Diurnal Cycle of Precipitation in the Amazon: Contrasting Observationally Constrained Cloud-System Resolving and Global Climate Models

Sheng-Lun Tai[1], Zhe Feng[1], Po-Lun Ma[1], Courtney Schumacher[2], and Jerome Fast[3]

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

November 22, 2022

Abstract

The ability of an observationally-constrained cloud-system resolving model (Weather Research and Forecasting; WRF, 4-km grid spacing) and a global climate model (Energy Exascale Earth System Model; E3SM, 1-degree grid spacing) to represent the precipitation diurnal cycle over the Amazon basin during the 2014 wet season is assessed. The month-long period is divided into days with and without the presence of observed propagating mesoscale convective systems (MCSs) over the central Amazon. The MCSs are strongly associated with rain amounts over the basin and also control the observed spatial variability of the diurnal rain rate. WRF model coupled with a 3-D variational data assimilation scheme reproduces the spatial variability of the precipitation diurnal cycle over the basin and the lifecycle of westward propagating MCSs initiated by the coastal sea-breeze front. In contrast, a single morning peak in rainfall is produced by E3SM for simulations with and without nudging the large-scale winds towards global reanalysis, indicating precipitation in E3SM is largely controlled by local convection associated with diurnal heating. Both models produce contrast in easterly wind profiles between days with and without MCS that are similar to data collected by U.S. DOE Atmospheric Radiation Measurement (ARM) facility during the Green Ocean Amazon (GoAmazon2014/5) campaign and other operational radiosondes. A multivariate perturbation analysis indicates the dryness of low-level air transported from ocean to inland has higher impact on the formation and maintenance of MCS in the Amazon than other processes.

Diurnal Cycle of Precipitation in the Amazon: Contrasting Observationally Constrained Cloud-System Resolving and Global Climate Models

[1]Sheng-Lun Tai, [1]Zhe Feng, [1]Po-Lun Ma, [2]Courtney Schumacher, [1]Jerome D. Fast

[1]Pacific Northwest National Laboratory, Richland, Washington, USA
[2]Department of Atmospheric Sciences, Texas A&M University, College Station, Texas, USA

Corresponding author: Dr. Sheng-Lun Tai

Atmospheric Sciences and Global Change Division

Pacific Northwest National Laboratory

902 Battelle Boulevard, Richland, WA 99352

E-mail: sheng-lun.tai@pnnl.gov

Key Points :

- 1 Spatial variability of the precipitation diurnal variation in the Amazon is mostly reproduced by WRF but not E3SM
- 2 Ambient environments are better simulated by WRF than E3SM as the convective processes have significant impact and are resolved by WRF
- 3 Intrusion of cooler and dryer sea breeze front into Amazon supports formation of propagating convection

Abstract

The ability of an observationally-constrained cloud-system resolving model (Weather Research and Forecasting; WRF, 4-km grid spacing) and a global climate model (Energy Exascale Earth System Model; E3SM, 1-degree grid spacing) to represent the precipitation diurnal cycle over the Amazon basin during the 2014 wet season is assessed. The month-long period is divided into days with and without the presence of observed propagating mesoscale convective systems (MCSs) over the central Amazon. The MCSs are strongly associated with rain amounts over the basin and also control the observed spatial variability of the diurnal rain rate. WRF model coupled with a 3-D variational data assimilation scheme reproduces the spatial variability of the precipitation diurnal cycle over the basin and the lifecycle of westward propagating MCSs initiated by the coastal sea-breeze front. In contrast, a single morning peak in rainfall is produced by E3SM for simulations with and without nudging the large-scale winds towards global reanalysis, indicating precipitation in E3SM is largely controlled by local convection associated with diurnal heating. Both models produce contrast in easterly wind profiles between days with and without MCS that are similar to data collected by U.S. DOE Atmospheric Radiation Measurement (ARM) facility during the Green Ocean Amazon (GoAmazon2014/5) campaign and other operational radiosondes. A multivariate perturbation analysis indicates the dryness of low-level air transported from ocean to inland has higher impact on the formation and maintenance of MCS in the Amazon than other processes.

Plain Language Summary

The Amazon basin in South America is one of the regions over land that has the highest occurrence of large-size and deep cloud systems (also called "Mescoscale Convective System" (MCS)). Since they have a wide coverage and produce much heavier rainfall than the other types of cloud, the regional climate and even the earth system are tied closely with their behaviors. However, current global atmospheric models are unable to reproduce realistic diurnal variation of precipitation in the Amazon and the poor representation of those MCSs is responsible for the deficiency. We use various observations as the reference to understand how accurate the physical processes related to MCS are represented by both the cloud-system resolving (higher-resolution) and global climate (lower-resolution) models. The results show the diurnal variation of local precipitation in the basin is mostly reproduced by cloud-system resolving model but not the global climate model, because the propagating MCSs and many related processes can only be simulated by using higher-resolution model. It is also found the dryness of low-level air transported from ocean to inland has the highest impact on the formation and maintenance of MCS in the Amazon.

1. Introduction

The Amazonia region in South America is recognized as one of the world's hot spots of convective activity (Nesbitt et al. 2000; Liu and Zipser 2013; Houze 2004). Since cloud populations associated with organized convective systems have significant impacts on radiation as well as produce significant amount of rainfall, the Earth's energy budget and water cycle is strongly modulated by deep convection in the Amazon. The Amazon basin is one of the regions over land with the most frequent mesoscale convective systems (MCSs), comparable to the Maritime Continents (Feng et al. 2021). MCSs account for over 50% of annual rainfall in the Amazon, during the wet season (December – May) the percentage over central and western Amazon increase to over 60%. It is therefore important to accurately simulate characteristics of convective precipitation, particularly those associated with MCSs in the Amazon, including the diurnal variations in intensity, frequency, and duration to better understand regional and global climate.

A long-standing issue of current global atmospheric models, however, is their inability to reproduce realistic diurnal variations of precipitation (Khairoutdinov and Randall 2006; Dai 2006; Dai and Trenberth 2004; Betts 2002; Xie et al. 2019; Rasch et al. 2019; Suhas and Zhang 2014; Guichard et al. 2004). For example, global climate models have a tendency to simulate an early onset of convective precipitation over land, which produces a peak rain intensity around local noon as opposed to late afternoon that is commonly observed in many locations (e.g. Bechtold et al. 2004; Xie et al. 2004; Betts and Jakob 2002; Fiedler et al. 2020). The early release of convective potential available energy in models could also lead to greatly reduced rainfall intensity. Several studies have suggested that exaggerated coupling between surface heating and convection is one factor responsible for causing errors in the timing of peak precipitation rates (Xie et al. 2019;

Zhang 2003, 2002). This factor may subsequently lead to a failure to capture elevated nocturnal precipitation which is decoupled from near-surface processes (Marsham et al. 2011).

Higher-resolution models such as convection-permitting and cloud-resolving models have been used to simulate the diurnal cycle of precipitation over regional spatial scales (Gao et al. 2017; Pearson et al. 2014; Hassim et al. 2016; Konduru and Takahashi 2020; Zhang et al. 2016; Clark et al. 2007; Love et al. 2011). Compared to global models, regional-scale models have more detailed and realistic representation of cloud processes and are able to better resolve the evolution of organized convection, which leads to improved precipitation simulations. Analysis of regional-scale simulations coupled with observations are also useful to inform how to optimize convective parameterizations used by global models. Diurnal variation of precipitation simulated by regional models is shown to be generally improved for regions affected by organized propagating convection; nevertheless, this result is not guaranteed for all cases. For instance, there are issues of excessive production of strong small-scale convection when convection is resolved explicitly (Kendon et al. 2012; Roberts and Lean 2008; Arnold et al. 2020). Moreover, Arnold et al. (2020) also reported problems of insufficient growth of cloud clusters in comparison with observations when conducting simulations using a global nonhydrostatic model with a grid spacing of 3.5 km. A number of remaining modeling issues have been raised, including misrepresentations of terrain, microphysical processes, and cold pool evolution that suggest accurate modeling of diurnal precipitation cycle is still a challenge for convection-permitting or cloud-resolving models.

In terms of improved treatments of atmospheric processes in models, observations play a crucial role to inform what processes are missing or poorly parameterized. Due to the vastness and relative inaccessibility of the Amazon tropical rainforest, in situ measurements of atmospheric properties needed for model assessment has been a challenging task. Improvements in remote satellite retrievals over the past two decades thus play a key role in quantifying and understanding environmental conditions in the Amazon. For example, the Global Precipitation Measurement (GPM, Huffman et al., 2014), successor of the Tropical Rainfall Measuring Mission (TRMM, Huffman et al. 2007), has been successful in providing quasi-global, high-quality, and fine resolution rainfall estimates. Specifically, Tan et al. (2019b) recently demonstrated an improvement in quantifying the diurnal precipitation cycle by using the latest version of Integrated Multi-satellitE Retrievals for GPM (IMERG). The global coverage of IMERG provides an opportunity to investigate precipitation characteristics over remote regions, including the Amazonian rainforest.

Despite the logistical challenges, several field campaigns have been conducted at different locations in the Amazon (Harriss et al. 1990; Dias et al. 2002). The Green Ocean Amazon experiment (GoAmazon2014/5) is among the most recent campaigns which collected extensive atmospheric observations by using a range of advanced instruments (Martin et al. 2016). Studies that use observational data from GoAmazon2014/5, including (Schiro et al. 2018; Burleyson et al. 2016; Ghate and Kollias 2016; Collow et al. 2016; Giangrande et al. 2017; Zhuang et al. 2017; Tang et al. 2016) have provided insights on the cloud life cycle, diurnal precipitation cycle, and large-scale environmental control on clouds over the central Amazon. Some of these studies (e.g., Burleyson et al. 2016) coupled field-campaign periods with long-term satellite observations to provide insights into the climatology of the spatiotemporal variability of convection over the central Amazon basin.

Global reanalysis data is often treated as an integrative "observation" used to interpret large-scale environmental conditions when measurements are not readily available (Espinoza et al. 2012; Rehbein et al. 2019; Anselmo et al. 2020; Oliveira and Oyama 2015). Since the reanalysis product is obtained by blending observations with coarse resolution global atmospheric model predictions, it may not be adequate to represent physical processes at spatial and temporal scales relevant to convective clouds. Previous studies that evaluated diurnal precipitation cycles over Amazon region with reanalysis data and large-scale atmospheric models (Betts and Jakob 2002; Itterly and Taylor 2014; Itterly et al. 2018) reported consistent deficiencies in reanalysis data such as peak rainfall occurring too early in the day, much weaker amplitude of diurnal cycle, and lack of propagating convection systems. For example, Itterly and Taylor (2014) showed that the error associated with convective precipitation was linked to errors in the top of atmosphere radiative flux of the reanalysis products. This suggests the necessity of using higher-resolution convection-permitting models to better examine convective precipitation processes.

Several earlier studies have conducted convection-permitting simulations encompassing portions of the Amazon basin. However, most of these studies focused on responses to climate change (Langenbrunner et al. 2019), aerosol-radiation-cloud interactions (Archer-Nicholls et al. 2016), air pollutants transport (Rafee et al. 2017), or atmospheric chemistry (Shrivastava et al. 2019), rather than examining the physical processes associated with propagating convective systems. To adequately simulate convective-scale processes over the entire Amazon for a long period requires a sufficiently large domain to represent the coupling of large-scale environmental conditions and lifecycle of propagating convection to avoid contamination of artificial disturbances generated near the model lateral boundary. On the other hand, for a large domain, ambient environmental conditions could slowly drift away from real states over several days without any constraint. A common workaround is to reinitialize models every one or two days, but this strategy has drawbacks such as (1) an adequate spin-up integration period is needed for each forecast and (2) hydrometeors are reinitialized and discontinuity in clouds and precipitation often occur when analyzing concatenated simulations that hamper interpretations.

To address these issues, we take advantage of a data assimilation technique to constrain the large-scale atmospheric conditions in the Weather Research and Forecasting (WRF, Skamarock et al. 2008) model at convection-permitting scales to simulate the precipitation diurnal cycle resulting from propagating convection across the Amazon basin during the wet season. The cloud hydrometeors are also retained whenever the simulation is reinitialized by optimized meteorological fields produced from data assimilation. The impact of data assimilation on predictions of organized convection and precipitation has been examined by numerous studies primarily for mid-latitude regions with strong synoptic forcing (Schwartz and Liu 2014; Schwartz et al. 2015; Tai et al. 2020; Clark et al. 2016; Bauer et al. 2015; Gustafsson et al. 2018; Trier et al. 2015), but fewer studies have assessed its impact in tropical regions with relatively weak synoptic forcing and where the density and frequency of in situ measurements is smaller. As mentioned earlier, convective cloud processes and the precipitation diurnal cycle are not adequately represented by global atmospheric models due to their coarse resolution and oversimplified parameterization (Xie et al. 2019), leading to uncertainties in the global energy budget and water cycle (Betts and Jakob 2002; Itterly and Taylor 2014; Genio and Wu 2010; Bergman and Salby 1996). Therefore, we also examine the performance of the new U.S. DOE Energy Exascale Earth System Model (E3SM, Golaz et al. 2019) in simulating the diurnal cycle over the Amazon basin which has not yet been assessed in detail.

This paper is organized as follows. The analysis domain and period along with sources of observational data are described in section 2. A brief description of the WRF and E3SM models as well as details of corresponding experiments conducted for this study are provided in section 3.

In section 4, analysis of the simulated precipitation diurnal cycle is presented in the context of the observations. Characteristics and spatiotemporal variability of the ambient flows in relation to the associated precipitation cycle are also investigated. Finally, a summary of the findings is given in section 5.

2. Domain and observations

2.1 Area and period of study

The study area encompasses most of the Amazon basin with the center located at Manaus, Brazil as depicted in Figure 1. Since MCSs in Amazonian region typically propagate westward across the Amazon basin (Feng et al. 2021) with its origin at the northeastern coast of Brazil and within the central Amazon basin, our domain is large enough to alleviate possible issues introduced at the domain boundary as well as allowing the spin-up and propagation of weather systems originating from the Atlantic Ocean. The simulation period is from March 11 to April 10 during the wet season of 2014 that includes numerous intense rainfall events. As shown by previous climatological studies (e.g., Tanaka et al. 2014; Burleyson et al. 2016), March and April are usually the months with the highest accumulated precipitation during the year.

2.2 Observational datasets

2.2.1 IMERG V06

The National Aeronautics and Space Administration (NASA) Integrated Multi-Satellite Retrievals for Global Precipitation Measurement (IMERG) V06 (Huffman et al. 2019) is one observational precipitation product used in this study. In IMERG V06 several improvements are introduced to address some of the issues discovered in earlier products, including the changes to the time-interpolation algorithm (Tan et al. 2019a,b). This product has a grid spacing of 0.1° and is 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). While the IMERG product has been widely used in hydrological and atmospheric research and has been demonstrated its crucial role in many precipitation-related studies (Moazami and Najafi 2021; Oliveira et al. 2016; Mandapaka and Lo 2020; Derin et al. 2019; Huang et al. 2020), there are still uncertainties associated with propagating precipitation influenced by orography as well as temporal interpolation that primarily relies on the frequency and quality of satellite data (Tan et al. 2019b).

2.2.2 SIPAM S-band radar rainfall estimation

The rainfall estimates from the distributed System for the Protection of Amazon (SIPAM) S-band conventional Doppler radar at Manaus, processed by Texas A & M University, are also used in this study (Schumacher and Funk 2018). The rainfall estimates were obtained through the CAPPI (Constant Altitude Plan Position Indicator) product which has a maximum detecting range of 240 km (2-km horizontal grid spacing), vertical levels between 0.5 and 20 km (0.5-km vertical grid spacing), and is available at ~12 minute intervals. Spurious data such as noise and ground clutter of the reflectivity field has been corrected. A time-dependent calibration constant derived from a comparison with TRMM precipitation radar data was applied to the CAPPI files at different periods. Rainfall estimates were then generated using the radar corrected reflectivity data at the 2.5-km CAPPI level within a radius of 160 km of coverage. A power law relation between radar reflectivity and rain rate was fitted based on disdrometer observations during GoAmazon2014/5 such that:

$$Z = 174.8\,R^{1.56}$$

where Z is the radar reflectivity factor ($mm^6\ m^{-3}$) and R is the radar rain rate ($mm\ h^{-1}$). Some of the radar beams were found to be contaminated by ground clutter; therefore, unrealistic rainfall estimations on those beams were removed using quality control masks. An hourly radar rainfall product was also produced from the 12-min data with a horizontal grid spacing of 2 km.

