed from examination of many MCS properties.

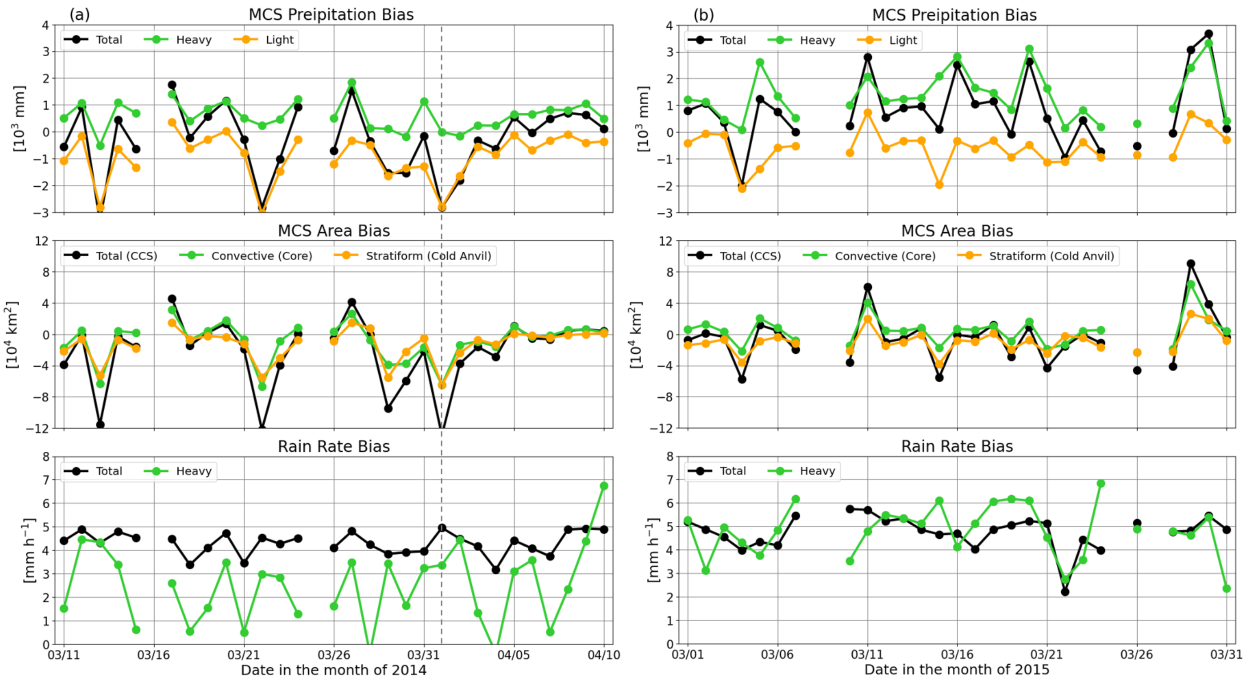

Figure 8 Similar to Figure 7 but for biases in MCS precipitation amount, rain rate, and cloud area. The black lines denote results for the total biases. Green lines illustrate fractional results for the heavy rainfall (≥ 10 mm h^{-1}) and convective-type (core) clouds. Orange lines represent results for the part with light rain (< 10 mm h^{-1}) and stratiform-type (cold anvil) clouds. The dashed line denoted in (a) identifies the date (April 1, 2014) selected for case study (Section 3.4).

3.4 Case study: April 1, 2014

Following the discussion in the previous section, we further investigate the contrasts between observed and simulated MCS clouds and precipitation by highlighting a case study for April 1, 2014. On this day, a MCS (denoted by white arrows at first row in Figure 9) initiated over northeastern corner of the analysis domain and experienced upscale growth through decay as it propagated westward (Figure 9). In general, the model reasonably reproduces the MCS's lifecycle with relatively accurate placement of clouds in time, allowing us to examine the evolution of clouds (Figure 9) and corresponding precipitation processes (Figure 10) with confidence. In this case, the model tends to simulate smaller cloud cover regardless of cloud types. The simulated stratiform clouds dissipated faster than observed after 12 UTC, particularly on the southeastern flank of the MCS, as the brightness temperature of cloud top significantly increases. Nevertheless, more isolated convective initiation remains active on the leading edge of the propagating MCS. Therefore, only a relatively narrow core cloud band is sustained in simulations. In reality, satellite observations suggest a much wider MCS cloud patch propagating toward the southwest through the domain.

The contrast between observed and simulated heavy/light rain distribution is shown in Figure 10. Though observations suggest heavy rainfall patches are scattered and mostly located in the center of a much wider light rain area, simulated heavy rainfall patches tend to have a much larger fractional area than observed and appear on the leading edge of relatively narrow cloud bands. This structural difference in the precipitation features implies that, most likely there are model deficiencies in representing the dynamics and/or microphysics within the simulated MCS causing the issue and warrants in-depth investigation in a future study to provide further insight.

The diurnal variations of cloud and precipitation type relative area fractions are shown in Figure 11. The evolution of simulated core cloud fraction is overall aligned with the observed trend until 12 UTC. Similar fractions (~20 – 30%) are obtained for cold anvil clouds. While the observed cold anvil cloud fraction is consistently around 30%, simulated cold anvil cloud fraction dropped to ~20% at the end of the day. Cold anvil clouds started dissipating at 16 UTC, which is about 3–4 hours later than the dissipation of core clouds.

The area fraction of heavy rain is consistent between observations and simulations and does not exceed 6% over the course of the day. However, we see a notable contrast in the fraction of light rain area. Overall, the observed light rain area covers more than twice of what it does in simulations. Hourly rain rate comparison indicates light rain rates in both observations and simulations are never greater than 5 mm hr^{-1}; whereas simulated heavy rain rate is mostly near 20 mm hr^{-1}, approximately 4 mm hr^{-1} more than observed, on average. Given much larger negative bias in anvil cloud cover and relatively small heavy rain rate bias during this event, the total MCS precipitation is thus primarily driven by the significant under-prediction of stratiform cloud cover. However, it is worth noting that validations of IMERG data against ground based observations (either radar or rain gauges) reveal that IMERG tends to significantly overestimate the frequency of weak precipitation (1-2 mm h^{-1}) while underestimating intense precipitation, particularly over land (Cui et al. 2020; Zhang et al. 2021; Ayat et al. 2021). Moreover, the actual resolution of IMERG is significantly coarser than its grid spacing (Guilloteau & Foufoula-Georgiou, 2020). Thus, associated rainrate model biases themselves may be overestimated.

