JGR Atmospheres

RESEARCH ARTICLE
10.1029/2024JD042204

Mesoscale Convective Systems Tracking Method Intercomparison (MCSMIP): Application to DYAMOND Global km-Scale Simulations

Zhe Feng[1], Andreas F. Prein[2,3], Julia Kukulies[2], Thomas Fiolleau[4], William K. Jones[5], Ben Maybee[6], Zachary L. Moon[2,7], Kelly M. Núñez Ocasio[2,8], Wenhao Dong[9,10], Maria J. Molina[2,11], Mary Grace Albright[12], Manikandan Rajagopal[13], Vanessa Robledo[14], Jinyan Song[15], Fengfei Song[15,16], L. Ruby Leung[1], Adam C. Varble[1], Cornelia Klein[17], Remy Roca[4], Ran Feng[12], and John F. Mejia[18]

[1]Pacific Northwest National Laboratory, Richland, WA, USA, [2]NSF National Center for Atmospheric Research, Boulder, CO, USA, [3]Institute for Atmospheric and Climate Science, ETH Zürich, Zurich, Switzerland, [4]LEGOS, CNES, CNRS, UPS, IRD, University of Toulouse, Toulouse, France, [5]University of Oxford, Oxford, UK, [6]University of Leeds, Leeds, UK, [7]Earth Resources Technology, Inc., Laurel, MD, USA, [8]Texas A&M University, College Station, TX, USA, [9]University Corporation for Atmospheric Research, Boulder, CO, USA, [10]NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA, [11]University of Maryland, College Park, MD, USA, [12]University of Connecticut, Storrs, CT, USA, [13]University of Utah, Salt Lake City, UT, USA, [14]University of Iowa, Iowa City, IA, USA, [15]Ocean University of China, Qingdao, China, [16]Laoshan Laboratory, Qingdao, China, [17]UK Centre for Ecology & Hydrology, Wallingford, UK, [18]Desert Research Institute, Reno, NV, USA

Abstract Global kilometer-scale models represent the future of Earth system modeling, enabling explicit simulation of organized convective storms and their associated extreme weather. Here, we comprehensively evaluate tropical mesoscale convective system (MCS) characteristics in the DYAMOND (DYnamics of the atmospheric general circulation modeled on non-hydrostatic domains) simulations for both summer and winter phases. Using 10 different feature trackers applied to simulations and satellite observations, we assess MCS frequency, precipitation, and other key characteristics. Substantial differences (a factor of 2–3) arise among trackers in observed MCS frequency and their precipitation contribution, but model-observation differences in MCS statistics are more consistent across trackers. DYAMOND models are generally skillful in simulating tropical mean MCS frequency, with multi-model mean biases ranging from −2%–8% over land and −8%–8% over ocean (summer vs. winter). However, most DYAMOND models underestimate MCS precipitation amount (23%) and their contribution to total precipitation (17%). Biases in precipitation contributions are generally smaller over land (13%) than over ocean (21%), with moderate inter-model variability. While models better simulate MCS diurnal cycles and cloud shield characteristics, they overestimate MCS precipitation intensity and underestimate stratiform rain contributions (up to a factor of 2), particularly over land, albeit observational uncertainties exist. Additionally, models exhibit a wide range of precipitable water in the tropics compared to reanalysis and satellite observations, with many models showing exaggerated sensitivity of MCS precipitation intensity to precipitable water. The MCS metrics developed here provide process-oriented diagnostics to guide future model development.

Plain Language Summary Global storm-resolving models represent a significant advancement in predicting extreme weather events, as they can directly simulate convective storms and their impacts. These models are crucial for understanding how extreme weather might change in a warming climate. This study evaluates the performance of these advanced models in predicting large tropical storms, known as mesoscale convective systems, which are key drivers of heavy rainfall and severe weather. The research used 10 different methods to track storms in model simulations and satellite observations. We found that different tracking methods can yield varying results regarding storm frequency and rainfall. However, when comparing simulated storms to observed ones, the results are more consistent across the different trackers. The models generally perform well in predicting the average frequency of tropical storms. But they underestimate the rainfall amount from these storms by about 23% and their contribution to total precipitation by about 17%. Additionally, most models predicted heavier rainfall for a given amount of atmospheric water vapor compared to observations. The multi-method tracking analysis offers valuable insights for improving future model development efforts, helping them better predict extreme weather events and address societal needs in a changing climate.

Key Points:

- Mesoscale convective systems (MCSs) in global km-scale models are evaluated against satellite observations using 10 feature trackers
- Most models underestimate MCS precipitation amount and their contribution to total precipitation, with larger biases over ocean
- Models overestimate MCS precipitation intensity and its sensitivity to precipitable water while underestimating stratiform rain contribution

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

Z. Feng,
zhe.feng@pnnl.gov

Citation:

Feng, Z., Prein, A. F., Kukulies, J., Fiolleau, T., Jones, W. K., Maybee, B., et al. (2025). Mesoscale convective systems tracking method intercomparison (MCSMIP): Application to DYAMOND global km-scale simulations. *Journal of Geophysical Research: Atmospheres*, *130*, e2024JD042204. https://doi.org/10.1029/2024JD042204

Received 13 AUG 2024
Accepted 26 MAR 2025

Author Contributions:

Conceptualization: Zhe Feng
Data curation: Zhe Feng, Andreas F. Prein, Julia Kukulies, Thomas Fiolleau, William K. Jones, Ben Maybee, Zachary L. Moon, Wenhao Dong, Mary Grace Albright, Manikandan Rajagopal, Vanessa Robledo, Jinyan Song
Formal analysis: Zhe Feng
Funding acquisition: L. Ruby Leung
Investigation: Zhe Feng

Methodology: Zhe Feng, Andreas F. Prein, Julia Kukulies, Thomas Fiolleau, William K. Jones, Ben Maybee, Zachary L. Moon, Kelly M. Núñez Ocasio, Wenhao Dong, Maria J. Molina, Mary Grace Albright, Manikandan Rajagopal, Vanessa Robledo, Ran Feng
Resources: Zhe Feng
Software: Zhe Feng
Supervision: Zhe Feng
Validation: Zhe Feng
Visualization: Zhe Feng
Writing – original draft: Zhe Feng
Writing – review & editing: Zhe Feng, Julia Kukulies, Ben Maybee, Kelly M. Núñez Ocasio, Wenhao Dong, Maria J. Molina, Mary Grace Albright, Manikandan Rajagopal, Vanessa Robledo, Fengfei Song, L. Ruby Leung, Adam C. Varble, Cornelia Klein, Remy Roca, Ran Feng, John F. Mejia

1. Introduction

Organized moist convection that is clustered in space and time plays a critical role in the Earth's energy budget and hydrological cycle (Stephens et al., 2023). A mesoscale convective system (MCS) is a form of highly organized moist convection with a cluster of cumulonimbus clouds spanning at least 100 km horizontally and lasting many hours to more than a day (Houze, 2014). Decades of research, field campaigns, and remote sensing technology advancements, along with high-resolution modeling and theoretical insights have led to a fundamental understanding of the structure and evolution of MCSs as well as their impacts on weather and climate (see reviews by Houze (2018), Schumacher and Rasmussen (2020), and Chakraborty et al. (2023)).

MCSs contribute to over half of the total precipitation in the tropics (Feng, Leung, et al., 2021; Nesbitt et al., 2006; Yuan & Houze, 2010) and several midlatitude regions (Feng et al., 2019; Haberlie & Ashley, 2019; Kukulies et al., 2021; Prein et al., 2024). When tropical convection is more clustered, the surrounding free troposphere becomes drier with reduced cloudiness, increasing emission to space, and thus enhancing radiative cooling (Bony et al., 2020). Tropical MCSs are often coupled with large-scale modes of variability. A dominant intraseasonal mode is the Madden-Julian Oscillation (Madden & Julian, 1971, 1972), which has significant impacts on sub-seasonal prediction (Kim et al., 2023). MCSs were considered as a potential building block of large-scale tropical variability such as the MJO (Chen et al., 1996; Mapes et al., 2006). Recent work showed that active phases of large-scale convectively coupled equatorial waves enhanced MCS frequency, size, and precipitation throughout the tropics, especially for Kelvin waves, tropical depressions, and the MJO (Cheng et al., 2023). Over Africa, the propagation and growth of African easterly waves is largely modulated by MCSs (Núñez Ocasio et al., 2020a, 2020b). The interaction of MCSs with African easterly waves over Africa and the Atlantic has proven intrinsic in tropical cyclone formation (Núñez Ocasio et al., 2024; Rajasree et al., 2023). These studies point to the importance of properly simulating the processes that produce MCSs and their upscale effects on large-scale circulations in numerical models (Yang et al., 2019) for more accurate sub-seasonal predictions.