2.2.3 Meteorological observations

As one of the major participants of GoAmazon2014/5, the Atmospheric Radiation Measurement (ARM) Mobile Facility (AMF, Miller et al. 2016) collected a unique set of observations of meteorology and aerosol properties near Manacapuru, west of Manaus in the central Amazon basin. Its geographic location is very close to the white dot within subregion "M" in Figure 1 and named the "T3" site hereafter as in Martin et al. 2016. Observations at the T3 site were collected between February 2014 to December 2015. We use ARM radiosonde, Doppler lidar, and surface station measurements for model assessment and analysis of environmental conditions associated with propagating convection described in Section 4.3.

During the GoAmazon2014/5 campaign period, radiosondes were launched at the T3 site at 6-hour intervals (06, 12, 18, 00 UTC; or 02, 08, 14 and 20 LT) to obtain tropospheric profiles of wind, temperature, and humidity. During the intensive observational periods (IOP), one additional radiosonde was launched at 15 UTC (11 LT) to enhance diurnal coverage. In addition to ARM's radiosondes, meteorological profiles measured at three other sites (also denoted in Figure 1) from NOAA's National Centers for Environmental Information (NCEI) Integrated Global Radiosonde Archive (IGRA, Durre et al. 2006) are used for data assimilation simulations and also assessment of model performance. These data have lower temporal frequency and vertical resolution compared to ARM's radiosonde profiles. Surface meteorological data was also collected at the T3 site throughout the campaign. Surface horizontal wind components, temperature, and specific humidity are used to quantify average diurnal variability during the wet season IOP. Table 1 summarizes all types of observations used in this study.

3. Model descriptions

3.1 Weather Research and Forecasting (WRF) model

3.1.1 Model configuration

The WRF model version 3.9.1 (ARW, Skamarock et al. (2008)) is used to conduct the regional atmospheric simulations of convective precipitation. The domain shown in Figure 1 encompasses the continent of South America north of 20°S and includes the entire Amazon basin and adjacent oceans. The domain uses a horizontal grid spacing of 4 km (1450 x 950 grid points) and a stretched vertical coordinate with 60 levels up to the model top at 100 hPa. There are approximately 12 model levels between the surface and 2 km. The model simulations use the 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). A cumulus parametrization scheme is not activated because the model's horizontal grid spacing (4 km) is considered capable of resolving MCSs at the storm system level (Prein et al. 2021), which are the targets of current study.

The NCEP FNL operational model global tropospheric analysis with a 1° grid spacing is used to initialize the model's atmosphere and soil variables. Analyses at 6-hour intervals are linearly interpolated in time for the model's lateral boundary conditions. Land-use data is obtained from the Moderate Resolution Imaging Spectroradiometer (MODIS)-based dataset available at a 1-km grid spacing and using the International Geosphere-Biosphere Programme (IGBP) land cover type classification.

3.1.2 Data assimilation and experiments

The version 3.6 Community Gridpoint Statistical Interpolation (GSI, Shao et al. 2015) software package is utilized to assimilate available observational data including conventional observations (e.g., radiosonde profiles, surface meteorology, aircraft, ship and others) and satellite radiances and derived properties. The GSI software package includes several techniques such as three-dimensional variational (3DVar; Wu et al. 2002), three-dimensional and four-dimensional ensemble-variational hybrid (3DEnVar, 4DEnVar; (Hamill et al. 2011; Wang 2010; Wang and Lei 2014)), and the Ensemble Kalman Filter (EnKF; Zhu et al. 2013). The 3DVar technique is chosen for this study since it has a much lower computational cost than other data assimilation techniques. Earlier studies show that ensemble-variational hybrid data assimilation techniques (3DEnVar and 4DEnVar) overall outperform 3DVar, but the performance of our own test case using 3DEnVar was only a slight improvement over 3DVar. We argue this is likely due to much lower density and frequency of available observations in the Amazon than in other data-rich regions such as North America. In 3DVar, the background error is static and computed globally using forecasts from the NCEP's North American Mesoscale Forecast System (NAM) model. Default localization parameters suggested in GSI are applied in adjustment for final background error covariances.

The schematic diagram in Figure 2 displays how WRF simulations are carried out, with the DA-coupled simulation named "WRF_DA" illustrated in the upper part of the figure. WRF is initialized at 00 UTC of March 11, 2014 and then runs for 12 hours. The first data assimilation is performed at 12 UTC of March 11 by blending the simulation with observations from NCEP GDAS data stream (http://rda.ucar.edu/datasets/ds337.0/) as well as the ARM T3 site radiosonde profiles. The yellow dots denoted in Figure 1 highlight how assimilated radiosonde profiles are distributed within the model domain in the example at 12 UTC of March 12, 2014. Note the number of observed profiles varies with time as not every site launches radiosondes at the same frequency. The zonal and meridional winds, specific humidity, temperature, and pressure are updated based on analyzed increments and the fields are used for the reinitialization of the subsequent 12-hour forecast. Therefore, the entire simulation period is partially reinitialized every 12 hours until 00 UTC of April 10. We note that ARM also launched radiosondes each day at 6, 15, and 18 UTC that were not included in the "WRF_DA" simulation since radiosondes were not available for other locations at these times. Instead, the additional ARM radiosondes are used for independent model evaluation purposes (Table 1).

To understand the impact of DA, a control "WRF_noDA" simulation that does not involve any DA is also performed as indicated in the bottom part of Figure 2. To prevent the simulated synoptic meteorology from drifting too far from the observed conditions, the entire month-long simulation is comprised of a series of overlapping short-term forecasts. For example, short-term 36-h forecasts are produced each day that are initialized at 00 UTC. To avoid spin-up issues, the first 12-hours of the simulation is discarded. Then, the entire simulation period is assembled by piecing together the remaining simulations as denoted by the blue shading in Figure 2. In this case, the discontinuity in hydrometeors variables may be noticeable between two adjacent simulations.

3.2 Energy Exascale Earth System Model (E3SM)

3.2.1 General description of model and configuration

We used the U.S. Department of Energy's Energy Exascale Earth System Model version 1 (E3SMv1) (Caldwell et al. 2019; Golaz et al. 2019) in this study to assess its ability to represent the diurnal variability of precipitation in the Amazon basin. Parameterization schemes employed in the E3SM atmosphere model (EAM) version 1 (EAMv1) (Rasch et al. 2019) are summarized in Table 2. The emissions of aerosols and their precursors (Hoesly et al. 2018) prepared for the

Coupled Model Intercomparison Project Phase 6 (CMIP6) (Eyring et al. 2016) are used in EAMv1 except that the emissions of marine dimethyl sulfide are based on Elliott (2009) and Wang et al. (2015).

EAMv1 uses a spectral element dynamical core to solve primitive equations on a cubed sphere grid (Dennis et al. 2012; Laprise 1992; Taylor and Fournier 2010). The model is configured to run at ne30np4 resolution, which means that it has 30x30 spectral elements in each cube face and each spectral element has 4x4 Gauss-Lobatto-Legendre points. The resulting equivalent horizontal grid spacing is ~1°. The model has 72 levels in the vertical using a hybrid sigma-pressure coordinate. The lowest model level is about 25 m thick at sea level and the model top is set to be 0.1 hPa.

We configured the model following the Atmospheric Model Intercomparison Project (AMIP) protocol (Gates et al. 1999), where the evolution of the atmosphere and land are simulated based on prescribed monthly mean SSTs and sea ice cover from observations. The horizontal winds in the atmosphere are nudged toward the Modern-Era Retrospective analysis for Research and Applications Version 2 (MERRA-2) with a 6-hour relaxation time scale (Sun et al. 2019; Zhang et al. 2014). The simulation named “E3SM_v1_nudge” starts from 2013-01-01 and continues until the end of 2015. To determine the impact of horizontal winds nudging, a similar simulation named “ES3M_v1_free” is performed without using analysis nudging during model integration.

4. Results and discussion

4.1 Diurnal cycle of precipitation over Amazon basin

4.1.1 General characteristics

To quantify the diurnal cycle of precipitation over a large portion of the Amazon basin (as shown by the dashed-rectangle area in Figure 1), the mean hourly rain rate is computed first from the IMERG product as well as the WRF and E3SM simulations between March 11 and April 10, 2014. Since many factors such as large-scale convergence, orographic lifting, MCSs and local surface-heating can possibly induce precipitation at any particular time, the mean rain rate is expected to filter weak signals while leaving stronger or more frequent signals. For simplicity in visualization, hourly results from the three sources are further averaged over four intervals 08 – 13, 14 – 19, 20 – 01, and 02 – 07, LT (00 – 05, 06 – 11, 12 –17, and 18 – 23 UTC) and illustrated in Figure 3.

In Figure 3a (08 – 13 LT), the IMERG results show that convection associated with the sea breeze (labeled as 1) forms along the northeastern coast of Brazil near local noon. Meanwhile, convection triggered within the Amazon basin or some remnants from the sea breeze systems formed on previous day propagates southwestward toward the T3 site (labeled as 2). During 14 – 19 LT (Figure 3b), precipitation associated with the sea breeze convection increases and begins propagating inland (labeled as 1), driven by the northeasterly trade winds. Precipitation also increases over much of the Amazon as daytime heating enhances convection across the basin. During the evening and early morning hours (20-07 LT, Figures 3c and d) the regions of convection that formed over the coast and intensified over central Amazon during the day mature and propagate to the southwest, while overall precipitation amounts decrease, particularly over the central Amazon basin where the T3 site locates within. It should be noted that not all the convective systems that form on the coast during the day propagate across the Amazon during the evening hours (e.g., Garstang et al. 1994, Cohen et al. 1995) as a number of studies have shown that coastal systems may decay within a few hundred kilometers of the coast (Greco et al. 1990; Alcântara et al. 2011; Anselmo et al. 2021). It is most likely the disturbances generated by the sea breeze system can occasionally propagate across the basin when the environmental conditions favor the sustainability of convective growth and produce precipitation as the disturbances propagate.

Simulated hourly rain rates from both the WRF model and E3SM are processed by adopting the identical procedure as the IMERG data and illustrated in Figure 3. Here we show only results from the WRF_DA and E3SM_v1_nudge simulations as the overall pattern are similar within the same modeling system compared to WRF_noDA and E3SM_v1_free, respectively. The comparisons between the IMERG and WRF_DA panels indicate the major characteristics of diurnal precipitation cycle are qualitatively simulated by the cloud-system resolving model. Specifically, the precipitation characteristics simulated by WRF are similar to IMERG in terms of the intensity, spatial pattern, and diurnal phase shifts. In contrast, E3SM simulates much less precipitation with little spatial variation than IMERG and WRF. Instead of producing a propagating rainfall pattern, precipitation from E3SM has the overall peak rainfall rate occurring during morning hours (Figure 3i), which is likely due to the dominant role of local convection driven by solar heating in the global climate model. The representation of propagating convective systems is missing in the current E3SM configuration.

We also compute the accumulated rainfall over the entire study period (March 11 to April 10 of 2014) to illustrate the performance of these two models in reproducing rainfall amounts over the month-long period. Figures S1a, b, and c in supporting information display the total accumulated rainfall over the study domain from IMERG, WRF_DA, and E3SM_v1_nudge, respectively. The minimum in rainfall over the northern part of the domain is reproduced by both WRF and E3SM. Outside of this region, the total rainfall from IMERG varies from 200 to nearly 600 mm. Both models produce a similar range in rainfall accumulation, although there are localized differences in the spatial distribution. This comparison suggests that E3SM has skill in estimating longer-term accumulated precipitation in this region during the wet season despite struggling to generate propagating convective systems that contribute to the diurnal cycle of precipitation.

To visualize the spatial distribution of those differences, IMERG is subtracted from WRF and E3SM rainfall in Figures S1d and e, respectively. Since the three sources have different spatial resolutions, the WRF and E3SM results were reapportioned onto the IMERG grid. In addition, the IMERG product also has its own potential bias at any particular location and time of a day. Thus, some caution is warranted in the magnitude of the rainfall biases at a given location. While the WRF model overestimates precipitation near the western edge and southeast corner of the domain, the simulated precipitation is underestimated over most of the rest of the domain compared to the IMERG data (Figure S1d). While the E3SM overestimation extends more into the central part of the domain (Figure S1e), the spatial distribution of the WRF and E3SM bias is qualitatively similar, and the domain averaged rainfall amount is comparable among the three data sources.

4.1.2 Difference between days with and without propagating convection

Based on the observational analysis in the previous section, models need to adequately represent the lifecycle of propagating mesoscale convective systems (MCSs) in the Amazon to realistically reproduce the observed diurnal precipitation cycle. To further investigate the influence of propagating MCS on precipitation in Amazon basin, we split all days in the period into two groups that consist of days with and without MCS that pass over Manaus near the domain center. After a subjective examination of IMERG hourly rain rate maps, there are 14 days (March 11 to 16, 18, 26, 28 to 30, and April 4, 6, and 7) identified without occurrence of MCS (named "noMCS" hereafter). The rest of the 17 days are then classified as the "MCS" group. This classification of days is applied to the WRF and E3SM simulations for consistency.

To provide another perspective of the observed and simulated diurnal precipitation cycle, we now examine the spatial distribution of the time of day when the peak precipitation is produced. First, the mean hourly rain rate distributions are computed separately over the MCS and noMCS periods. Then the hour of day with the maximum rain rate is identified for each grid cell over the domain. Identical processing is applied to IMERG, WRF, and E3SM datasets and the composite distributions are given in Figure 4 for the time of maximum rain rates.

A well-defined rainbow pattern can be identified in in Figure 4a, similar to previous studies of diurnal patterns across the Amazon (e.g., Yang and Slingo 2001, Dupuis and Schumacher 2018). This result implies that the observed diurnal cycle of precipitation highly correlates with propagating MCS which is in association with disturbances that triggered by sea breeze and advected inland. In contrast, the IMERG noMCS group (Figure 4b) exhibits more irregular spatial distributions for the daily maximum hour. The diurnal precipitation cycle in the basin is therefore more complicated and random during periods of weaker synoptic forcing without propagating MCSs.

The maximum hour of precipitation from the WRF simulation for the MCS group is very similar to the IMERG distribution (Figure 4a), suggesting that the model is capable of reproducing the major features of the diurnal precipitation cycle associated with propagating convective systems over the Amazon basin. However, for the noMCS group there are larger differences between the WRF and IMERG distributions (Figure 4b). This implies precipitation tends to be more randomly distributed when synoptic forcing is weak, and therefore more unpredictable.

For the E3SM model, the hour of maximum precipitation agrees with the IMERG product only over the northeastern corner of the domain (Figure 4). Over the rest of the domain, the spatial variability is significantly reduced, and the simulated maximum rain rate occurs almost exclusively between 8 and 14 LT. Furthermore, the contrast between MCS and noMCS groups is less evident in the E3SM simulation than in the IMERG product and the WRF simulation. This suggests that precipitation in E3SM is an outcome of the convective parameterization that preferentially triggers convective clouds around local noon during the highest local heating associated with the diurnal cycle of incoming solar radiation. Since convection near the coast associated with the sea breeze occurs during the afternoon, the E3SM predictions agree better with IMERG product over the northeastern corner of the domain. In addition, the convective parameterization does not contain any memory of convective clouds so propagating systems cannot be represented. A recent study (Xie et al. 2019) proposed a revision to the convective triggering function implemented in the Zhang-MacFarlane scheme (1995), which may have positive impact on this metric, but this revision is not included in this study. While revised triggering function had a positive impact on the overall timing of maximum convective precipitation rate in many tropical regions, its impact on representing propagating convective systems in the Amazon basin was not investigated. With respect to all the panels in Figure 4, we found that the diurnal cycle of precipitation over the northeastern corner of domain near the coast is more predictable by the models since this region is dominated by local land-sea contrasts and propagating MCS are of lesser importance.