To elucidate how environment conditions may influence differences in this MCS's lifecycle, we examine the pre-storm environment (before 12 UTC) as observed by radiosonde profiles at the AMF T3 site and simulated profiles within a 2° by 2° box centered at the AMF site (Figure 1). Vertical interpolation with an interval of 0.1 km was carried out for both radiosonde and model profiles. To exclude profiles affected by convective clouds, only the model profiles with column maximum reflectivity less than 0 dBZ are sampled. Resulting wind profiles valid at 00, 06, and 12 UTC are illustrated in Figure 12.

There are two jets evident in the observed wind speed profiles; one peaking at z = 2–3 km, and another above mid-troposphere (z = 7–9 km) (Figure 12a). Owing to small vertical heterogeneity in the meridional wind, these jets are primarily a result of variations in the zonal wind (Figure 12 b-c). Wind conditions do not vary significantly in the pre-storm environment. Although the model simulated wind conditions over the period are qualitatively similar to observations, simulated maximum wind speed of the lower jet is consistently smaller (~2–3 m s^{-1}

less than radiosonde). Furthermore, at 12 UTC, the lower jet descends to height below 1 km in the model. This may imply a shallower and weaker low-level jet being simulated, leading to much weaker moisture transport and convergence in the lower troposphere. Results of convective available potential energy (CAPE) and convective inhibition (CIN) further demonstrate while the available energy for convective growth increases from 00 to 12 UTC in reality, the model simulates a completely opposite trend (Figure 13), where CAPE dropped from nearly 2000 to 500 J kg^{-1} over the 12-hour period. Despite good agreement in CIN values, the simulated environment does not favor convective growth as observed. This evidence may at least partially explain why the simulated MCS quickly dissipates after 12 UTC and thus has a much smaller area cover by stratiform clouds.

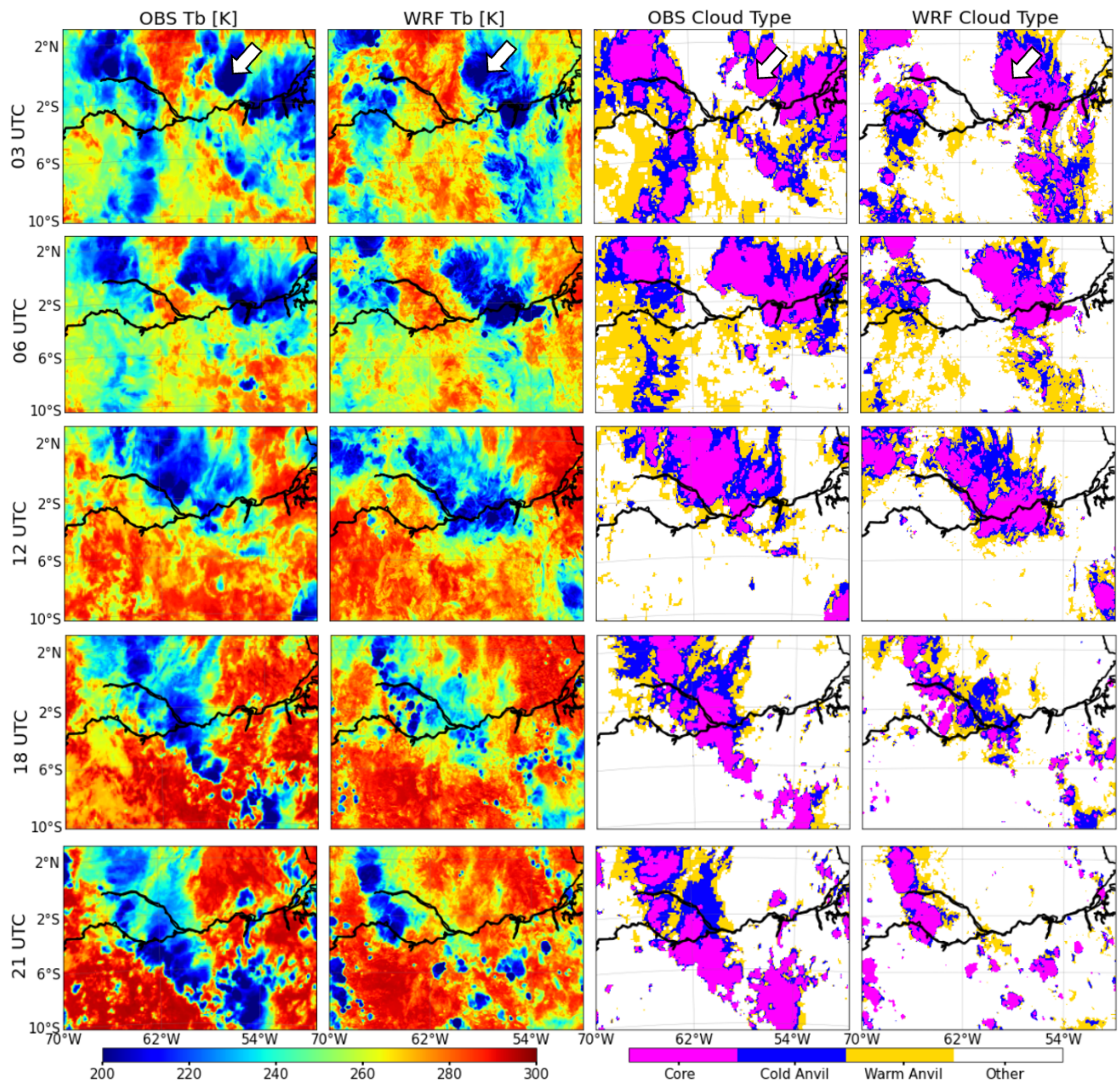

Figure 9 Similar to Figure 2a – 2d, but for 03 to 21 UTC on April 1, 2014. White arrows on the panels for 03 UTC point to the initiating MCS of interest.