Over land, MCSs were linked to a myriad of hazardous weather events. Extreme precipitation is disproportionately produced by infrequent and long-lived MCSs (Núñez Ocasio et al., 2020a; Prein et al., 2023; Roca & Fiolleau, 2020; Stevenson & Schumacher, 2014; Wu et al., 2024). While numerous case studies have documented flooding associated with individual MCS events in both tropics and midlatitudes, some recent studies have quantified the dominant role of MCSs in producing warm-season floods in the United States and East Asia (Brunner & Dougherty, 2022; Ding et al., 2024; Hu et al., 2021). Certain types of MCSs can cause widespread straight-line wind damage, sometimes referred to as derechos, which were extensively observed in the United States and Europe (Ashley & Mote, 2005; Bentley & Mote, 1998; Coniglio et al., 2006; Fery & Faranda, 2024; Surowiecki & Taszarek, 2020). Although relatively rare, MCSs can also produce large hail and tornadoes (Thompson et al., 2012; Wang et al., 2023).

Despite their importance to both weather and climate, simulating MCSs in general circulation models (GCMs) with coarse horizontal grid spacing (∼100 km) and parameterized convection has been a long-standing challenge (Feng, Song, et al., 2021; Hsu et al., 2023; Wu et al., 2024). Some recent studies showed that higher resolution GCMs (25–50 km grid spacing), such as those from the High Resolution Model Intercomparison Project (HighResMIP) for CMIP6 (Haarsma et al., 2016), also struggled to properly simulate organized convection and their associated precipitation (Dong et al., 2021; Feng, Song, et al., 2021; Lin et al., 2022; Zhang et al., 2024), as exemplified by biases in the diurnal cycle of precipitation (Dong et al., 2023; Song et al., 2024) and surface temperature (Lin et al., 2017; Ma et al., 2018; Morcrette et al., 2018).

In contrast, convection-permitting models (CPMs) with km-scale horizontal grid spacing can explicitly simulate deep convection. Many studies have demonstrated that CPMs have significantly improved skill in simulating various aspects of MCSs in observations, including the frequency, size, duration, diurnal cycle, and precipitation distribution in different geographical regions (Chen et al., 2021; Crook et al., 2019; Feng et al., 2018; Müller et al., 2023; Prein, Liu, Ikeda, Bullock, et al., 2017; Prein et al., 2024; Zhang et al., 2021). Compared to models with parameterized convection, km-scale models more reliably simulate potential changes of MCSs in a warmer climate because they explicitly simulate processes that facilitate upscale growth and organization of convection, such as cold pools and convective dynamics interactions with cloud microphysics, the surface and surrounding atmosphere. CPMs were used to quantify the effects of warming on high-impact MCS hazards including extreme

precipitation (Bao et al., 2024; Fitzpatrick et al., 2020; Prein, Liu, Ikeda, Trier, et al., 2017), flooding (Dougherty & Rasmussen, 2020; Feng, Chen, & Leung, 2024), and severe winds (González-Alemán et al., 2023; Lasher-Trapp et al., 2023; Li et al., 2023; Prein, 2023).

While a km-scale model is a promising tool to address urgent needs for decision-making and climate change impact assessments related to extremes at the regional scale (Gutowski et al., 2020), many important biases exist in these models. For example, convective updrafts and precipitation intensity are consistently overestimated (Fan et al., 2017; Müller et al., 2023; Prein et al., 2021; Varble et al., 2014a, 2020; Wang et al., 2020) while stratiform precipitation is underestimated (Feng et al., 2018; Hagos et al., 2014; Han et al., 2019; Varble et al., 2014b). These biases highlight the need for more comprehensive evaluation and improvements for CPMs, particularly regarding their representation of MCSs.

With increasing computational power, global CPM is now a reality (Satoh et al., 2019). The DYAMOND (DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains) project (Stevens et al., 2019) provided an opportunity to examine state-of-the-art global models that explicitly simulate clouds and convection. Some global km-scale models have already been used to produce short-term climate simulations (Cheng et al., 2022; Hohenegger et al., 2023). Feng, Leung, et al. (2023) found a surprisingly large spread in the simulated deep convection and MCS frequencies in the DYAMOND models, though large, long-lived MCSs are generally underestimated compared to observations. However, only a subset of the DYAMOND Winter (Phase-II) ensemble was examined with a single tracking algorithm in that study. Prein et al. (2024) highlighted significant differences among different feature trackers in estimating MCS characteristics such as frequency, size, and duration for both observations and km-scale regional climate simulation in South America. These studies raise further questions: (a) How sensitive are the DYAMOND simulated MCS characteristics to different tracker formulations? (b) What km-scale model biases are robust among trackers and how do the biases relate to environmental moisture?

In this study, we conduct a more comprehensive evaluation of simulated MCS characteristics in all available DYAMOND models for both summer and winter phases by applying 10 different feature trackers to the simulations and satellite observations. We illustrate metrics that are robust among trackers to evaluate MCSs in global km-scale models to guide future model evaluation efforts, and also propose a framework for future MCS tracking method intercomparison (MCSMIP) efforts. The rest of the paper is organized as follows: Section 2 describes the data sets and tracking methods; Section 3 provides an overview of the simulated convective clouds, precipitation, and moisture; Section 4 presents results for the MCS tracking intercomparison; discussions and conclusions are provided in Section 5.

2. Data and Methods

2.1. Observational Data

The primary global satellite data used for tracking MCS in observations are the NOAA CPC/NCEP global merged infrared brightness temperature (T_b) data (Janowiak et al., 2017). The data set merge multiple geostationary satellite T_b data into a global (60°S–60°N) data set with 4 km pixel resolution at 30-min intervals. Occasionally, certain regions have missing T_b data possibly due to satellite sensor issues. Most of the regions between 60°S and 60°N have more than 98% valid data for both DYAMOND phases, except for the southeastern Pacific dropping to ∼80% during summer and east Asia dropping to ∼96% during winter (Figure S1 in Supporting Information S1). Fortunately, the southeastern Pacific during summer has relatively low MCS frequency (Feng, Leung, et al., 2021); thus the impact of the missing T_b data on MCS tracking is expected to be small.

The Integrated Multi-satellitE Retrievals for GPM (IMERG) precipitation product is also used for identifying MCSs. The IMERG product has $0.1° \times 0.1°$ (∼10×10 km) resolution at 30 min intervals. The IMERG V06B data (Huffman et al., 2019) have been used in our previous work to produce long-term global MCS tracking data set (Feng, Leung, et al., 2021). The new IMERG V07B data (Huffman et al., 2023) have been released recently. Numerous changes were made in the V07B to improve precipitation detection, and systematic and random biases (see Huffman et al. (2023) Technical Documentation for more details). Quantitative evaluation of the IMERG V07B data set is currently underway. We included the V07B data set in many of our analyses as an estimate of the uncertainty in the precipitation retrieval.

Similar to our previous approach (Feng, Leung, et al., 2021), the 4-km T_b data set are conservatively regridded to match the IMERG 0.1° × 0.1° resolution. The 00-min T_b snapshot is used to represent a given hour, and the two 30-min IMERG precipitation retrievals are averaged to obtain hourly precipitation rate (e.g., 01:00 and 01:30 are averaged to 01:00). The instantaneous T_b snapshot is comparable to the model simulated outgoing longwave radiation (OLR), while the average hourly precipitation reduces the artificial temporal variability in the 30-min IMERG data due to spatial resolution differences from various satellite microwave sensors (Rajagopal et al., 2021). The T_b and IMERG precipitation data sets are then combined into a single data set for MCS tracking in observations. The combined T_b and IMERG data set covers the two DYAMOND phases: 1 August 2016–10 September 2016 (Summer, Phase-I) and 20 January 2020–28 February 2020 (Winter, Phase-II).

We used the GPM DPR (Dual-Frequency Precipitation Radar) Ku-band precipitation profile 2A V07 data set (Iguchi & Meneghini, 2021) to further provide uncertainty estimates for observed precipitation intensity. We consider the DPR precipitation data set to be more accurate than IMERG because the DPR retrieval uses active radar sensing that measures reflectivity vertical profiles used to retrieve near-surface precipitation, while IMERG relies on passive microwave sensing sensitive to column-integrated hydrometeor scattering signatures combined with cloud-top infrared temperature retrievals and spatiotemporal morphing. However, DPR is limited in spatiotemporal coverage and cannot be used for tracking MCSs. The DPR Ku-band precipitation profile product is a single-frequency retrieval from the Ku-band radar and has a 5 × 5 km footprint at nadir with a 245-km swath width. The Ku-band product provides a snapshot of precipitation profiles over the swath and completes an orbit every ∼1.5 hr. The near-surface precipitation rate from the DPR product is used in this study. To facilitate comparison with IMERG, the DPR precipitation data set is conservatively regridded to match the IMERG 0.1° × 0.1° grid. The DPR data set used in this study covers the same period as the T_b and IMERG data set.