4.2 Spatial variability of diurnal precipitation cycle

Errors in the simulated propagating speed of the MCS passage contribute to errors in not only the diurnal cycle of precipitation, but also in rain rate intensity that depends on diurnally varying ambient boundary layer conditions. To quantitatively assess the spatiotemporal variability of simulated precipitation in the basin, five subregions (2E, E, M, W, and 2W) with its width of 2-degrees in each dimension are defined to represent different locations in the basin as illustrated in Figure 1. The five subregions are arranged in a NE-SW orientation to align with the approximate southwestward MCS propagation and the subregion-mean hourly rain rate is quantified for the MCS and noMCS groups. To better distinguish the difference between sensitivity simulations conducted by the same model, the following are separate discussions for the WRF and E3SM models.

4.2.1 Result of WRF simulation

The observed and WRF-simulated diurnal precipitation cycles at the five locations are given in Figure 5. For the coastal 2E subregion, the observed peak rain rate of ~1.4 mm hr^{-1} occurs around 17 LT (20 UTC) on days with MCS propagation. During noMCS days, the observed rain rate is overall slightly lower than MCS days with a peak of ~1.0 mm hr^{-1} and the peak hour is about two hours earlier. The diurnal variation of precipitation from the WRF_DA and WRF_noDA simulations are similar to each other, suggesting that the impact of DA is relatively small at this coastal location. They both reproduce peak rain rate earlier than what was observed in both the MCS and noMCS groups and the rain rate is lower than observed for the MCS group. In subregion E, which is located in between the Atlantic coast and Manaus, the observed rain rate peak of 0.8 mm hr^{-1} for MCS days shifts to 05 LT (08 UTC) as the coastal systems age and propagate westward. In addition, the impact of DA is more evident. While WRF_DA has a comparable diurnal cycle to IMERG for the MCS group, there is no peak rain rate from WRF_noDA so the curve is essentially flat. However, neither simulation is able to capture the observed nocturnal precipitation peak for the noMCS days.

For subregion M that encompasses Manaus and the surrounding area, rainfall estimation based on SIPAM radar reflectivity data is included as an additional observational data source. The magnitude and diurnal variability of the rain rate from SIPAM and IMERG are quite similar. A single broad rainfall peak near 1.0 mm hr^{-1} occurs between 08 and 12 LT (12 and 16 UTC) for MCS days, likely due to precipitation contributed by both MCS and locally forced convection (Burleyson et al. 2016; Giangrande et al. 2017; Tang et al. 2016). The WRF_DA reproduces both the observed magnitude and diurnal variation of rain rate. In contrast, WRF_noDA has a higher and narrower peak about one hour later than WRF_DA. A notable reduction of precipitation rate from both IMERG and SIPAM is observed for the noMCS group between 06 and 14 LT (10 and 18 UTC) over this region, likely due to a higher frequency of less organized and weaker local convection on noMCS days. The later peak is also consistent with Tian et al. (2021), who showed that pre-existing disturbances (either within the region or propagating through the region) cause an earlier diurnal peak over Manaus. The reduction in precipitation is reproduced in both the WRF_DA and WRF_noDA simulations; however, nocturnal precipitation is substantially underestimated in both simulations. As a result, the mean daily rain rate is much more underestimated for days in the noMCS group than MCS days over this subregion.

Downwind of Manaus in subregion W, the observed peak rainfall rate of ~1.4 mm hr^{-1} occurs at 17 LT (21 UTC) for the MCS group as convective systems initiated near Manaus grow and propagate westward. Both the WRF_DA and WRF_noDA simulations capture the timing and magnitude of this peak precipitation. A weaker secondary peak between 03 and 05 LT (07 and 09 UTC) is evident in the IMERG product. However, this secondary peak is not captured by either WRF simulation. For this subregion, the rain rate for noMCS days is generally much lower than MCS days. While the diurnal rain rate from WRF_DA is similar to IMERG, WRF_noDA significantly overpredicts nocturnal rainfall from 23 to 05 LT (03 to 09 UTC). Further west, in subregion 2W, the IMERG diurnal precipitation cycle for MCS days is quite similar to subregion E as both of them exhibit a peak rain rate of 0.8 mm hr^{-1} around 03 LT (08 UTC). This is consistent with the pattern of the hour of rainfall maximum as shown in Figure 4. Both WRF simulations overamplify the intensity of propagating convection and the simulated peak rain rates are about double the IMERG amount. The two simulations diverge after 06 LT (11 UTC), with WRF_DA being closer to IMERG than WRF_noDA. On noMCS days, the diurnal variation in precipitation rate is better represented when the model is constrained by data assimilation. Overall, the comparison of mean daily precipitation in each subregion indicates the presence of MCS increases rainfall amount in many local regions of Amazon basin.

4.2.2 Result of E3SM simulation

Two E3SM simulations (E3SM_v1_free and E3SM_v1_nudge) were performed and the results within the five subregions over the Amazon basin are shown in Figure 6. Since the difference between results of the MCS and noMCS groups was not significant, diurnal cycles demonstrated here are computed for all days.

While the IMERG product over subregion 2E near the coast indicates that rain rate increased from 11 LT (14 UTC) and reached its maximum at 17 LT (20 UTC) (Figure 6), the two E3SM simulations have a much smaller amplitude in diurnal variation and therefore significantly underestimate the rain rate during the afternoon. This leads to underestimation of simulated mean daily rain rates as given in Figure 6. In subregion E, diurnal cycles of IMERG and E3SM are in opposite phases. The observed peak rate occurred around 02 LT (05 UTC) whereas the two E3SM simulations have higher rain rates from noon to afternoon (~12 to 17 LT).

Near Manaus in subregion M, the simulated diurnal cycle is more in phase with the observations and both simulations reasonably reproduce the observed rain rate with slightly higher daily mean rain rate than IMERG (Figure 6). However, the simulated diurnal cycle in the two western subregions (W and 2W) is similar to that over Manaus (subregion M), implying the spatial variability of the diurnal precipitation cycle is quite small in E3SM simulations, especially west of Manaus. In these three subregions, simulated hourly rain rate starts to increase right after sunrise and reaches its peak near noon. This result is consistent with Figure 4, demonstrating that E3SM simulated precipitation over the western part of the Amazon is primarily triggered by land surface heating in the convective parameterization (Xie et al. 2019). With unstable atmospheric conditions frequently occur over the Amazon basin during the wet season, convective parameterizations used in E3SM as well as other global climate models would then trigger the development of deep convection as soon as the solar radiation heats the surface sufficiently to produce positive buoyancy.

In summary, the nudging of the large-scale wind field in E3SM exhibits a limited impact on the overall spatial variability of simulated precipitation amount and diurnal cycle over the Amazon basin. This is because the development of convective clouds also depends on temperature and humidity gradients which are not nudged towards global analyses in this study. In addition, any improvements in the simulated wind field is expected to have limited impact on predicted diurnal precipitation cycle, since the parameterized convection is triggered locally and there is no mechanism in the model to propagate unresolved convective activity with the winds.

4.3 Characteristics of ambient flow and its variability

Several observational studies (e.g., Kousky 1980; Garstang et al. 1994; Cohen et al. 1995; Alcântara et al. 2011) describe certain environmental factors, such as the intensity of the low-level easterly jet, associated with the occurrence of well-organized convection that propagates over the Amazon basin. However, how meteorological states evolve across the area in the context of the basin-scale convective diurnal cycle has not been addressed in detail by most studies because of the limited number and spatiotemporal distribution of meteorological measurements. To address this issue, we examine the ambient environments associated with and without propagating convective precipitation by leveraging the model's relatively high temporal and spatial coverage along with the independent high temporal resolution ARM sounding measurements at the T3 site and low-temporal measurements made at three other locations in the basin (Figure 1).

4.3.1 Comparison with observations

Radiosondes launched at four sites within the analysis domain are used to verify the simulated spatial variability of the ambient environment including horizontal wind velocities, temperature, and specific humidity. In addition to radiosondes at the ARM T3 site (-3.1492°N, -59.992°E) near Manaus, profiles collected at three other sites were acquired from NOAA's IGRA, including 82099 (0.05°N, -51.0667°E), 82244 (-2.433°N, -54.7167°E), and 82411 (-4.25°N, -69.9333°E). As shown in Figure 1, the four sites are near or inside the corresponding subregions (2E, E, M, 2W) designed for the precipitation analysis in Section 4.2. While ARM radiosondes at T3 site are available five times per day, radiosondes at sites 82099 and 82244 are launched every 12 hours (00 and 12 UTC). Site 82017 has only one profile per day at 12 UTC.

To facilitate model-observation comparisons, the simulated profiles located within a 2-degree by 2-degree subregion centered at each site are interpolated vertically from the surface to 10 km

MSL at intervals of 0.1 km. Then, the mean profile at each given hour (00, 06, 12, 15, 18 UTC) is obtained by taking horizontal average over the subregion. These profiles are temporally averaged over the MCS and noMCS group days. Only simulations WRF_DA and E3SM_v1_nudge are selected here to represent each model since the differences among the simulations from the same model are relatively small.

The composite plot for the zonal wind profiles at the four sites is given in Figure 7. The comparison of the observed and simulated profiles show that: (1) the spatial variability of the zonal wind (U) profile is well represented in both WRF and E3SM simulations; (2) while the low-level easterly jet has a peak wind speed of ~12 m-s^{-1} at an altitude of ~3 km for the MCS group, the height and speed of peak zonal wind speed is relatively lower and weaker, respectively, in the noMCS group (Alcântara et al. 2011; Anselmo et al. 2020); (3) the vertical wind shear becomes progressively weaker from the coastal to inland locations; (4) only near Manaus (subregion M) and roughly below 1-km altitude does the easterly wind from the noMCS group exceed the velocity from the MCS group; and (5) while wind profile varies diurnally within the MCS and noMCS groups, the differences in the zonal wind profiles between the MCS and noMCS groups is far larger.

To highlight how well the two simulations can reproduce the ambient flow in a quantitative manner, profiles of the bias in the zonal wind component, temperature, and specific humidity for each simulation and for each subregion are computed and displayed in Figure 8. These results show that: (1) WRF has an overall better agreement with the radiosonde observations for every variable at each location in the basin; (2) while the bias of zonal wind in WRF and E3SM has no consistent sign at the four locations, E3SM has robust cold and dry biases in the troposphere below 10 km height that are not evident in the WRF simulation. The relatively large temperature bias in the upper troposphere is most likely introduced by the deep convection parameterization. Nevertheless, it should be noted an overall 2-3 K cold bias is considered minor for assessment of global climate models (Rasch et al. 2019).

To supplement model validation, the average diurnal variability of the surface meteorological observations collected at the T3 site for the MCS and noMCS groups are computed separately and compared with the corresponding model simulations. Figure 9 shows that the observed 10-m zonal wind is consistently easterly (negative) throughout the day and has a maximum velocity of -2 m s^{-1} and -3.2 m s^{-1} for the MCS and noMCS groups, respectively. While the WRF simulation qualitatively reproduces the observed diurnal cycle for both the MCS and noMCS groups, the E3SM simulation significantly underestimates the variability during the day as well as the contrast between the days with and without MCS. This indicates that E3SM does not reproduce the near surface wind fluctuation even when the large-scale wind field are nudged. While both models reproduce the observed diurnal temperature variability, both WRF and E3SM have a cold bias. The peak daytime temperature from E3SM is closer to observed on average, but WRF is closer to the observations at other times in a day. Both models also have a dry bias in specific humidity throughout the day, but WRF is closer to the observations with an overall bias less than 1 g kg^{-1}. E3SM has much larger dry bias in general, and the atmosphere becomes significantly drier between 10 to 17 LT when simulated convective precipitation is at its peak.

4.3.2 Diurnal variation of low-level flow

The convective diurnal cycle (CDC) was defined in Itterly et al. (2018) as *the response of convection and its related processes to the daily cycle of solar insolation regulating the timing and intensity of clouds and convective rainfall*. That said, convective-scale processes may influence the diurnal cycle and large-scale atmospheric state but are not explicitly resolved by global climate models. Therefore, the behavior of convection is commonly approximated by using its statistical relationship with the resolved large-scale state. Since our WRF simulation can better resolve the response at cloud-system level such as vertical momentum transport and cold pools taking place over the Amazon basin, we are motivated to understand how these two models represent the atmospheric flow changes in the presence of propagating MCS.

The "perturbed" state from the WRF and E3SM simulations is examined here. The full state contains distinct variations with height as well as a diurnal signal that is stronger than other smaller-scale responses; therefore, the full state is not necessarily useful in describing convection-induced responses. To be consistent with the precipitation analysis discussed in the earlier sections, the same domain over the Amazon basin is used for the following analysis. To compute perturbed states, domain-mean vertical profiles are obtained at each hour and height level for the zonal wind, temperature, and specific humidity. With the horizontal mean state as a reference, the perturbed state can be then obtained by subtracting the horizontal mean state from the full state and given by

$$X'_{ijk} = X_{ijk} - \bar{X}_k.$$

where X represents full model state of a given variable at one particular hour, the overbar denotes the horizontal average, the prime denotes the perturbed state, and the subscript indicates the corresponding dimensions. Note the diurnal signal is filtered as the horizontal domain mean at each hour is computed and subtracted when calculating the perturbed state.

To visualize how those perturbations vary diurnally, we first examine their low-level (below 3 km MSL) mean. The perturbations are first averaged at each column in the domain. Then, similar to what is done for the daily maximum rain rate shown in Figure 4, the hour (UTC) of daily maximum rain mixing ratio, easterly wind perturbation, negative temperature perturbation, and specific humidity are displayed instead. This approach is performed separately for both the MCS and noMCS groups from WRF and E3SM simulations.

While the spatial distribution of the low-level mean rain mixing ratio maximum for the MCS group in the WRF simulation (Figure 10) is somewhat noisier than the rain rate maximum due to the nature of the variable's spatial scale, the spatial distribution is very similar to the propagating precipitation shown in Figure 4. The orange to red region (equivalent to 11 to 17 LT) over the domain center in both the easterly wind and negative temperature perturbation distribution indicates the strongest low-level easterly flow and coldest air occurs primarily during the afternoon near Manaus. This timing is correlated with the precipitation propagation (Figure 4) in this region despite some differences near the domain's southeastern corner. The diurnal variation of low-level moisture perturbation is also in phase with propagating precipitating systems. Based on what we find here and the results as will be shown later, the easterly wind and negative temperature perturbations are most likely the footprints of cold pools induced by the organized propagating convection that is superimposed on the large-scale ambient easterly trade winds.

Compared to the MCS group, the rain mixing ratio (precipitation) perturbation on the noMCS days is more random over the domain reflecting the lack of organized propagating convection (Figure 10). The perturbations for the easterly wind, negative temperature, and specific humidity are also quite different in the central Amazon than their counterparts on MCS days. For example, the easterly wind perturbations occur in the late evening and early morning in the vicinity of Manaus on noMCS days as opposed to the late afternoon and early evening on MCS days. The specific humidity perturbations are shifted much later in the day as well. These simulation results demonstrate that the occurrence of MCS systems in the Amazon basin not only alter the precipitation diurnal cycle, but also the fluctuation of low-level flow states.

The corresponding results from the E3SM simulation (Figure S2) show that the diurnal perturbation distributions of rain mixing ratio and specific humidity have less correlation with the precipitation pattern, except for the northeastern corner of the domain (Figure 4). This rain behavior for both MCS and noMCS days and the environmental fields on MCS days are different from WRF (Figure 10). Nevertheless, the diurnal perturbation distributions of specific humidity, easterly wind and negative temperature for the noMCS group in WRF and E3SM are similar. This suggests that E3SM has a better skill in simulating flow patterns when MCS systems are absent but has difficulty in reproducing the observed diurnal flow variations in days with MCS.

4.3.3 Diurnal variation of perturbed flow in vertical cross-section

We now present the WRF cross-sections of zonal wind, temperature, and specific humidity perturbations as they evolve in time for the MCS (Figure 11) and noMCS (Figure 12) groups along a northeast-southwest oriented plane as denoted by the red line in Figure 1. The cross-sections of each perturbed state at each hour are obtained through three-dimensional interpolation. To assist in the analysis, the information on precipitation occurrence is also provided. We compute the occurrence frequency (in percentage) of simulated reflectivity is greater than 15 dBZ to indicate the occurrence of large hydrometeor particles in time and space.