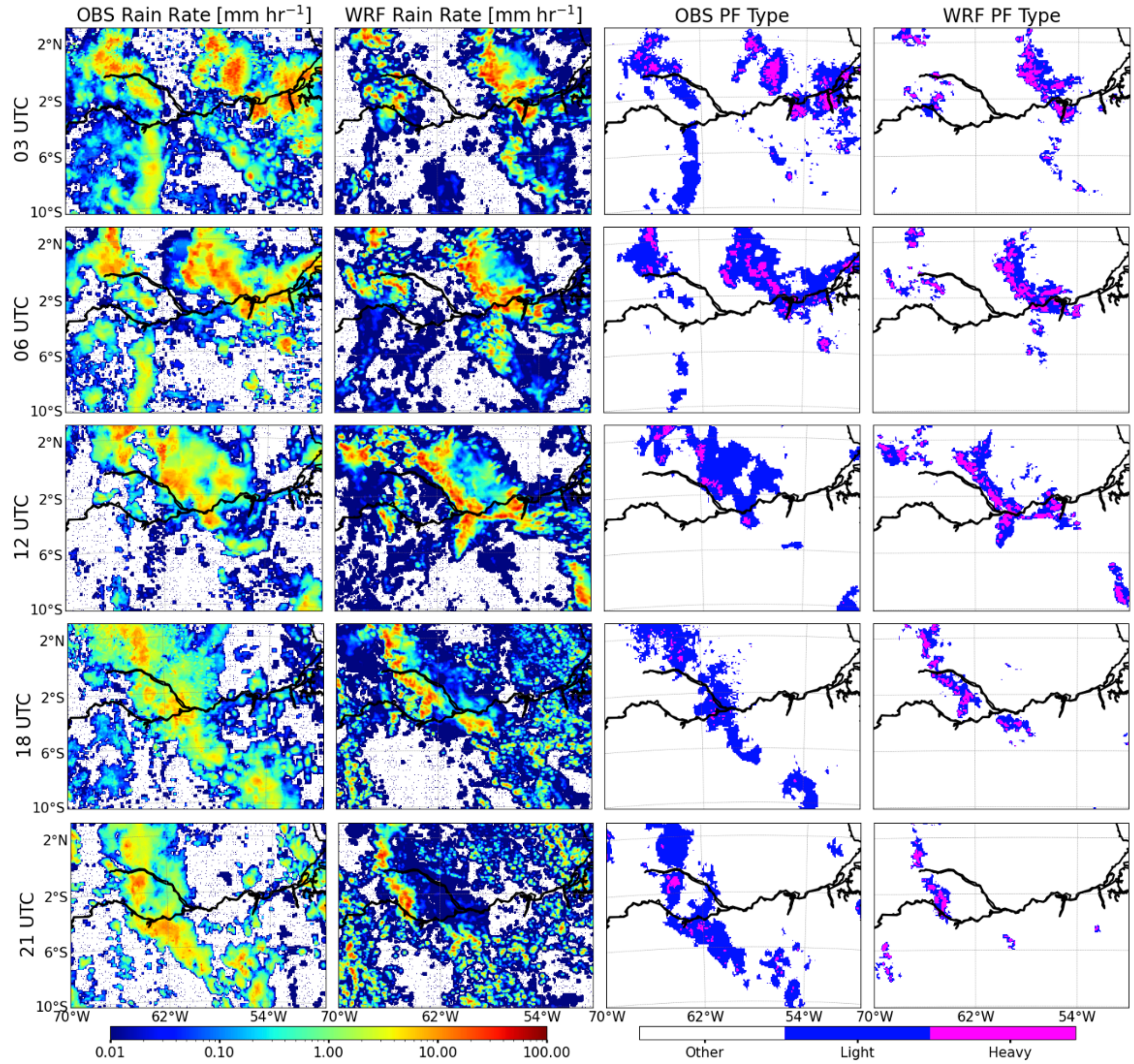

Figure 10 Similar to Figure 2e – 2h, but for 03 to 21 UTC on April 1, 2014.