The satellite-derived precipitable water (PW, i.e., total column water vapor) product is obtained from the Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite Data (HOAPS, Andersson et al. (2021)). The HOAPS PW is a gridded data set with 0.5° × 0.5° horizontal resolution and 6 hourly temporal resolution. The HOAPS uses microwave measurements from the Special Sensor Microwave Imager (SSM/I) and the Special Sensor Microwave Imager Sounder (SSMIS) sensors to retrieve PW over ice-free ocean surfaces.

PW from the ERA5 reanalysis (Hersbach et al., 2020) is used to characterize the observed MCS environments. The ERA5 PW data has 0.25° × 0.25° horizontal resolution and hourly temporal resolution. We compared PW from ERA5 with the HOAPS product over the tropical ocean in Section 3 to quantify the uncertainty of the ERA5 data set.

2.2. DYAMOND Models

We evaluated most of the available DYAMOND models from both Summer and Winter phases with horizontal grid spacing at or finer than 5 km (see Table S1 in Supporting Information S1). Several models were excluded from our analysis due to errors either in the outputs or model implementations. Models were initialized on 1 August 2016 (Summer phase) and 20 January 2020 (Winter phase), respectively, and integrated for 40 days. In this study, only the simulation outputs for the last 30 days were used to avoid the model spin up period.

The model simulated OLR and surface precipitation were used to track MCS consistently with observations. All simulation outputs were first regridded to match the IMERG 0.1° × 0.1° horizontal grid between 60°S and 60°N. Variables with finer temporal resolution (e.g., precipitation, typically at 15-min intervals) were further averaged to hourly. The simulated hourly instantaneous OLR was empirically converted to T_b following Yang and Slingo (2001). The combined OLR and precipitation at hourly resolution are used for MCS tracking. Example snapshots of observed and simulated convective clouds (using $T_b < 241$ K as proxy) and precipitation from the Winter phase are shown in Figure 1. Visual inspection reveals that simulated deep convective clouds vary substantially in size and depth, particularly in the tropics. For example, some models predominantly simulated isolated deep convection (e.g., ICON, SCREAMv0) while some others produced very large and deep cloud clusters (e.g., MPAS, SCREAMv1). The characteristics of the simulated organized tropical deep convection are the subject of detailed investigation in Section 4.

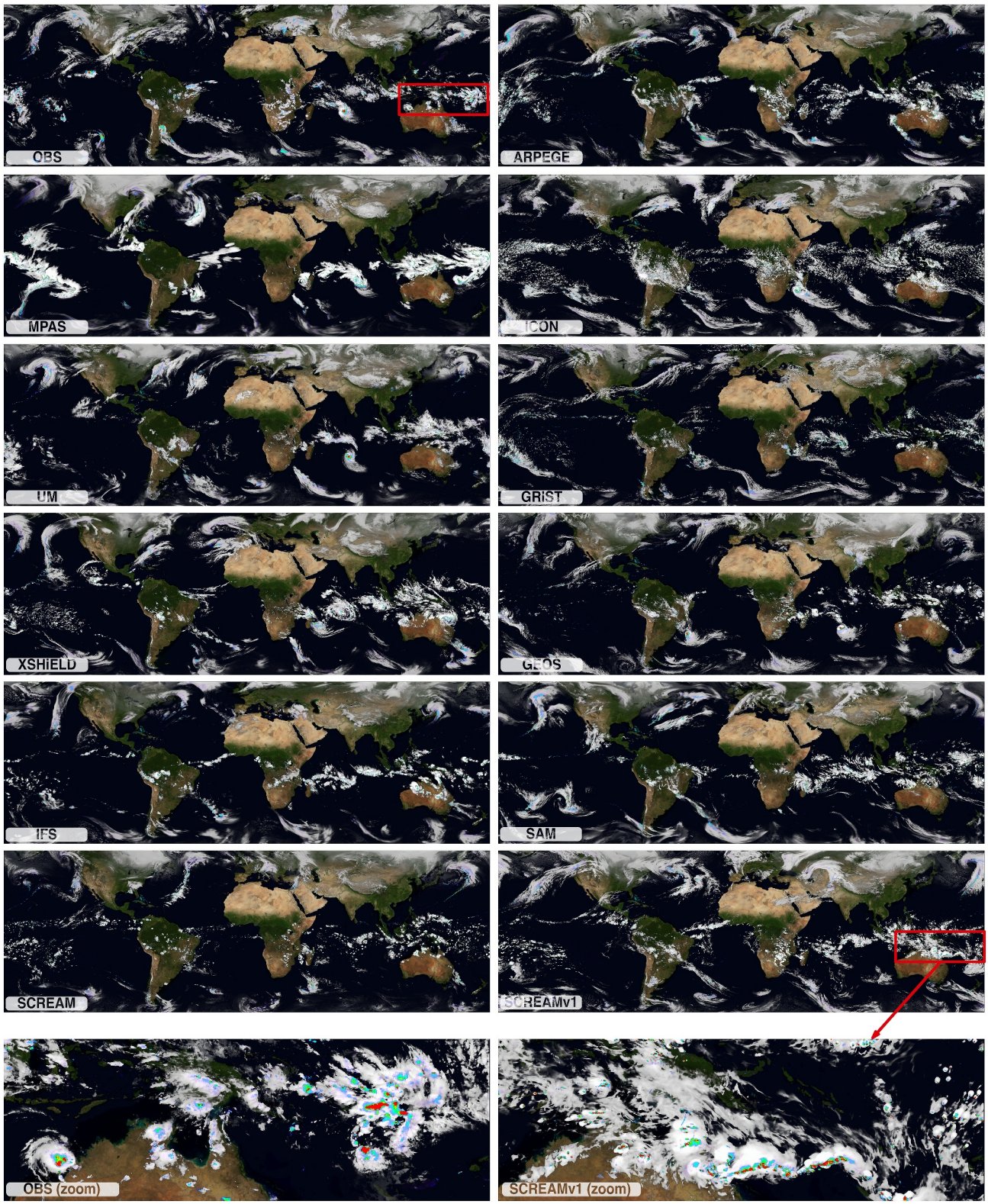

Figure 1. Example snapshots of cold clouds (T_b < 241K, white shading) and precipitation (color shading) from observations and DYAMOND winter models at 13 UTC on 5 February 2020 (15 days from model initialization). Brighter shading of the clouds denotes lower T_b associated with deep convection with higher cloud tops. A zoom-in region over the West Pacific (red boxes) from OBS and SCREAMv1 is shown in the bottom row. The example is meant for qualitative comparisons of cloud and precipitation features rather than their exact locations.

2.3. MCS Definition

In this study, MCS is defined using both T_b-identified cloud shield and precipitation characteristics following Prein et al. (2024):

1. Cloud shield area with $T_b < 241$ K must be at least 40,000 km^2 and persist for at least 4 continuous hours, and must contain $T_b < 225$ K at least once during those hours.
2. Must contain minimum peak precipitation of 10 mm hr^{-1} inside the cloud shield for 4 continuous hours.
3. Must reach a minimum rainfall volume of 20,000 km^2 mm h^{-1} (e.g., 100 km × 100 km × 2 mm hr^{-1}) at least once during the lifetime.

Most previous studies use T_b-only, typically ranging from 220 to 240 K, to track MCSs with varying area and duration criteria (Dong et al., 2021; Fiolleau & Roca, 2013; Laing & Fritsch, 1997; Machado et al., 1998; Núñez Ocasio et al., 2020a; Williams & Houze, 1987). Such a methodology works reasonably well in the tropics but becomes problematic in the extratropics where many types of large and long-lived cloud systems are often not entirely MCSs (e.g., extratropical cyclones, frontal systems/atmospheric rivers), although MCSs can be embedded in them. T_b-only methods can result in substantially higher MCS frequency than methods that consider both T_b and precipitation characteristics (Feng, Leung, et al., 2021; Kukulies et al., 2023; Leung et al., 2022). Therefore, using both cloud shield and precipitation is important to correctly identify MCSs across different climate regimes. The peak hourly rainfall rate threshold (10 mm hr^{-1} in criterion 2) serves to identify storms with heavy convective precipitation, while the rainfall volume threshold (20,000 km^2 mm h^{-1} in criterion 3) ensures the storm precipitation footprint achieves mesoscale dimensions, capturing the broader spatial extent with contributions from moderate stratiform rainfall. We note that tropical cyclones (TCs) can be aliased as MCS because the definition used here does not exclude TCs. However, TCs are relatively rare in the short period examined in this study.

