Evaluation of Mesoscale Convective Systems in High Resolution E3SMv2

Meng Zhang[1], Shaocheng Xie[1*], Zhe Feng[2], Christopher R. Terai[1], Wuyin Lin[3], Chih-Chieh-Jack Chen[4], Jiwen Fan[5], Jean-Christophe Golaz[1], L. Ruby Leung[2], Jadwiga H. Richter[4], Yunpeng Shan[2], Xiaoliang Song[6], Qi Tang[1], Guang Zhang[6]

[1] Lawrence Livermore National Laboratory, [2] Pacific Northwest National Laboratory, [3] Brookhaven National Laboratory, [4] National Center for Atmospheric Research, [5] Argonne National Laboratory, [6] Scripps Institution of Oceanography

Corresponding author: Shaocheng Xie, xie2@llnl.gov

Key points:

- Simulated MCS precipitation and occurrences are substantially underestimated in E3SMv2 over both tropical and CONUS regions.
- MCS defined by both cloud shield and surface precipitation provides a more stringent assessment on the model capability in simulating MCSs.
- Simulated MCS properties in E3SMv2 are not significantly improved with the new cloud and convection parameterizations developed for E3SMv3.

Abstract

Mesoscale convective systems (MCSs) play an important role in modulating the global hydrological cycle, general circulation, and radiative energy budget. In this study, we evaluate MCS simulations in the second version of U.S. Department of Energy (DOE) Energy Exascale Earth System Model (E3SMv2). E3SMv2 atmosphere model (EAMv2) is run at the uniform 0.25° horizontal resolution. We track MCSs consistently in the model and observations using the PyFLEXTRKR algorithm, which defines MCS based on both cloud-top brightness temperature (T_b) and surface precipitation. Results from using T_b only to define MCS, commonly used in previous studies, are also discussed. Furthermore, sensitivity experiments are performed to examine the impact of new cloud and convection parameterizations developed for EAMv3 on simulated MCSs.

Our results show that EAMv2 simulated MCS precipitation is largely underestimated in the tropics and contiguous United States. This is mainly attributed to the underestimated precipitation intensity in EAMv2. In contrast, the simulated MCS frequency becomes more comparable to observations if MCSs are defined only based on cloud-top T_b. The T_b-based MCS tracking method, however, includes many cloud systems with very weak precipitation which conflicts with the MCS definition. This result illustrates the importance of accounting for precipitation in evaluating simulated MCSs. We also find that the new physics parameterizations help increase the relative contribution of convective precipitation to total precipitation in the tropics, but the simulated MCS properties are overall not significantly improved. This suggests that simulating MCSs will remain a challenge for the next version of E3SM.

Plain Language Summary

Mesoscale convective systems (MCSs) are one of the largest forms of deep convective storms, which play an important role in the earth system. It is imperative for global climate models to reasonably simulate the MCS properties. This study aims to evaluate the simulated MCS properties in the second version of U.S. Department of Energy (DOE) Energy Exascale Earth System Model (E3SMv2). We utilized two different approaches to define and track MCSs in the model and observations for consistent comparisons. Our results show that the E3SMv2 model underestimates MCS precipitation in the tropics and contiguous United States regions. The too weak precipitation intensity is the primary reason for this MCS precipitation bias. The simulated MCS number becomes more comparable to the observations when precipitation features are not included in the MCS definition. However, many cloud systems with precipitation characteristics not associated with MCSs are falsely included. Therefore, this comparison illustrates the importance of accounting for precipitation features in evaluating simulated MCSs. In addition, by examining the impact of new physics parameterizations that are developed for the next generation of E3SM model on the MCS simulation, we find simulating MCSs will remain a challenge for the next version of E3SM model.

1. Introduction

Mesoscale convective systems (MCSs) are the largest form of cumulonimbus cloud aggregates that cover a horizontal scale of hundreds of kilometers with lifetimes that can last more than 24 hours (Houze et al., 2004, 2018). Observations show that MCSs are ubiquitous over the tropics (Nesbitt et al., 2006; Yuan & Houze, 2010) and contribute to more than 50% of total precipitation in the tropical region (Nesbitt et al., 2006; Feng et al., 2021a), as well as over the Great Plains east of the Rocky Mountains in the contiguous United States (CONUS) region (Feng et al., 2016; Haberlie & Ashley, 2019). MCSs contain both active convective towers and extensive stratiform clouds, which differentiates them from ordinary convective storms (Houze et al., 2018). The presence of robust stratiform anvil clouds and precipitation in MCSs produce top-heavy heating profiles that impact global circulations (Schumacher et al., 2004) and feedback on the evolution of MCS lifecycle (Yang et al., 2017, 2023). Using long-term observations over the past decades, the frequency and intensity of springtime MCSs are found to increase in the central U.S., which is associated with a strengthening of the southerly low-level jet and associated moisture transport in the Central and Northern Great Plains (Feng et al., 2016; You & Deng, 2023). Such an increase in MCS frequency and intensity also suggests a potential future increase in extreme rainfall occurrence in the warming climate (Prein et al., 2017). Therefore, MCS plays an important role in the global hydrological cycle, large-scale state environments, and global energy budget.

To better understand the changes of MCSs in the future climate, it is imperative to accurately represent their key characteristics in regional and global climate models. However, there are large uncertainties in current numerical models with respect to the representation of essential cloud processes associated with MCS formation and development (e.g., Fan et al.,

2017; Moncrieff, 2019). This is particularly true for conventional general circulation models (GCMs) with a coarser horizontal resolution and convective parameterizations to simulate the multiscale interactions in MCSs (Feng et al., 2021b; Hsu et al., 2023). For GCMs that participate in the Coupled Model Intercomparison Project Phase 5 (CMIP5), most models simulate a severe underestimation of summertime precipitation over the central U.S. (Lin et al., 2017; Mueller & Seneviratne, 2014), which is a longstanding issue that is believed to be associated with the failure to capture strong precipitation events produced by MCSs (Klein et al., 2006; Van Weverberg et al., 2018; Xie et al., 2019; Zheng et al., 2019). However, low-resolution GCMs are computationally efficient tools currently used in century-long climate projections and to understand changes in global convection and cloud properties under future climate scenarios. Therefore, it remains important to understand and improve the representation of convection and MCS processes in GCMs with coarser horizontal resolutions.

Previous studies show that GCMs can simulate MCSs that are comparable to observations at ~50 km horizontal resolution on the global scale. Dong et al. (2021) compared the characteristics of tropical MCSs using high resolution (~50 km) Geophysical Fluid Dynamics Laboratory (GFDL) AM4 model (C192AM4, Zhao, 2020) with a comprehensive long-term observational dataset. They showed that the spatial distribution of MCSs as well as the seasonality and interannual variability of MCS frequency over different land and oceanic regions are reasonably simulated. Dong et al. (2023) additionally suggested that the spatial distribution and seasonality of genesis frequency of MCSs during spring to early summer are also broadly in agreement with observations over the central U.S. However, the identified MCSs in these two studies are purely based on the cloud-top brightness temperature (T_b) data (Huang et al., 2018). More recently, a new MCS tracking algorithm has been developed that uses both T_b and surface

precipitation characteristics (Feng et al., 2023). By comparing the difference in tracked MCSs using the two approaches on the global scale, Feng et al. (2023) found that the T_b-only tracking method produces more MCS occurrences in the midlatitudes compared with the T_b and precipitation method. The false MCS identification by the T_b-only method is related to cloud systems that have long lifetime and cover a large area but generate very low surface precipitation intensity. These cloud systems are more likely to mainly contain stratiform-type precipitation associated with synoptic systems which is inconsistent with the typical MCS precipitation characteristics. Therefore, combining both T_b and precipitation features to track MCS should be more accurate in terms of capturing the essential MCS characteristics.

This more advanced MCS tracking algorithm has been used in Wang et al. (2021) to evaluate the simulation of MCS in the Department of Energy's Energy Exascale Earth System Model version 1 (E3SMv1) (Golaz et al., 2019). Wang et al. (2021) found that the E3SM atmosphere model (EAMv1) (Rasch et al., 2019; Xie et al., 2018) can reasonably capture the observed spatial pattern of spring season total precipitation in the CONUS region with a regional refined model (RRM) setup featuring 0.25° model resolution over the CONUS (Tang et al., 2019). However, the model greatly underestimates heavy precipitation over the southern states in the CONUS, and thus underestimates the MCS precipitation and MCS occurrences compared to the observations. Similar underestimation in MCS precipitation was found in the central U.S. and Indo-Pacific region when evaluating the global 0.25° E3SMv1 results (Xie et al., 2020). In addition, the underestimation of MCS precipitation still exists over the CONUS in summertime, even though a cloud resolving model is coupled in E3SMv1 using the super-parameterization approach (Lin et al., 2022), which suggests the deficiencies in model capability to simulate MCS events in E3SMv1.

The second version of E3SM along with its North American RRM version has recently been released (Golaz et al., 2022; Tang et al., 2023). E3SMv2 includes minor improvements in its physics parameterizations but with significantly retuned cloud and convection related parameters. For its updated physical parameterizations, a new convective trigger described in Xie et al. (2019) is implemented in E3SMv2 to improve its simulation of precipitation and its diurnal cycle (Golaz et al., 2022; Tang et al., 2022; Tao et al., 2022, 2023). The new trigger and the re-adjusted model parameters have also led to considerable improvements in the cloud simulation compared to E3SMv1 (M. Zhang et al., 2022, 2023; Y. Zhang et al., 2023*;* Qin et al., 2023). In this study, we perform a comprehensive evaluation of E3SMv2's capability to simulate MCS by using the MCS tracking algorithm developed in Feng et al. (2023). To make the model resolution more relevant to the horizontal scales of MCS, the global 0.25° horizontal resolution is used in this study for a global MCS evaluation. In addition, a new set of cloud and convection parameterizations that are developed for the third version of E3SM (E3SMv3) is also tested in this study to examine their impacts on the simulated MCS. To demonstrate the impact of different MCS tracking methods on the model evaluation, we apply both the T_b-only tracking and combined T_b and surface precipitation tracking in our evaluations.

The paper is organized as below. Section 2 introduces the default model physics parameterizations in E3SMv2 and the new convection and cloud microphysics parameterizations that are developed for E3SMv3. Observational dataset used for the model evaluation and the MCS tracking method are also described in this section. Section 3 discusses the E3SMv2 model evaluation results, the impact of different MCS tracking methods on the global scale and over the CONUS region, and the impact of new physics parameterizations on the MCS simulation. The summary and discussion are provided in section 4.

2. EAMv2, numerical experiments, observations, and MCS tracking

2.1. EAMv2 model

EAMv2 features a few notable changes in the atmospheric physics and significantly recalibrated tuning parameters compared to EAMv1 (Golaz et al., 2022). Specifically, the dCAPE_ULL convective trigger described in Xie et al. (2019) was implemented in the deep convection scheme (Zhang & McFarlane, 1995, ZM hereafter) in EAMv2 to improve the simulated precipitation and its diurnal cycle. The new convective trigger combines the dynamical Convective Available Potential Energy (CAPE) (dCAPE) trigger developed in Xie and Zhang (2000) to prevent CAPE from being released simultaneously after its generation and the Unrestricted air parcel Launch Level (ULL) method described in Wang et al. (2015) to allow convective instability to be detected above the boundary layer for elevated nocturnal convections. In addition, a new linearized ozone scheme is used for stratospheric ozone (Tang et al., 2021). The treatments of other physical processes are the same as in EAMv1, which include the Cloud Layers Unified by Binormals (CLUBB, Golaz et al., 2002; Larson, 2017) parameterization for shallow convection, cloud macrophysics, and boundary layer turbulence; the second version of Morrison and Gettelman (MG2, Gettelman & Morrison, 2015; Gettelman et al., 2015) cloud microphysics; the four-mode Modal Aerosol Model (MAM4, Liu et al., 2016; Wang et al., 2020), and the gravity wave parameterization following Richter et al. (2010) with updated treatments (Beres et al., 2004; Richter et al., 2019). In addition, significant re-adjustments were made to a number of parameters used in cloud microphysics, CLUBB, and deep convection schemes to improve the cloud and precipitation simulation and cloud radiative forcing (Ma et al., 2022). In this study, we run EAMv2 at 0.25° horizontal resolution globally

with a 900 second time step. Note that this 0.25° horizontal resolution model configuration is not officially supported. However, the cloud and precipitation climatology remains reasonable in our model validation compared with the standard low resolution version of E3SMv2. Thus, it is suitable for this study.

2.2. Numerical experiments

2.2.1. Simulation setup

In this study, EAMv2 is run at a global uniform 0.25° horizontal resolution from 2004 to 2009. The sea surface temperature and sea ice are prescribed by weekly observational data from the National Oceanic and Atmospheric Administration (NOAA) Optimum Interpolation Sea Surface Temperature version 2 (OISST v2) product (Huang et al., 2021). The model simulation results between 2005 and 2009 are used for the MCS evaluation with the first-year results discarded for model spin-up. Hourly outputs of global surface precipitation flux and outgoing longwave radiation are saved and used for the MCS tracking. The MCS tracking approach is introduced in Section 2.4.

2.2.2. Sensitivity tests with cloud and convection physical parameterizations

To examine the impact of model physics on the simulation of MCS, we perform a set of sensitivity experiments with different cloud and convection parameterizations developed for the next version of E3SM atmosphere model (i.e., EAMv3). One of the major developments is the use of predicted particle properties (P3) cloud microphysics scheme (Milbrandt & Morrison, 2016; Milbrandt et al., 2021; Wang et al., 2021) to replace the MG2 stratiform cloud microphysics that was used in EAMv1 and EAMv2. Convective cloud parameterizations

200 received several significant updates during the model development. First, the two-moment
201 convective cloud microphysics parameterization (Song & Zhang, 2011; Song et al., 2012) is
202 implemented to more physically represent the evolution of convective cloud hydrometeors and
203 their interactions with large-scale stratiform clouds and aerosols. Second, the multiscale coherent
204 structure parameterization (MCSP, Chen et al., 2021; Moncrieff et al., 2017; Moncrieff, 2019) is
205 introduced to simulate the physical and dynamical effects of organized convection that are
206 currently missed in EAMv2. Third, a cloud base mass flux adjustment described in Song et al.
207 (2023) is incorporated into the ZM scheme to improve the coupling of deep convection and its
208 associated large-scale environment. In this study, we perform four sensitivity experiments to
209 examine their individual impacts from all four new features. The control run with the default
210 EAMv2 and the sensitivity tests are summarized in Table 1. Below, we provide more details on
211 the tested new parameterizations.
212
213 **1) P3 microphysics**
214 P3 is a new bulk cloud microphysics scheme that represents the evolution of physical
215 properties of various hydrometeors in space and time (Milbrandt & Morrison, 2016; Milbrandt et
216 al., 2021). Unlike the MG2 cloud microphysics scheme used in the default EAMv2 model which
217 artificially defines separate hydrometeor categories for different ice species, P3 represents the
218 evolution of ice particle properties from ice crystals to snow and rimed particles (e.g., graupel)
219 by prognosing rimed mass and volume. This method avoids the impact of artificial separation of
220 ice species on the simulation of ice particle microphysical processes, thus improving the
221 representation of physical evolution of ice particles in the model. Considering rimed particles,
222 which is important for MCS precipitation, is another advantage of P3 compared with MG2, in

which only cloud ice and snow are considered. The current P3 implemented in E3SM is two-moment, which prognoses the total ice mass mixing ratio and ice number concentration with the predicted ice mass from riming growth and the rimed volume to track the particle growth processes (Wang et al., 2021). Note that the single ice category is used in E3SM, meaning that there is only one single type of ice particle predicted at a given time in one model grid. For liquid phase hydrometeors, a two-moment bulk scheme is used to prognose the mass mixing ratio and number concentration of cloud droplet and rain drop in their evolution. By comparing P3 with the default MG2 cloud microphysics EAMv1 RRM simulation, Wang et al. (2021) showed that P3 microphysics greatly improves the simulation of precipitation statistics over the CONUS region. The higher hourly rain rate simulated by P3 results in 20% more MCS occurrence and stronger total MCS precipitation than MG2, which agrees better with the observations.

2) Convective cloud microphysics

The convective cloud microphysics developed by Song and Zhang (2011) prognoses the mass mixing ratio and number concentration of cloud droplet, cloud ice, rain, and snow in the ZM parameterization. Cloud microphysical processes including autoconversion, collection between hydrometeor species, self-collection, freezing, ice nucleation, droplet activation, and sedimentation are represented to simulate the evolution and interaction between different hydrometeor species. Previous studies showed that this convective cloud microphysics scheme enables a more accurate convection and precipitation simulation (Song et al., 2012). The interactions between convective clouds and aerosols and large-scale clouds are also better represented in the model, in particular, the aerosol impacts on convective clouds can be examined in GCMs with this new convective cloud microphysics.