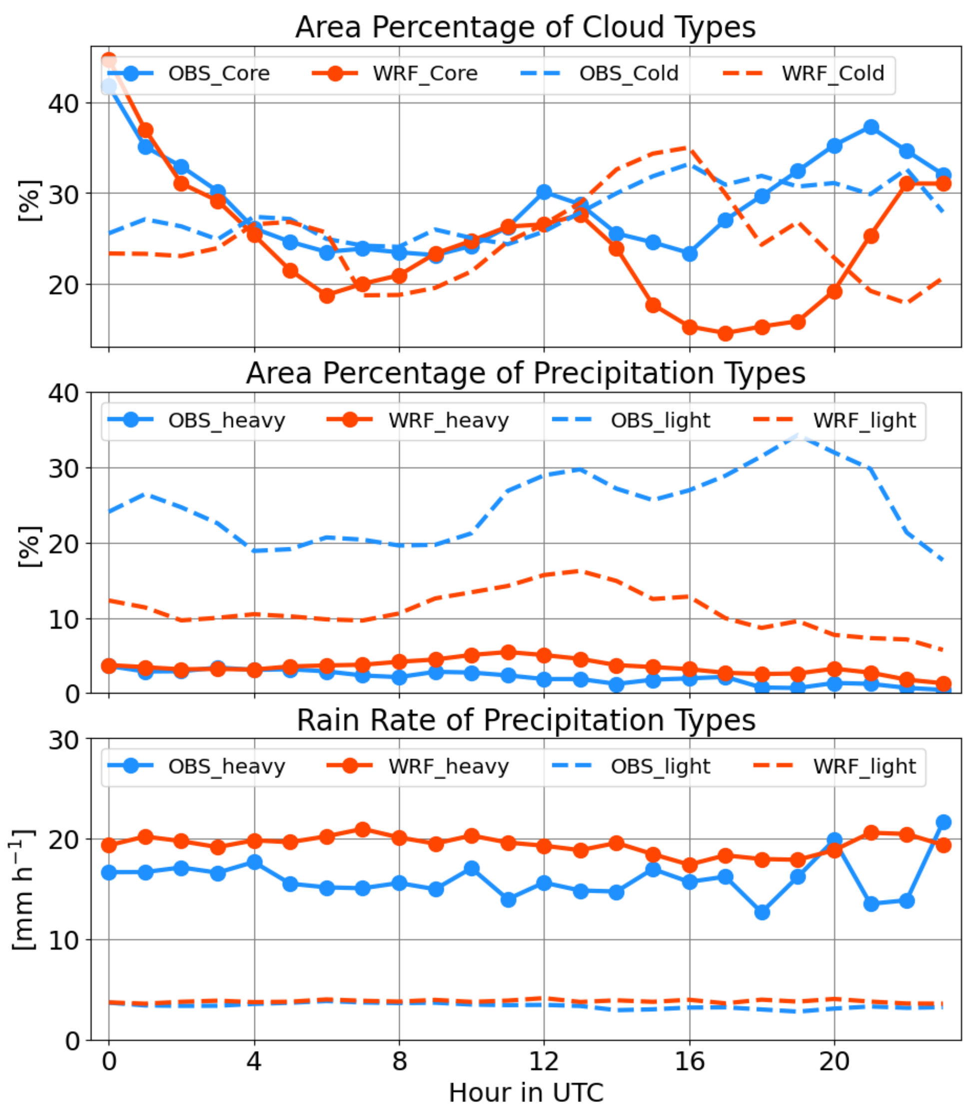

Figure 11 Comparison in diurnal varations of area percentages of cloud (top panel) and precipitation (middle panel) types as computed over the analysis domain on April 1, 2014. Correspondng rain rate comparison in dependency of precipitation type is displayed in the bottom panel.

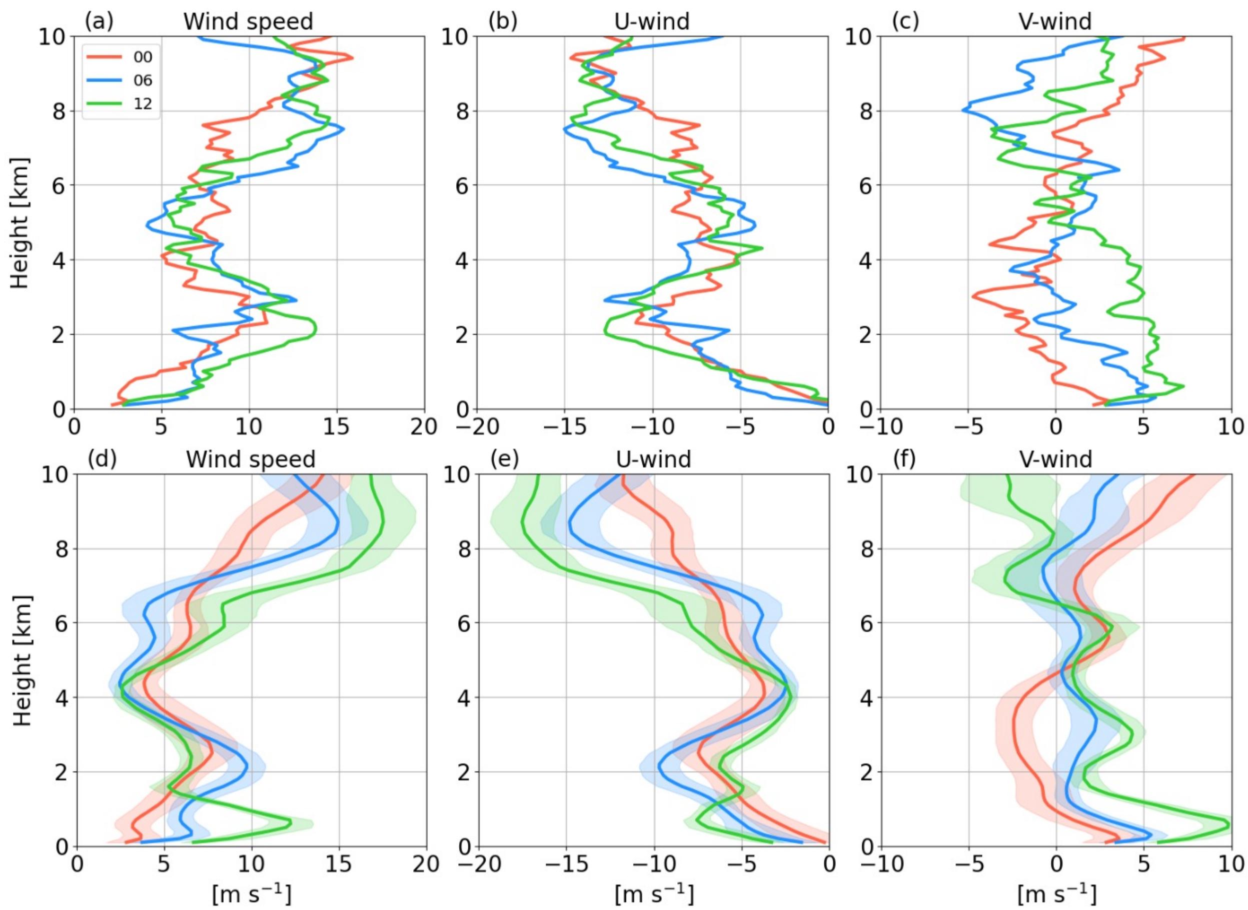

Figure 12 Wind speed, zonal (U-) wind, and meridional (V-) wind profiles as observed by radiosondes launched (a, b, and c) and simulated by WRF (d, e, and f) at the location of ARM T3 site. Colors denoted in legend indicate results for 00, 06, and 12 UTC of April 1, 2014. Swath of each line in d, e, and f represents the range within ± 1 standard deviation among the samples.