The MCS definition used in the current study reflects a consensus among experts in the MCS tracking community, aiming to balance capturing the physical characteristics of MCSs with the simplicity required for application across different tracking methods. As expected, some sensitivity of identified MCS properties to threshold choices exists. For instance, higher T_b thresholds yield larger cloud areas and longer durations, while higher precipitation thresholds reduce the frequency of identified MCSs. The peak rain rate threshold (criterion 2) is highly sensitive to spatial resolution (~10 km in this study), whereas the rainfall volume threshold (criterion 3) is less affected by grid spacing. This study focuses on examining the sensitivity of MCS statistics to variations in tracker formulations. Future studies should further investigate how individual trackers respond to varying thresholds, including potential shifts in the distribution of T_b and precipitation under a changing climate.

2.4. MCS Feature Trackers

There are a total of 10 feature trackers participating in this study. Some trackers have specific functionality to track MCSs while others are general feature trackers that require post-processing to identify MCSs. Most trackers have separate procedures for spatial segmentation of cloud objects and temporal linking (Table 1). The most significant differences among the trackers lie in the spatial segmentation, which could result in substantial differences in the number of cloud objects before they are linked in time. Most trackers use spatial overlap between consecutive timesteps as a temporal linking method, though detailed implementations (e.g., fraction of overlap between time steps) could differ. A brief description of each tracker is provided in the Supporting Information S1. Five out of the 10 trackers were used in Prein et al. (2024) and further details can be found in that study.

Considering the diversity and different built-in functionality of the trackers, each of them was run with their preferred configuration to track convective clouds using T_b (various thresholds are provided in Table 1. Afterward, each tracker identified MCSs following the same criteria described in Section 2.3. Subsequently, masks that contain unique MCSs were recorded from each tracker at the 0.1° × 0.1° grid for each hour, which is then standardized to a common netCDF file for each DYAMOND model and observations. Lastly, MCS lifecycle statistics such as size, duration, and precipitation characteristics were derived from the standardized MCS mask files for further analysis.

3. Evaluation of Simulated Clouds, Precipitation, and Moisture

We first compare the overall distribution of the simulated T_b and precipitation in the tropics against observations to understand how these variables may affect MCS tracking. Figure 2 shows that the PDF of the simulated T_b from

Table 1
Key Characteristics of MCS Tracking Algorithms and Thresholds Used in the Current Intercomparison Study

	Spatial segmentation	Temporal linking method	Merging/splitting treatment	Programming language
PyFLEXTRKR (Feng, Hardin, et al., 2023)	2D detect-and-spread cold core: $T_b < 225$ K, spread to $T_b < 241$ K	Overlap (0.5)	Yes	Python
TIMPS (Rajagopal et al., 2023; Russell et al., 2024)	2D detect-and-spread multi-thresholds with outer $T_b < 241$ K	Overlap (max overlap)	No	Python, C++
simpleTrack (Stein et al., 2015; Crook et al., 2019)	Threshold-connectivity $T_b < 241$ K	Overlap (0.6)	Yes	Python
KFyAO (Dong et al., 2021; Huang et al., 2018)	Threshold-connectivity $T_b < 241$ K	Overlap and Kalman filter (0.15)	No	Matlab
tobac (Heikenfeld et al., 2019; Sokolowsky et al., 2024)	Multiple threshold-connectivity $T_b < 241$ K, 233 K, 225 K	Parcels, propagation speed	Yes	Python
DL + TempestExtremes (Ronneberger et al., 2015; Ullrich et al., 2021; and this study)	Convolutional neural network (U-Net) Trained with PyFLEXTRKR data	Overlap (uses TempestExtremes) (0.5)	No	Python, C++
TAMS (Núñez Ocasio & Moon, 2024; Núñez Ocasio et al., 2020a)	Contour (polygon) $T_b < 241$ K	Overlap (0.5)	No	Python
ATRACKCS (Robledo et al., 2024)	Contour (polygon) $T_b < 241$ K	Overlap (0.25)	No	Python
MOAAP (Prein et al., 2023)	2D water-shedding	Overlap (one gridcell)	Yes	Python
TOOCAN (Fiolleau & Roca, 2013)	3D region growing (space + time) T_b dilation from 190 to 241 K, +2 K/step		No	C

Note. Spatial segmentation refers to how a cloud field is segmented into individual cloud objects, and temporal linking refers to the procedure of connecting cloud objects in time. Merging/splitting treatment refers to whether an MCS starting as a split from an existing system or ending by merging with another system is explicitly tagged, hence separating them from "naturally" initiating or dissipating systems.

most DYAMOND models agrees reasonably well with observations, particularly for $T_b < 241$ K (blue shaded region in Figures 2a and 2b), which is the T_b range to identify deep convective clouds. Several models have noticeable deviations from observations, for example, for the summer period, UM and ARPEGE underestimate low T_b frequency; during the winter period, MPAS and ICON overestimate low T_b frequency while UM and

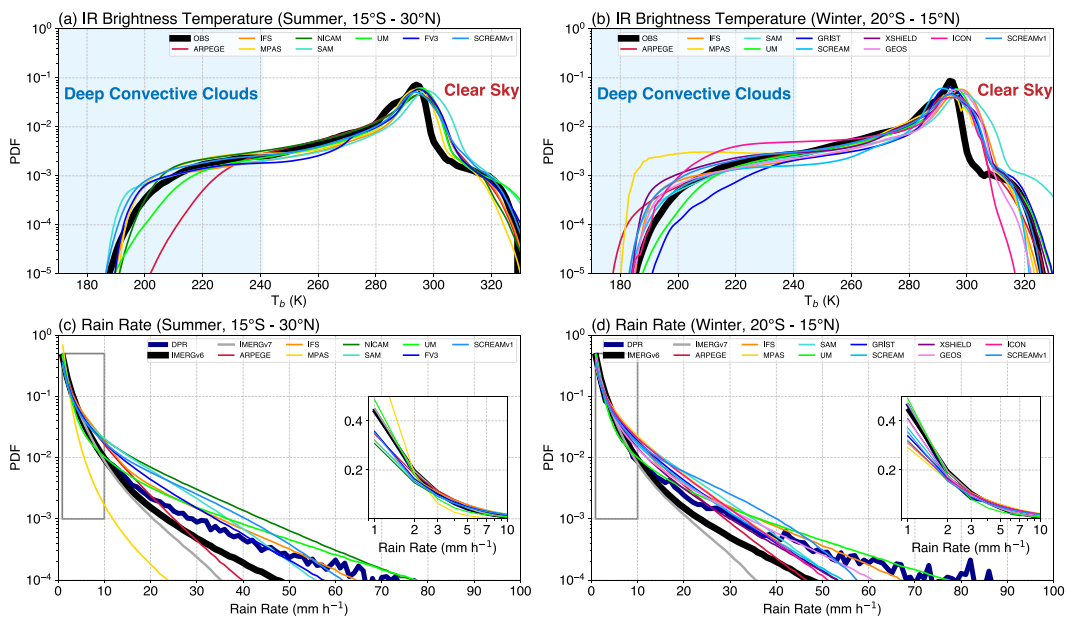

Figure 2. PDFs of (a, b) infrared T_b and (c, d) hourly rain rate in the tropics for summer (15°S–30°N) and winter (20°S–15°N). In (c, d), thick black lines are IMERG v6, gray lines are IMERG v7, and thick dark blue lines are GPM DPR retrieval. The blue shaded region in (a, b) is the T_b range for tracking deep convective clouds. The insets in (c, d) show PDFs of weak rain rates (gray box) between 1–10 mm hr^{-1}.