246

247 **3) MCSP**

248 To account for the important mesoscale heating associated with convective organization,
249 Chen et al. (2021) implemented the MCSP parameterization (Moncrieff et al., 2017; Moncrieff,
250 2019) in E3SM. The MCSP simulates the heating effects of the slantwise overturning structure
251 typically organized by MCSs. The heating component of MCSP contains a temperature tendency
252 of multiscale convective systems that is added to the temperature tendency simulated by the
253 existing ZM parameterization. The heating profile is represented as a top-heavy second
254 baroclinic normal mode and its amplitude is a function of the vertically averaged convective
255 heating induced by the convective parameterization. Chen et al. (2021) showed that the MCSP
256 improves the representation of the Madden-Julian Oscillation (MJO) and reduces the
257 precipitation biases over the tropical Pacific region in E3SMv1.

258

259 **4) Cloud base mass flux adjustment**

260 To represent the dynamical effects of large-scale vertical motion on the convection
261 development, model simulated cloud base mass flux adjustment is introduced in EAMv2 (Song
262 et al., 2023). The cloud base mass flux is adjusted by subtracting the grid-scale pressure vertical
263 velocity at the PBL top from the cloud base mass flux determined in the CAPE closure in the ZM
264 scheme. In this case, the moisture transported through the PBL top by large-scale vertical motion
265 becomes fully available for the convective cloud development. As simulated convection is
266 directly modulated by the large-scale dynamical circulation, such a cloud base mass flux
267 adjustment enables the ZM scheme to better represent convection generation in the low-CAPE
268 environment. Song et al. (2023) indicated that the cloud base mass flux adjustment substantially

improves the climate variability across multiple scales, from the precipitation diurnal cycle to the MJO.

2.3. Observations

The observational dataset used to evaluate global MCS properties are the NASA Global Merged IR V1 infrared T_b product (Janowiak et al., 2017) and the Global Precipitation Measurement (GPM) Integrated Multi-satelliteE Retrievals (IMERG) V06B precipitation data (Huffman et al., 2019a, 2019b, 2019c). The global T_b data are derived from geostationary satellites that cover the region between 60°S-60°N latitudes at 4 km pixel resolution. The IMERG precipitation is estimated from various precipitation-relevant satellite passive microwave sensors at 10 km horizontal resolution. The hourly global T_b and precipitation satellite observational datasets are regridded to 0.25° (~25 km) horizontal resolution to match the model grid spacing.

Over the CONUS region, in addition to the coverage of satellite T_b and precipitation observations, radar reflectivity from the National Weather Service Next-Generation Radar (NEXRAD) and the Stage IV multi-sensor precipitation datasets are also available. Following Feng et al. (2019), the NEXRAD radar reflectivity data, the Stage IV precipitation estimates, and the Merged IR V1 infrared T_b product are combined to derive the MCS tracking product (Feng, 2019). Note that the original horizontal resolution of this MCS product over the CONUS is 4 km. Same as the global satellite MCS data, this radar and rain gauge observation dataset is also regridded to 25 km, consistent with the model simulations. The MCS products from 2005 to 2009 are used to evaluate EAMv2 model simulations. Table 2 summarizes the observational datasets used for the MCS tracking in this study.

292

293 **2.4. MCS tracking method**

294 This study uses the PyFLEXTRKR (Python FLEXible object TRacKeR) software
295 package (Feng et al., 2022) to identify and track the time evolution and spatial distribution of
296 MCS and calculate the statistics of MCS properties in the observation and EAMv2 model
297 simulations. PyFLEXTRKR (Feng et al., 2023) is a flexible atmospheric feature tracking
298 software package with specific capabilities to track MCS features based on T_b and precipitation
299 characteristics. Note that the EAMv2 model outputs the outgoing longwave radiation (OLR)
300 instead of T_b. To consistently define simulated MCS as the observations, OLR is converted to T_b
301 in PyFLEXTRKR following the empirical method by Yang and Slingo (2001).

302 The detailed workflow for tracking MCSs is described in Feng et al. (2023). The first step
303 of MCS tracking is to identify the cold cloud system (CCS) associated with deep convective
304 events in the observation and model simulations. The CCS is detected by iteratively growing a
305 cold cloud core with $T_b < 225$ K outwards to 241 K. After each CCS segments are defined, if the
306 CCS from two consecutive time steps (1 hour apart) overlaps for more than 50% of their area,
307 the CCS pairs are linked to track their temporal evolution. Along the CCS temporal evolution, if
308 the CCS area exceeds 4×10^4 km^2 and the duration of CCS is longer than 6 hours, the track is
309 then considered as MCS. This T_b based MCS definition is similar to that used in Dong et al.
310 (2021, 2023) for the GFDL C192AM4 model evaluation, although their MCS tracking method
311 (Huang et al., 2018) was based on the thresholds of T_b and a minimum area coverage. In
312 additional to the T_b defined MCS, PyFLEXTRKR has an option to further consider surface
313 precipitation characteristics as an additional criterion for defining MCSs. For the MCS tracking
314 with precipitation, precipitation feature (PF) statistics are calculated over the regions where

precipitation rate is greater than 0.5 mm hr^{-1} underneath the CCS. Calculated PF statistics include PF centroids, area, major axis length, mean and maximum rain rate, rain rate skewness, and total and heavy rain volume. Three PF parameters are used to identify and track robust MCSs, which are the PF area, PF mean rain rate, and PF rain rate skewness. If these three PF parameters of a CCS track exceed their corresponding thresholds, that track is defined as a robust MCS. The corresponding tunable threshold values follow a linear function of duration when the largest PF major axis length is greater than 100 km. In the following discussion, if not explicitly mentioned, we refer "MCS" as the MCSs detected by the combined method using both T_b and surface precipitation. Table A1 lists all the parameters used in the MCS tracking and sampling. Note that these tunable parameters are sensitive to the data resolution, particularly for the thresholds related to precipitation. We adjusted these thresholds based on the previous work that matched MCS tracking statistics between coarse resolution (0.25°-0.5°) and high resolution (0.04°) datasets over the CONUS region (Feng et al., 2021b). More details of the tracking method can be found in Feng et al. (2023).

In this study, our evaluation primarily focuses on the robust MCSs defined with both T_b and precipitation. MCSs defined by the T_b-only method are also examined to demonstrate the impact of different MCS tracking approaches on evaluating the model capability in capturing MCS properties, which helps explain the differences between the current study and previous literatures relevant to the MCS evaluation in GCMs (e.g., Dong et al., 2021, 2023; Hsu et al. 2023).

3. Results

3.1. MCS simulated in EAMv2

3.1.1. Global MCS

We evaluate the global MCS properties simulated with the default EAMv2 model physics against observations. Figure 1 compares the global spatial distribution of annual mean total and MCS precipitation amount between 60°S and 60°N. For total precipitation, the default EAMv2 model reasonably simulates the large precipitation amount along the intertropical convergence zone (ITCZ) region. Precipitation in the South Pacific convergence zone (SPCZ) is also comparable to the IMERG observation. However, simulated total precipitation in the tropical Indian Ocean and the Maritime Continent region is underestimated by ~30% in EAMv2 (Figure 2a). Over the Amazon region, the model also slightly underestimates total precipitation compared to the observation, but strong precipitation peaks are found in the coastal region of Colombia near the equatorial eastern Pacific Ocean. In terms of the MCS precipitation (Figures 1c and 1d), observations show that MCSs greatly contribute to the total precipitation in the tropics. Observed MCS precipitation well co-locates with the spatial patterns of total precipitation occurrences, where the ITCZ, SPCZ, tropical Indian Ocean, and Maritime Continent regions have the largest MCS precipitation amount. In EAMv2, it is encouraging that the spatial distribution of simulated MCS precipitation is overall reasonably simulated. For example, simulated MCSs are dominant over the ITCZ, SPCZ, and tropical Indian Ocean. However, the simulated MCS precipitation amount is underestimated by more than 60% compared to the observation (Figure 2b). The underestimation is most substantial over the Maritime Continent and tropical Indian Ocean, while the underestimation is also noticeable over the tropical lands (e.g., Africa and Amazon) and midlatitude storm tracks.

We note that earlier literatures stated that GCMs at 50 km horizontal resolution can reasonably capture observed MCSs in the tropics (Dong et al., 2021), which seems to be

inconsistent with our current analysis. One possible reason for this is a different MCS tracking method used in Dong et al. (2021), which only used T_b. To understand the impact of different MCS tracking methods on the GCM evaluation, we additionally show the MCS precipitation defined only using T_b in both observation and EAMv2 (Figures 1e and 1f) and the MCS precipitation amount difference between two definitions (Figures 2d and 2e). Compared with the MCS precipitation defined with T_b and precipitation, Figure 2c indicates that the low bias of MCS precipitation is slightly alleviated using the T_b-only tracking method. The larger MCS precipitation from MCSs defined with T_b are more noticeable over the tropical land areas (e.g., Africa, Amazon, and Maritime Continent) and to a lesser degree over midlatitude storm tracks (Figure 2e). On the other hand, the precipitation difference between two MCS tracking methods is substantially smaller for the IMERG observation, particularly in the tropics. The small impact from MCS tracking method on the observed MCS precipitation suggests that the mesoscale cloud structures associated with observed MCSs nearly always contain heavy precipitation features. However, this is not the case for the model. This comparison suggests that using the T_b-only method could possibly overestimate the model's capability in simulating MCSs.

The contribution of MCS precipitation to total precipitation is evaluated in Figure 3 by examining the MCS precipitation fraction between EAMv2 and the IMERG data. In the tropics, the observed annual mean MCS precipitation is found to contribute to up to 90% of total precipitation, and the largest MCS precipitation contribution is found in the Indo-Pacific region. This is consistent with previous studies (Nesbitt et al. 2006; Feng et al., 2021a). In contrast, the default EAMv2 simulated MCS precipitation fraction is significantly lower than observed (Figure 3e). The simulated MCS precipitation fraction rarely reaches 80%. Using the T_b-only tracking method, the MCS precipitation fraction increases largely in the tropical land (Africa and

Amazon) and midlatitude storm tracks compared to MCSs defined with combined T_b and precipitation in EAMv2 (Figure 3f). However, the simulated MCS precipitation fraction remains underestimated compared to observations using the T_b-only method.

Figure 4 compares the MCS precipitation frequency between EAMv2 and the observations. The MCS precipitation frequency is calculated as the number of hours MCS precipitation occurred (rain rate > 0.5 mm h^{-1}) divided by total number of hours in the five years period. In the observation, the large MCS precipitation frequency appears in the same regions where observed MCS precipitation is large. There are also relatively high MCS precipitation occurrences in the midlatitude storm track region in the northern hemisphere. Compared to the observations, the EAMv2 simulated MCS frequency is largely underestimated over the tropical Indo-Pacific warm pools and midlatitude regions (Figure 4e). This is consistent with the underestimated MCS precipitation in Figure 2b. However, over the central Pacific Ocean, the simulated MCS occurrence becomes comparable to the observations, while over the tropical eastern Pacific Ocean, EAMv2 shows higher MCS frequency even though its MCS precipitation is slightly underestimated. The overestimated frequency of occurrence could imply either the simulated MCSs occur too frequently, or they are overly long-lived. In addition, simulated MCS precipitation frequency and precipitation amount are both overestimated over the tropical Andes and east African highlands.

To provide more insights on the MCS occurrence, Figure 5 shows the annual mean MCS number in both the model and observation. The MCS number is counted as the number of unique latitude/longitude pairs of each MCS track within a 5° × 10° latitude/longitude grid. The difference between MCS number and MCS frequency (Figure 4) is that the impact of MCS lifetime and area footprint is excluded from the MCS counts in Figure 5. For example, if an MCS

takes a few hours to slowly pass one grid, the MCS number in this grid is counted once but the occurrence of frequency of MCS is the total number of hours this MCS takes to move over the grid. In Figure 5, the location of large observed MCS number is overall consistent with those regions having large MCS frequency and MCS precipitation amount. However, the number of MCSs defined using both T_b and precipitation in EAMv2 is largely underestimated globally except over the central Pacific Ocean and SPCZ region (Figure 5e). Such an underestimation of MCS number suggests that the overestimated MCS precipitation frequency in the tropical eastern Pacific Ocean (Figure 4e) may be caused by the overly long-lived and/or larger MCSs. Further analysis shows that the underestimated MCS number is the result of underestimated MCS genesis in the tropics (not shown).

Consistent with MCS precipitation amount, the T_b-only MCS tracking method results in a substantial increase in MCS precipitation frequency and MCS number compared to MCSs tracked using the combined method in EAMv2, particularly over Congo and Amazon (Figures 4f and 5f). Additional increase in the simulated MCS number in the Maritime Continent, tropical Indian Ocean, and midlatitude continents are also found, leading to a more comparable MCS number spatial distribution to the observations. Note that such an increase in MCS frequency and number is less noticeable for the observations (not shown). Dong et al. (2021) also found that MCS defined using only T_b well produces the resemblance of observed tropical MCS number in the GFDL C192AM4 model, where they concluded that GCM is capable of simulating MCS characteristics at 50 km horizontal resolution. However, our current analysis shows that if surface precipitation is included in the MCS tracking to sample more robust MCSs, it is still challenging for EAMv2 to simulate MCSs at 25 km horizontal grid spacing.

To further evaluate the MCS characteristics in EAMv2, Figure 6 compares the probability density function (PDF) of simulated and observed MCS properties over the Indo-Pacific region to understand the reasons for the biased MCS precipitation in EAMv2 and the impact of different MCS definitions on MCS evaluation. The choice of Indo-Pacific region is because of (1) the largely underestimated MCS precipitation amount in simulated MCSs; (2) the substantial increase in MCS number between two tracking methods in EAMv2, while the IMERG observation presents negligible difference.

Figure 6 shows that most of the simulated MCS properties differ substantially from the observation for MCSs tracked using combined T_b and precipitation, with the exceptions in the maximum CCS area and maximum PF area. Compared to the observation, EAMv2 simulated MCSs tend to have longer CCS lifetime, warmer minimum cloud-top T_b (i.e., lower convective cloud-top height), and weaker mean rain rate within the PF area. Although the model overestimates the probabilities of total rain volume and heavy rain volume between 10^6 and 10^7 kg, the heavy rain ratio (i.e., heavy rain volume divided by total rain volume) in EAMv2 peaks at a lower value (~50%) than the observation (~75%). The lower heavy rain ratio is probably associated with the underestimated convection strength, which is indicated by the large occurrences of warm cloud-top T_b in the model (minimum T_b warmer than 205 K). These warm T_b occurrences suggest the simulated MCS in EAMv2 is less penetrative compared to the observed MCS, implying the issue in representing MCS development in the model physical parameterizations. Moreover, by comparing the PDFs of CCS area and PF area to the PDFs of PF features, it is likely that the weaker precipitation intensity, rather than the areas of convective clouds or surface precipitation, explain the underestimated MCS number (Figure 5) and their associated precipitation (Figures 1-2). This speculation is supported by the fact that in the eastern

Pacific Ocean, where the default EAMv2 better simulates MCS precipitation amount, the PDF of heavy rain ratio tends to peak at ~78% which matches the IMERG observation (~80%) more than any other examined locations (not shown).

For MCSs defined with two different methods, it is shown that the PDFs of MCS properties are nearly identical in the observation over the Indo-Pacific region. The comparable PDFs between MCSs defined with T_b and combined T_b and precipitation are also found in other regions (i.e., eastern Pacific Ocean, Africa and Amazon tropical lands, not shown). In other words, the MCS precipitation features are nearly always generated in these mesoscale cloud structures, especially over the tropics. Such a feature indicates the robustness of observed MCS characteristics using the cloud-top T_b tracking. However, the differences in MCS properties between the two MCS definitions are substantially larger in EAMv2. The largest discrepancies are found in the lifetime minimum T_b and heavy rain volume ratio. For example, while the minimum T_b of MCSs defined using combined T_b and precipitation peaks at ~198 K, which is comparable to the observation, the largest occurrence of minimum T_b for MCSs defined with T_b locates at ~215 K, suggesting cloud clusters with much weaker convective strength are included. Meanwhile, the largest occurrence for heavy rain volume ratio locates below 5% for MCSs defined with T_b, whereas it is at ~50% for MCSs defined with both T_b and precipitation. The lower heavy rain ratio from the MCSs defined with T_b is caused by the more occurrences of weak precipitation under the CCS cloud shields (Figures 6e-6g). These weak precipitation events are more likely to be associated with stratiform-type precipitation rather than convective-type strong precipitation in EAMv2, which is not the characteristics of MCS precipitation as in the observations. The difference in simulated MCS precipitation also impacts the CCS lifetime

simulation. With the weaker simulated MCS strength, the CCS lifetime is also shorter for MCSs defined with T_b than those defined with both T_b and precipitation.