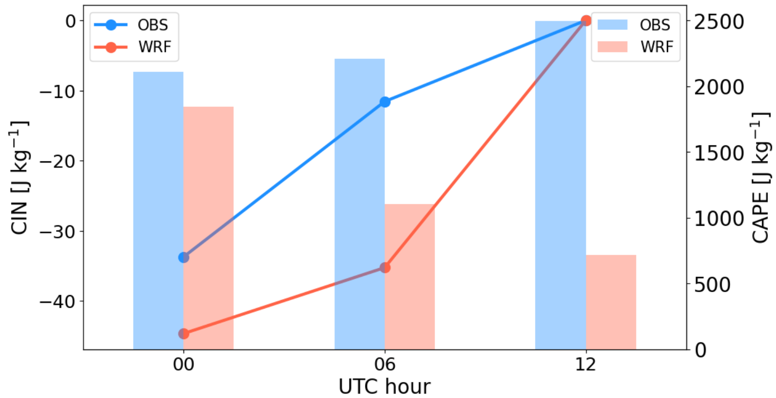

Figure 13 CIN (line with dots; left y-axis) and CAPE (bar; righy y-axis) values computed using ARM T3 radiosonde and corresponding WRF-simulated profiles at 00, 06, and 12 UTC of April 1, 2014.

4 Summary and conclusions

Mesoscale convective systems (MCSs) are responsible for a large fraction of the total precipitation in the Amazon. However, various uncertainties in state-of-the-art atmospheric models hinder them from reproducing a realistic lifecycle and morphology of MCSs within this region. To facilitate comprehensive characterization of MCS precipitation in the central Amazon, the Python FLEXible object TRacKeR (PyFLEXTRKR) is employed to track individual MCSs that are simulated by convection permitting (4 km grid) Weather Research and Forecasting (WRF) simulations. The WRF simulations are performed over two separate months during the 2014 and 2015 Amazon wet seasons. A 3DVar data assimilation scheme is used to constrain environmental conditions throughout simulations. These results are then compared to observed satellite analogs to examine possible mechanisms of MCS model biases.

First, we examined the MCS occurrence and its relation to accumulated precipitation. Overall, the model tended to produce fewer MCSs than were observed within the study area. While only 8% fewer MCSs are reproduced in 2014 period, we observe a difference of -22% for the month in 2015. The heterogenous precipitation bias distribution is closely tied with how well the MCS track density was reproduced. A distinct feature of dry biases along the Amazon river is identified and found to be well explained by model error in reproducing realistic MCS occurrence near the river.

Analysis of monthly means of tracked MCS characteristics further reveal the contrasts between observed and modeled MCS properties in general. Although simulated MCSs are generally smaller than observed ones, they produced far more rain and propagated farther than observed. Moreover, we find the model-observation discrepencies in various MCS properties must be considered when accounting for the sources of MCS total precipitation bias. For example, in 2014 period, MCS total precipitation is underestimated by the model due in part to relatively large negative bias in MCS size and minor positive rain rate bias. Whereas in 2015, while model bias in MCS size is relatively small, substantial positive bias in rain rate results in severe overpredication of MCS total precipitation. Aside from biases in total precipitation, we also show that the model has difficulty in reproducing realistic fraction of heavy/light precipitation.

The model errors in MCS number, rain rate, and precipitation (MCS and domain-mean) notably increase from 2014 to 2015 (Figures 3, 4, and 5). We find the degraded model performance is most likely driven by the availability of observational data for assimilation. The amount of assimilated radiosonde moisture data in 2015 dropped to only half of what was available in 2014 despite consistent assimilation of radiosonde profiles collected at the ARM T3 site in central Amazon. This reinforces the importance of additional observation sites to constrain simulated synoptic environments over the continent.

We further break down the statistics by each day and demonstrate that model skill in reproducing MCS properties, including number, size, and distance traveled, could vary significantly from day to day. On many days, the bias in total precipitation can be attributed to the wet bias in heavy rainfall, which result in overall overpredicted precipitation. However, light rainfall may occasionally drive the total precipitation error. Such events happened when both convective and stratiform cloud cover are under-predicted, hence both contribute to dry biases in precipitation. This suggests that it is critical to validate simulated precipitation by considering its dependency per rainfall regime because the biases sourced in different regimes may imply unrealistic model representations of various dynamical and/or microphysics processes. Analysis of daily bias provides more details in terms of model biases in MCS characteristics.

Finally, an analysis of an MCS on April 1, 2014 is provided to illustrate how differently the clouds and precipitation are resolved in both observational and model data. We showed that in this particular event, while relatively small wet bias in heavy rainfall is analyzed, the large dry bias in light rain controls the total precipitation bias. This is mainly caused by significant under-prediction in area cover of light rain. Examination of the pre-storm environment suggests the jet in lower troposphere is relatively shallow and weak in the simulations compared to observations. This could lead to insufficient moisture transport and hence weaker convergence that are essential for convective growth and sustainability. Moreover, weaker simulated CAPE also indicated unfavorable conditions for convective growth. Given the evidence, we conclude the environmental conditions may be causing the early dissipation of MCSs and significant negative bias in stratiform cloud cover.

In addition to environmental conditions as discussed in Section 3.4, potential sources of model uncertainties in reproducing observed MCS clouds/precipitation may also relate to 1) model resolution, which directly influences how MCS's dynamic structure (e.g. vertical motion) may be resolved and thus alters the secondary circulation accordingly (Varble et al. 2020); and 2) paramterization of microphysical processes. For instance, the magnitude of simulated stratiform precipitation is found to be associated with ice particle mass fluxes as predicted by the employed microphysics schemes (Han et al. 2019). Heating profiles could be changed drastically by replacing one microphysics scheme by another (Feng et al. 2018).