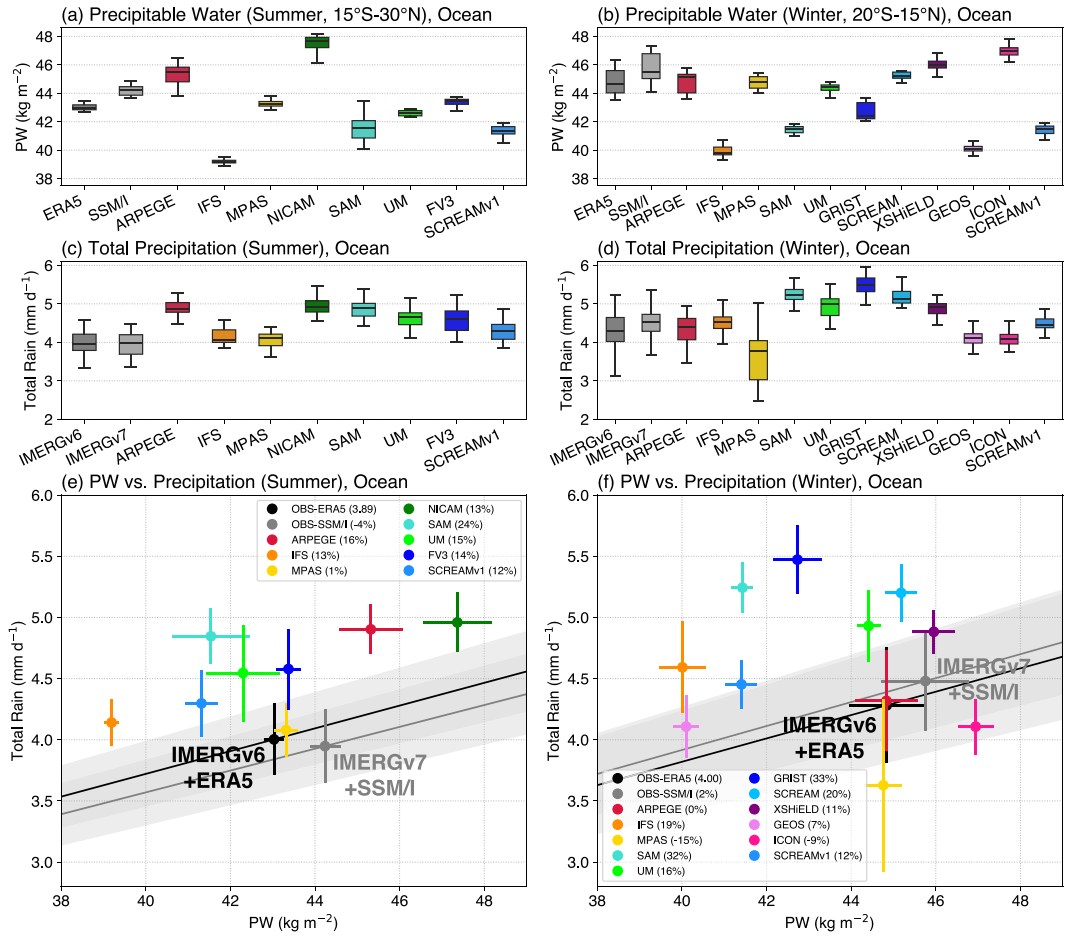

Figure 3. Box-whisker plots of daily mean (a, b) precipitable water (PW), (c, d) total precipitation over tropical ocean, and (e, f) mean tropical ocean PW versus precipitation. In (a–d), boxes are interquartile ranges and horizontal lines are median values, in (e, f), circles are mean values and lines are standard deviations. The diagonal lines represent constant precipitation-to-PW ratios from the two observational data sets, and the shaded areas encompass the range of observed standard deviations. Legends in (e, f) show mean values of the ratio between precipitation and PW for OBS, and relative difference (%) between other sources and OBS. Only the last 30 days of the simulations are included.

GRIST underestimate low T_b frequency. These biases in T_b imply that MCS frequencies from those models may have larger discrepancies with observations.

Both versions of IMERG have a similar frequency of light rain rates (<3 mm hr^{-1}) but a lower frequency of heavy rain rates (>20 mm hr^{-1}) compared to DPR, with IMERG V7 deviating further than V6 (Figures 2c and 2d). These results are consistent with the literature that IMERG underestimates heavy precipitation and vice versa for light precipitation (Ayat et al., 2021; Cui et al., 2020; Rajagopal et al., 2021; Zhang et al., 2021). We note that the hourly rain rates from DPR may represent an instantaneous estimate from the GPM core satellite overpass, while those from IMERG represent rain rates averaged over an hour. Therefore, we may expect that the rain rates from IMERG should be lower than those from DPR, although the exact differences are difficult to quantify. The rain rate PDFs from DPR and two versions of IMERG should be treated as a range of observational uncertainties. The simulated rain rates exhibit a larger spread among the models than the simulated T_b, with most models overestimating observed heavy rain rate frequencies between 10 and 40 mm hr^{-1}, while underestimating light rain rate frequencies. The large spread in simulated precipitation rate may also affect MCS identification.

We further compare daily mean PW and accumulated precipitation over the tropical ocean to assess the model precipitation sensitivity to environmental moisture. Global maps of time-mean simulated PW and precipitation differences with observations are provided in Figures S2 and S3 in Supporting Information S1. Compared to the HOAPS satellite retrieval, ERA5 has a slightly lower PW in both seasons by ∼1 kg m^{-2} (Figures 3a and 3b).

DYAMOND models have widespread PW ranging from overestimation by up to 3 kg m^{-2} (e.g., NICAM) to underestimation by 3–5 kg m^{-2} (IFS, GEOS, SCREAMv1, SAM). In contrast, models produce similar or higher precipitation than both versions of IMERG, with the lone exception being MPAS in Winter (Figures 3c and 3d). As a result, most models (13 out of 19) have higher precipitation-to-PW ratios than observations (points above the shaded diagonal lines in Figures 3e and 3f), with biases ranging from ∼10% to 33%. These results suggest most DYAMOND models overestimate the sensitivity of precipitation to environmental moisture over the tropical ocean.

4. MCS Tracking Intercomparison

4.1. MCS Tracker Differences in Observations

Before assessing the model representation of MCSs and their sensitivity to tracker formulations, we first compare MCS tracking in observations among the participated trackers. An example of MCS tracks from observations over the western Pacific during summer is shown in Figure 4. The example contains several very large cloud clusters (100s–1,000 km) with complex evolution that are common over tropical oceans. Substantially different MCS tracks are seen across the trackers (Figures 4a–4j). Some trackers produce a few very large MCSs (e.g., simpeTrack, TAMS, DL, ATRACKCS). In contrast, TOOCAN and MOAAP produce numerous small MCSs, while the rest of the trackers produce MCSs of intermediate size. This difference among trackers has been highlighted by Prein et al. (2024) (Five of the six trackers in that work are also used in the current study) and was attributed to several factors, including differences in the spatial segmentation procedures (Table 1), and treatment of merging and splitting. This example highlights that the *number of MCS tracks* is highly sensitive to tracker formulation, and it may not be a robust metric to evaluate model simulations, despite its simplicity.

Rather than counting the number of MCS tracks, we make use of the MCS mask outputs from the trackers to compute the hour counts for each grid point (the sum of all MCS masks covering the grid point over a period), regardless of the number of individual MCS tracks. This metric can be converted to MCS frequency of occurrence (hereafter, MCS frequency) by simply dividing by the total number of hours over the period. Results show that MCS frequency is more comparable among the trackers than the number of MCS tracks (Figures 4k–4t) because it eliminates the ambiguity in determining between one large MCS and many small MCSs.

Next, we compare observed MCS contributions to total precipitation among the trackers. Figures 5a and 5b shows the global distribution from PyFLEXTRKR for the two DYAMOND phases and Figures 5c–5f shows the regional, tropical, and global mean values from each tracker. Global maps from individual trackers are provided in Figure S4 in Supporting Information S1. Large tracker differences (up to a factor of 2) in MCS contribution to total precipitation are seen over most regions for both seasons. Four trackers produce estimates of ∼60% or higher over tropical ocean (PyFLEXTRKR, TAMS, simpleTrack, DL) while two trackers' (TOOCAN, tobac) estimates are ∼40% or less. Inter-tracker variability in MCS cloud frequency is also quite large (up to a factor of 3, Figure S5 in Supporting Information S1), consistent with MCS rainfall contribution. Similar contrasts are found over tropical land and global averages.

The lower MCS frequency and precipitation fraction in TOOCAN, tobac, and MOAAP trackers are potentially due to their more complex spatial segmentation and temporal tracking techniques, such as water-shedding and parcel tracking (Table 1), which segment large cloud clusters into smaller, shorter-lived systems. A fraction of these smaller systems does not meet the MCS criteria and is therefore excluded, whereas simpler threshold-connectivity or contour-based methods used by other trackers often retain them as part of larger clusters. These methodological differences contribute to the observed discrepancies in MCS frequency and precipitation fractions.

Such large differences highlight that the frequently referenced statement "MCSs contribute to over 50% of total precipitation" is highly sensitive to tracker formulations, especially across major MCS production regions. Our results are qualitatively consistent with Prein et al. (2024) for South America. Caution is therefore needed when using a particular tracker to quantify MCS frequency and rainfall contributions. On the other hand, the consistent discrepancies among trackers across geographic regions and seasons suggest that regional tracker comparisons, which are computationally lighter and easier for process level investigation, may be broadly applicable globally.