We note that the simulated precipitation features (i.e., mean PF rain rate, total and heavy rain volume) of tracked MCSs are more comparable between the model and the observation in the Indo-Pacific region for MCSs defined with both T_b and precipitation than the T_b-only definition, even though the annual mean MCS number and precipitation rate are substantially underestimated. Although Figure 5 shows that the number of T_b-only defined MCSs is overall comparable to the observed MCS number, the use of T_b-only method in MCS tracking could include many weak convective events (i.e., suggested by the warm minimum T_b and low heavy rain ratio). The inclusion of these weak convective systems can ultimately result in an underestimation of the severity of MCS extreme precipitation by introducing a severe low bias in the precipitation counted as MCSs.

3.1.2. MCS over CONUS

MCS precipitation was found to contribute more than 50% of total precipitation over the CONUS region and can reach 70% in the central U.S. during the warm season (Feng et al., 2018, 2019). It is therefore imperative for GCMs to accurately simulate MCS precipitation in order to understand and examine the impact of future climate change on MCS precipitation over the CONUS. In this section, both the IMERG satellite precipitation data and the ground-based Stage IV rain gauge measurements are analyzed to address the potential uncertainties in observational datasets.

Figure 7 compares the mean total and MCS precipitation amount from March to August over the CONUS region. Both IMERG and ground-based Stage IV observations indicate that

497 total precipitation peaks in the central CONUS (i.e., Kansas, Missouri, Oklahoma, and
498 Arkansas). Strong total precipitation is also found in southeast U.S. (i.e., Florida) in both
499 observations. Note that the strong total precipitation in the central U.S. is observed in both boreal
500 spring (MAM) and summer (JJA), while the southeast strong precipitation mainly occurs in the
501 summertime. Due to the higher horizontal resolution (native resolution of 4 km), the ground-
502 based data shows more fine-scale precipitation variability than the satellite data, but the general
503 precipitation patterns and magnitudes are similar in both datasets. For observed MCS
504 precipitation, both datasets present the peak MCS precipitation around the same regions where
505 strong total precipitation occurs in the central U.S. (i.e., the border of Kansas, Missouri,
506 Oklahoma, and Arkansas). Note that the difference between two MCS tracking methods is
507 negligible for observed MCS precipitation amount (Figures 7d-7e and Figures 7g-7h), same as
508 the tropical MCSs.

509 Compared to observations, the total precipitation peak simulated by the EAMv2 is mostly
510 located in the southeast U.S. and Florida. The simulated total precipitation is overestimated along
511 the southeast coasts, but the model significantly underestimates the strong precipitation in the
512 central U.S. This dry bias in the central U.S. is consistent with earlier studies (Cheruy et al.,
513 2014; Klein et al. 2006; Morcrette et al., 2018; Zheng et al., 2019). With the biased total
514 precipitation amount in EAMv2, simulated MCS precipitation defined with combined T_b and
515 precipitation is substantially underestimated in the CONUS region. Not only the MCS
516 precipitation magnitude is significantly weaker than the observed MCS precipitation, the spatial
517 coverage of MCS precipitation is also much smaller. Similar to the results from tropical MCS,
518 the MCS precipitation increases largely when using the T_b-only tracking method, in particular in

519 the southeast U.S. However, MCS remains significantly underestimated in the central CONUS
520 regardless which tracking method is used.

521 The biased MCS precipitation in EAMv2 is also reflected in the MCS precipitation
522 frequency over the CONUS (Figure 8). The simulated MCS precipitation frequency defined by
523 T_b and precipitation is substantially underestimated in the central U.S., while the frequency is
524 comparable to the observations in the southeast U.S. Although the impact of different MCS
525 tracking methods is small for observations, the MCS precipitation frequency is substantially
526 larger for MCSs from the T_b-only tracking in EAMv2. It is also shown in Figure 8f that the
527 frequency of occurrence of simulated MCSs defined with T_b becomes largely overestimated in
528 the eastern U.S.

529 Figure 9 compares the MCS number between EAMv2 and the observations in the
530 CONUS region. It is unsurprising to find the significant underestimation of MCS number in
531 EAMv2 over the entire CONUS when MCSs are defined using the combined method. On the
532 other hand, the impact of MCS definition on simulated and observed MCS number is again
533 consistent with the global MCS number in Figure 5. For example, the MCS number over the
534 CONUS region in both the IMERG and Stage-IV observations remains similar between two
535 definitions (Figures 9a-9b and 9d-9e), but the simulated MCS number from T_b-only method
536 becomes more comparable to the observation in EAMv2, although the location of MCS number
537 peaks is misrepresented. However, we note that the good agreement is a result of less accurate
538 tracking of MCSs.

539 The PDFs of MCS properties over the central CONUS region are shown in Figure 10. In
540 general, the statistics of MCS properties are comparable between the two observations. This is
541 the case especially for cloud shield related properties such as CCS lifetime, maximum CCS area,

and minimum cloud-top T_b throughout the lifetime, which implies the small uncertainty in these observed CCS properties in different datasets. On the other hand, larger differences are found in the PDFs of precipitation related properties (i.e., maximum PF area, mean rain rate, rain volume, and heavy rain ratio). For instance, the IMERG data shows larger PF area and higher rain volume, but lower mean rain rate and heavy rain ratio compared to the ground-based measurements, consistent with previous studies (Cui et al., 2020; Zhang et al., 2021; Ayat et al., 2021). Compared to the observations, it is similar to the tropical MCSs that the simulated CCS lifetime is longer than observed (Figure 10a) and the simulated cloud-top minimum T_b is warmer (Figure 10c) over the CONUS region. EAMv2 simulated MCSs contain substantially weaker rain rate (Figure 10e) and lower heavy rain ratio (Figure 10h) than observations. However, different from the tropical region, the simulated maximum CCS area and PF area are larger than both observations (Figures 10b and 10d), which leads to larger total and heavy rain volume within the MCS cloud shields (Figures 10f and 10g) over the CONUS. We note that the biases in rain rate and convective strength (i.e., implied by the minimum T_b) are the primary reasons for the underestimated MCS precipitation in the CONUS region.

Figure 10 shows that the differences in observed MCS properties between two tracking methods are again insignificant, similar to the Indo-Pacific region (Figure 6). This indicates the robustness of the MCS tracking methods in identifying warm season MCSs in observations. On the other hand, the statistics of model simulated MCS properties are more sensitive to the MCS definition with precipitation. In contrast to the Indo-Pacific region, all the simulated MCS properties show large sensitivities to the MCS definition. The difference from the Indo-Pacific region exists in the maximum CCS area and PF area, which suggests the impact of MCS definition on CCS identification in the central U.S. But again, the simulated PF characteristics

such as the PF rain rate and heavy rain ratio are significantly weaker for MCSs defined with T_b than the combined method. Note that our analysis suggests that EAMv2 does not show an improvement in simulating MCSs compared to EAMv1, in which the MCS precipitation amount, frequency of occurrence, and number are all underestimated compared to observations from March to May (Wang et al., 2021).

Similar MCS evaluation was made in Dong et al. (2023) to evaluate the GCM (uniform 50 km horizontal resolution) simulated MCS properties in the CONUS region. Based on the MCS samples tracked using only T_b, they found that the model well reproduces the spatial distribution of occurrence frequency of MCS and the MCS duration, MCS strength, size, and movement speed. However, given the differences between two MCS tracking methods analyzed in the current study, without considering surface precipitation features in the MCS tracking might lead to different conclusions in the evaluation of model skills in simulating MCSs and associated precipitation characteristics. For example, large-scale predominantly stratiform precipitation associated with synoptical-scale cloud bands (e.g., low pressure or frontal systems) may be included in the sampled MCSs when surface precipitation is not accounted for. Together with the similar findings in the global analysis, our study suggests the importance of including precipitation characteristics in the MCS definition and tracking when evaluating MCS properties in GCMs.

3.2. Impact of cloud and convection parameterizations

The previous sections have shown that EAMv2 is not capable of reproducing the observed MCS properties in the tropics and the CONUS region. In this section, we examine four new cloud and convection parameterizations (described in Section 2.2) that are developed for

588 EAMv3 to see if these new developments will lead to an improvement in the E3SM simulated
589 MCSs. The sensitivity tests on each of the four new features are examined to understand their
590 individual impacts on the MCS simulation. Note that the sensitivity test with all new features
591 combined was also examined. However, because the combined impact of all new features on
592 MCS simulation is dominated by P3 and MAdj, we do not include this sensitivity experiment in
593 this discussion. Based on the MCS evaluation in the earlier sections, we find that the MCS
594 tracking method using only T_b overestimates the model's capability in capturing MCSs for
595 E3SMv2. Therefore, we focus our discussion on results using the more stringent MCS tracking
596 method with both T_b and surface precipitation in this section.

597 The spatial distribution of annual mean total precipitation and MCS precipitation
598 differences between EAMv2 sensitivity experiments and the default EAMv2 physics (CTL
599 hereafter) is shown in Figure 11. In general, the impact of these new cloud and convection
600 schemes on the simulated total precipitation is minor over most regions for the MCSP and
601 ZMmicro experiments, but noticeable impacts are found in the sensitivity experiments of P3 and
602 MAdj. For example, using P3 cloud microphysics largely increases the total precipitation
603 simulated over the subtropical western Pacific Ocean and SPCZ compared to CTL, but the total
604 precipitation is decreased in the central and eastern Pacific Ocean and tropical Indian Ocean.
605 Similar effects on total precipitation are also found with the cloud base mass flux adjustment test
606 in the tropical ocean, but the changed precipitation amount is much smaller than P3. Meanwhile,
607 the deep convective cloud microphysics in the ZM scheme only slightly increases tropical total
608 precipitation, and the impact of MCSP is minimal.

609 Similar impact is seen in the simulated MCS precipitation. For example, P3 largely
610 increases the MCS precipitation in the subtropical western Pacific Ocean while it leads to a

611 reduction in the Indian Ocean and equatorial central and eastern Pacific Ocean. It is interesting
612 that MAdj leads to a reduction of MCSs over oceans globally, particularly over the ITCZ and
613 SPCZ regions. The increased and decreased MCS precipitation due to P3 and MAdj both enlarge
614 the existing MCS precipitation biases identified in CTL shown in Figure 2. We note that the
615 reduced MCS precipitation in the ITCZ region is more significant for the MCSs defined using
616 combined T_b and surface precipitation than the MCSs defined with T_b (not shown). By
617 examining the PDF of hourly precipitation rate in the tropics where MCS precipitation
618 substantially increased for P3 (Figures 12e), we indeed find that compared with CTL where
619 MG2 is used, P3 tends to largely increase the occurrences of heavy precipitation rate (rain rate >
620 5 mm hr^{-1}) because of the accounted riming (Wang et al., 2021). The increased heavy
621 precipitation rate becomes most comparable to observations among all sensitivity experiments.
622 However, P3 presents less degree of improvement in simulated heavy rain rate compared to CTL
623 over other regions (Figures 12d). This suggests that the change in simulated MCS precipitation is
624 primarily driven by the variation in simulated heavy precipitation rate. On the other hand, the
625 simulated frequency of heavy rain rate is significantly lower in MAdj than CTL. This explains
626 why MCS precipitation becomes substantially weaker in the tropics in MAdj. The reduced heavy
627 rain frequency in MAdj is possibly the result of more convection formation in the low CAPE
628 environment using the new cloud base mass flux adjustment treatment. Both MCSP and
629 ZMmicro show little impacts on MCS precipitation compared to CTL.

630 To further diagnose the reasons of model behavior change in simulating tropical
631 precipitation, Figure 13 shows the global annual mean relative contribution of convective
632 precipitation to total precipitation in CTL and sensitivity experiments. It is shown in Figure 13a
633 that the total precipitation is primarily generated by large-scale precipitation in the western

Pacific Ocean and tropical Indian Ocean, while the convective precipitation mainly contributes to the precipitation in the tropical eastern Pacific Ocean. The substantial contribution of large-scale precipitation to total precipitation in CTL is somehow counterintuitive, particularly for the tropical convective systems. It also probably explains the underestimated heavy precipitation in simulated MCSs. In sensitivity experiments, it is interesting to note that both P3 and MAdj largely increase the convective precipitation fractions in the tropics compared to CTL. Although ZMmicro slightly decreases the convective precipitation fraction over the subtropical ocean in the southern hemisphere, the impacts of MCSP and ZMmicro on convective precipitation fraction is overall insignificant in the tropics. In addition, with larger convective precipitation fraction in P3 and MAdj, the occurrences of large convective precipitation rate are also increased compared to CTL (Figures 12c and 12f). However, the magnitude of convective precipitation rate remains lower than large-scale precipitation rate by more than a factor of 5. The increased convective precipitation fraction but the weak convective precipitation from P3 and MAdj suggest that the precipitation formation in convective scheme is likely not strong enough to produce sufficient heavy precipitation to be counted as MCSs, which therefore causes the underestimated MCS precipitation in the model.

Figure 14 shows the individual impacts of new physics features on the total and MCS precipitation simulation over the CONUS region. In general, P3 microphysics tends to further worsen the underestimated total precipitation in the central U.S., except for Arkansas, Louisiana, and Mississippi, while all other three schemes generally show positive effects to increase simulated total precipitation compared to CTL. In terms of the simulated MCS precipitation, although P3, MCSP, and ZMmicro all tend to increase the MCS precipitation rate, the magnitude is too small to have a meaningful impact on the largely underestimated MCS in EAMv2.

In summary, the sensitivity experiments suggest that these newly implemented cloud and convection features are likely not to help improve the simulation of MCS in the next version of E3SM. A higher model resolution (i.e., km-scale) for the model to better resolve heavy precipitation processes for MCSs or the better representation of mesoscale dynamics and physics in cloud and convection parameterizations are needed.

4. Summary and discussion

This study evaluates the MCS simulation in EAMv2 model using uniform high resolution (~0.25°) model configuration. We use the recently developed PyFLEXTRKR MCS tracking algorithm, which considers both cloud-top T_b and surface precipitation to track global MCS evolution and evaluate the statistics of MCS properties simulated in EAMv2 against observational datasets. The MCSs defined purely based on cloud-top T_b, which is commonly used in previous studies, are also examined to understand the impact of different MCS definitions on the MCS evaluation.

In the tropical region, EAMv2 reasonably simulates the total precipitation in the equatorial central and eastern Pacific Ocean, but it underestimates the total precipitation over the tropical Indian Ocean and Maritime Continent region. For the simulated MCS precipitation, EAMv2 largely underestimates the tropical MCS precipitation compared to the observations. The underestimation is more substantial in the tropical Indian Ocean and Maritime Continent, indicating that the dry bias in the total precipitation over the region is primarily due to the lack of MCS precipitation. Simulated MCS precipitation fraction is thus also substantially underestimated. EAMv2 shows that the simulated frequency of occurrence is comparable to the observations in the central Pacific Ocean along the equator, while it substantially underestimates

MCS occurrences in the tropical Indian Ocean and Maritime Continent. The simulated MCS frequency of occurrence in the eastern equatorial Pacific Ocean, on the other hand, is found slightly higher than observations even though the MCS precipitation rate over the region is slightly underestimated. This is because the model CCS lifetime is longer than the observed MCSs.

Over the CONUS region, EAMv2 also substantially underestimates the MCS precipitation rate and MCS precipitation occurrences in the central U.S. in both spring and summer seasons. Note that EAMv2 shows problems in simulating MCSs in both magnitude and location. We find that EAMv2 simulated MCS is dominant in the southeast U.S., but the model significantly misses the MCS occurrence in the central U.S.

Our analysis also shows that the MCS number is substantially higher by using the T_b-only tracking, which includes cloud systems with weak surface precipitation that may not be associated with MCSs. As a result, the MCS precipitation amount and frequency are also largely increased by using the T_b tracking. The largest impact is found in the tropical Africa and Amazon lands and the southeast U.S., where simulated MCS number detected by T_b becomes comparable to the observations. This indicates using the T_b tracking method could overestimate the model's capability in simulating MCSs. It is thus important to include precipitation characteristics in MCS definition when evaluating MCS properties in GCMs. It also points out that the biases in simulating MCSs in EAMv2 is mainly due to the underestimation of MCS precipitation intensity. Additionally, we also find that the simulated MCS number is slightly increased in the tropics when we reduce the lifetime-dependent tunable PF parameters (i.e., smaller slopes of coefs_pf_area, coefs_pf_rr, and coefs_pf_skew in Table A1) used in the tracking (not shown). This further suggests the sensitivity of simulated MCS number to the

703 precipitation associated tracking parameters. Compared to the MCS precipitation, the model can
704 reasonably simulate the mesoscale cloud shield structures (CCS area) at 0.25° horizontal
705 resolution.