For the MCS group, an easterly wind (negative) perturbation persists below 5 km over the eastern part of the cross-sections at all times (Figure 11a). However, the thickness and strength of easterly wind coming from the Atlantic Ocean varies diurnally. Meanwhile, colder and drier air is advected by the easterly wind near the coast (Figures 11b and c), suggesting the air advected from the ocean is relatively colder and drier than the air over continent. Embedded in the deck of easterly wind perturbation, the sea breeze near the northeastern coast can be identified by the patch of negative temperature perturbation as shown in Figure 11b. At local noon (12 LT), a relatively weak and shallow negative temperature perturbation appears near the bottom right corner can be inferred as the time of landfall of sea breeze. It intrudes inland in the afternoon hours and stays in similar horizontal extension overnight (from 20 to 08 LT), indicating the sea breeze front can only directly influence region within ~400 km distance from coast (east of -55° longitude). Together with the occurrence frequency of reflectivity > 15 dBZ (Figure 11d), it implies the sea breeze's high-density flow starts to trigger shallow convection at its front near the coast at local noon. Then, as it propagates inland, the convection grows deeper, and moisture is transported upward from lower troposphere to higher altitudes (Figure 11c and 11d). A nearly persistent pattern of divergent flow is found in Figure 11a in the central basin above ~7 km through the day, suggesting the conditions favorable for growth of convective systems. As the convection intensifies and transitions to a mature stage, rain evaporation most likely takes place below the convective clouds to cool the lower troposphere (Figures 11b and 11d). From 12 to 16 LT, while the colder air in association with sea breeze becomes relatively weak, an extensive pool of colder air (6 to 7 degrees wide in longitude and ~1.5 km deep in vertical) is formed. The westward outflow from the pool of colder air is potentially responsible for maintenance and propagation of MCSs in the afternoon. Long-lasting precipitation is also produced near the southwest end of cross-sections and denoted by a rectangle in Figure 11d. This may be relevant to a persistent convergence as denoted in Figure 11a.

Figures 12a show that the low-level easterly wind perturbation in the noMCS group is generally weaker than what is shown for MCS group (Figure 11a), which is associated with an eastward-shifted convergence below 5 km. Moreover, the upper-tropospheric divergence is weaker than on MCS days and becomes hardly distinguishable. While temperature perturbations in the noMCS groups show similar evolution of the sea breeze as the MCS group (i.e., negative perturbations over the lower right corners in Figure 12b), the most distinct difference between days is found in the specific humidity perturbation. For the noMCS group, a layer of dry air between 1 to 4 km penetrates westward inland (Figure 12c) and its front edge is well collocated with the deep convergence denoted by the dashed long arrow in Figure 12a as well as the position of deep convection denoted by the long arrow in Figure 12d. These results imply the intensity of sea breeze may have less of an impact on the formation of MCS in Amazon basin. Instead, it is the dry air advected from ocean to the central basin that suppresses the formation of MCS which leads to much less rainfall near Manaus.

The results for E3SM are given in the supporting information (Figures S3 and S4). Several features of the mesoscale environment such as strong low-level easterly wind, representation of a sea breeze (i.e., negative temperature perturbation near northeastern coast), relatively cold and dry air advected from the Atlantic Ocean are represented in E3SM simulations. While E3SM has a notable warm and dry bias when compared against radiosonde profiles (Figure 8), it does show stronger westward penetration of dry air for the noMCS group which agrees with the WRF simulation (Figures 12c, S3c and S4c). Nevertheless, other features closely related to a propagating MCS system that we saw in the WRF simulations are not found in the E3SM results. While the effects associated with land-sea contrasts are represented in E3SM, the physical processes associate with MCS are not reproduced. Furthermore, an almost stagnant pattern is found in almost every perturbed state variable, implying the diurnal variation of convective activity is subtle.

5. Summary and conclusion

A cloud-system resolving model (WRF) that explicitly represents the lifecycle of convective systems and a global climate model (E3SM) that parameterizes deep convective clouds are used to better understand the processes that affect the diurnal precipitation cycle over the Amazon basin during the wet season of 2014. These simulations are combined with unique meteorological observations collected during the GoAmazon2014/15 campaign as well as operational in situ and satellite datasets. Through a comprehensive intercomparison among models and observations, the primary findings include:

1. Impact of data assimilation: Our analysis shows that by using an observationally constrained cloud-system resolving model, the overall spatiotemporal variability of the precipitation diurnal cycle in the Amazon basin during the 2014 wet season is similar to what was observed. Larger differences between the observed and simulated diurnal rainfall rates over many locations in the basin are produced when data assimilation is not used. It also shows lack of observations over the Atlantic Ocean may limit the optimization of simulated convective activity over the coastal region. This suggests that adequately representing the large-scale environmental conditions in the tropics, where synoptic forcing is relatively weak, is one factor needed to adequately represent the formation and propagation of convective systems over the Amazon basin.

2. Role of MCS in diurnal rainfall distribution: Analysis of IMERG precipitation data reveals the frequent southwestward propagation of MCSs triggered by the coastal sea breeze front and over the central Amazon basin during the wet season IOP (Figure 3). When days are separated into groups with and without MCS propagation, the MCS's major role in contributing to mean diurnal precipitation cycle that varies over the basin is revealed. Close to the coast, the local diurnal precipitation cycle is controlled primarily by the sea breeze associated with land-sea temperature contrasts. Further inland, MCS takes over the dominant role (Figure 4).

3. Characteristics of simulated precipitation over the Amazon basin: Precipitation simulated by WRF agrees reasonably well with IMERG data in terms of the phase change of diurnal cycle as well as the peak rain rate intensity during the MCS and noMCS periods. There are larger uncertainties in predicting precipitation for the noMCS group than the MCS group because weaker and more isolated convections are harder to predict. In contrast, E3SM does not produce propagating convective systems over the Amazon basin. While the total precipitation amount is similar to observed at many locations in the basin during March and April 2014 (Figure S1), the simulated precipitation diurnal cycle is often out of phase with observations. The exception is near Manaus where the peak of average rainfall rate is associated with both locally forced and propagating convection that occur at about the same time. E3SM tends to underestimate precipitation in the Amazon basin during other months (not shown), which is a common bias in global climate models. This bias may be due, in part, to the inability to adequately represent propagating organized convective systems that are longer-lived and produce more intense precipitation than isolated convection.

4. Reproduction of variability in ambient environment: A comparison of the simulated tropospheric meteorological profiles with radiosonde observations at four sites across the Amazon basin shows that both WRF and E3SM simulations reasonably represent the spatial characteristics of tropospheric winds (below 10 km) such as the intensity of low-level easterly winds and differences in the vertical wind shear on days with and without a propagating MCS (Figure 7). This analysis also indicates diurnal variation of wind profiles appears to be much smaller than the day-to-day variation of tropospheric wind speed in general. While both models constrain their simulations with observed winds in some way, the profiles of bias indicate that the WRF simulations agree better with the radiosonde observations than E3SM, especially for temperature and moisture profiles (Figure 8). While WRF has small vertical variations in the biases, significant cold and dry biases are produced by E3SM in the lower troposphere that are as large as ~2 K and ~1 g kg^{-1} for temperature and specific humidity, respectively.

5. Impact of MCS on the environment: A 4-D multivariate perturbation analysis, which has not been applied by earlier studies, is performed to better understand how the meteorological states vary between the MCS and noMCS days. Results suggest the responses induced by MCS substantially change the diurnal cycle of local meteorological states in the basin except near the Atlantic coast where the sea breeze front is primarily responsible for the convective initiation (Figures 10). Consistent with the precipitation analysis, MCS-induced perturbations produced by WRF are essentially absent in the E3SM simulation. Vertical cross-sections of perturbed states show that easterly wind perturbation and upper-level divergence are enhanced when propagating MCSs occur in the WRF simulation (Figures 11 and 12). Furthermore, the presence of large pools of colder air over the central Amazon basin highlights the role of organized deep convection associated with the propagating precipitation. While the negative temperature perturbations related to the sea breeze do not vary much between the MCS and noMCS groups, the phase of the diurnal cycle in moisture is essentially opposite in the central basin. A westward propagating positive moisture perturbation occurs in the MCS group, whereas a persistent negative moisture perturbation occurs over the central basin for the noMCS group. This suggests that the intensity of the sea breeze has a smaller impact on the formation and maintenance of organized convection in the Amazon basin than the dryness of low-level air transported from ocean to inland.

Not surprisingly, the multivariate perturbation analysis shows that in E3SM the diurnal variation of convective perturbation for each state variable is almost stagnant (Figures S3 and S4) since the pattern for each perturbed state does not shift in time. This confirms that the unrealistic representation of the convective diurnal cycle in E3SM simulations may contribute to off-phase diurnal variation in precipitation. Aside from misrepresentation of ambient environment of tropospheric flow in E3SM simulations, convective parameterizations that are strongly coupled with surface heating most likely suppress the spatial variability of the precipitation diurnal cycle (Xie et al. 2019). A future study is needed to assess the new trigger functions proposed in Xie et al. (2019) by using the current metrics as well as to quantify the improvement in the representation of spatial and diurnal variability in convective rainfall across the Amazon basin.

Acknowledgements

The GoAmazon2014/5 field campaign was supported by the Atmospheric Radiation Measurement (ARM) Climate Research Facility and the Environmental Molecular Science Laboratory (EMSL), both are U.S. Department of Energy (DOE) Office of Science User Facilities sponsored by the Office of Biological and Environmental Research. The modeling studies were performed on PNNL's Research Computing cluster. The WRF community model is available from the National Center for Atmospheric Research (NCAR) at http://www2.mmm.ucar.edu/wrf/users/. 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. GoAmazon2014/15 data used in this manuscript are freely available from the ARM data archive (https://www.arm.gov/data). The IMERG Final-Run V06B precipitation products used in this study were accessed from http://pmm.nasa.gov/data-access/downloads/gpm/. Radiosonde data from NOAA's Integrated Global Radiosonde Archive was access from https://www.ncdc.noaa.gov/data-access/weather-balloon/integrated-global-radiosonde-archive. NCEP PREPBUFR data are available from https://doi.org/10.5065/Z83F-N512 (National Centers

for Environmental Prediction/National Weather Service/NOAA/U.S. Department of Commerce, 2008). The NCEP FNL analyses were access from a National Center for Atmospheric Research (NCAR) archive at http://dss.ucar.edu/datasets/ds083.2/. 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

Alcântara, C. R., M. A. F. Silva Dias, E. P. Souza, and J. C. P. Cohen, 2011: Verification of the role of the low level jets in Amazon squall lines. *Atmospheric Res.*, **100**, 36–44, https://doi.org/10.1016/j.atmosres.2010.12.023.

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. Weather Rev.*, **148**, 4083–4099, https://doi.org/10.1175/MWR-D-19-0414.1.

Arnold, N. P., W. M. Putman, and S. R. Freitas, 2020: Impact of Resolution and Parameterized Convection on the Diurnal Cycle of Precipitation in a Global Nonhydrostatic Model. *J. Meteorol. Soc. Jpn. Ser II*, **98**, 1279–1304, https://doi.org/10.2151/jmsj.2020-066.

Bauer, H.-S., T. Schwitalla, V. Wulfmeyer, A. Bakhshaii, U. Ehret, M. Neuper, and O. Caumont, 2015: Quantitative precipitation estimation based on high-resolution numerical weather prediction and data assimilation with WRF – a performance test. *Tellus Dyn. Meteorol. Oceanogr.*, **67**, 25047, https://doi.org/10.3402/tellusa.v67.25047.

Bechtold, P., J.-P. Chaboureau, A. Beljaars, A. K. Betts, M. Köhler, M. Miller, and J.-L. Redelsperger, 2004: The simulation of the diurnal cycle of convective precipitation over land in a global model. *Q. J. R. Meteorol. Soc.*, **130**, 3119–3137, https://doi.org/10.1256/qj.03.103.

Bergman, J. W., and M. L. Salby, 1996: Diurnal Variations of Cloud Cover and Their Relationship to Climatological Conditions. *J. Clim.*, **9**, 2802–2820, https://doi.org/10.1175/1520-0442(1996)009<2802:DVOCCA>2.0.CO;2.

Betts, A. K., 2002: Surface diurnal cycle and boundary layer structure over Rondônia during the rainy season. *J. Geophys. Res.*, **107**, https://doi.org/10.1029/2001JD000356.

——, and C. Jakob, 2002: Evaluation of the diurnal cycle of precipitation, surface thermodynamics, and surface fluxes in the ECMWF model using LBA data. *J. Geophys. Res. Atmospheres*, **107**, LBA 12-1-LBA 12-8, https://doi.org/10.1029/2001JD000427.

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. Meteorol. Climatol.*, **55**, 1579–1598, https://doi.org/10.1175/JAMC-D-15-0229.1.

Caldwell, P. M., and Coauthors, 2019: The DOE E3SM Coupled Model Version 1: Description and Results at High Resolution. *J. Adv. Model. Earth Syst.*, **11**, 4095–4146, https://doi.org/10.1029/2019MS001870.

Chen, F., and J. Dudhia, 2001: Coupling an Advanced Land Surface–Hydrology Model with the Penn State–NCAR MM5 Modeling System. Part I: Model Implementation and Sensitivity. *Mon. Weather Rev.*, **129**, 569–585, https://doi.org/10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2.

Clark, A. J., W. A. Gallus, and T.-C. Chen, 2007: Comparison of the Diurnal Precipitation Cycle in Convection-Resolving and Non-Convection-Resolving Mesoscale Models. *Mon. Weather Rev.*, **135**, 3456–3473, https://doi.org/10.1175/MWR3467.1.

Clark, P., N. Roberts, H. Lean, S. P. Ballard, and C. Charlton-Perez, 2016: Convection-permitting models: a step-change in rainfall forecasting. *Meteorol. Appl.*, **23**, 165–181, https://doi.org/10.1002/met.1538.

Cohen, J. C. P., M. A. F. Silva Dias, and C. A. Nobre, 1995: Environmental Conditions Associated with Amazonian Squall Lines: A Case Study. *Mon. Weather Rev.*, **123**, 3163–3174, https://doi.org/10.1175/1520-0493(1995)123<3163:ECAWAS>2.0.CO;2.

Collow, A. B. M., M. A. Miller, and L. C. Trabachino, 2016: Cloudiness over the Amazon rainforest: Meteorology and thermodynamics. *J. Geophys. Res. Atmospheres*, **121**, 7990–8005, https://doi.org/10.1002/2016JD024848.

Dai, A., 2006: Precipitation Characteristics in Eighteen Coupled Climate Models. *J. Clim.*, **19**, 4605–4630, https://doi.org/10.1175/JCLI3884.1.

——, and K. E. Trenberth, 2004: The Diurnal Cycle and Its Depiction in the Community Climate System Model. *J. Clim.*, **17**, 930–951, https://doi.org/10.1175/1520-0442(2004)017<0930:TDCAID>2.0.CO;2.

Derin, Y., and Coauthors, 2019: Evaluation of GPM-era Global Satellite Precipitation Products over Multiple Complex Terrain Regions. *Remote Sens.*, **11**, 2936, https://doi.org/10.3390/rs11242936.

Dias, M. A. F. S., and Coauthors, 2002: Cloud and rain processes in a biosphere-atmosphere interaction context in the Amazon Region. *J. Geophys. Res. Atmospheres*, **107**, LBA 39-1-LBA 39-18, https://doi.org/10.1029/2001JD000335.

Durre, I., R. S. Vose, and D. B. Wuertz, 2006: Overview of the Integrated Global Radiosonde Archive. *J. Clim.*, **19**, 53–68, https://doi.org/10.1175/JCLI3594.1.

Espinoza, J. C., M. Lengaigne, J. Ronchail, and S. Janicot, 2012: Large-scale circulation patterns and related rainfall in the Amazon Basin: a neuronal networks approach. *Clim. Dyn.*, **38**, 121–140, https://doi.org/10.1007/s00382-011-1010-8.

Fiedler, S., and Coauthors, 2020: Simulated Tropical Precipitation Assessed across Three Major Phases of the Coupled Model Intercomparison Project (CMIP). *Mon. Weather Rev.*, **148**, 3653–3680, https://doi.org/10.1175/MWR-D-19-0404.1.

Gao, Y., L. R. Leung, C. Zhao, and S. Hagos, 2017: Sensitivity of U.S. summer precipitation to model resolution and convective parameterizations across gray zone resolutions. *J. Geophys. Res. Atmospheres*, **122**, 2714–2733, https://doi.org/10.1002/2016JD025896.