Compared to a mesoscale model, climate models tend to simulate even more unrealistic representations of tropical precipitation features (e.g., Tai et al. 2021), due in part to coarse grid spacing and much more simplified physics parameterizations. Given the substantial increase in computational power, climate models are now more frequently run at cloud-resolving scales (e.g., Tang et al. 2021; Liu et al. 2023). Despite promise as seen in selected case studies (Liu et al. 2023), a high-resolution configuration does not always lead to distinct improvements in general precipitation features (e.g., diurnal cycle) and associated meteorological conditions. We note climate models should use mesoscale model (e.g., WRF) simulations as benchmarks when assessing their performance. In this way, the behaviors of state-of-the-art climate models can be constrained by both success and failure of relatively well-developed mesoscale models.

Acknowledgments

This research was supported by the Atmospheric Science Research (ASR) program as part of the DOE Office of Biological and Environmental Research. Pacific Northwest National Laboratory is operated by DOE by the Battelle Memorial Institute under contract DE-A06-76RLO 1830. The modeling computation was performed on PNNL's Research Computing clusters.

Data Availability Statement

The WRF community model is available from the National Ceter for Atmospheric Research (NCAR) at http://www2.mmm.ucar.edu/wrf/users/. The NCEP FNL operational model Global Tropospheric Analysis is accessible at https://rda.ucar.edu/datasets/ds083.2/. The NCEP ADP global upper air and surface weather obserbations (prepbufr format) can be downloaded from https://rda.ucar.edu/datasets/ds337.0/. The dataset of National Aeronautics and Space Administration (NASA) Global Merged IR V1 infrared brightness temperature can be accessed at https://disc.gsfc.nasa.gov/datasets/GPM_MERGIR_1/summary/. The IMERG Final-Run V06B precipitation products used in this study were acquired from http://pmm.nasa.gov/data-access/downloads/gpm/. The GoAmazon2014/15 data used in this manuscript are freely available from the ARM data archive (https://www.arm.gov/data). The WRF model outputs generated by the simulations in this study are saved on a long-term storage system at PNNL (rc-support@pnnl.gov)

References

Nesbitt, S. W., R. Cifelli, and S. A. Rutledge, 2006: Storm Morphology and Rainfall Characteristics of TRMM Precipitation Features. *Mon. Wea. Rev.*, **134**, 2702–2721, https://doi.org/10.1175/MWR3200.1.

Feng, Z., Leung, L. R., Liu, N., Wang, J., Houze, R. A., Li, J., et al. (2021). A global high-resolution mesoscale convective system database using satellite-derived cloud tops, surface precipitation, and tracking. *Journal of Geophysical Research: Atmospheres*, 126, e2020JD034202. https://doi.org/10.1029/2020JD034202

Sousa, A. C., L. A. Candido, and P. Satyamurty, 2021: Convective Cloud Clusters and Squall Lines along the Coastal Amazon. *Mon. Wea. Rev.*, **149**, 3589–3608, https://doi.org/10.1175/MWR-D-21-0045.1.

Schumacher, R.S., Rasmussen, K.L. The formation, character and changing nature of mesoscale convective systems. *Nat Rev Earth Environ* **1**, 300–314 (2020). https://doi.org/10.1038/s43017-020-0057-7

Feng, Z., Dong, X., Xi, B., Schumacher, C., Minnis, P., and Khaiyer, M. (2011), Top-of-atmosphere radiation budget of convective core/stratiform rain and anvil clouds from deep convective systems, *J. Geophys. Res.*, 116, D23202, doi:10.1029/2011JD016451.

Laurent, H., Machado, L. A. T., Morales, C. A., and Durieux, L., Characteristics of the Amazonian mesoscale convective systems observed from satellite and radar during the WETAMC/LBA experiment, *J. Geophys. Res.*, 107(D20), 8054, doi:10.1029/2001JD000337, 2002.

Petersen, W. A., S. W. Nesbitt, R. J. Blakeslee, R. Cifelli, P. Hein, and S. A. Rutledge, 2002: TRMM Observations of Intraseasonal Variability in Convective Regimes over the Amazon. *J. Climate*, **15**, 1278–1294, https://doi.org/10.1175/1520-0442(2002)015<1278:TOOIVI>2.0.CO;2.

Cifelli, R., Petersen, W. A., Carey, L. D., Rutledge, S. A., and da Silva Dias, M. A. F., Radar observations of the kinematic, microphysical, and precipitation characteristics of two MCSs in TRMM LBA, J. Geophys. Res., 107(D20), 8077, doi: ref="info:doi/doi:10.1029/2000JD000264">doi:10.1029/2000JD000264 2002.

Da Silva, N. A., & Haerter, J. O. (2023). The precipitation characteristics of mesoscale convective systems over Europe. Journal of Geophysical Research: Atmospheres, 128, e2023JD039045. https://doi.org/10.1029/2023JD039045.

Rickenbach, T. M., Ferreira, R. N., Halverson, J. B., Herdies, D. L., and Silva Dias, M. A. F., Modulation of convection in the southwestern Amazon basin by extratropical stationary fronts, J. Geophys. Res., 107(D20), doi:10.1029/2000JD000263, 2002.

Machado, L. A. T., Morales, C., & Avissar, R. (2004). The Large-Scale Biosphere-Atmosphere Experiment in Amazonia (LBA): Insights and future research needs. Journal of Geophysical Research: Atmospheres, 109(D20).

Burleyson, C. D., Z. Feng, S. M. Hagos, J. Fast, L. A. T. Machado, and S. T. Martin, 2016: Spatial Variability of the Background Diurnal Cycle of Deep Convection around the GoAmazon2014/5 Field Campaign Sites. *J. Appl. Meteor. Climatol.*, **55**, 1579–1598, https://doi.org/10.1175/JAMC-D-15-0229.1.

Rehbein, A, Ambrizzi, T, Mechoso, CR, Espinosa, SAI, Myers, TA. Mesoscale convective systems over the Amazon basin: The GoAmazon2014/5 program. *Int J Climatol*. 2019; 39: 5599–5618. https://doi.org/10.1002/joc.6173

Giangrande, S. E., Wang, D., and Mechem, D. B.: Cloud regimes over the Amazon Basin: perspectives from the GoAmazon2014/5 campaign, Atmos. Chem. Phys., 20, 7489–7507, https://doi.org/10.5194/acp-20-7489-2020, 2020.