706 To examine the impact of model physics on the simulation of MCS, four sensitivity
707 experiments are performed with the new cloud and convection parameterizations that are
708 developed for EAMv3. These include the P3 microphysics, MCSP, convective cloud
709 microphysics, and cloud base mass flux adjustment. Our analysis shows little impact on the
710 simulation of MCSs using these new features at 0.25° horizontal resolution. Only the P3 cloud
711 microphysics scheme presents a notable improvement in the simulated MCS precipitation in the
712 subtropical western Pacific Ocean. This indicates that the MCS simulation will likely remain a
713 challenge in the next version of EAM at the mesoscale resolution. However, it is interesting to
714 note that both P3 and cloud base mass flux adjustment treatment largely increase the convective
715 precipitation contribution to total precipitation in the tropics. This increased convective
716 precipitation fraction is physically more reasonable, but the remaining issue in simulating heavy
717 precipitation in MCSs is likely caused by the insufficient strong precipitation formation in the
718 convective scheme.

719 Over the CONUS region, P3, MCSP, and deep convective cloud microphysics all tend to
720 enhance the simulated MCS precipitation in the central U.S., but the change is overall minimal.
721 However, we note that Wang et al. (2021) suggested that the use of P3 scheme substantially
722 improves the MCS number and precipitation over the CONUS region in the RRM version of
723 E3SMv1. The minimal impact of P3 based on EAMv2 model configuration tested in this study
724 suggests that the performance of parameterizations could also depend on model configurations
725 and other physics parameterizations used in the model. Furthermore, our results are consistent

with Feng et al. (2021b) who also tracked MCS using both T_b and precipitation feature and found weak MCS precipitation intensity and underestimated MCS number over CONUS in simulations at 25 km and 50 km resolution. Tracking MCS using both T_b and precipitation in the NICAM (Nonhydrostatic ICosahedral Atmospheric Model) simulations at 14 km resolution, Na et al. (2022) also found underestimated MCS number over the CONUS during summer. Without the use of any cumulus parameterizations, the NICAM simulation produced stronger MCS precipitation, smaller precipitation area, and larger cold cloud system than those observed. Both Feng et al. (2021b) and Na et al. (2022) attributed the MCS biases over CONUS to the dry bias in the lower atmosphere. Recent studies suggested that the dry bias in the atmospheric boundary layer could result from biases in land surface models and land-atmosphere interactions (Barlage et al., 2021; Qin et al., 2023). Future direction of improving MCS simulation in E3SM involves both increasing model resolution to better resolve key dynamical processes and improving model physics to better represent MCSs.

Acknowledgment: This research was supported as part of the Energy Exascale Earth System Model (E3SM) Science Focus Area, funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research Earth System Model Development program area. Zhe Feng was supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research Regional and Global Model Analysis program area. Portions of this study were supported by the Regional and Global Model Analysis (RGMA) component of the Earth and Environmental System Modeling Program of the U.S. Department of Energy's Office of Biological & Environmental Research (BER) under Award Number DE-

SC0022070, and by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation (NSF) under Cooperative Agreement No. 1852977. Work at LLNL was performed under the auspices of the U.S. DOE by Lawrence Livermore National Laboratory under contract No. DE-AC52-07NA27344. PNNL is operated for the Department of Energy by Battelle Memorial Institute under contract DE-AC05-76RL01830. Argonne National Laboratory is operated for the DOE by UChicago Argonne, LLC, under contract DE-AC02-06CH11357. We gratefully acknowledge the computing resources provided on Bebop (and/or Swing and/or Blues), a high-performance computing cluster operated by the Laboratory Computing Resource Center at Argonne National Laboratory. This research also used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No.DE-AC02-05CH11231.

Data Availability Statement: The U.S. DOE E3SMv2 (E3SM Project, DOE, 2021, September 29) model was used in the creation of this manuscript. The model and observational data used in this study can be accessible at https://doi.org/10.5281/zenodo.10521291 (Zhang et al., 2024). The PyFLEXTRKR tracking codes are available on GitHub (https://github.com/FlexTRKR/PyFLEXTRKR).

References

Ayat, H., Evans, J. P., Sherwood, S., & Behrangi, A. (2021). Are Storm Characteristics the Same When Viewed Using Merged Surface Radars or a Merged Satellite Product?. *Journal of Hydrometeorology*, 22(1), 43-62. https://doi.org/10.1175/JHM-D-20-0187.1

Barlage, M., Chen, F., Rasmussen, R., Zhang, Z., & Miguez-Macho, G. (2021). The importance of scale-dependent groundwater processes in land-atmosphere interactions over the central United States. Geophysical Research Letters, 48(5), e2020GL092171. https://doi.org/10.1029/2020GL092171

Beres, J. H., Alexander, M. J., & Holton, J. R. (2004). A Method of Specifying the Gravity Wave Spectrum above Convection Based on Latent Heating Properties and Background Wind, *Journal of the Atmospheric Sciences*, 61, 324–337. https://doi.org/10.1175/1520-0469(2004)061<0324:AMOSTG>2.0.CO;2

Chen, C.-C., Richter, J. H., Liu, C., Moncrieff, M. W., Tang, Q., Lin, W., et al. (2021). Effects of organized convection parameterization on the MJO and precipitation in E3SMv1. Part I: Mesoscale heating. *Journal of Advances in Modeling Earth Systems*, 13, e2020MS002401. https://doi.org/10.1029/2020MS002401

Cheruy, F., Dufresne, J. L., Hourdin, F., & Ducharne, A. (2014). Role of clouds and land-atmosphere coupling in midlatitude continental summer warm biases and climate change

amplification in CMIP5 simulations. *Geophysical Research Letters*, 41, 6493–6500. https://doi.org/10.1002/2014GL061145

Cui, W., Dong, X., Xi, B., Feng, Z., & Fan, J. (2020). Can the GPM IMERG Final Product Accurately Represent MCSs' Precipitation Characteristics over the Central and Eastern United States?. *Journal of Hydrometeorology*, 21(1), 39-57. https://doi.org/10.1175/JHM-D-19-0123.1

Dong, W., Zhao, M., Ming, Y., Krasting, J. P., & Ramaswamy, V. (2023). Simulation of United States Mesoscale Convective Systems using GFDL's New High-Resolution General Circulation Model. *Journal of Climate*, *36*(19), 6967-6990. https://doi.org/10.1175/JCLI-D-22-0529.1

Dong, W., Zhao, M., Ming, Y., & Ramaswamy, V. (2021). Representation of Tropical Mesoscale Convective Systems in a General Circulation Model: Climatology and Response to Global Warming. *Journal of Climate*, *34*(14), 5657-5671. https://doi.org/10.1175/JCLI-D-20-0535.1

E3SM Project, DOE. (2021, September 29). Energy Exascale Earth System Model v2.0. 733 [Software]. https://github.com/E3SM-Project/E3SM/releases/tag/v2.0.0. 734. https://doi.org/10.11578/E3SM/dc.20210927.1

Fan, J., Han, B., Varble, A., Morrison, H., North, K., Kollias, P., et al. (2017). Cloud-resolving model intercomparison of an MC3E squall line case: Part I—Convective updrafts. *Journal of*

Geophysical Research: Atmospheres, 122(17), 9351–9378.
https://doi.org/10.1002/2017JD026622

Feng, Z. (2019). Mesoscale convective system (MCS) database over United States (V2). [Dataset]. Retrieved from: https://doi.org/10.5439/1571643

Feng, Z., Hardin, J., Barnes, H., Li, J., Leung, L. R., Varble, A., & Zhang, Z. (2022). PyFLEXTRKR: a flexible feature tracking Python software for convective cloud analysis (Version 2023.10.0) [Computer software]. https://doi.org/10.5194/gmd-16-2753-2023

Feng, Z., Hardin, J., Barnes, H. C., Li, J., Leung, L. R., Varble, A., & Zhang, Z. (2023). PyFLEXTRKR: a flexible feature tracking Python software for convective cloud analysis, *Geoscientific Model Development*, 16, 2753–2776, https://doi.org/10.5194/gmd-16-2753-2023

Feng, Z., Houze, R. A., Leung, L. R., Song, F., Hardin, J. C., Wang, J., et al. (2019). Spatiotemporal characteristics and large-scale environments of mesoscale convective systems east of the Rocky Mountains. *Journal of Climate*, 32(21), 7303–7328. https://doi.org/10.1175/JCLI-D-19-0137.1

Feng, Z., Leung, L. R., Hagos, S., Houze, R. A., Burleyson, C. D., & Balaguru, K. (2016). More frequent intense and long-lived storms dominate the springtime trend in central US rainfall. *Nature Communications*, 7, 13429. https://doi.org/10.1038/ncomms13429

834 Feng, Z., Leung, L. R., Houze, R. A., Hagos, S., Hardin, J., Yang, Q., et al. (2018). Structure and
835 evolution of mesoscale convective systems: Sensitivity to cloud microphysics in convection-
836 permitting simulations over the United States. *Journal of Advances in Modeling Earth*
837 *Systems*, 10(7), 1470–1494. https://doi.org/10.1029/2018MS001305
838
839 Feng, Z., Leung, L. R., Liu, N., Wang, J., Houze, R. A., Li, J., et al. (2021a). A global high-
840 resolution mesoscale convective system database using satellite-derived cloud tops, surface
841 precipitation, and tracking. *Journal of Geophysical Research: Atmospheres*, 126(8).
842 https://doi.org/10.1029/2020JD034202
843
844 Feng, Z., Song, F., Sakaguchi, K., & Leung, L. R. (2021b). Evaluation of Mesoscale Convective
845 Systems in Climate Simulations: Methodological Development and Results from MPAS-CAM
846 over the United States. *Journal of Climate*, *34*(7), 2611-2633. https://doi.org/10.1175/JCLI-D-
847 20-0136.1
848
849 Gettelman, A., & Morrison, H. (2015). Advanced two-moment bulk microphysics for global
850 models. Part I: Off-line tests and comparison with other schemes. *Journal of Climate*, 28(3),
851 1268–1287. https://doi.org/10.1175/JCLI-D-14-00102.1
852
853 Gettelman, A., Morrison, H., Santos, S., Bogenschutz, P., & Caldwell, P. M. (2015). Advanced
854 two-moment bulk microphysics for global models. Part II: Global model solutions and aerosol-
855 cloud interactions. *Journal of Climate*, 28(3), 1288–1307.
856 https://doi.org/10.1175/jcli-d-14-00103.1

Golaz, J.-C., Caldwell, P. M., Van Roekel, L. P., Petersen, M. R., Tang, Q., Wolfe, J. D., et al. (2019). The DOE E3SM coupled model version 1: Overview and evaluation at standard resolution. *Journal of Advances in Modeling Earth Systems*, 11, 2089–2129. https://doi.org/10.1029/2018MS001603

Golaz, J.-C., Larson, V. E., & Cotton, W. R. (2002). A PDF-based model for boundary layer clouds. Part I: Method and model description. *Journal of the Atmospheric Sciences*, 59(24), 3540–3551. https://doi.org/10.1175/1520-0469(2002)059<3540:APBMFB>2.0.CO;2

Golaz, J.-C., Van Roekel, L. P., Zheng, X., Roberts, A. F., Wolfe, J. D., Lin, W., et al. (2022). The DOE E3SM Model version 2: Overview of the physical model and initial model evaluation. *Journal of Advances in Modeling Earth Systems*, 14, e2022MS003156. https://doi.org/10.1029/2022MS003156

Haberlie, A. M., & Ashley, W. S. (2019). A radar-based climatology of mesoscale convective systems in the United States. *Journal of Climate*, 32(5), 1591–1606. https://doi.org/10.1175/jcli-d-18-0559.1

Houze, R. A. (2004). Mesoscale convective systems. *Reviews of Geophysics*, 42(4), RG4003. https://doi.org/10.1029/2004RG000150

879 Houze, R. A. (2018). 100 years of research on mesoscale convective systems. *Meteorological*
880 *Monographs*, 59, 17.1–17.54. https://doi.org/10.1175/amsmonographs-d-18-0001.1
881
882 Hsu, W.-C., Kooperman, G. J., Hannah, W. M., Reed, K. A., Akinsanola, A. A., & Pendergrass,
883 A. G. (2023). Evaluating mesoscale convective systems over the US in conventional and
884 multiscale modeling framework configurations of E3SMv1. *Journal of Geophysical Research:*
885 *Atmospheres*, 128, e2023JD038740. https://doi.org/10.1029/2023JD038740
886
887 Huang, B., Liu, C., Banzon, V., Freeman, E., Graham, G., Hankins, B., et al. (2021).
888 Improvements of the Daily Optimum Interpolation Sea Surface Temperature (DOISST) Version
889 2.1. *Journal of Climate*, 34, 2923–2939. https://doi.org/10.1175/JCLI-D-20-0166.1
890
891 Huang, X. M., Hu, C. Q., Huang, X., Chu, Y., Tseng, Y. H., Zhang, G. J., & Lin, Y. L. (2018). A
892 long-term tropical mesoscale convective systems dataset based on a novel objective automatic
893 tracking algorithm. *Climate Dynamics*, 51, 3145–3159, https://doi.org/10.1007/s00382-018-
894 4071-0
895
896 Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., et al. (2019a).
897 Algorithm Theoretical Basis Document (ATBD) Version 06. NASA Global Precipitation
898 Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG) (Algorithm
899 Theoretical Basis Document). Retrieved from
900 https://docserver.gesdisc.eosdis.nasa.gov/public/project/GPM/IMERG_ATBD_V06.pdf
901

902 Huffman, G. J., Bolvin, D. T., Nelkin, E. J., & Tan, J. (2019b). Integrated Multi-satellitE
903 Retrievals for GPM (IMERG) technical documentation (technical documentation). Retrieved
904 from https://docserver.gesdisc.eosdis.nasa.gov/public/project/GPM/IMERG_doc.06.pdf
905
906 Huffman, G. J., Stocker, E. F., Bolvin, D. T., Nelkin, E. J., & Tan, J. (2019c). GPM IMERG
907 final precipitation L3 half hourly 0.1 degree x 0.1 degree V06. Retrieved from Greenbelt, MD:
908 Goddard Earth Sciences Data and Information Services Center (GES DISC). https://doi.
909 org/10.5067/GPM/IMERG/3B-HH/06
910
911 Janowiak, J., Joyce, B., & Xie, P. (2017). NCEP/CPC L3 half hourly 4km global (60S - 60N)
912 merged IR V1. Retrieved from https://doi.org/10.5067/P4HZB9N27EKU
913
914 Klein, S. A., Jiang, X., Boyle, J., Malyshev, S., & Xie, S. (2006). Diagnosis of the summertime
915 warm and dry bias over the U.S. Southern Great Plains in the GFDL climate model using a
916 weather forecasting approach. *Geophysical Research Letters*, 33, L18805. https://doi.
917 org/10.1029/2006GL027567
918
919 Larson, V. E. (2017). CLUBB-SILHS: A parameterization of subgrid variability in the
920 atmosphere. arXiv preprint arXiv:1711.03675v2
921
922 Lin, G., Jones, C. R., Leung, L. R., Feng, Z., & Ovchinnikov, M. (2022). Mesoscale convective
923 systems in a superparameterized E3SM simulation at high resolution. *Journal of Advances in*
924 *Modeling Earth Systems*, 14, e2021MS002660. https://doi.org/10.1029/2021MS002660

Lin, Y., Dong, W., Zhang, M., Xie, Y., Xue, W., Huang, J., & Luo, Y. (2017). Causes of model dry and warm bias over central U.S. and impact on climate projections. *Nature Communications*, 8(1), 881. https://doi.org/10.1038/s41467-017-01040-2

Liu, X., Ma, P.-L., Wang, H., Tilmes, S., Singh, B., Easter, R. C., et al. (2016). Description and evaluation of a new four-mode version of the Modal Aerosol Module (MAM4) within version 5.3 of the Community Atmosphere Model. *Geoscientific Model Development*, 9(2), 505–522. https://doi.org/10.5194/gmd-9-505-2016

Ma, P.-L., Harrop, B. E., Larson, V. E., Neale, R., Gettelman, A., Morrison, H., et al. (2022). Better calibration of cloud parameterizations and subgrid effects increases the fidelity of E3SM Atmosphere Model version 1, *Geoscientific Model Development*. 15(7), 2881–2916. https://doi.org/10.5194/gmd-15-2881-2022