Garstang, M., and Coauthors, 1990: The Amazon Boundary-Layer Experiment (ABLE 2B): A Meteorological Perspective. *Bull. Am. Meteorol. Soc.*, **71**, 19–32, https://doi.org/10.1175/1520-0477(1990)071<0019:TABLEA>2.0.CO;2.

——, H. L. Massie, J. Halverson, S. Greco, and J. Scala, 1994: Amazon Coastal Squall Lines. Part I: Structure and Kinematics. *Mon. Weather Rev.*, **122**, 608–622, https://doi.org/10.1175/1520-0493(1994)122<0608:ACSLPI>2.0.CO;2.

Genio, A. D. D., and J. Wu, 2010: The Role of Entrainment in the Diurnal Cycle of Continental Convection. *J. Clim.*, **23**, 2722–2738, https://doi.org/10.1175/2009JCLI3340.1.

Ghate, V. P., and P. Kollias, 2016: On the Controls of Daytime Precipitation in the Amazonian Dry Season. *J. Hydrometeorol.*, **17**, 3079–3097, https://doi.org/10.1175/JHM-D-16-0101.1.

Giangrande, S. E., and Coauthors, 2017: Cloud characteristics, thermodynamic controls and radiative impacts during the Observations and Modeling of the Green Ocean Amazon (GoAmazon2014/5) experiment. *Atmospheric Chem. Phys.*, **17**, 14519–14541, https://doi.org/10.5194/acp-17-14519-2017.

Golaz, J.-C., and Coauthors, 2019: The DOE E3SM Coupled Model Version 1: Overview and Evaluation at Standard Resolution. *J. Adv. Model. Earth Syst.*, **11**, 2089–2129, https://doi.org/10.1029/2018MS001603.

Guichard, F., and Coauthors, 2004: Modelling the diurnal cycle of deep precipitating convection over land with cloud-resolving models and single-column models. *Q. J. R. Meteorol. Soc.*, **130**, 3139–3172, https://doi.org/10.1256/qj.03.145.

Gustafsson, N., and Coauthors, 2018: Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. *Q. J. R. Meteorol. Soc.*, **144**, 1218–1256, https://doi.org/10.1002/qj.3179.

Hamill, T. M., J. S. Whitaker, D. T. Kleist, M. Fiorino, and S. G. Benjamin, 2011: Predictions of 2010's Tropical Cyclones Using the GFS and Ensemble-Based Data Assimilation Methods. *Mon. Weather Rev.*, **139**, 3243–3247, https://doi.org/10.1175/MWR-D-11-00079.1.

Harriss, R. C., and Coauthors, 1990: The Amazon Boundary Layer Experiment: Wet season 1987. *J. Geophys. Res. Atmospheres*, **95**, 16721–16736, https://doi.org/10.1029/JD095iD10p16721.

Hassim, M. E. E., T. P. Lane, and W. W. Grabowski, 2016: The diurnal cycle of rainfall over New Guinea in convection-permitting WRF simulations. *Atmospheric Chem. Phys.*, **16**, 161–175, https://doi.org/10.5194/acp-16-161-2016.

Hou, A. Y., and Coauthors, 2014: The Global Precipitation Measurement Mission. *Bull. Am. Meteorol. Soc.*, **95**, 701–722, https://doi.org/10.1175/BAMS-D-13-00164.1.

Huang, W.-R., P.-Y. Liu, Y.-H. Chang, and C.-Y. Liu, 2020: Evaluation and Application of Satellite Precipitation Products in Studying the Summer Precipitation Variations over Taiwan. *Remote Sens.*, **12**, 347, https://doi.org/10.3390/rs12030347.

Huffman, G. J., and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. *J. Hydrometeorol.*, **8**, 38–55, https://doi.org/10.1175/JHM560.1.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, 2008: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res. Atmospheres*, **113**, https://doi.org/10.1029/2008JD009944.

Itterly, K. F., and P. C. Taylor, 2014: Evaluation of the Tropical TOA Flux Diurnal Cycle in MERRA and ERA-Interim Retrospective Analyses. *J. Clim.*, **27**, 4781–4796, https://doi.org/10.1175/JCLI-D-13-00737.1.

——, ——, J. B. Dodson, and A. B. Tawfik, 2016: On the sensitivity of the diurnal cycle in the Amazon to convective intensity. *J. Geophys. Res. Atmospheres*, **121**, 8186–8208, https://doi.org/10.1002/2016JD025039.

——, ——, and J. B. Dodson, 2018: Sensitivity of the Amazonian Convective Diurnal Cycle to Its Environment in Observations and Reanalysis. *J. Geophys. Res. Atmospheres*, **123**, 12,621-12,646, https://doi.org/10.1029/2018JD029251.

Janjić, Z. I., 2001: Nonsingular Implementation of the Mellor-Yamada Level 2.5 Scheme in the NCEP Meso model. *Met Off. Note*, **437**, 61.

Kendon, E. J., N. M. Roberts, C. A. Senior, and M. J. Roberts, 2012: Realism of Rainfall in a Very High-Resolution Regional Climate Model. *J. Clim.*, **25**, 5791–5806, https://doi.org/10.1175/JCLI-D-11-00562.1.

Khairoutdinov, M., and D. Randall, 2006: High-Resolution Simulation of Shallow-to-Deep Convection Transition over Land. *J. Atmospheric Sci.*, **63**, 3421–3436, https://doi.org/10.1175/JAS3810.1.

Konduru, R. T., and H. G. Takahashi, 2020: Effects of Convection Representation and Model Resolution on Diurnal Precipitation Cycle Over the Indian Monsoon Region: Toward a Convection-Permitting Regional Climate Simulation. *J. Geophys. Res. Atmospheres*, **125**, e2019JD032150, https://doi.org/10.1029/2019JD032150.

Kousky, V. E., 1980: Diurnal Rainfall Variation in Northeast Brazil. *Mon. Weather Rev.*, **108**, 488–498, https://doi.org/10.1175/1520-0493(1980)108<0488:DRVINB>2.0.CO;2.

Langenbrunner, B., M. S. Pritchard, G. J. Kooperman, and J. T. Randerson, 2019: Why Does Amazon Precipitation Decrease When Tropical Forests Respond to Increasing CO2? *Earths Future*, **7**, 450–468, https://doi.org/10.1029/2018EF001026.

Lou, S., and Coauthors, 2020: New SOA Treatments Within the Energy Exascale Earth System Model (E3SM): Strong Production and Sinks Govern Atmospheric SOA Distributions and Radiative Forcing. *J. Adv. Model. Earth Syst.*, **12**, e2020MS002266, https://doi.org/10.1029/2020MS002266.

Love, B. S., A. J. Matthews, and G. M. S. Lister, 2011: The diurnal cycle of precipitation over the Maritime Continent in a high-resolution atmospheric model. *Q. J. R. Meteorol. Soc.*, **137**, 934–947, https://doi.org/10.1002/qj.809.

Mandapaka, P. V., and E. Y. M. Lo, 2020: Evaluation of GPM IMERG Rainfall Estimates in Singapore and Assessing Spatial Sampling Errors in Ground Reference. *J. Hydrometeorol.*, **21**, 2963–2977, https://doi.org/10.1175/JHM-D-20-0135.1.

Marsham, J. H., S. B. Trier, T. M. Weckwerth, and J. W. Wilson, 2011: Observations of Elevated Convection Initiation Leading to a Surface-Based Squall Line during 13 June IHOP_2002. *Mon. Weather Rev.*, **139**, 247–271, https://doi.org/10.1175/2010MWR3422.1.

Martin, S. T., and Coauthors, 2016: The Green Ocean Amazon Experiment (GoAmazon2014/5) Observes Pollution Affecting Gases, Aerosols, Clouds, and Rainfall over the Rain Forest. *Bull. Am. Meteorol. Soc.*, **98**, 981–997, https://doi.org/10.1175/BAMS-D-15-00221.1.

Miller, M. A., K. Nitschke, T. P. Ackerman, W. R. Ferrell, N. Hickmon, and M. Ivey, 2016: The ARM Mobile Facilities. *Meteorol. Monogr.*, **57**, 9.1-9.15, https://doi.org/10.1175/AMSMONOGRAPHS-D-15-0051.1.

Moazami, S., and M. R. Najafi, 2021: A comprehensive evaluation of GPM-IMERG V06 and MRMS with hourly ground-based precipitation observations across Canada. *J. Hydrol.*, **594**, 125929, https://doi.org/10.1016/j.jhydrol.2020.125929.

Nakanishi, M., and H. Niino, 2009: Development of an Improved Turbulence Closure Model for the Atmospheric Boundary Layer. *J. Meteorol. Soc. Jpn. Ser II*, **87**, 895–912, https://doi.org/10.2151/jmsj.87.895.

Oliveira, F. P. de, and M. D. Oyama, 2015: Antecedent Atmospheric Conditions Related to Squall-Line Initiation over the Northern Coast of Brazil in July. *Weather Forecast.*, **30**, 1254–1264, https://doi.org/10.1175/WAF-D-14-00120.1.

Oliveira, R., V. Maggioni, D. Vila, and C. Morales, 2016: Characteristics and Diurnal Cycle of GPM Rainfall Estimates over the Central Amazon Region. *Remote Sens.*, **8**, 544, https://doi.org/10.3390/rs8070544.

Pearson, K. J., G. M. S. Lister, C. E. Birch, R. P. Allan, R. J. Hogan, and S. J. Woolnough, 2014: Modelling the diurnal cycle of tropical convection across the 'grey zone.' *Q. J. R. Meteorol. Soc.*, **140**, 491–499, https://doi.org/10.1002/qj.2145.

Prein, A. F., Rasmussen, R. M., Wang, D., & Giangrande, S. E. (2021). Sensitivity of organized convective storms to model grid spacing in current and future climates. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 379*(2195), 20190546. https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2019.0546

Rafee, S. A. A., and Coauthors, 2017: Contributions of mobile, stationary and biogenic sources to air pollution in the Amazon rainforest: a numerical study with the WRF-Chem model. *Atmos Chem Phys*, 19.

Rasch, P. J., and Coauthors, 2019: An Overview of the Atmospheric Component of the Energy Exascale Earth System Model. *J. Adv. Model. Earth Syst.*, **11**, 2377–2411, https://doi.org/10.1029/2019MS001629.

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

Roberts, N. M., and H. W. Lean, 2008: Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events. *Mon. Weather Rev.*, **136**, 78–97, https://doi.org/10.1175/2007MWR2123.1.

Schiro, K. A., F. Ahmed, S. E. Giangrande, and J. D. Neelin, 2018: GoAmazon2014/5 campaign points to deep-inflow approach to deep convection across scales. *Proc. Natl. Acad. Sci.*, **115**, 4577–4582, https://doi.org/10.1073/pnas.1719842115.

Schumacher, C., and A. Funk, 2018: *GoAmazon2014/5 Rain Rates from the SIPAM Manaus S-band Radar*. Oak Ridge National Lab. (ORNL), Oak Ridge, TN (United States). Atmospheric Radiation Measurement (ARM) Archive,.

Schwartz, C. S., and Z. Liu, 2014: Convection-Permitting Forecasts Initialized with Continuously Cycling Limited-Area 3DVAR, Ensemble Kalman Filter, and "Hybrid" Variational–Ensemble Data Assimilation Systems. *Mon. Weather Rev.*, **142**, 716–738, https://doi.org/10.1175/MWR-D-13-00100.1.

——, G. S. Romine, R. A. Sobash, K. R. Fossell, and M. L. Weisman, 2015: NCAR's Experimental Real-Time Convection-Allowing Ensemble Prediction System. *Weather Forecast.*, **30**, 1645–1654, https://doi.org/10.1175/WAF-D-15-0103.1.

Shao, H., and Coauthors, 2015: Bridging Research to Operations Transitions: Status and Plans of Community GSI. *Bull. Am. Meteorol. Soc.*, **97**, 1427–1440, https://doi.org/10.1175/BAMS-D-13-00245.1.

Suhas, E., and G. J. Zhang, 2014: Evaluation of Trigger Functions for Convective Parameterization Schemes Using Observations. *J. Clim.*, **27**, 7647–7666, https://doi.org/10.1175/JCLI-D-13-00718.1.

Tai, S.-L., J. D. Fast, W. I. Gustafson, D. Chand, B. Gaudet, Z. Feng, and R. Newsom, 2020: Simulation of Continental Shallow Cumulus Populations Using an Observation-Constrained Cloud-System Resolving Model. *J. Adv. Model. Earth Syst.*, **12**, e2020MS002091, https://doi.org/10.1029/2020MS002091.

Tan, J., G. J. Huffman, D. T. Bolvin, and E. J. Nelkin, 2019a: Diurnal Cycle of IMERG V06 Precipitation. *Geophys. Res. Lett.*, **46**, 13584–13592, https://doi.org/10.1029/2019GL085395.

——, ——, ——, and ——, 2019b: IMERG V06: Changes to the Morphing Algorithm. *J. Atmospheric Ocean. Technol.*, **36**, 2471–2482, https://doi.org/10.1175/JTECH-D-19-0114.1.
Tanaka, L. M. d S., P. Satyamurty, and L. a. T. Machado, 2014: Diurnal variation of precipitation in central Amazon Basin. *Int. J. Climatol.*, **34**, 3574–3584, https://doi.org/10.1002/joc.3929.

Tang, S., and Coauthors, 2016: Large-scale vertical velocity, diabatic heating and drying profiles associated with seasonal and diurnal variations of convective systems observed in the GoAmazon2014/5 experiment. *Atmospheric Chem. Phys.*, **16**, 14249–14264, https://doi.org/10.5194/acp-16-14249-2016.

Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization. *Mon. Weather Rev.*, **136**, 5095–5115, https://doi.org/10.1175/2008MWR2387.1.

Trier, S. B., G. S. Romine, D. A. Ahijevych, R. J. Trapp, R. S. Schumacher, M. C. Coniglio, and D. J. Stensrud, 2015: Mesoscale Thermodynamic Influences on Convection Initiation near a Surface Dryline in a Convection-Permitting Ensemble. *Mon. Weather Rev.*, **143**, 3726–3753, https://doi.org/10.1175/MWR-D-15-0133.1.

Wang, X., 2010: Incorporating Ensemble Covariance in the Gridpoint Statistical Interpolation Variational Minimization: A Mathematical Framework. *Mon. Weather Rev.*, **138**, 2990–2995, https://doi.org/10.1175/2010MWR3245.1.

——, and T. Lei, 2014: GSI-Based Four-Dimensional Ensemble–Variational (4DEnsVar) Data Assimilation: Formulation and Single-Resolution Experiments with Real Data for NCEP Global Forecast System. *Mon. Weather Rev.*, **142**, 3303–3325, https://doi.org/10.1175/MWR-D-13-00303.1.

Wu, W.-S., R. J. Purser, and D. F. Parrish, 2002: Three-Dimensional Variational Analysis with Spatially Inhomogeneous Covariances. *Mon. Weather Rev.*, **130**, 2905–2916, https://doi.org/10.1175/1520-0493(2002)130<2905:TDVAWS>2.0.CO;2.

Xie, S., R. T. Cederwall, and M. Zhang, 2004: Developing long-term single-column model/cloud system–resolving model forcing data using numerical weather prediction products constrained by surface and top of the atmosphere observations. *J. Geophys. Res. Atmospheres*, **109**, https://doi.org/10.1029/2003JD004045.

——, and Coauthors, 2019: Improved Diurnal Cycle of Precipitation in E3SM With a Revised Convective Triggering Function. *J. Adv. Model. Earth Syst.*, **11**, 2290–2310, https://doi.org/10.1029/2019MS001702.

Zhang, G., K. H. Cook, and E. K. Vizy, 2016: The Diurnal Cycle of Warm Season Rainfall over West Africa. Part II: Convection-Permitting Simulations. *J. Clim.*, **29**, 8439–8454, https://doi.org/10.1175/JCLI-D-15-0875.1.

Zhang, G. J., 2002: Convective quasi-equilibrium in midlatitude continental environment and its effect on convective parameterization. *J. Geophys. Res. Atmospheres*, **107**, ACL 12-1-ACL 12-16, https://doi.org/10.1029/2001JD001005.