Wang, D., Giangrande, S. E., Feng, Z., Hardin, J. C., & Prein, A. F. (2020). Updraft and downdraft core size and intensity as revealed by radar wind profilers: MCS observations and idealized model comparisons. *Journal of Geophysical Research: Atmospheres*, 125, e2019JD031774. https://doi.org/10.1029/2019JD031774

Tian, Y., Zhang, Y., Klein, S. A., & Schumacher, C. (2021). Interpreting the diurnal cycle of clouds and precipitation in the ARM GoAmazon observations: Shallow to deep convection transition. *Journal of Geophysical Research: Atmospheres*, 126, e2020JD033766. https://doi.org/10.1029/2020JD033766

Anselmo, E. M., C. Schumacher, and L. A. T. Machado, 2020: The Amazonian Low-Level Jet and Its Connection to Convective Cloud Propagation and Evolution. *Mon. Wea. Rev.*, **148**, 4083–4099, https://doi.org/10.1175/MWR-D-19-0414.1.

Carvalho, L. M. V., Silva Dias, M. A. F., & Fitzjarrald, D. R. (2002). Atmospheric conditions for deep convective clouds in the Amazon. Journal of Geophysical Research: Atmospheres, 107(D20).

Machado, L. A. T., Laurent, H., Gherboudj, I., Lima, A. P., Morales, C., Blakeslee, R., ... & Cifelli, R. (2004). The convective, turbulent, and microphysical characteristics of the Amazon boundary layer. Monthly Weather Review, 132(11), 2919-2931.

Prein, A. F., Ge, M., Valle, A. R., Wang, D., & Giangrande, S. E. (2022). Towards a unified setup to simulate mid-latitude and tropical mesoscale convective systems at kilometer-scales. Earth and Space Science, 9, e2022EA002295. https://doi.org/10.1029/2022EA002295

Paccini, L., & Stevens, B. (2023). Assessing precipitation over the Amazon basin as simulated by a storm-resolving model. *Journal of Geophysical Research: Atmospheres*, 128, e2022JD037436. https://doi.org/10.1029/2022JD037436

Silva Dias, M. A. F., et al., A case study of convective organization into precipitating lines in the Southwest Amazon during the WETAMC and TRMM-LBA, *J. Geophys. Res.*, 107(D20), 8078, doi:10.1029/2001JD000375, 2002.

Tai, S.-L., Feng, Z., Ma, P.-L., Schumacher, C., & Fast, J. D. (2021). Representations of precipitation diurnal cycle in the amazon as simulated by observationally constrained cloud-system resolving and global climate models. Journal of Advances in Modeling Earth Systems, 13, e2021MS002586. https://doi.org/10.1029/2021MS002586

Feng, Z., Leung, L. R., Houze, R. A., Hagos, S., Hardin, J., Yang, Q., et al. (2018). Structure and evolution of mesoscale convective systems: Sensitivity to cloud microphysics in convection-permitting simulations over the United States. *Journal of Advances in Modeling Earth Systems*, 10, 1470–1494. https://doi.org/10.1029/2018MS001305

Prein, A. F., Ge, M., Valle, A. R., Wang, D., & Giangrande, S. E. (2022). Towards a unified setup to simulate mid-latitude and tropical mesoscale convective systems at kilometer-scales. *Earth and Space Science*, 9, e2022EA002295. https://doi.org/10.1029/2022EA002295

Prein A. F., Rasmussen R. M., Wang D. and Giangrande S. E. (2021). Sensitivity of organized convective storms to model grid spacing in current and future climates. Phil. Trans. R. Soc. A. 379:20190546.

Ramos-Valle, A. N., Prein, A. F., Ge, M., Wang, D., & Giangrande, S. E. (2023). Grid spacing sensitivities of simulated mid-latitude and tropical mesoscale convective systems in the convective gray zone. *Journal of Geophysical Research: Atmospheres*, 128, e2022JD037043. https://doi.org/10.1029/2022JD037043

Yang, Z., Varble, A., Berg, L. K., Qian, Y., Tai, S.-L., Chen, J., et al. (2023). Sensitivity of precipitation displacement of a simulated MCS to changes in land Surface conditions. *Journal of Geophysical Research: Atmospheres*, 128, e2022JD037642. https://doi.org/10.1029/2022JD037642

Laurent, H., Machado, L. A. T., Morales, C. A., and Durieux, L., Characteristics of the Amazonian mesoscale convective systems observed from satellite and radar during the WETAMC/LBA experiment, *J. Geophys. Res.*, 107(D20), 8054, doi:10.1029/2001JD000337, 2002.

Anselmo, EM, Machado, LAT, Schumacher, C, Kiladis, GN. Amazonian mesoscale convective systems: Life cycle and propagation characteristics. Int J Climatol. 2021; 41: 3968–3981. https://doi.org/10.1002/joc.7053

Huang, X., Hu, C., Huang, X. *et al.* A long-term tropical mesoscale convective systems dataset based on a novel objective automatic tracking algorithm. *Clim Dyn* **51**, 3145–3159 (2018). https://doi.org/10.1007/s00382-018-4071-0

Feng Z., L. Leung, N. Liu, J. Wang, R.A. Houze, J. Li, and J.C. Hardin, et al. 2021. "A Global High-resolution Mesoscale Convective System Database using Satellite-derived Cloud Tops, Surface Precipitation, and Tracking." *Journal of Geophysical Research: Atmospheres* 126, no. 8:e2020JD034202. PNNL-SA-157800. doi:10.1029/2020JD034202

Feng Z., J.C. Hardin, H.C. Barnes, J. Li, L. Leung, A.C. Varble, and Z. Zhang. 2023. "PyFLEXTRKR: a Flexible Feature Tracking Python Software for Convective Cloud Analysis." *Geoscientific Model Development* 16, no. 10:2753-2776. PNNL-SA-179219. doi:10.5194/gmd-16-2753-2023