Milbrandt, J. A., & Morrison, H. (2016). Parameterization of cloud microphysics based on the prediction of bulk ice particle properties. Part III: Introduction of multiple free categories. *Journal of the Atmospheric Sciences*, 73, 975–995. https://doi.org/10.1175/JAS-D-15-0204.1

Milbrandt, J. A., Morrison, H., Dawson, D. T., II, & Paukert, M. (2021). A triple-moment representation of ice in the Predicted Particle Properties (P3) microphysics scheme. *Journal of the Atmospheric Sciences*, 78, 439–458. https://doi.org/10.1175/JAS-D-20-0084.1

948 Moncrieff, M. W. (2019). Toward a dynamical foundation for organized convection
949 parameterization in GCMs. *Geophysical Research Letters*, 46, 14103–14108.
950 https://doi.org/10.1029/2019GL085316
951
952 Moncrieff, M. W., Liu, C., & Bogenschutz, P. (2017). Simulation, modeling, and dynamically
953 based parameterization of organized tropical convection for global climate models. *Journal of*
954 *the Atmospheric Sciences*, 74, 1363–1380. https://doi.org/10.1175/JAS-D-16-0166.1
955
956 Morcrette, C. J., Van Weverberg, K., Ma, H.-Y., Ahlgrimm, M., Bazile, E., Berg, L. K., et al.
957 (2018). Introduction to CAUSES: Description of weather and climate models and their near-
958 surface temperature errors in 5 day hindcasts near the Southern Great Plains. *Journal of*
959 *Geophysical Research: Atmospheres*, 123, 2655–2683. https://doi.org/10.1002/2017JD027199
960
961 Mueller, B., & Seneviratne, S. I. (2014). Systematic land climate and evapotranspiration biases
962 in CMIP5 simulations. *Geophysical Research Letters*, 41, 128–134.
963 https://doi.org/10.1002/2013GL058055
964
965 Na, Y., Fu, Q., Leung, L. R., Kodama, C., & Lu, R. (2022). Mesoscale convective systems
966 simulated by a high-resolution global nonhydrostatic model over the United States and China.
967 *Journal of Geophysical Research: Atmospheres*, 127, e2021JD035916. https://doi.
968 org/10.1029/2021JD035916
969

970 Nesbitt, S. W., Cifelli, R., & Rutledge, S. A. (2006). Storm morphology and rainfall
971 characteristics of TRMM precipitation features. *Monthly Weather Review*, 134(10), 2702–2721.
972 https://doi.org/10.1175/mwr3200.1
973
974 Prein, A. F., Liu, C. H., Ikeda, K., Trier, S. B., Rasmussen, R. M., Holland, G. J., & Clark, M. P.
975 (2017). Increased rainfall volume from future convective storms in the US. *Nature Climate*
976 *Change*, 7(12), 880–884. https://doi.org/10.1038/s41558-017-0007-7
977
978 Qin, H., Klein, S. A., Ma, H.-Y., Van Weverberg, K., Feng, Z., Chen, X., et al. (2023).
979 Summertime near-surface temperature biases over the central United States in convection-
980 permitting simulations. *Journal of Geophysical Research: Atmospheres*, 128, e2023JD038624.
981 https://doi.org/10.1029/2023JD038624
982
983 Qin, Y., Zheng, X., Klein, S. A., Zelinka, M. D., Ma, P.-L., Golaz, J.-C., Xie, S. (2023). Causes
984 of Reduced Climate Sensitivity in E3SM from Version 1 to Version 2. *ESS Open Archive.*
985 https://doi.org/10.22541/essoar.168677205.57448605/v1
986
987 Rasch, P. J., Xie, S., Ma, P.-L., Lin, W., Wang, H., Tang, Q., et al. (2019). An overview of the
988 atmospheric component of the Energy Exascale Earth System Model. *Journal of Advances in*
989 *Modeling Earth Systems*, 11, 2377–2411. https://doi.org/10.1029/2019MS001629
990

991 Richter, J. H., Sassi, F., & Garcia, R. R. (2010). Toward a Physically Based Gravity Wave
992 Source Parameterization in a General Circulation Model, *Journal of the Atmospheric Sciences*,
993 67, 136–156. https://doi.org/10.1175/2009JAS3112.1
994
995 Richter, J. H., Chen, C.-C., Tang, Q., Xie, S., & Rasch, P. J. (2019). Improved Simulation of the
996 QBO in E3SMv1, *Journal of Advances in Modeling Earth Systems*, 11, 3403–3418.
997 https://doi.org/10.1029/2019MS001763
998
999 Schumacher, C., Houze, R. A., & Kraucunas, I. (2004). The tropical dynamical response to latent
1000 heating estimates derived from the TRMM precipitation radar. *Journal of the Atmospheric*
1001 *Sciences*, 61(12), 1341–1358. https://doi.org/10.1175/1520-
1002 0469(2004)061<1341:TTDRTL>2.0.CO;2
1003
1004 Song, X., & Zhang, G. J. (2011). Microphysics parameterization for convective clouds in a
1005 global climate model: Description and single-column model tests. *Journal of Geophysical*
1006 *Research: Atmospheres*, 116, D02201. https://doi.org/10.1029/2010JD014833
1007
1008 Song, X., Zhang, G. J., & Li, J. F. (2012). Evaluation of Microphysics Parameterization for
1009 Convective Clouds in the NCAR Community Atmosphere Model CAM5. *Journal of Climate*,
1010 *25*(24), 8568-8590. https://doi.org/10.1175/JCLI-D-11-00563.1
1011

Song, X., Zhang, G., Wan, H., & Xie, S. (2023). Incorporating the effect of large-scale vertical motion on convection through convective mass flux adjustment in E3SMv2. *Journal of Advances in Modeling Earth Systems*, 15, e2022MS003553. https://doi.org/10.1029/2022MS003553

Tang, Q., Golaz, J.-C., Van Roekel, L. P., Taylor, M. A., Lin, W., Hillman, B. R., et al. (2023). The fully coupled regionally refined model of E3SM version 2: overview of the atmosphere, land, and river results, *Geoscientific Model Development*, 16, 3953–3995. https://doi.org/10.5194/gmd-16-3953-2023

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

Tang, Q., Klein, S. A., Xie, S., Lin, W., Golaz, J.-C., Roesler, E. L., et al. (2019). Regionally refined test bed in E3SM atmosphere model version 1 (EAMv1) and applications for high-resolution modeling, *Geoscientific Model Development*, 12, 2679–2706. https://doi.org/10.5194/gmd-12-2679-2019

Tang, S., Xie, S., Guo, Z., Hong, S.-Y., Khouider, B., Klocke, D., et al. (2022). Long-term single-column model intercomparison of diurnal cycle of precipitation over midlatitude and tropical land. *Quarterly Journal of the Royal Meteorological Society*, 641(743), 669. https://doi.org/10.1002/qj.4222

Tao, C., Xie, S., Tang, S., Lee, J., Ma, H.-Y., Zhang, C., et al. (2022). Diurnal cycle of precipitation over global monsoon systems in CMIP6 simulations. *Climate Dynamics*, 60, 3947–3968. https://doi.org/10.1007/s00382-022-06546-0

Tao, C., Xie, S., Ma, H.-Y., Bechtold, P., Cui, Z., Vaillancourt, P.A., et al. (2023). Diurnal Cycle of Precipitation Over the Tropics and Central U.S.: GCM Intercomparison. *Quarterly Journal of the Royal Meteorological Society*. https://doi.org/10.1002/qj.4629

Van Weverberg, K., Morcrette, C. J., Petch, J., Klein, S. A., Ma, H.-Y., Zhang, C., et al. (2018). CAUSES: Attribution of surface radiation biases in NWP and climate models near the U.S. Southern Great Plains. *Journal of Geophysical Research: Atmospheres*, 123. https://doi.org/10.1002/2017JD027188

Wang, H., Easter, R. C., Zhang, R., Ma, P.-L., Singh, B., Zhang, K., et al. (2020). Aerosols in the E3SM Version 1: New developments and their impacts on radiative forcing. *Journal of Advances in Modeling Earth Systems*, 12, e2019MS001851. https://doi.org/10.1029/2019MS001851

Wang, J., Fan, J., Feng, Z., Zhang, K., Roesler, E., Hillman, B., et al. (2021). Impact of a new cloud microphysics parameterization on the simulations of mesoscale convective systems in E3SM. *Journal of Advances in Modeling Earth Systems*, 13, e2021MS002628. https://doi.org/10.1029/2021MS002628

1058 Wang, Y.-C., Pan, H.-L., & Hsu, H.-H. (2015). Impacts of the triggering function of cumulus
1059 parameterization on warm-season diurnal rainfall cycles at the Atmospheric Radiation
1060 Measurement Southern Great Plains Site. *Journal of Geophysical Research: Atmospheres*, 120,
1061 10681–10702. https://doi.org/10.1002/2015JD023337
1062
1063 Xie, S., L.-Y. Leung, Z. Feng, W. Lin, C.-C. Chen, J. Richter, & J. Fan. (2020). Fy2020 fourth
1064 quarter performance metric: Evaluate improvement in simulations of mesoscale convective
1065 systems from new parameterization developments in E3SM. *Technical Report*. LLNL-TR-
1066 814618, 13 pp., https://doi.org/10.2172/1661028
1067
1068 Xie, S., Lin, W., Rasch, P. J., Ma, P.-L., Neale, R., Larson, V. E., et al. (2018). Understanding
1069 cloud and convective characteristics in version 1 of the E3SM atmosphere model. *Journal of*
1070 *Advances in Modeling Earth Systems*, 10, 2618–2644. https://doi.org/10.1029/2018MS001350
1071
1072 Xie, S., Wang, Y.-C., Lin, W., Ma, H.-Y., Tang, Q., Tang, S., et al. (2019). Improved diurnal
1073 cycle of precipitation in E3SM with a revised convective triggering function. *Journal of*
1074 *Advances in Modeling Earth Systems*, 11. https://doi.org/10.1029/2019MS001702
1075
1076 Xie, S., & Zhang, M. (2000). Impact of the convection triggering function on single-column
1077 model simulations. *Journal of Geophysical Research: Atmospheres*, 105(D11), 14983–14996.
1078 https://doi.org/10.1029/2000JD900170
1079

Yang, Q., Houze, R. A., Jr, Leung, L. R., & Feng, Z. (2017). Environments of long-lived mesoscale convective systems over the central United States in convection permitting climate simulations. *Journal of Geophysical Research: Atmospheres*,122, 13, 288–13, 307. https://doi.org/10.1002/2017JD027033

Yang, Q., Leung, L. R., Feng, Z., & Chen, X. (2023). A Moist Potential Vorticity Model for Midlatitude Long-Lived Mesoscale Convective Systems over Land. *Journal of the Atmospheric Sciences*, 80(10), 2399-2418. https://doi.org/10.1175/JAS-D-22-0244.1

Yang, G. Y., & Slingo, J. (2001). The diurnal cycle in the Tropics. *Monthly Weather Review*, 129(4), 784–801. https://doi.org/10.1175/1520-0493(2001)129%3C0784:TDCITT%3E2.0.CO;2

You, Z., & Deng, Y. (2023). A Hierarchical Dissection of Multiscale Forcing on the Springtime Mesoscale Convective Systems in the United States. *Journal of Climate*, 36(1), 39-54. https://doi.org/10.1175/JCLI-D-22-0150.1

Yuan, X., & Houze, J. (2010). Global variability of mesoscale convective system anvil structure from A-Train satellite data. *Journal of Climate*, 23(21), 5864–5888. https://doi.org/10.1175/2010JCLI3671.1

Zhang, G., & McFarlane, N. A. (1995). Sensitivity of climate simulations to the parameterization of cumulus convection in the Canadian climate centre general circulation model. *Atmosphere-Ocean*, 33(3), 407–446. https://doi.org/10.1080/07055900.1995.9649539

Zhang, M., Xie, S., Feng, Z., Terai, C. R., Lin, W., Chen, C.-C.-J., Fan, J., Golaz, J.-C., Leung, L. R., Richter, J. H., Shan, Y., Song, X., Tang, Q., & Zhang, G. (2024). Evaluation of Mesoscale Convective Systems in High Resolution E3SMv2 [Data set]. Zenodo. https://doi.org/10.5281/zenodo.10521291

Zhang, M., Xie, S., Liu, X., Zhang, D., Lin, W., Zhang, K., et al. (2023). Evaluating EAMv2 simulated high latitude clouds using ARM measurements in the Northern and Southern Hemispheres. *Journal of Geophysical Research: Atmospheres*, 128, e2022JD038364. https://doi.org/10.1029/2022JD038364

Zhang, M., Xie, S., Liu, X., Lin, W., Zheng, X., Golaz, J.-C., & Zhang, Y. (2022). Cloud Phase Simulation at High Latitudes in EAMv2: Evaluation using CALIPSO Observations and Comparison with EAMv1. *Journal of Geophysical Research: Atmospheres*, 127, e2022JD037100. https://doi.org/10.1029/2022JD037100

Zhang, Y., Xie, S., Qin, Y., Lin, W., Golaz, J.-C., Zheng, X., et al. (2023). Understanding Changes in Cloud Simulations from E3SM Version 1 to Version 2, *EGUsphere [preprint]*, https://doi.org/10.5194/egusphere-2023-1263

Zhang, Z., Varble, A., Feng, Z., Hardin, J., & Zipser, E. (2021). Growth of Mesoscale Convective Systems in Observations and a Seasonal Convection-Permitting Simulation over Argentina. *Monthly Weather Review*, 149(10), 3469-3490. https://doi.org/10.1175/MWR-D-20-0411.1

Zhao, M. (2020). Simulations of Atmospheric Rivers, Their Variability, and Response to Global Warming Using GFDL's New High-Resolution General Circulation Model. *Journal of Climate*, *33*(23), 10287-10303. https://doi.org/10.1175/JCLI-D-20-0241.1

Zheng, X., Golaz, J.-C., Xie, S., Tang, Q., Lin, W., Zhang, M., et al. (2019). The summertime precipitation bias in E3SM Atmosphere Model version 1 over the Central United States. *Journal of Geophysical Research: Atmospheres*, 124, 8935–8952. https://doi.org/10.1029/2019JD030662

Table 1. Model simulations used in evaluation.

Model experiments	Description
CTL	Default EAMv2 model simulation at 0.25° horizontal resolution.
P3	Based on CTL, including P3 cloud microphysics scheme.
ZMmicro	Based on CTL, with two-moment cloud microphysics in ZM convection scheme.
MCSP	Based on CTL, with Multi-scale Coherent Structure Parameterization.
MAdj	Based on CTL, with cloud base mass flux adjustment treatment in ZM convection scheme.

Table 2. Observations used in evaluation.

Data name	Horizontal resolution	Coverage area	Data period	Note for MCS tracking
NASA Global Merged IR V1 infrared T_b product	4 km	60°S-60°N	2005-2009	Used as the T_b data source for both global and CONUS MCS tracking
GPM Integrated Multi-satelliteE Retrievals (IMERG) V06B precipitation data	10 km	60°S-60°N	2005-2009	Used as surface precipitation data source for global MCS tracking
NEXRAD radar and Stage IV multisensor precipitation data	4 km	CONUS	2005-2009	Used as surface precipitation data source for CONUS MCS tracking

Note: T_b and precipitation data are regridded to 0.25° (~25 km) horizontal resolution to match the model grid spacing.