——, 2003: Roles of tropospheric and boundary layer forcing in the diurnal cycle of convection in the U.S. southern great plains. *Geophys. Res. Lett.*, **30**, https://doi.org/10.1029/2003GL018554.

Zhu, K., Y. Pan, M. Xue, X. Wang, J. S. Whitaker, S. G. Benjamin, S. S. Weygandt, and M. Hu, 2013: A Regional GSI-Based Ensemble Kalman Filter Data Assimilation System for the Rapid Refresh Configuration: Testing at Reduced Resolution. *Mon. Weather Rev.*, **141**, 4118–4139, https://doi.org/10.1175/MWR-D-13-00039.1.

Zhuang, Y., R. Fu, J. A. Marengo, and H. Wang, 2017: Seasonal variation of shallow-to-deep convection transition and its link to the environmental conditions over the Central Amazon. *J. Geophys. Res. Atmospheres*, **122**, 2649–2666, https://doi.org/10.1002/2016JD025993.

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., & Xie, P. (2014). Integrated Multi-satellitE Retrievals for GPM (IMERG), version 4.4. *NASA's Precipitation Processing Center*, ftp://arthurhou.pps.eosdis.nasa.gov/gpmdata/

Huffman, G. J., Stocker, E. F., Bolvin, D. T., Nelkin, E. J., & Tan, J., 2019: GPM IMERG Final Precipitation L3 Half Hourly 0.1 degree x 0.1 degree V06. Retrieved from: ftp://arthurhou.pps.eosdis.nasa.gov/gpmdata/.

TABLES

Table 1. Summary of observations used for data assimilation and validation.

Source	Measurement/Instrument	Observed or retrieved properties	Assimilation	Validation
IMERG	Satellite	Rain rate		X
SIPAM	Doppler Radar	Rain rate		X
NCEP GDAS	Radiosonde, surface station, ship and satellite	Wind, humidity, temperature, pressure and radiance	X	
ARM T3	Radiosonde	Wind, humidity, temperature and pressure	X (00, 12 UTC)	X (06, 15, 18 UTC)
	Surface station	Wind, humidity and temperature		X
IGRA	Radiosonde	Wind, humidity and temperature		X

Table 2. Parameterization schemes and corresponding references employed in ES3M atmospheric model version 1 (EAMv1).

Parameterization	Description	Reference
Cloud Microphysics	2-moment cloud microphysics	Gettelman and Morris (2015); Gettleman et al. (2015)
Turbulence and Shallow Convection	Cloud Layer Unified by Binormals (CLUBB)	Bogenschutz et al. (2013); Golaz et al. (2002); Larson and Golaz (2005); Larson (2002)
Deep Convection	With addition of convective momentum transports and a modifies dilute plume calculation	Zhang and Mcfarlane (1995); Richter and Rasch 2008; Neal et al. (2008)
Aerosol	4-mode version of the modal aerosol module (MAM4) with improved treatments of sea spray aerosols, secondary organic aerosols and processes in scavenging, transport and microphysics	Liu et al. (2016); Burrows et al. (2014); Shrivastava et al. (2015); Lou et al. (2020); Wang et al. (2020)

FIGURES

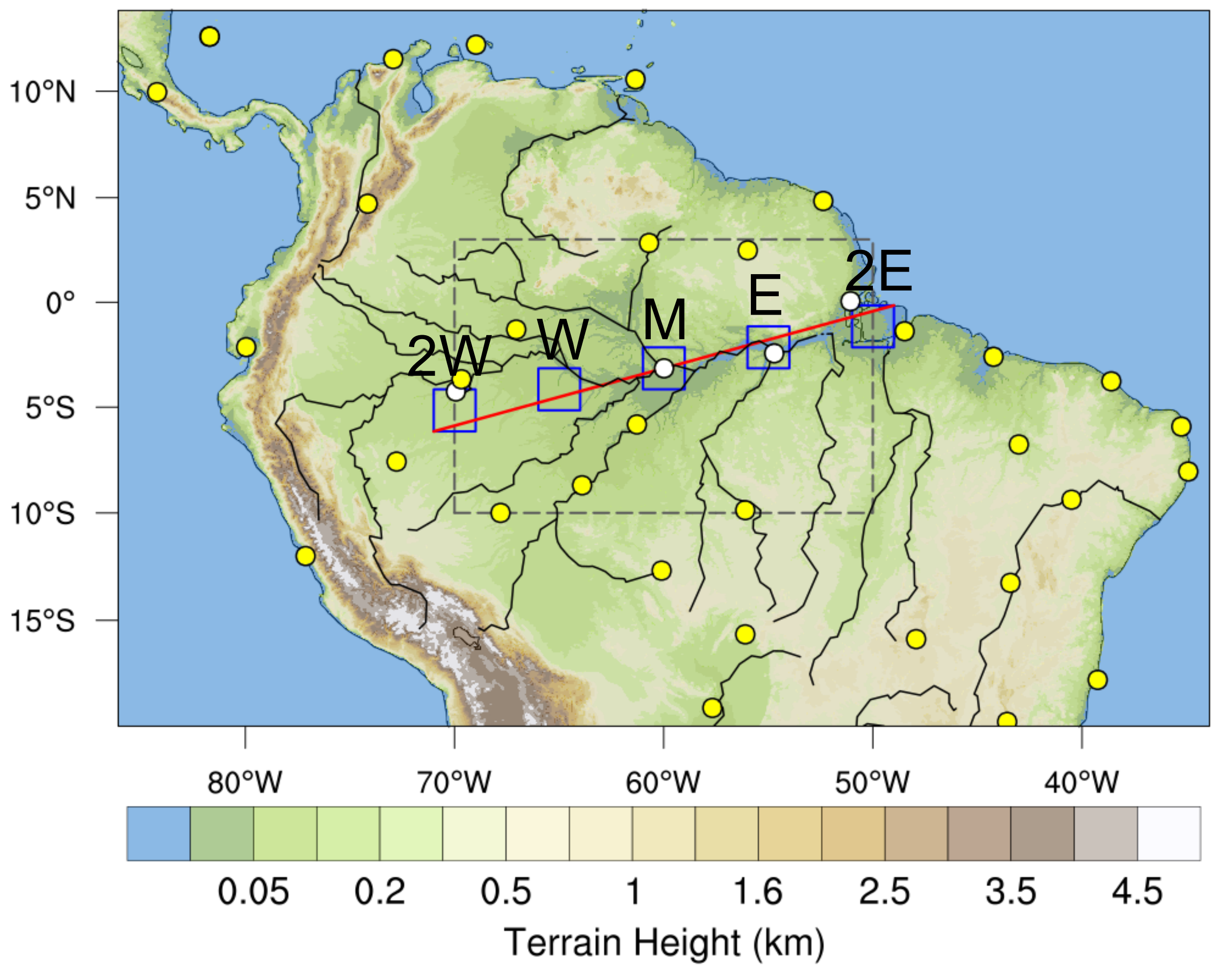


Figure 1. Simulation domain of the WRF model where color shading represents terrain height (km). Most of the analyses are performed within the region bounded by the dashed-line rectangle. The five blue squares are subregions used for precipitation analysis across the basin. The four white dots indicate where radiosonde sites associated with the subdomains are located. The red line is the location of a vertical cross-section to examine parameter associated with propagating convection. The yellow dots depict where the assimilated radiosonde profiles (in addition to the four white dots) are located (example from assimilation of 12 UTC on March 12, 2014).

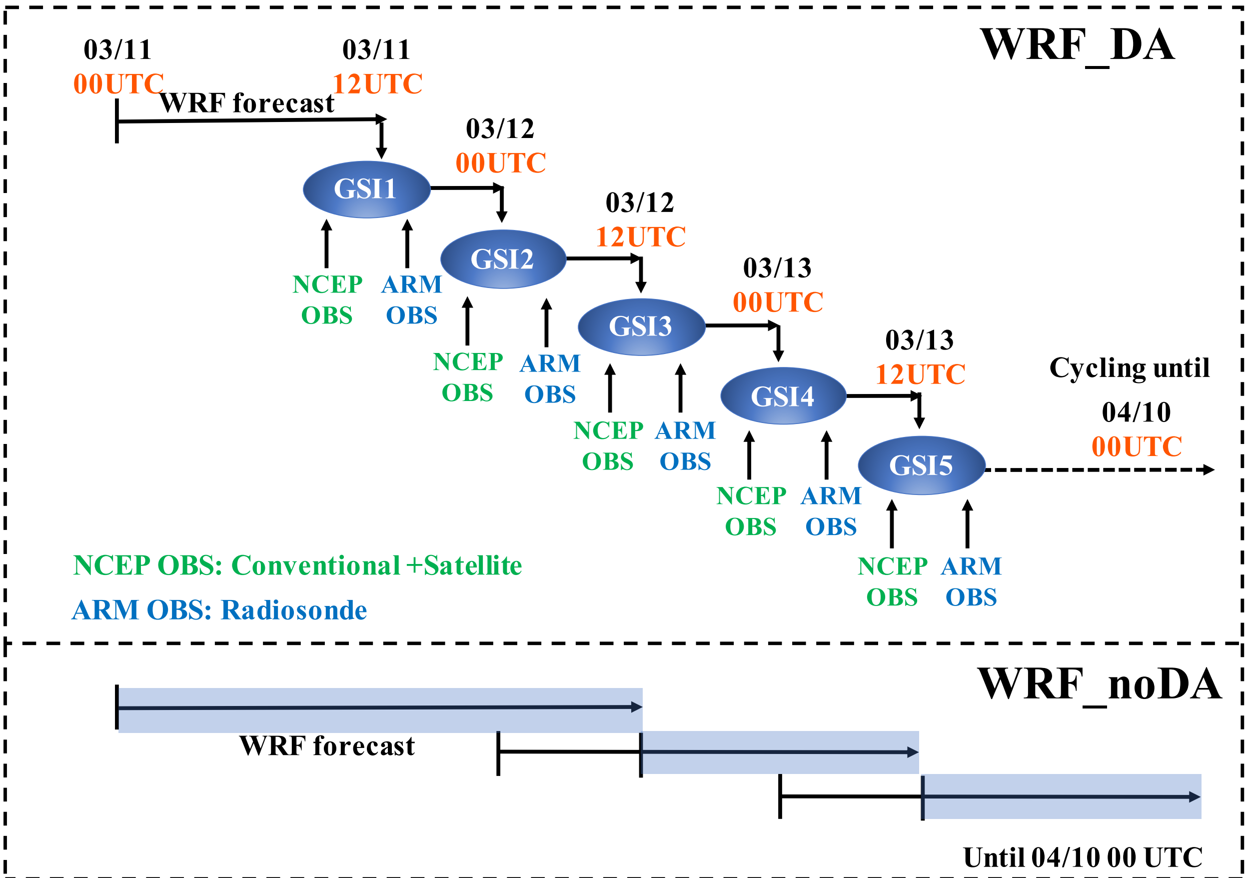


Figure 2. Schematic diagram to illustrate the data assimilation cycles and forecast periods for the WRF_DA and WRF_noDA simulations.

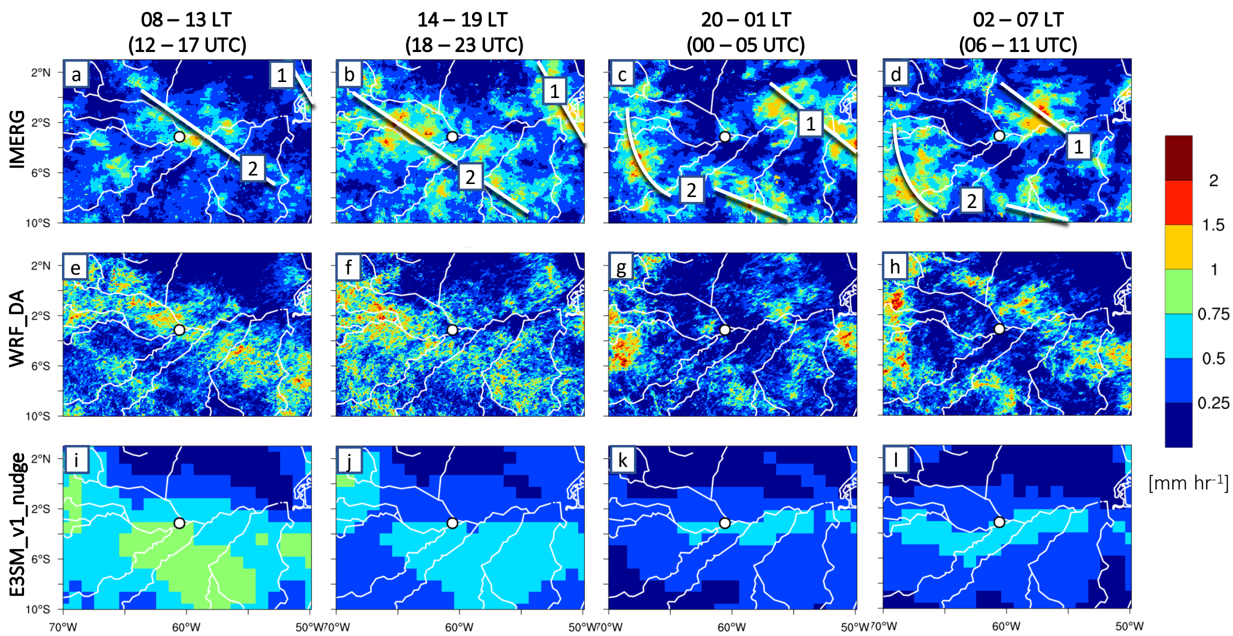


Figure 3. Mean rain rate maps over the central Amazon basin for four time periods including 08 – 13, 14 – 19, 20 – 01, and 02 – 07 LT. Note the time in LT is for Manaus. The IMERG, WRF_DA, and E3SM_v1_nudge results are given in (a) – (d), (e) – (h), and (i) – (l), respectively. In (a) – (d), the lines of precipitation are labeled by numbers for identification. The white dot is the T3 site and the white lines are rivers.

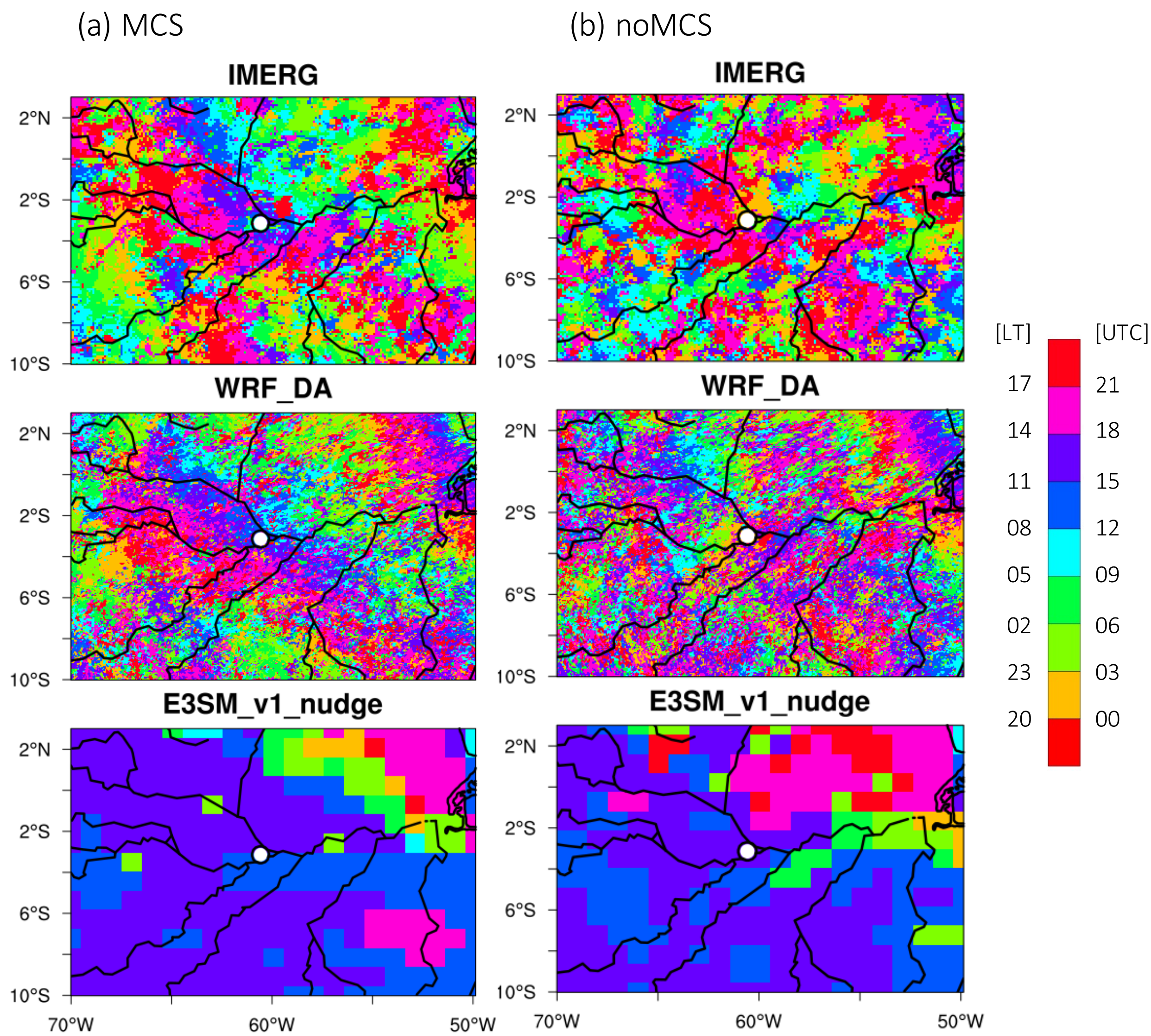


Figure 4. Spatial distribution of hour of maximum rain rate from IMERG, WRF_DA, E3SM_v1_nudge for (a) days with a propagating MCS (MCS group) and (c) days without a propagating MCS (noMCS group). The hours in LT (UTC) are given on the left (right) side of color bar. The white dot is the T3 site, and the black lines are rivers.