Rehbein, A., Ambrizzi, T. and Mechoso, C.R. (2018), Mesoscale convective systems over the Amazon basin. Part I: climatological aspects. Int. J. Climatol, 38: 215-229. https://doi.org/10.1002/joc.5171

Galarneau, T. J., Zeng, X., Dixon, R. D., Ouyed, A., Su, H., & Cui, W. (2023). Tropical mesoscale convective system formation environments. *Atmospheric Science Letters*, 24(5), e1152. https://doi.org/10.1002/asl.1152

Prein, A.F., Liu, C., Ikeda, K. *et al.* Simulating North American mesoscale convective systems with a convection-permitting climate model. *Clim Dyn* **55**, 95–110 (2020). https://doi.org/10.1007/s00382-017-3993-2

Emanuel, K., 2023: On the Physics of High CAPE. *J. Atmos. Sci.*, **80**, 2669–2683, https://doi.org/10.1175/JAS-D-23-0060.1.

Martin, S. T., Artaxo, P., Machado, L. A. T., Manzi, A. O., Souza, R. A. F., Schumacher, C., Wang, J., Andreae, M. O., Barbosa, H. M. J., Fan, J., Fisch, G., Goldstein, A. H., Guenther, A., Jimenez, J. L., Pöschl, U., Silva Dias, M. A., Smith, J. N., and Wendisch, M.: Introduction: Observations and Modeling of the Green Ocean Amazon (GoAmazon2014/5), Atmos. Chem. Phys., 16, 4785–4797, https://doi.org/10.5194/acp-16-4785-2016, 2016.

Na, Y., Fu, Q., Leung, L. R., Kodama, C., & Lu, R. (2022). Mesoscale convective systems simulated by a high-resolution global nonhydrostatic model over the United States and China. *Journal of Geophysical Research: Atmospheres*, 127, e2021JD035916. https://doi.org/10.1029/2021JD035916

Cheng, W. Y. Y., and W. R. Cotton, 2004: Sensitivity of a Cloud-Resolving Simulation of the Genesis of a Mesoscale Convective System to Horizontal Heterogeneities in Soil Moisture Initialization. *J. Hydrometeor.*, **5**, 934–958, https://doi.org/10.1175/1525-7541(2004)005<0934:SOACSO>2.0.CO;2.

National Centers for Environmental Prediction/National Weather Service/NOAA/U.S. Department of Commerce. (2000). NCEP FNL Operational Model Global Tropospheric Analyses, continuing from July 1999 (Updated daily) [Dataset]. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. https://doi.org/10.5065/D6M043C6. Accessed† 01 Apr 2020.

Andreas Franz Prein, Zhe Feng, Thomas fiolleau, et al. Km-Scale Simulations of Mesoscale Convective Systems (MCSs) Over South America - A Feature Tracker Intercomparison. *ESS Open Archive* . October 27, 2023. DOI: 10.22541/essoar.169841723.36785590/v1

Da Silva, N. A., & Haerter, J. O. (2023). The precipitation characteristics of mesoscale convective systems over Europe. *Journal of Geophysical Research: Atmospheres*, **128**, e2023JD039045. https://doi.org/10.1029/2023JD039045

Liu, W., Ullrich, P. A., Li, J., Zarzycki, C., Caldwell, P. M., Leung, L. R., & Qian, Y. (2023). The June 2012 North American derecho: A testbed for evaluating regional and global climate modeling systems at cloud-resolving scales. *Journal of Advances in Modeling Earth Systems*, **15**, e2022MS003358. https://doi.org/10.1029/2022MS003358

Tang, Q., Prather, M. J., Hsu, J., Ruiz, D. J., Cameron-Smith, P. J., Xie, S., and Golaz, J.-C. (2021). Evaluation of the interactive stratospheric ozone (O3v2) module in the E3SM version 1 Earth system model, *Geosci. Model Dev.*, **14**, 1219–1236, https://doi.org/10.5194/gmd-14-1219-2021

Han, B., Fan, J., Varble, A., Morrison, H., Williams, C. R., Chen, B., et al. (2019). Cloud-resolving model intercomparison of an MC3E squall line case: Part II. Stratiform precipitation properties. *Journal of Geophysical Research: Atmospheres*, **124**, 1090–1117. https://doi.org/10.1029/2018JD029596

Varble, A., H. Morrison, and E. Zipser, 2020: Effects of Under-Resolved Convective Dynamics on the Evolution of a Squall Line. *Mon. Wea. Rev.*, **148**, 289–311, https://doi.org/10.1175/MWR-D-19-0187.1.

Houze, R. A. (2014). Cloud Dynamics (2nd ed.): Elsevier/Academic Press, Oxford.

Cui, W., X. Dong, B. Xi, Z. Feng, and J. Fan, 2020: Can the GPM IMERG Final Product Accurately Represent MCSs' Precipitation Characteristics over the Central and Eastern United States?. J. Hydrometeor., 21, 39–57, https://doi.org/10.1175/JHM-D-19-0123.1.

Zhang, Z., A. Varble, Z. Feng, J. Hardin, and E. Zipser, 2021: Growth of Mesoscale Convective Systems in Observations and a Seasonal Convection-Permitting Simulation over Argentina. Mon. Wea. Rev., 149, 3469–3490, https://doi.org/10.1175/MWR-D-20-0411.1.

Ayat, H., J. P. Evans, S. Sherwood, and A. Behrangi, 2020: Are Storm Characteristics the Same When Viewed Using Merged Surface Radars or a Merged Satellite Product?. J. Hydrometeor., 22, 43–62, https://doi.org/10.1175/JHM-D-20-0187.1.

Guilloteau, C., Foufoula-Georgiou, E. (2020). Multiscale Evaluation of Satellite Precipitation Products: Effective Resolution of IMERG. In: Levizzani, V., Kidd, C., Kirschbaum, D., Kummerow, C., Nakamura, K., Turk, F. (eds) Satellite Precipitation Measurement. Advances in Global Change Research, vol 69. Springer, Cham. https://doi.org/10.1007/978-3-030-35798-6_5