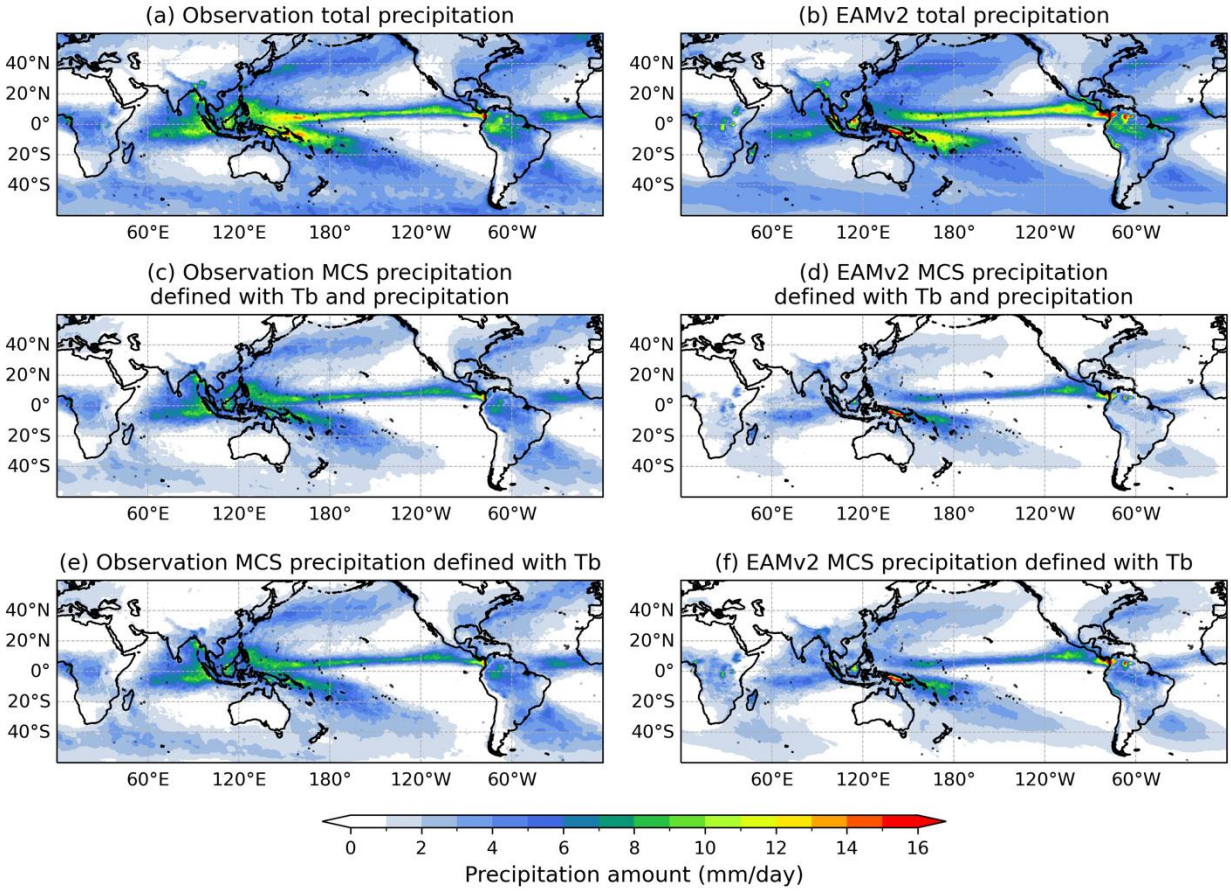

Figure 1. Maps of annual mean total precipitation amount (a, b), MCS precipitation amount defined with T_b and precipitation tracking method (c, d), MCS precipitation amount defined with T_b only tracking method (e, f), The IMERG observation is shown on the left while EAMv2 simulation is on the right. Model and observations cover between 2005 and 2009.

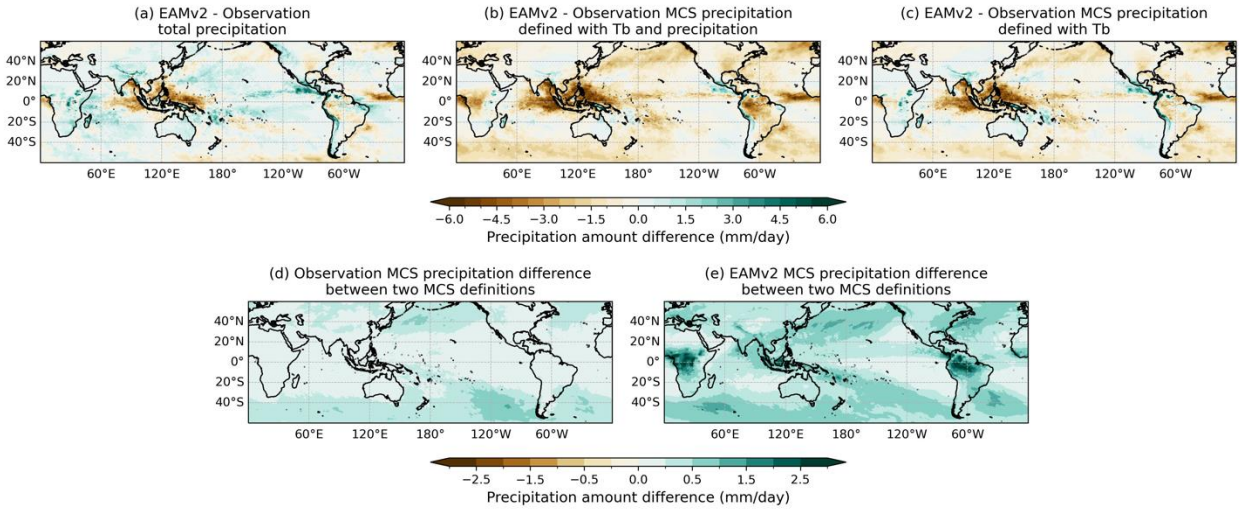

Figure 2. Maps of annual mean total precipitation bias (a), MCS precipitation bias defined using T_b and surface precipitation (b), MCS precipitation bias defined using T_b only in EAMv2 compared to the IMERG observation. The MCS precipitation difference between two tracking methods is shown in (d) for the IMERG observation and (e) for EAMv2. The MCS precipitation differences in (d) and (e) are calculated by subtracting T_b and precipitation combined method from T_b only method for tracked MCSs.

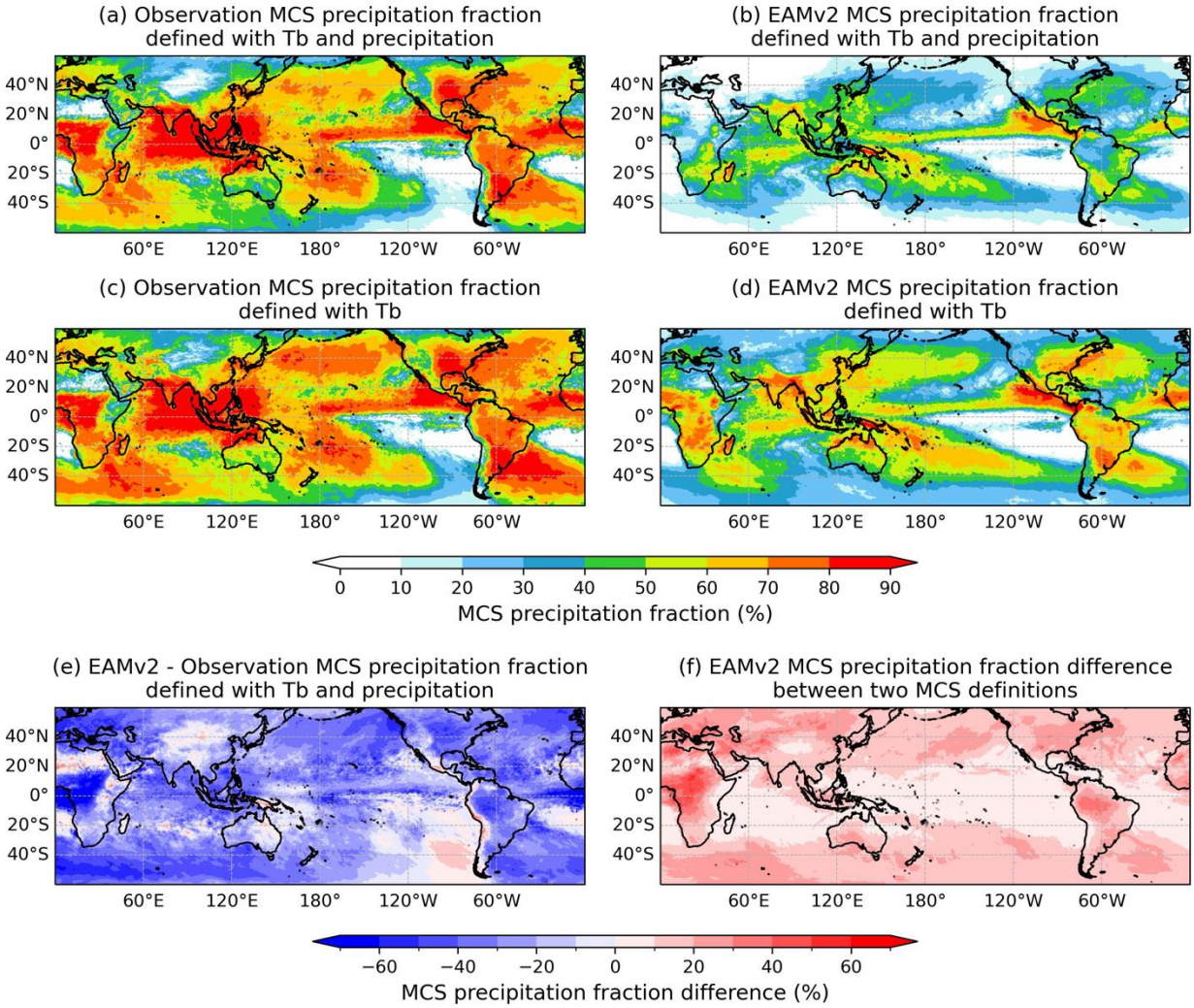

Figure 3. Maps of annual mean MCS precipitation fraction defined using T_b and surface precipitation (a, b) and MCS precipitation fraction defined using T_b only (c, d). The IMERG observation is shown on the left and EAMv2 simulation is shown on the right. (e) shows the MCS precipitation fraction bias defined using T_b and precipitation, which is (b) minus (a). (f) shows the MCS precipitation fraction difference between two tracking methods in EAMv2, which is (d) minus (b). The MCS precipitation fraction is calculated by dividing MCS precipitation by total precipitation.

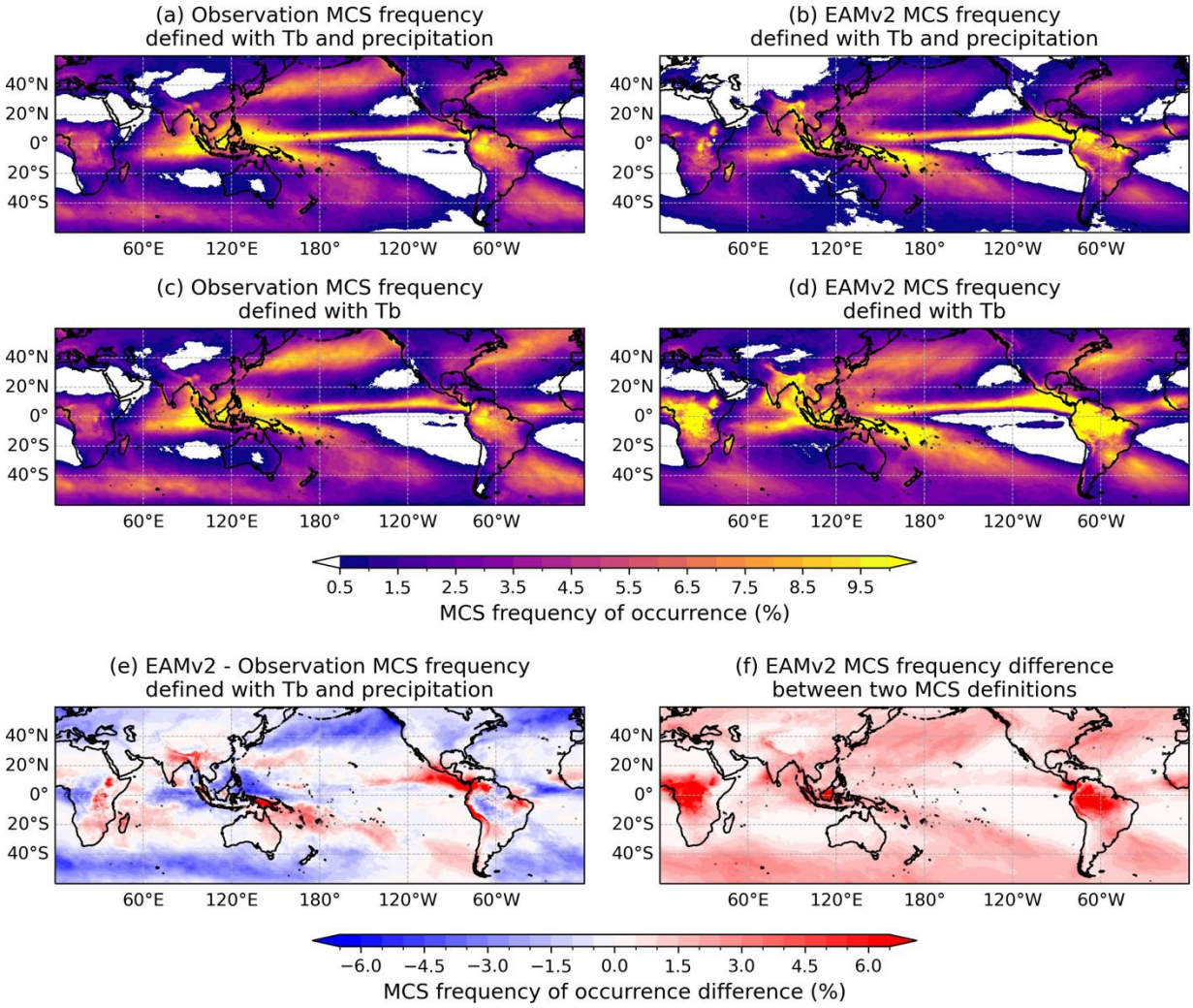

Figure 4. Maps of annual mean MCS precipitation frequency defined using T_b and surface precipitation (a, b) and MCS precipitation frequency defined using T_b only (c, d). The IMERG observation is shown on the left and EAMv2 simulation is shown on the right. (e) shows the MCS precipitation frequency bias defined using T_b and precipitation, which is (b) minus (a). (f) shows the MCS precipitation frequency difference between two tracking methods in EAMv2, which is (d) minus (b). MCS precipitation frequency is defined as the ratio of total hours of MCS precipitation to total hours between 2005-2009.

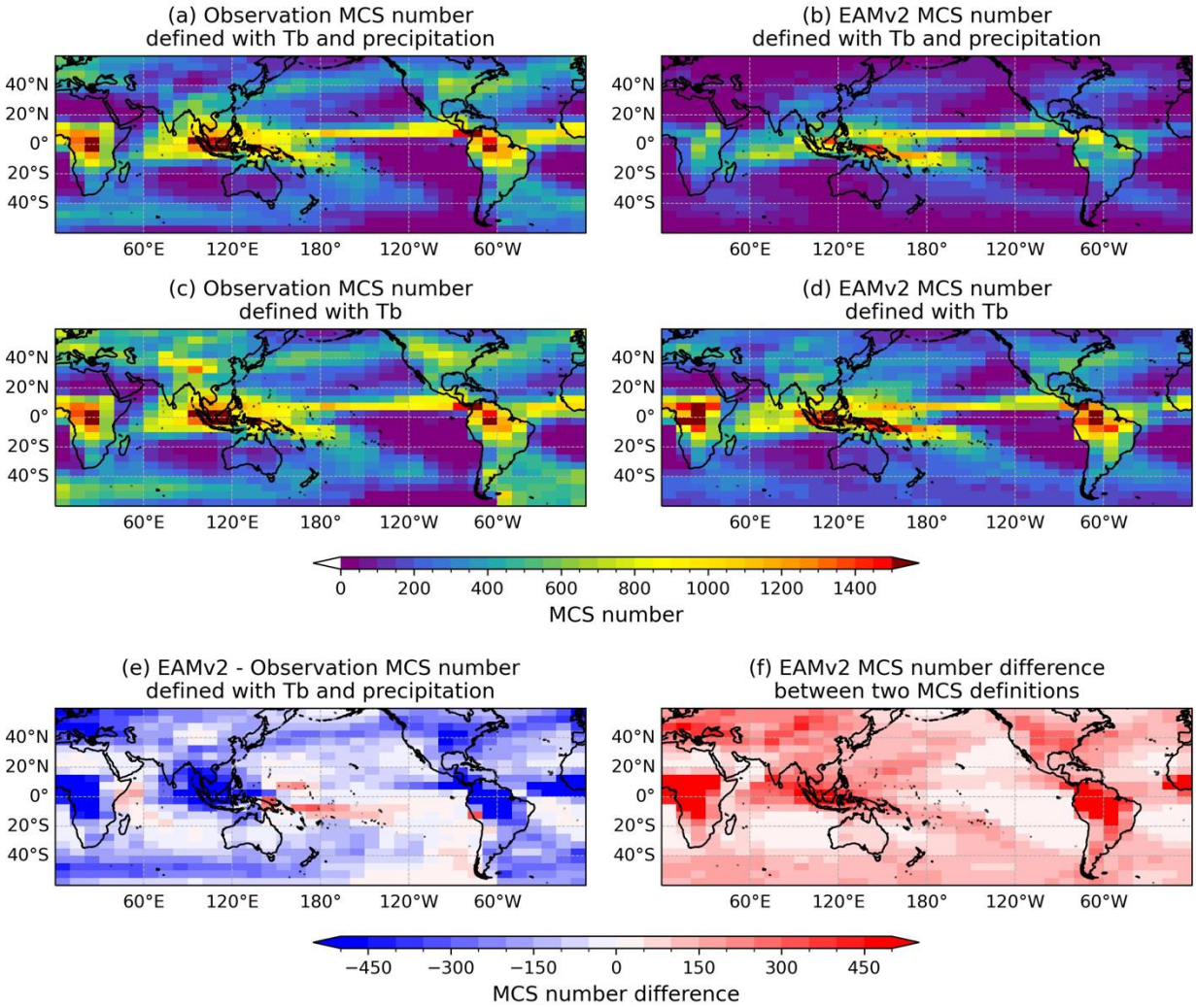

Figure 5. Maps of annual mean MCS number defined using T_b and surface precipitation (a, b) and MCS number defined using T_b only (c, d). The IMERG observation is shown on the left and EAMv2 simulation is shown on the right. (e) shows the MCS number bias defined using T_b and precipitation, which is (b) minus (a). (f) shows the MCS number difference between two tracking methods in EAMv2, which is (d) minus (b). MCS number is counted as the number of unique latitude/longitude pairs of each MCS track within the 5° × 10° latitude/longitude grids.