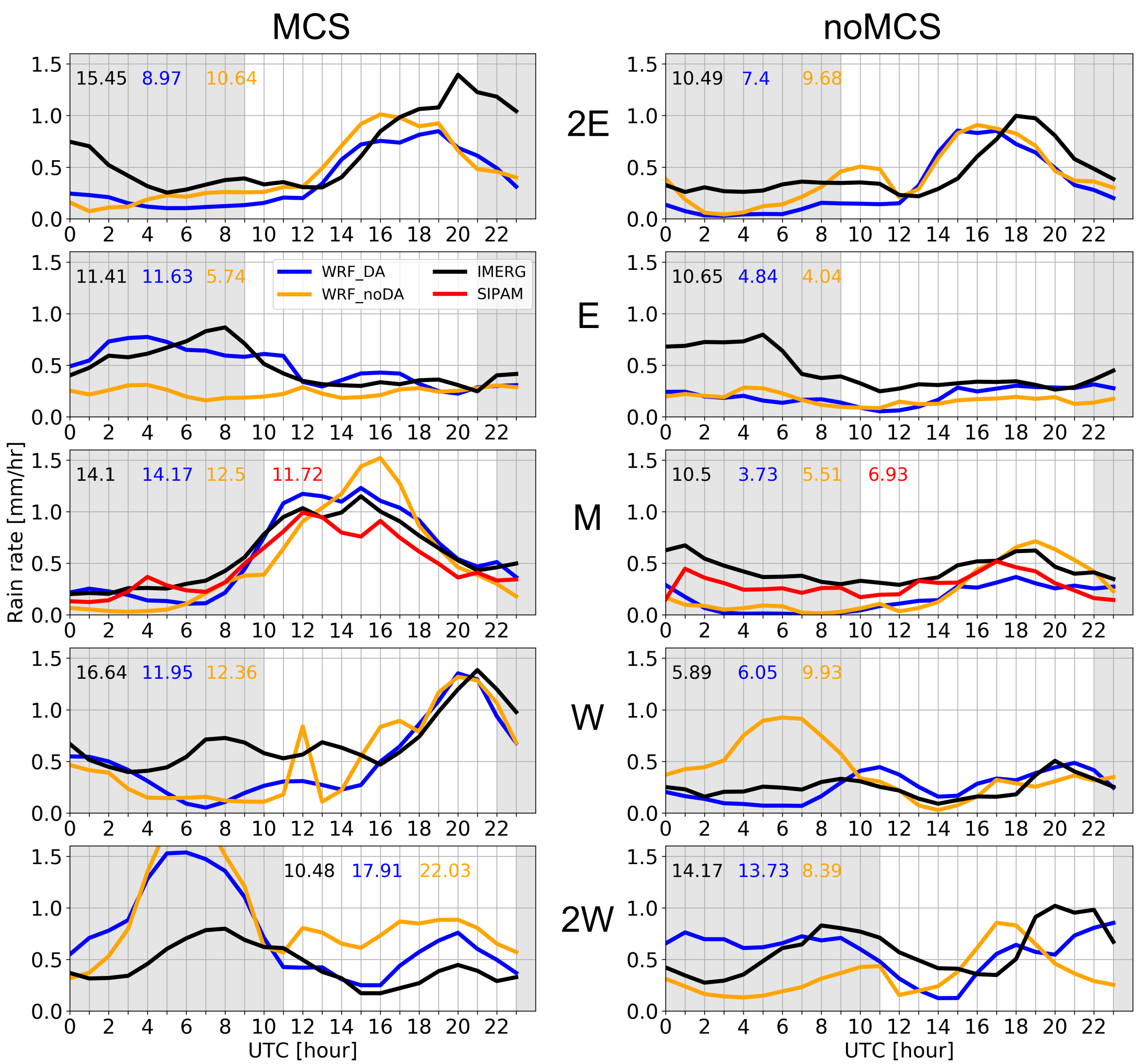


Figure 5. Diurnal cycle of mean rain rate (mm hr^{-1}) from the MCS and noMCS groups over five subdomains (2E, E, M, W, 2W) shown in Figure 1. The grey patch in each panel indicates the nighttime period (from 18 to 06 LT) at each location. IMERG and SIPAM (available only in subdomain M) observations denoted by black and red, respectively, while WRF_DA and WRF_noDA simulation denoted by blue and orange, respectively. Mean daily rain rate (mm day^{-1}) calculated from each diurnal cycle is given with corresponding color at upper left or right corner of each panel.

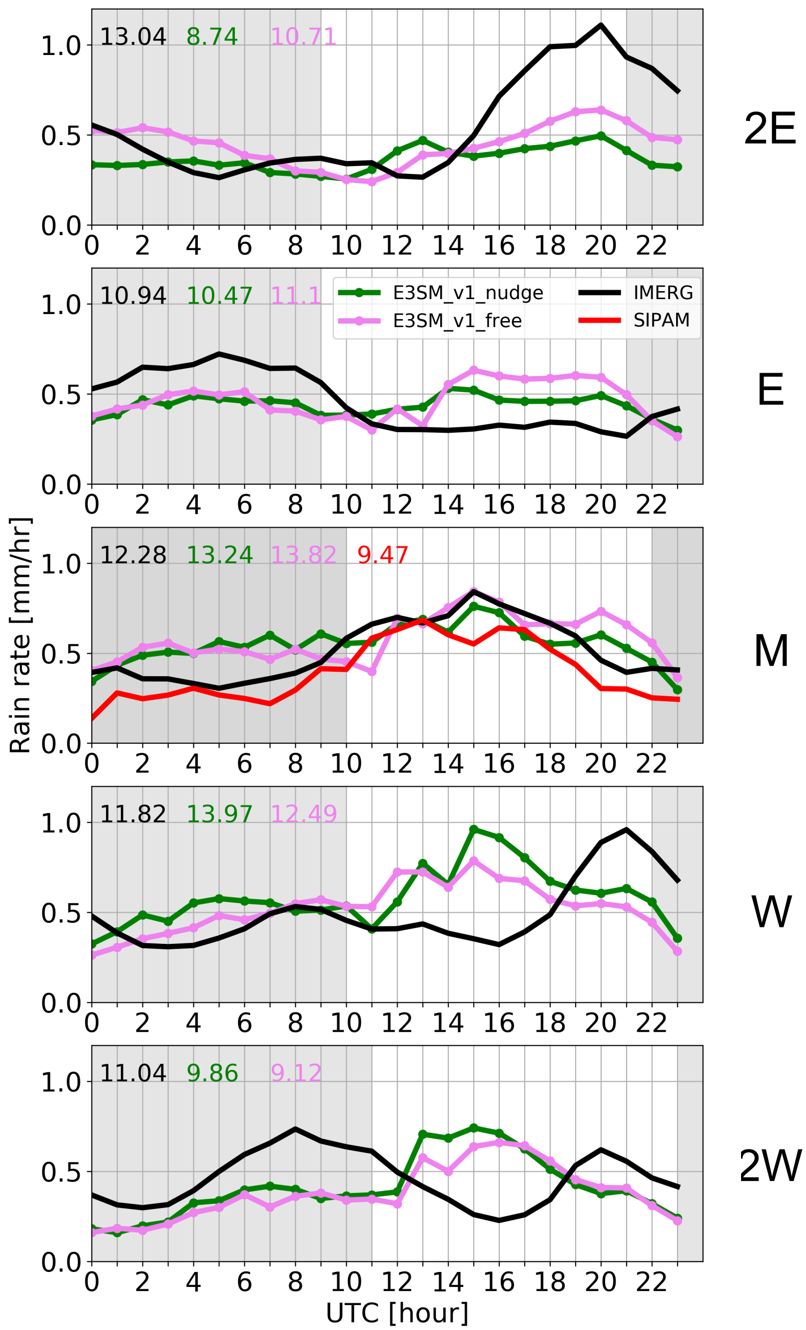


Figure 6. Diurnal cycle of mean rain rate (mm hr^{-1}) computed for the entire study period over five subdomains (2E, E, M, W, 2W) shown in Figure 1. The grey patch in each panel indicates the nighttime period (from 18 to 06 LT) at each location. IMERG and SIPAM (available only in subdomain M) observations denoted by black and red, respectively, while E3SM_v1_free and E3SM_v1_nudge, simulation denoted by violet and dark green, respectively. Mean daily rain rate (mm day^{-1}) calculated for each diurnal cycle is given with corresponding color at upper left corner of each panel.

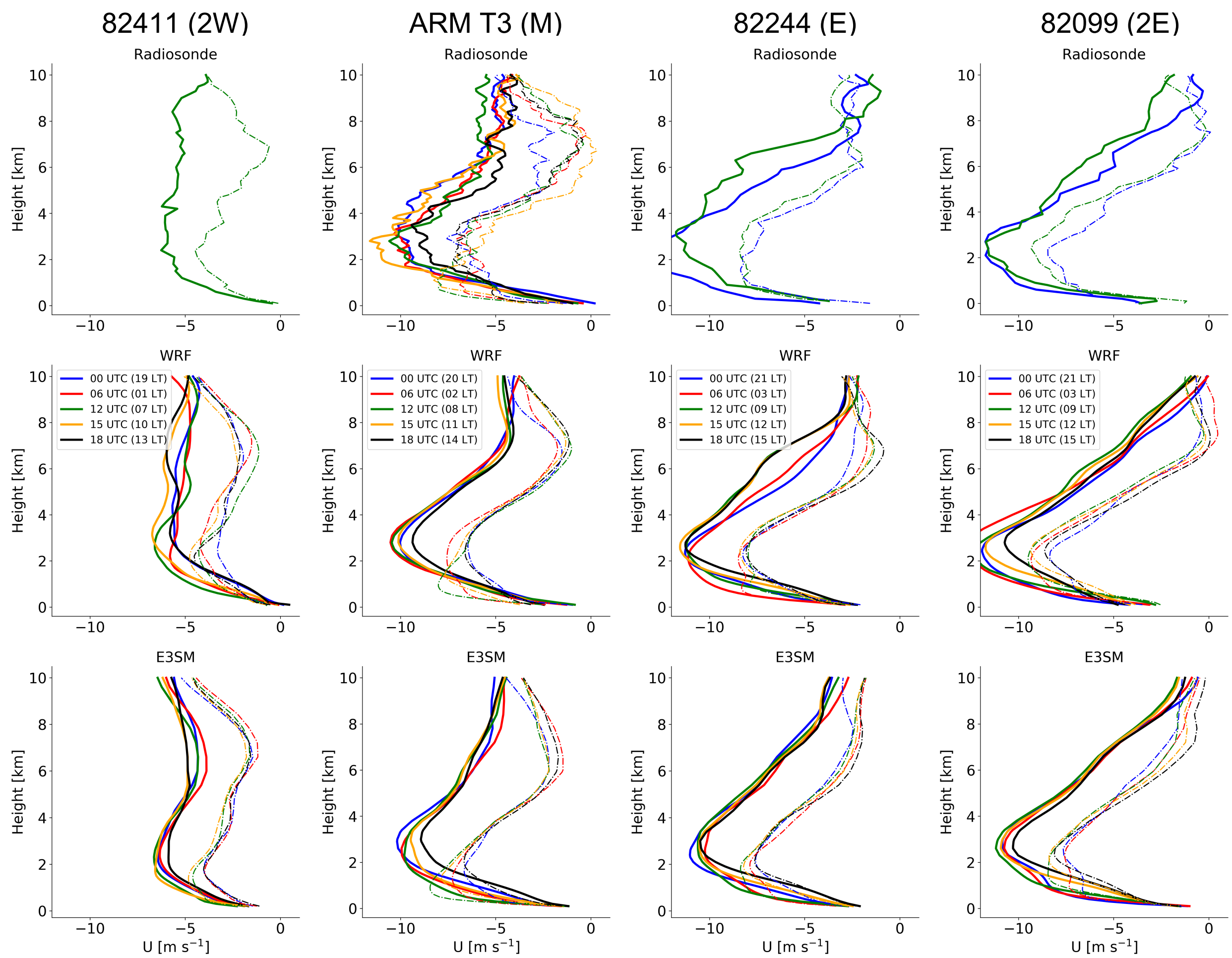


Figure 7. Comparison of observed mean zonal wind profiles at four locations along with the corresponding simulated mean zonal wind profiles from WRF and E3SM. Columns from left to right display results at observational sites 82411, ARM T3, 82244, and 82099, which correspond to nearby subdomains 2W, M, E, and 2E shown in Figure 1, respectively. Profiles obtained at different hours of a day can be distinguished by the line colors as indicated in the legend which is given in the middle row of each column. Solid and dashed lines represent profiles from the MCS and noMCS groups, respectively.