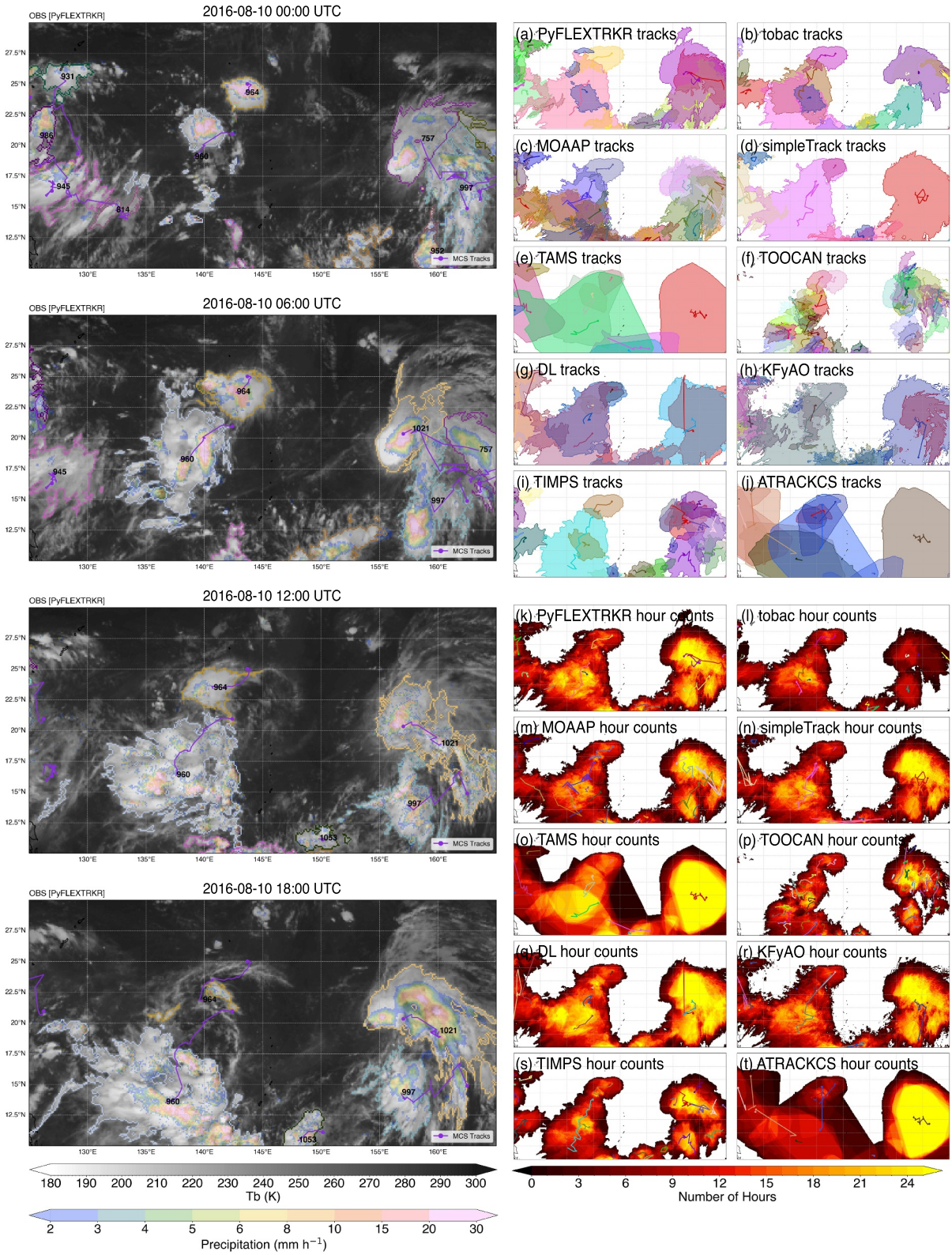

Figure 4. Examples of MCS tracker results during a single day for observations over the west Pacific (125°E–165°E, 10°N–30°N). The left column shows the 6-hourly evolution of the cloud shield (gray shading), precipitation (color shading), and the outlines of the detected MCSs based on results from PyFLEXTRKR. The purple circles show the MCS initiation points, and the purple lines show the MCS tracks. Panels (a–j) show the tracks and swaths of individual MCSs, with different colors indicating individual MCSs from each tracker. Panels (k–t) show the hour counts of MCS masks from each tracker.

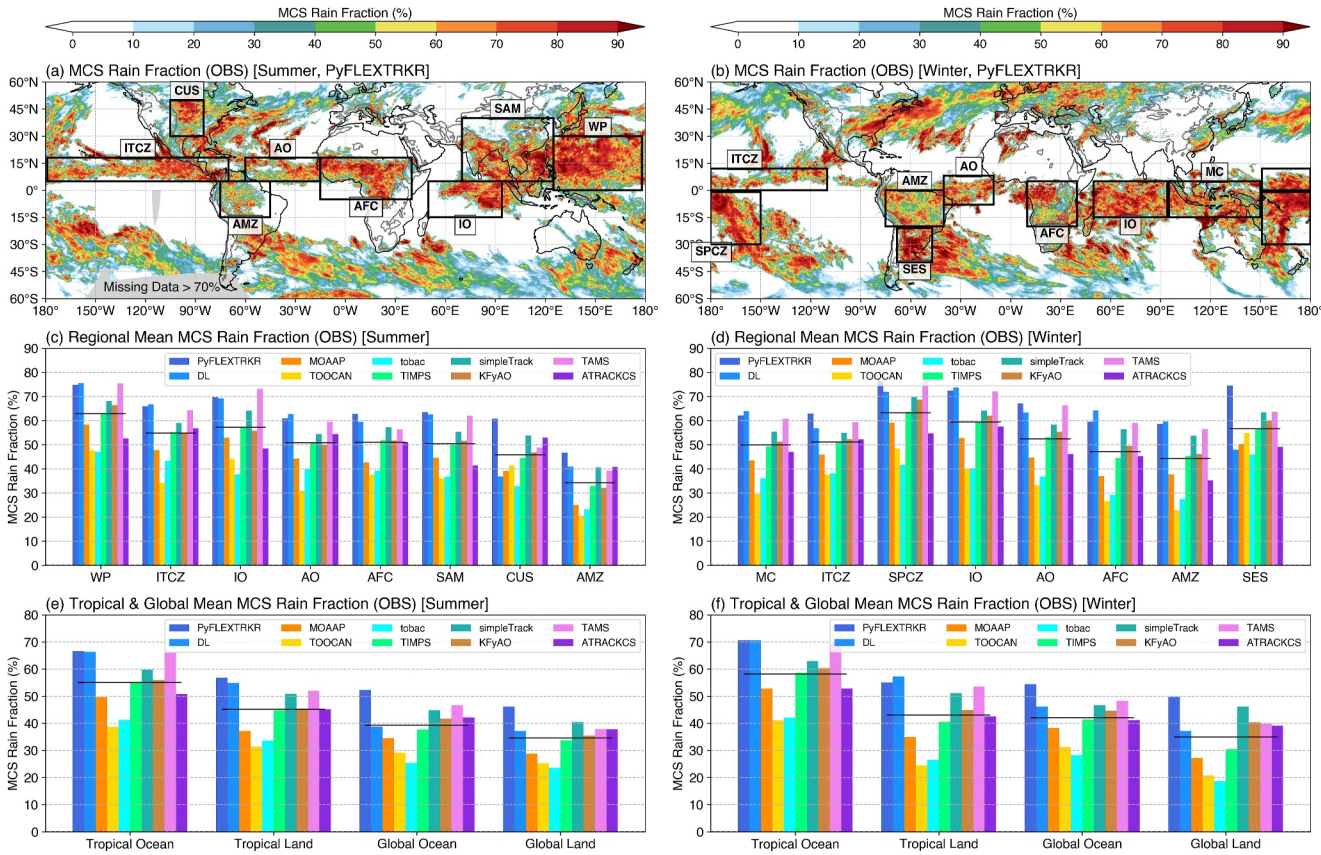

Figure 5. Comparisons of observed MCS contribution to total precipitation among the trackers. Global distribution of MCS contribution to total precipitation for (a) summer and (b) winter based on results from PyFLEXTRKR. Regional mean values of the trackers are shown in (c) summer and (d) winter, and global (60°S–60°N) mean values are shown in (e) summer and (f) winter. Black boxes show the regions in (a, c) WP (West Pacific), ITCZ (Intertropical Convergence Zone), IO (Indian Ocean), AO (Atlantic Ocean), AFC (Africa), SAM (South Asian Monsoon), CUS (Central United States), AMZ (Amazon); and in (b, d): MC (Maritime Continent), ITCZ, SPCZ (South Pacific Convergence Zone), IO, AO, AFC, AMZ, and SES (Southeast South America). Gray shaded areas in (a) are regions with missing T_b data >70%. Black horizontal lines in each group of bars in (c–f) denote averages from all trackers for that region. Mean values are calculated by dividing the MCS precipitation amount by the total precipitation amount in each region.