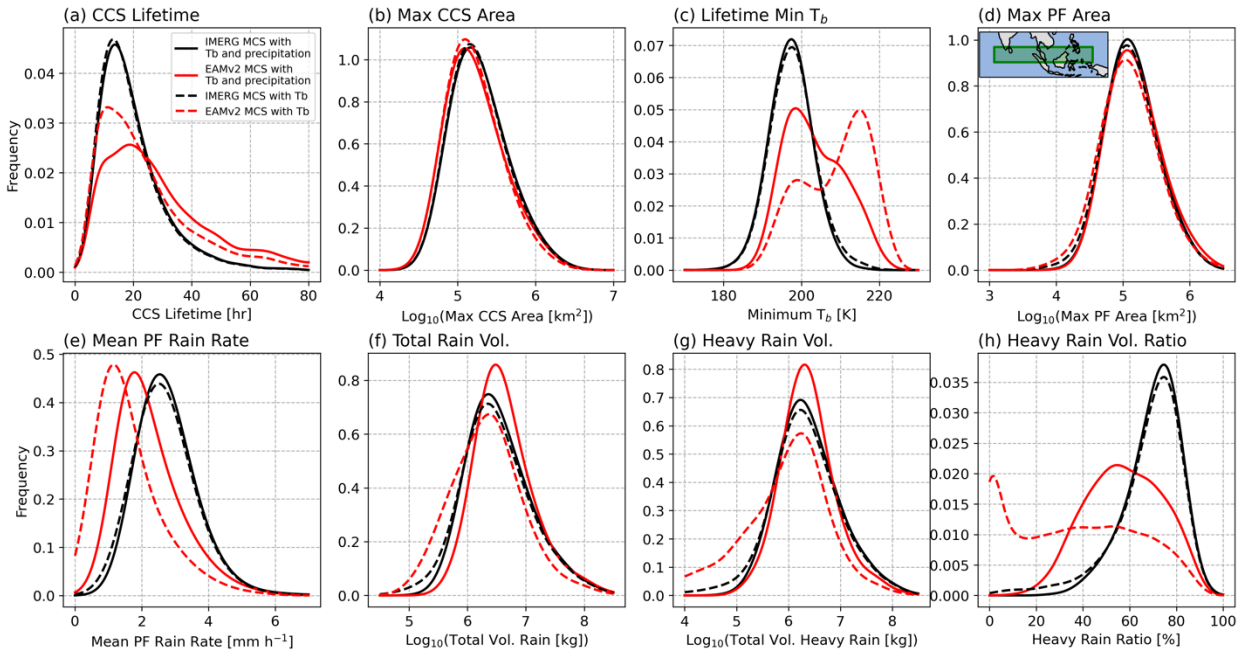

Figure 6. The PDFs of (a) cold cloud system lifetime, (b) maximum cold cloud system area throughout the MCS lifetime, (c) minimum cloud top brightness temperature throughout the MCS lifetime, (d) maximum precipitation feature area, (e) mean rain rate within the precipitation feature domain, (f) total rain volume, (g) heavy rain (rain rate > 2 mm h^{-1}) volume, and (h) heavy rain volume ratio over the Indo-Pacific region. Black lines represent the IMERG observation, and red lines represent the EAMv2 model simulation. Solid lines are for MCSs defined using combined T_b and surface precipitation, and dashed lines are for MCSs defined using only T_b. The PDFs are calculated in the Indo-Pacific domain indicated by the green box in panel (d).

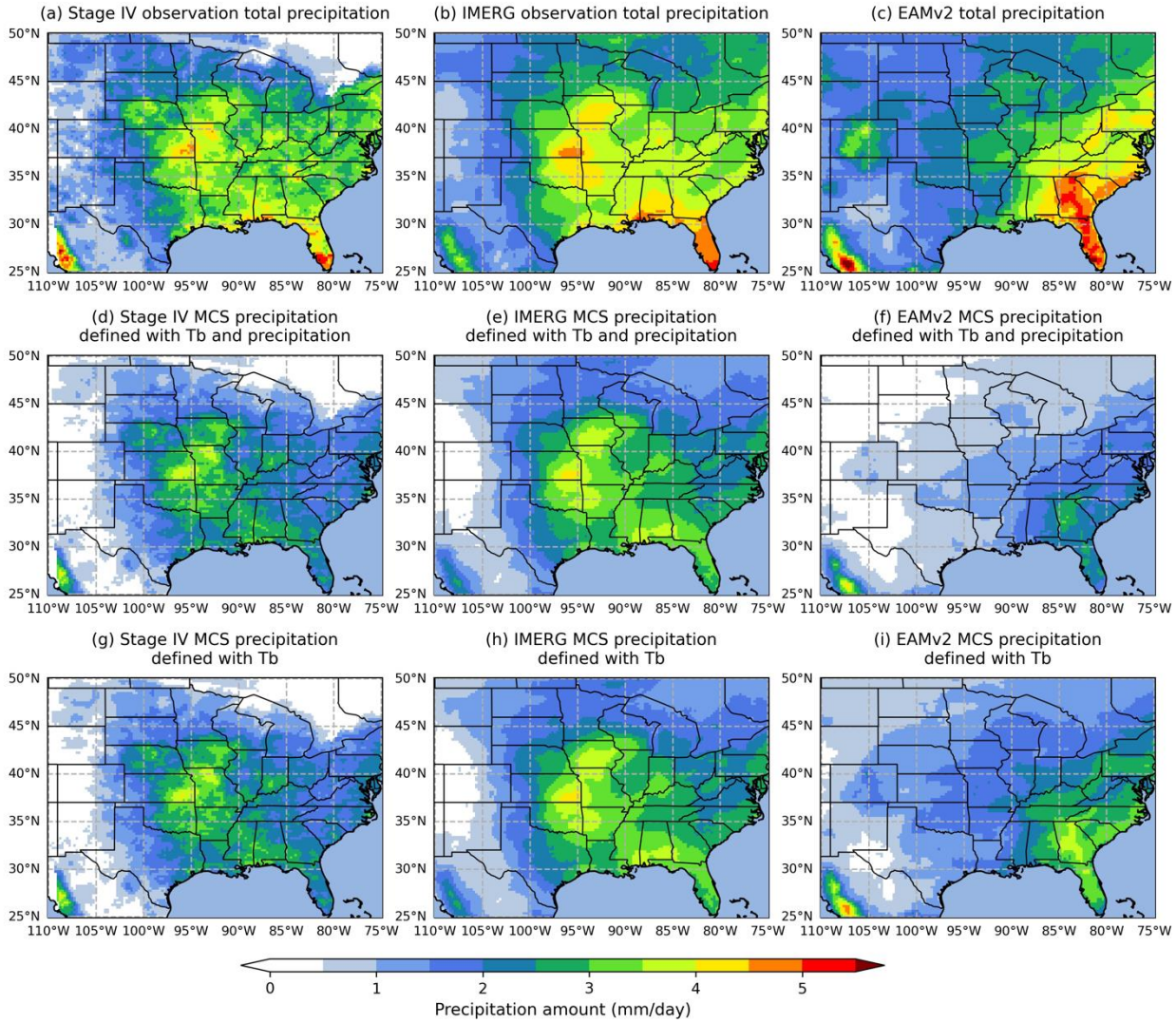

Figure 7. Maps of mean total precipitation amount (a-c), MCS precipitation amount defined using combined T_b and surface precipitation (d-f), and MCS precipitation amount defined using only T_b (g-i) from March to August between 2005 and 2009 over the CONUS region. Precipitation observations from the Stage IV precipitation gauge (left) and from the IMERG satellite retrieval (middle) are compared with the EAMv2 simulations (right).

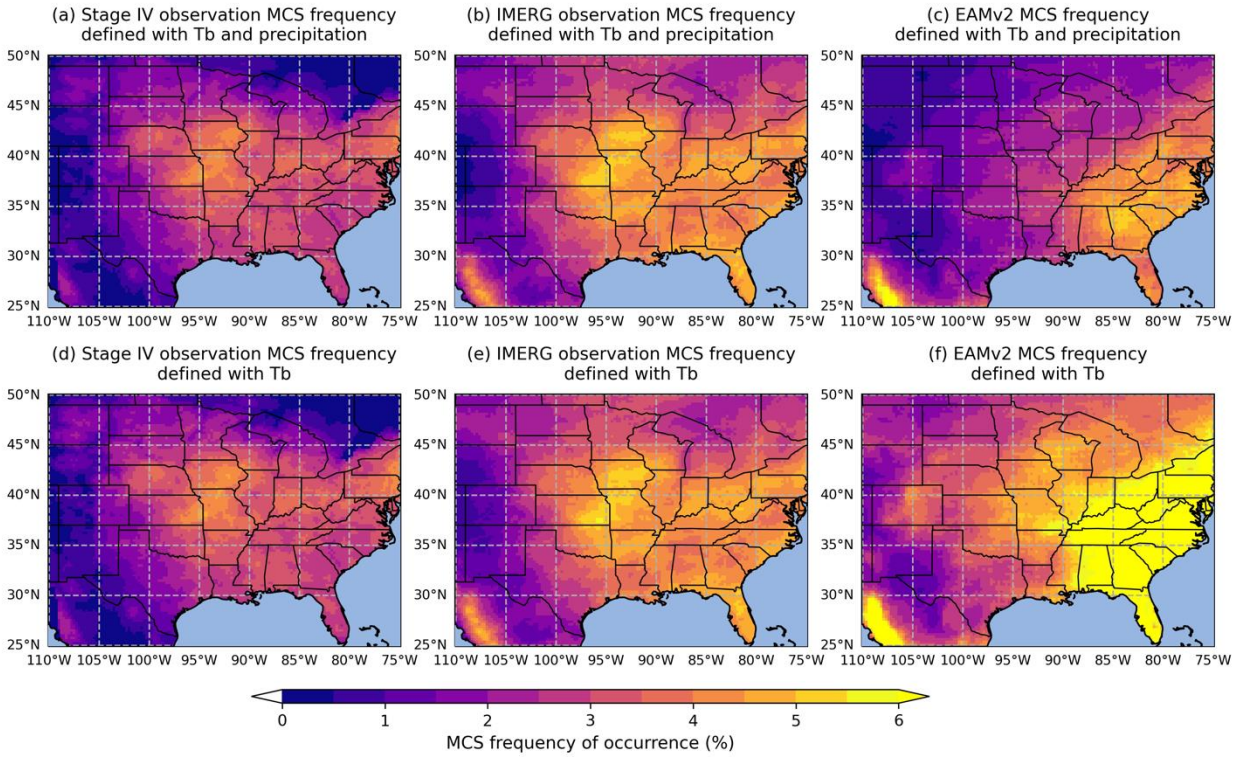

Figure 8. Maps of mean MCS precipitation frequency defined using combined T_b and surface precipitation (a-c) and MCS defined using only T_b (d-f) from March to August between 2005 and 2009 over the CONUS region. Observations from the Stage IV precipitation gauge (left) and from the IMERG satellite retrieval (middle) are compared with the EAMv2 simulations (right).

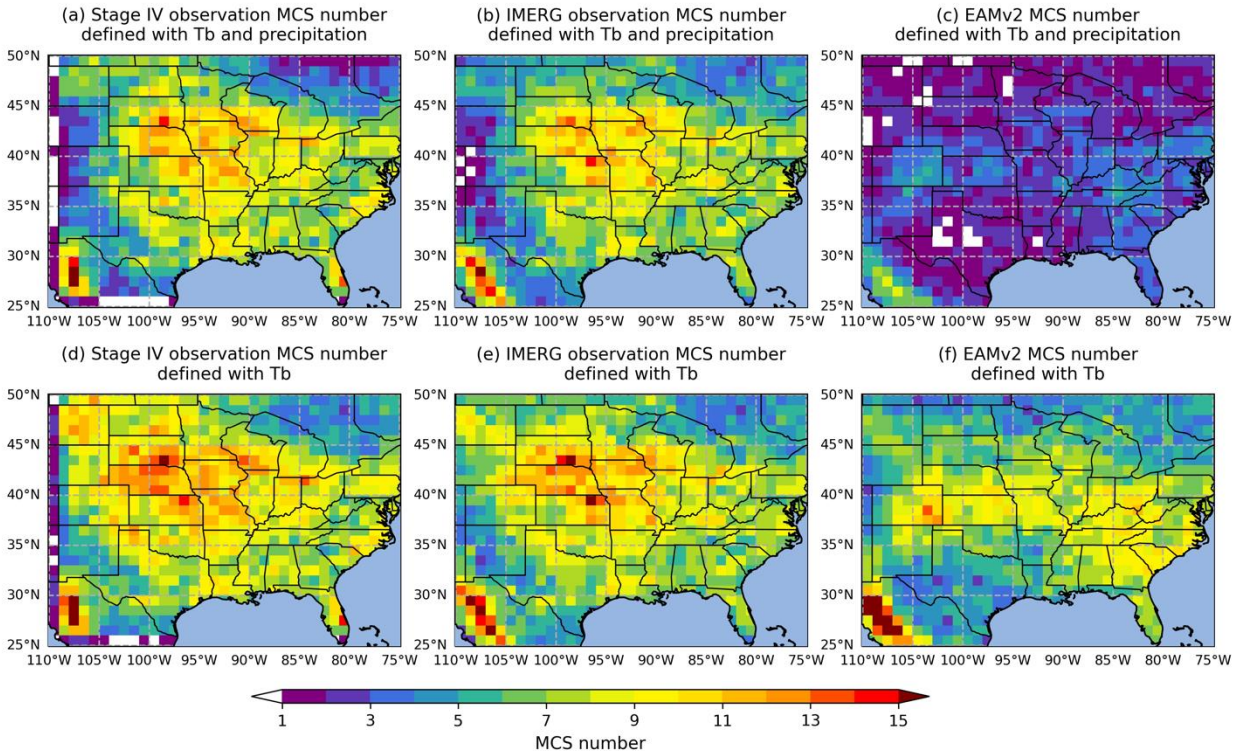

Figure 9. Maps of mean MCS number of MCS defined using combined T_b and surface precipitation (a-c) and MCS defined using only T_b (d-f) over the CONUS region. Observations from the Stage IV precipitation gauge (left) and from the IMERG satellite retrieval (middle) are compared with the EAMv2 simulations (right). The MCS number is calculated as the number of unique latitude/longitude pairs of each MCS track within the 1° × 1° latitude/longitude grids from March to August between 2005 and 2009.