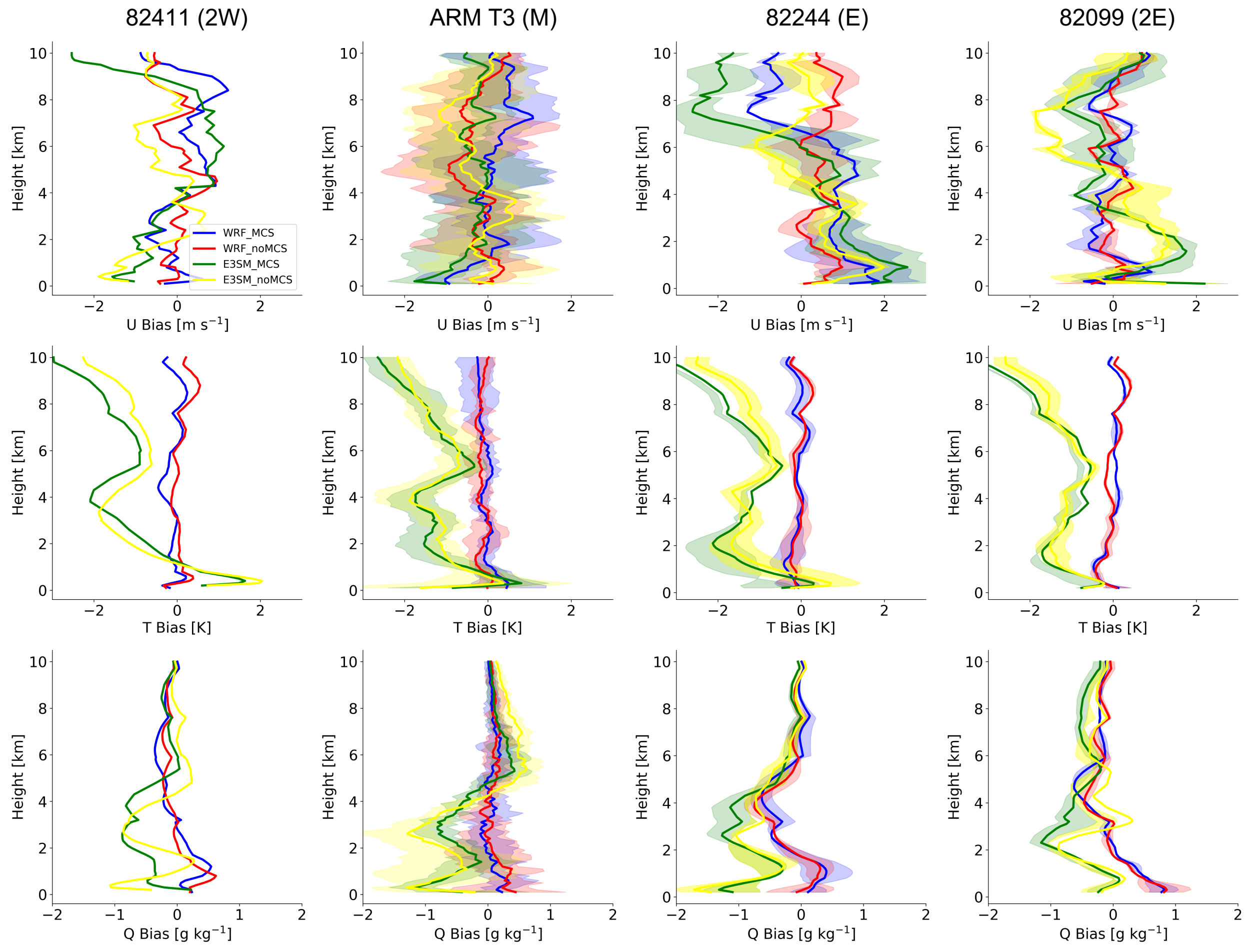
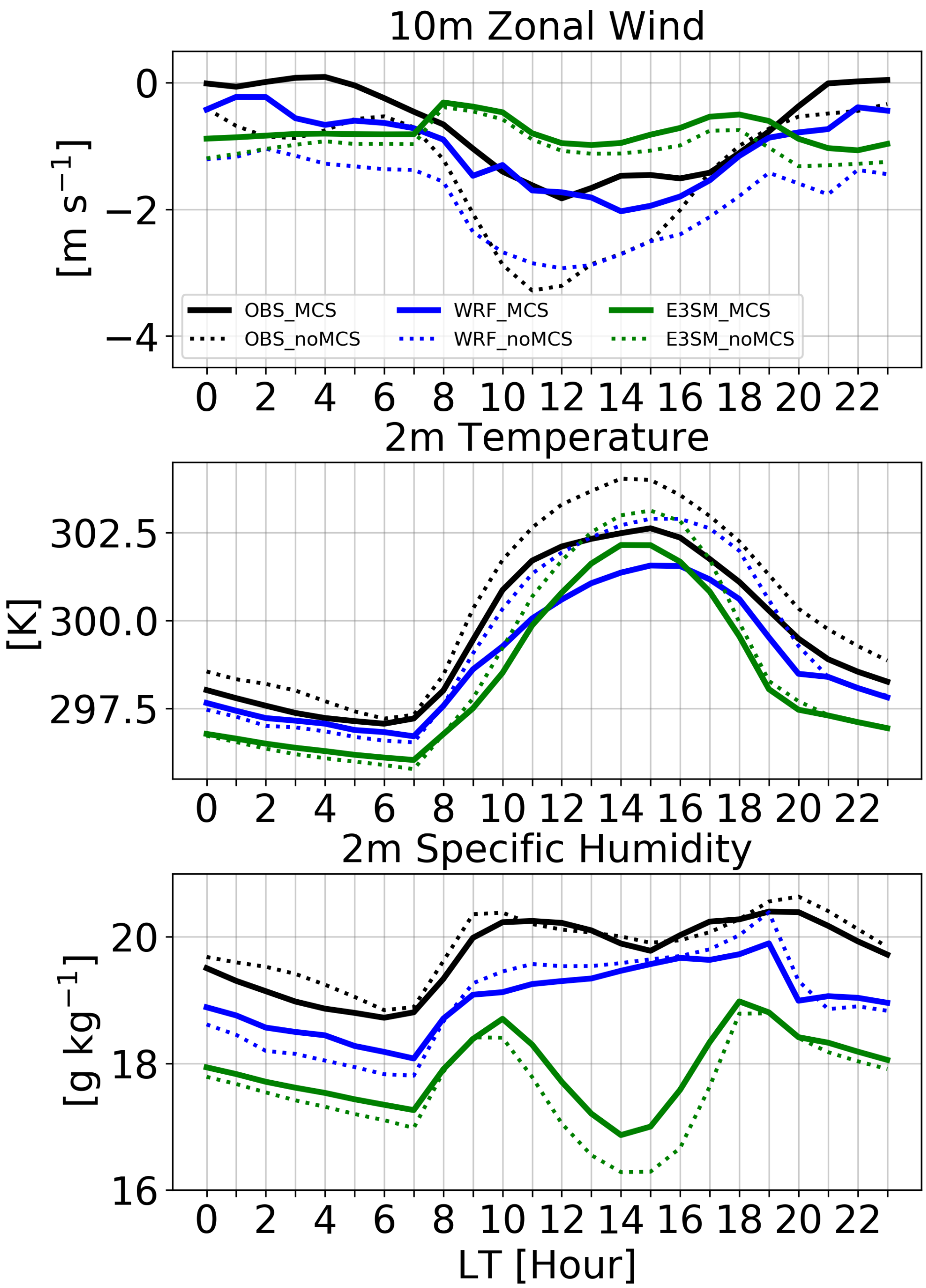


Figure 8. Comparison of WRF and E3SM bias profiles at four subregions. Rows from top to bottom show biases of zonal wind (U), temperature (T), and specific humidity (Q). Columns from left to right display results at observational sites 82411, ARM T3, 82244, and 82099, which correspond to nearby subdomains 2W, M, E, and 2E, respectively. Shading represents the variability among profiles obtained in a day, which varies by site.



1244
1245 Figure 9. Comparison of observe and simulated surface mean diurnal variation in a) zonal wind, b) temperature, and c) specific humidity surface meteorology at ARM's T3 site. Observations denoted in black, while WRF and E3SM simulation denoted by blue and green, respectively. The grey patch in each panel indicates the nighttime period (from 18 to 06 LT) at each location. Solid and dashed lines represent results from the MCS and noMCS groups, respectively.

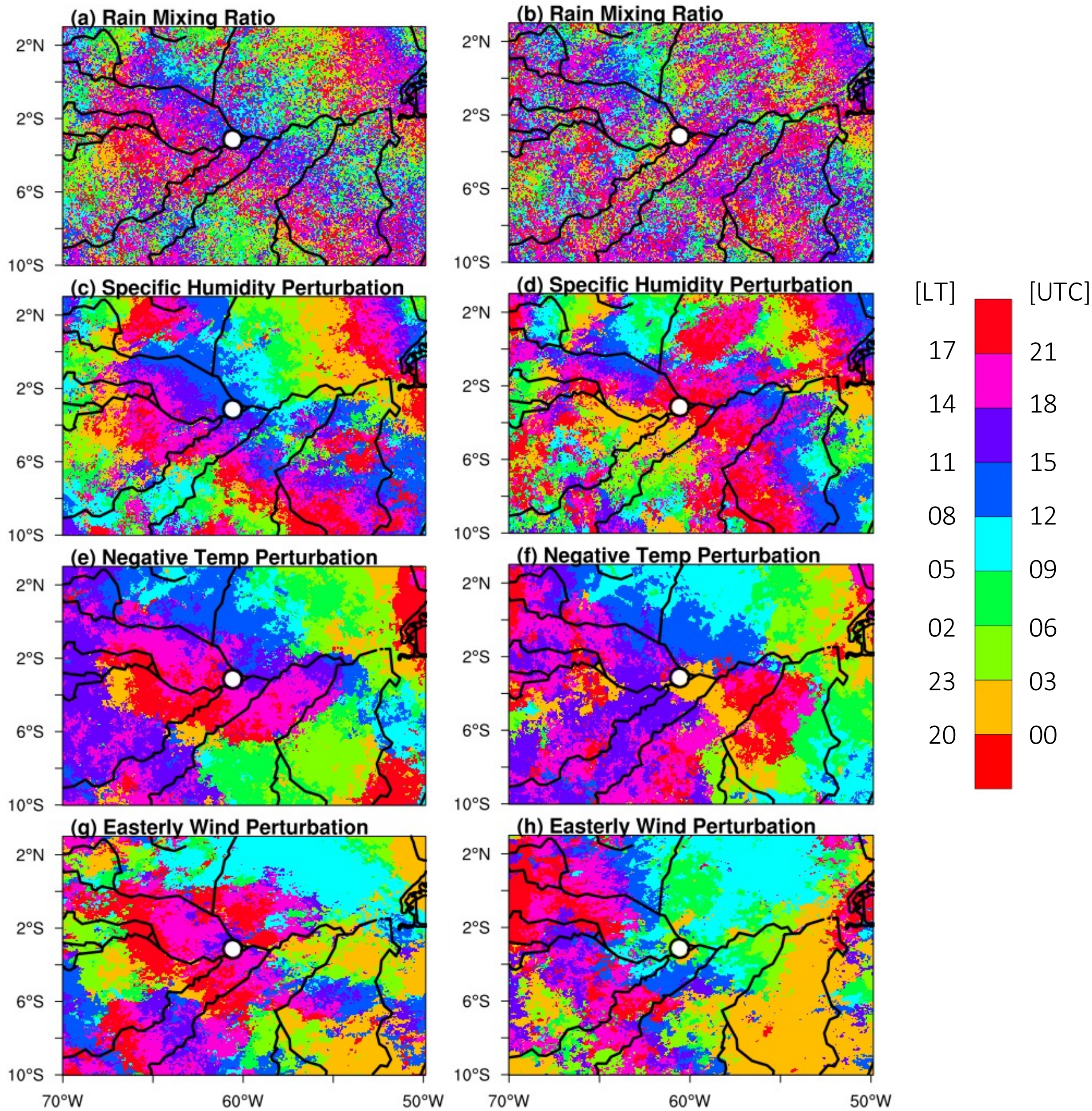


Figure 10. Spatial distribution of time with the local daily maximum of WRF-simulated rain mixing ratio along with the specific humidity, negative temperature, and easterly wind perturbations over the domain with respect to the MCS (a, c, e, and g) and noMCS (b, d, f, and h) groups. The hours in LT (UTC) are given on the left (right) side of color bar. White dot denotes the location of ARM T3 site.

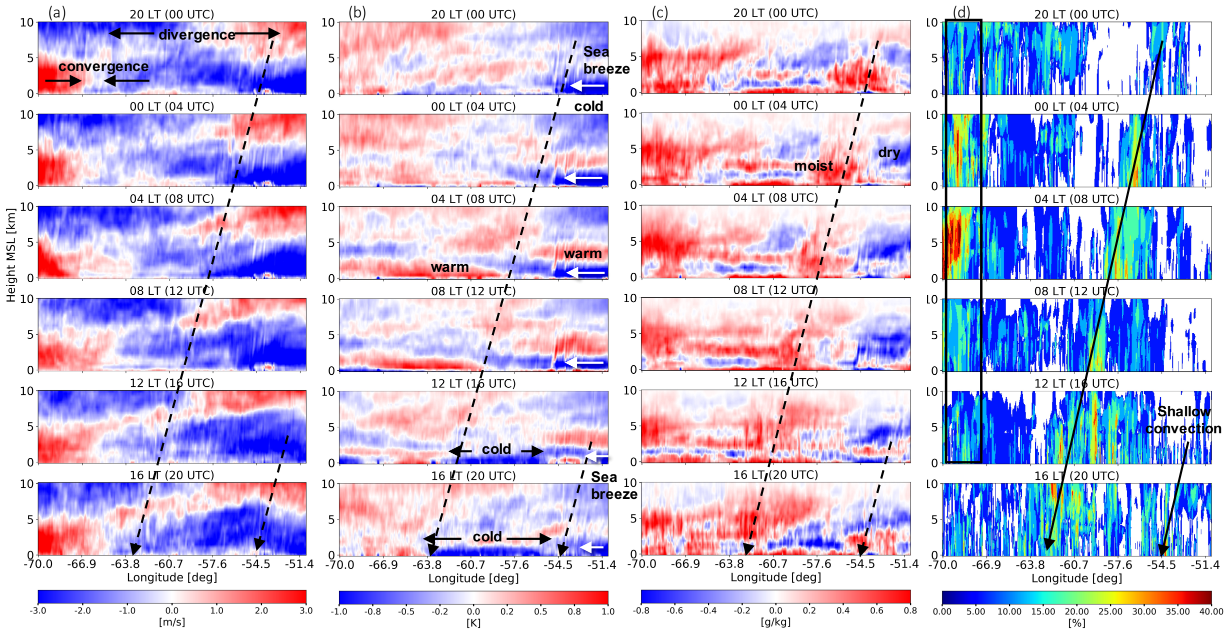


Figure 11. Vertical cross-sections to illustrate diurnal evolution of WRF-simulated tropospheric flow associate with convective activity along the path of MCS propagation, including (a) zonal wind (m s^{-1}), (b) temperature (K), (c) specific humidity (g kg^{-1}) perturbations, and (d) occurrence frequency (%) of reflectivity to be greater than 15 dBZ from the MCS group. The long arrows (solid and dashed) in each panel represent the front edge of propagating convective systems. White arrows in (b) denote the diurnal variation of sea breeze. The LT time given in the heading of each panel is for Manaus.

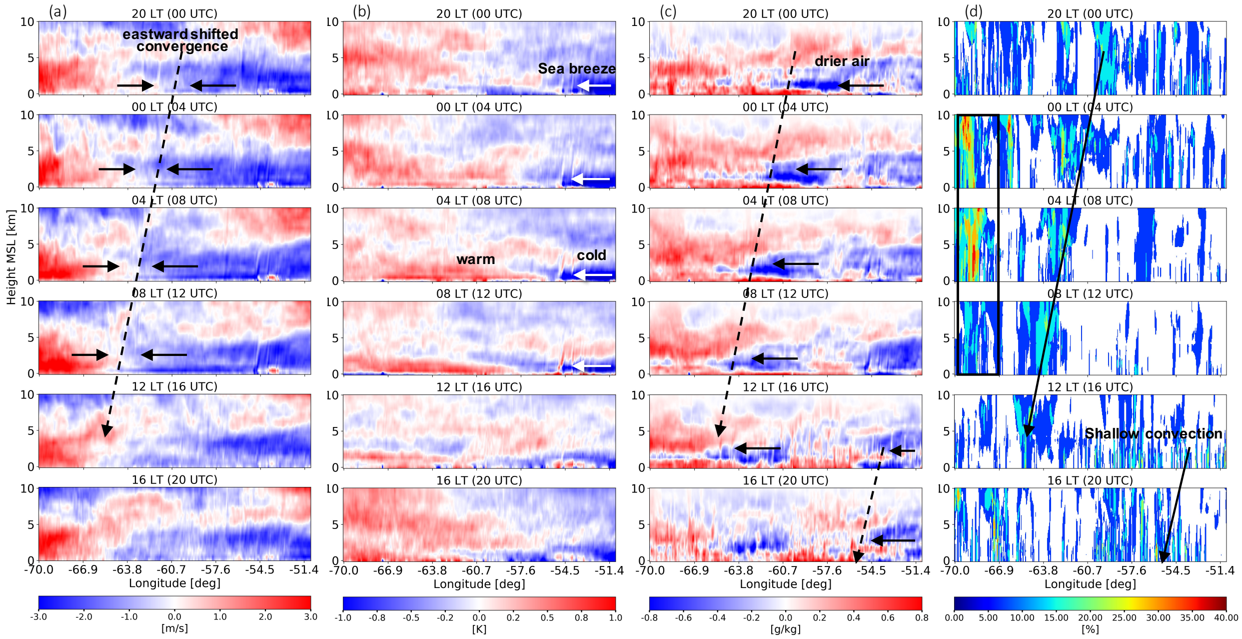


Figure 12. Similar to Figure 11 but for WRF simulations from the noMCS group.