4.2. MCS Frequency and Precipitation

In this subsection, we evaluate simulated MCS frequency and precipitation over tropical land and ocean separately. The tropical belt was selected for each season based on the latitude bands where MCSs most frequently occur in observations (Figures 6a and 6b). The simulated tropical mean MCS frequency biases are shown by individual models (rows) and trackers (columns) as heat maps, with multi-model mean biases shown in the top row (Figures 6c and 6d). Absolute values from each tracker are provided in Figures S6 and S7 in Supporting Information S1. The skill of simulating tropical MCS frequency varies widely among DYAMOND models, with some models substantially underestimating observed values over both land and ocean by up to 90% (e.g., ARPEGE, SCREAM, GRIST) while others overestimate by up to 300% (e.g., SCREAMv1, ICON, MPAS). The model biases are generally consistent among the trackers (similar color/values across a row) with a few exceptions (MPAS, ICON). The opposite behaviors in SCREAMv1 and SCREAM potentially arise from changes in the sub-grid enhancement of autoconversion and accretion, which were disabled in v1 to reduce the tendency to convert cloud water to rain too quickly (i.e., too much precipitating shallow convection), and the transition from a relative humidity-based to a mixing-ratio-based ice cloud fraction scheme in v1 that likely increases upper-level ice clouds (Donahue et al., 2024), impacting MCS tracking. Multi-model mean values show that DYAMOND models are generally skillful in simulating tropical mean MCS frequency, with biases ranging from −11% to 30% over land and from −18% to 20% over ocean, though regional differences are larger (not shown).

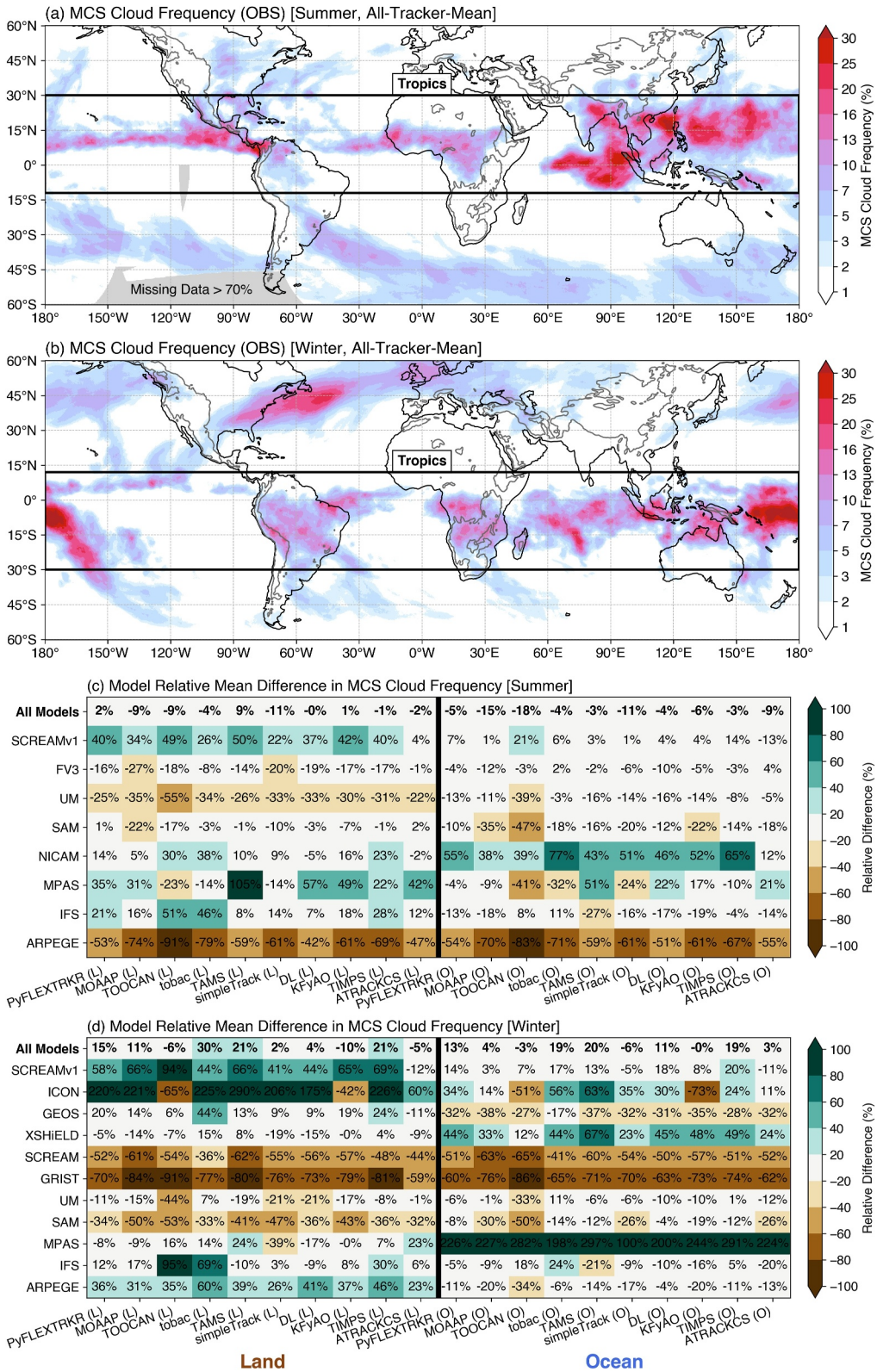

Figure 6. Comparisons of MCS frequency based on cloud shield occurrence. (a, b) Observed MCS cloud shield frequency based on average from all trackers, (c, d) relative difference (%) between simulations and observations for each tracker (*x*-axis) and for each model (*y*-axis), with the top row showing multi-model mean differences. Regions of the averaging results are shown in the black box in (a, b). Gray shaded areas in (a) are regions with missing T_b data >70%. In (c, d) the left half is for land MCS and the right half is for ocean MCS. Relative difference is computed by (simulation–observation)/observation.

Simulated MCS precipitation amount biases are more consistent than MCS frequency among the DYAMOND models and trackers (Figure 7). Nearly all models underestimate MCS precipitation amounts over the tropical ocean with a lesser underestimate by a lesser number of models over tropical land (exceptions over land being SCREAMv1, ICON, and IFS). ICON (land) and MPAS (ocean) biases differ in sign among trackers. Visual inspections reveal that ICON and MPAS simulated excessive extents of cold cloud tops with more numerous embedded small convective precipitation cores (∼10s km), morphologies that differ significantly from observations (i.e., typically a few large precipitation features spanning ∼100s km underneath a cold cloud shield, example animations are provided in the Supporting Information S1). These conditions cause the most disagreements in MCS identification among trackers. Further research is needed to better understand the cause of these morphological differences for individual models, which is beyond the scope of this study. Multi-model mean MCS precipitation amount biases among trackers range from −23% to 4% over land and from −45% to −17% over ocean.

Simulated MCS contributions to total precipitation show the most consistent difference among DYAMOND models and trackers (Figure 8). Nearly all models show lower MCS contribution than observations in both seasons except for MPAS over the ocean during winter. The multi-model mean biases are slightly smaller over land (−20% to −6%) than over ocean (−28% to −16%), although several models (ARPEGE, GRIST, SCREAM, ICON) have relatively larger biases (−30% to −50%). These results confirm the finding from Feng, Leung, et al. (2023) that MCS precipitation is generally underestimated in DYAMOND models. The agreement among the trackers suggests that although DYAMOND multi-model means are skillful in simulating MCS frequency, their associated precipitation is underestimated and more so over the ocean than land. Additionally, some individual models can produce much larger differences with observations and can be further improved.

4.3. MCS Diurnal Cycle

A recent study showed that the diurnal cycle of precipitation is better simulated in DYAMOND models than HighResMIP models, which was attributed to a better representation of MCSs (Song et al., 2024). Here, we compare the diurnal cycles of MCS initiation and mature stages (see figure caption for definitions) over tropical land among the trackers (Figure 9). Results show that most DYAMOND models can capture the diurnal cycle of the MCS mature phase well for both seasons, with peak amplitude occurring between late afternoon to early evening hours, consistent with the peak precipitation timing over land (Song et al., 2024). Most trackers show larger amplitudes of the simulated MCS mature phase than observations in winter, while better agreement is found in summer except for the ARPEGE model.

The diurnal cycle of MCS initiation differs more among the trackers. Several trackers show peak MCS initiation occurring around 13 LT (PyFLEXTRKR, TOOCAN, tobac, simpleTrack) while others peak ∼3 hr later around 16 LT (MOAAP, TAMS, DL, KFyAO). This may be partly related to the trackers in the latter group identifying MCSs only when they reach 40,000 km^2, while those in the former group start tracking MCSs from convection initiation (Prein et al., 2024). The diurnal cycle of oceanic MCSs is much weaker than those over land, and most DYAMOND models also captured the mature phase from observations well (Figure S8 in Supporting Information S1). Overall, the MCS diurnal cycle is reasonably simulated in most of the DYAMOND models, particularly their mature phase that corresponds to when most of the precipitation is produced (Song et al., 2024). Nuance among the trackers is whether convection initiation is included as part of the MCS lifecycle, which affects the timing of MCS initiation.