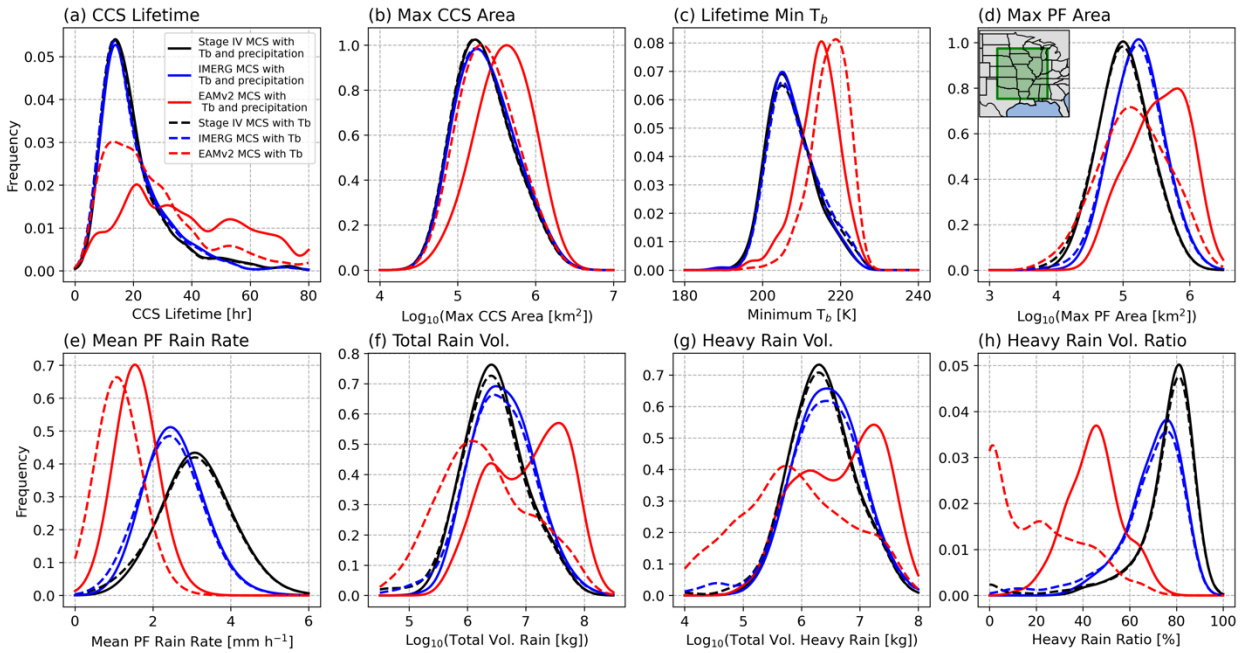

Figure 10. Same as Figure 6, but the PDFs are calculated in the central U.S. domain shown in panel (d). Black lines represent observations using ground-based measurements. Blue lines represent the IMERG precipitation data. Black lines represent EAMv2 model simulations.

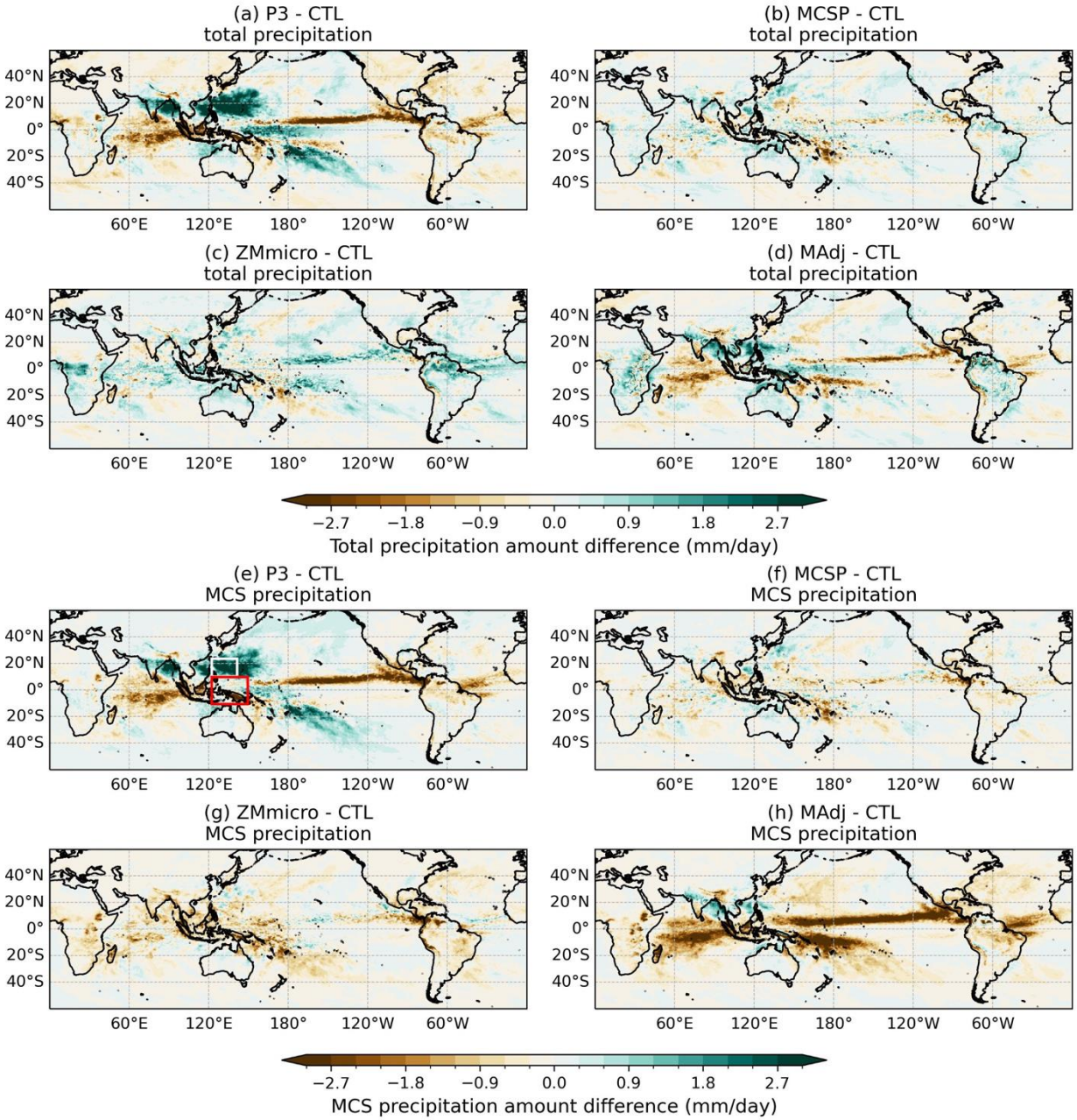

Figure 11. Maps of annual mean total precipitation rate difference (a-d) and MCS precipitation rate difference defined using combined T_b and surface precipitation (e-h) between 2005 and 2009. (a)-(d) and (e)-(h) are the differences between individual new physics feature (i.e., P3 cloud microphysics, MCSP, convective microphysics scheme in ZM, and cloud base mass flux adjustment) and CTL simulations, respectively. White and red boxes are the areas used to calculate rain rate PDFs in Figure 13.

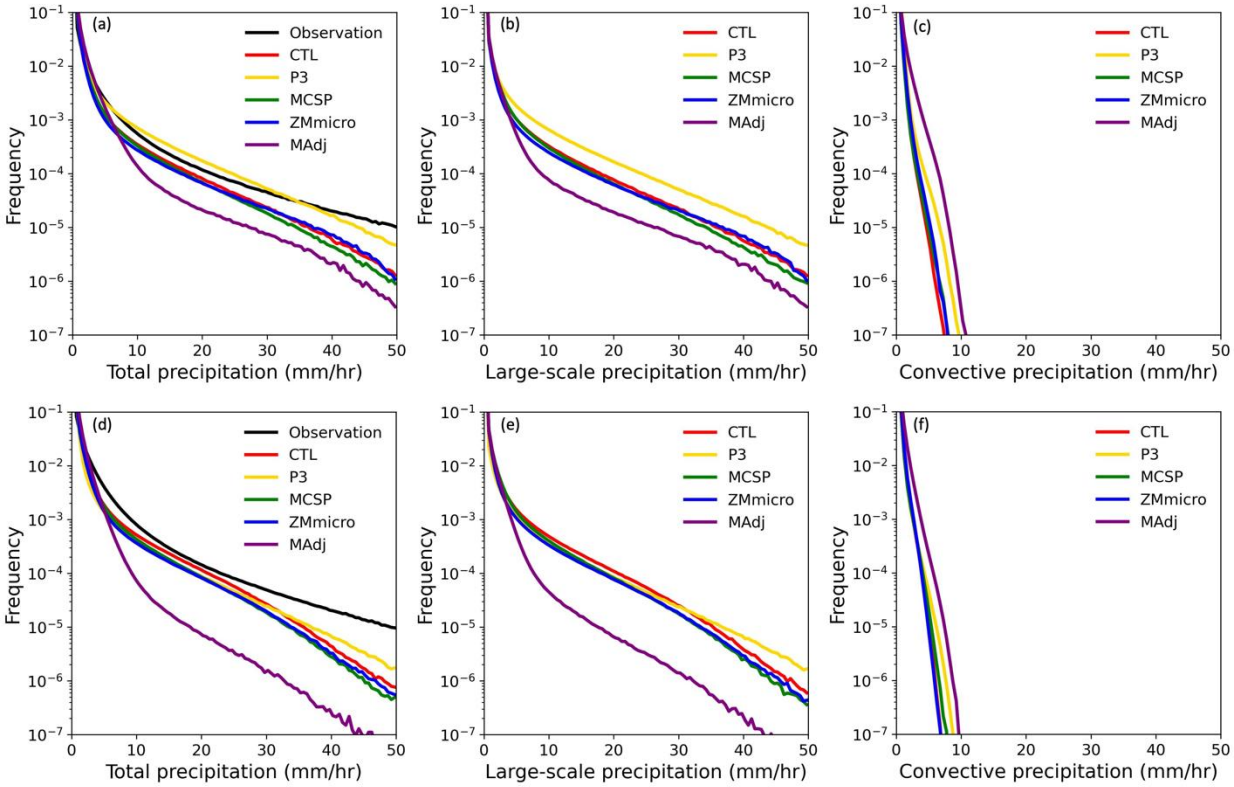

Figure 12. Probability density functions of hourly total, large-scale, and convective precipitation rates between 2005 and 2009 sampled over the white box domain (a-c) and red box domain (d-f) in Figure 11. The IMERG observation and EAMv2 model simulations with CTL and new physics features are shown.

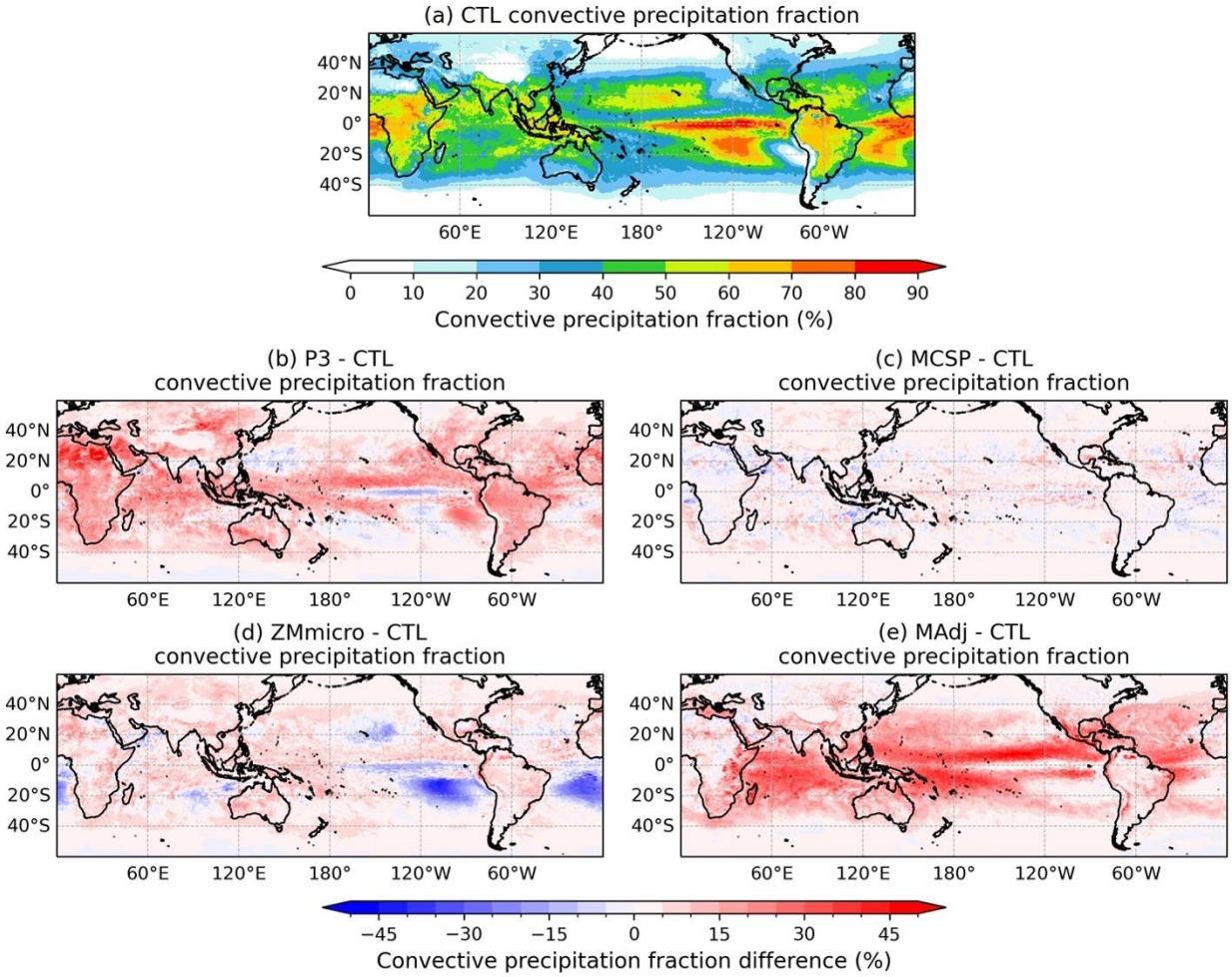

Figure 13. Maps of annual mean convective precipitation fraction in the default EAMv2 model (CTL) between 2005 and 2009, and the convective precipitation fraction difference (b-e) between individual new physics feature (i.e., P3 cloud microphysics, MCSP, convective microphysics scheme in ZM, and cloud base mass flux adjustment) and CTL simulations, respectively. Convective precipitation fraction is calculated as the contribution of hourly convective precipitation rate to total precipitation rate in the model.

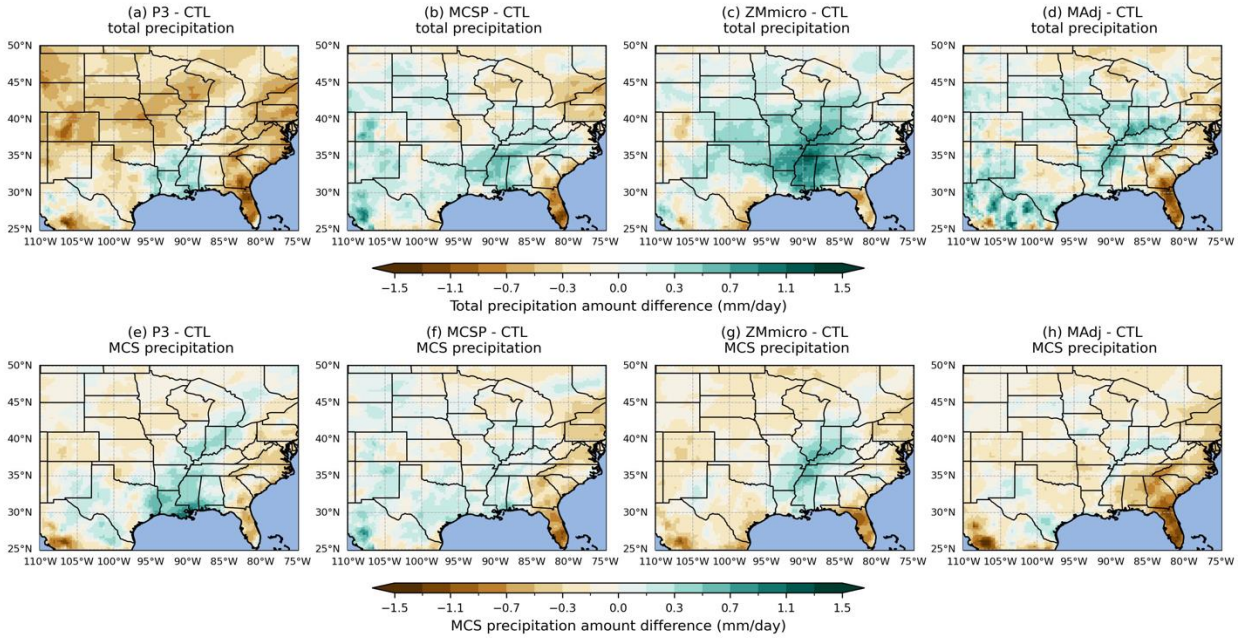

Figure 14. Same as Figure 11, but for the CONUS region from March to August between 2005 and 2009.