4.4. MCS Characteristics

In this subsection, we examine the models' skill in simulating the cloud and precipitation characteristics of tropical MCSs during their lifecycle. We performed lifecycle composite analysis for MCSs with median lifetimes among all models for each tracker. Figure 10 shows composites of tropical MCSs over land for both seasons. Similar composites for tropical oceanic MCSs are provided in Figure S9 in Supporting Information S1. A typical MCS lifecycle is denoted by an inverted V shaped cloud shield area and an asymmetric check mark (tick) shaped minimum T_b peaking earlier in the life cycle (Figures 10a, 10b, 10f, and 10g), corresponding to the growth phase (area expansion and cloud-top deepening) and the decay phase (area contraction and cloud-top lowering). Most DYAMOND models captured this canonical evolution, although the magnitude and rate of change differ among the trackers.

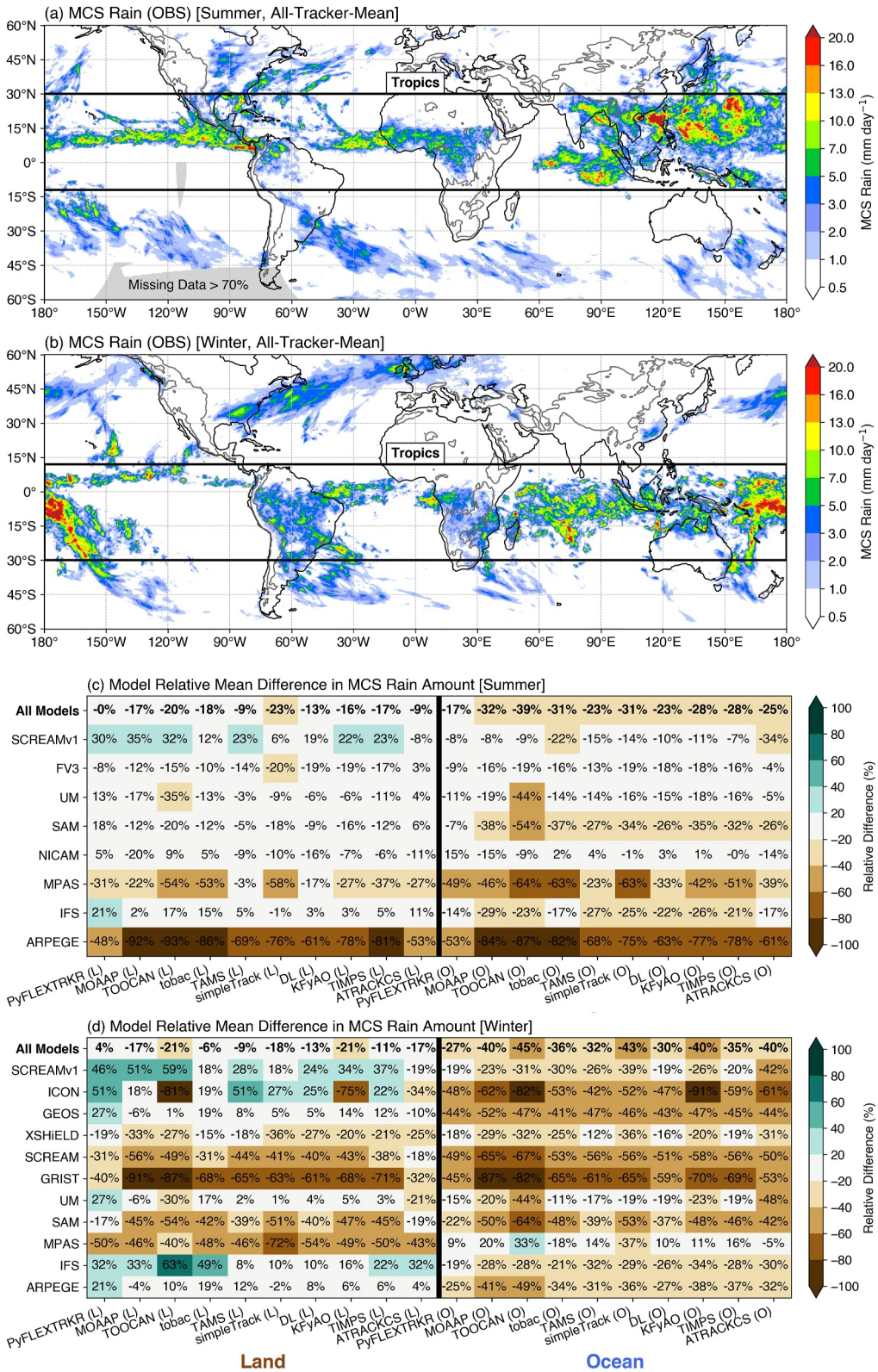

Figure 7. Same as Figure 6 except for MCS precipitation amount.

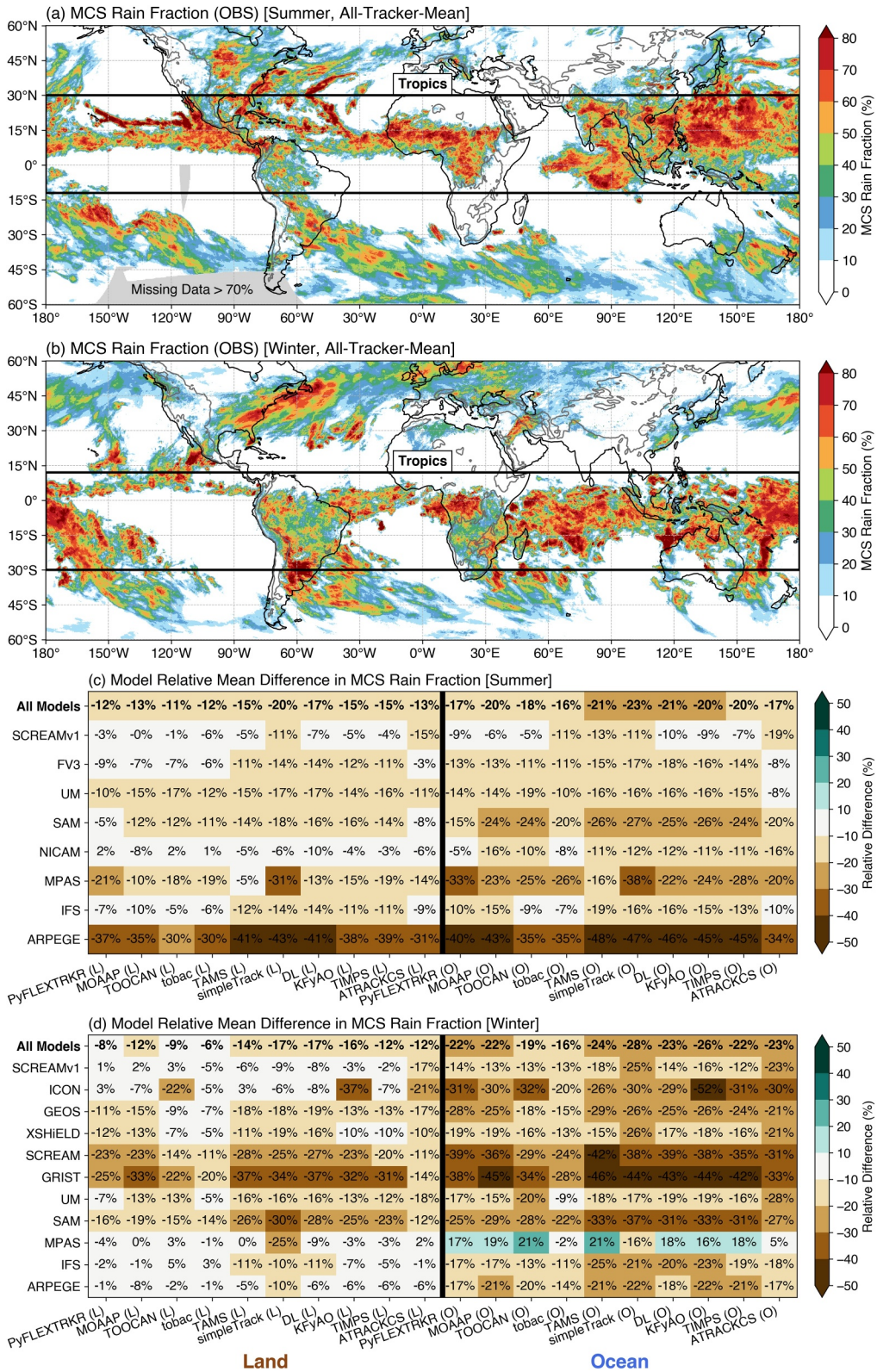

Figure 8. Same as Figure 6 except for MCS contribution to total precipitation amount. Relative difference is computed by (simulation–observation).

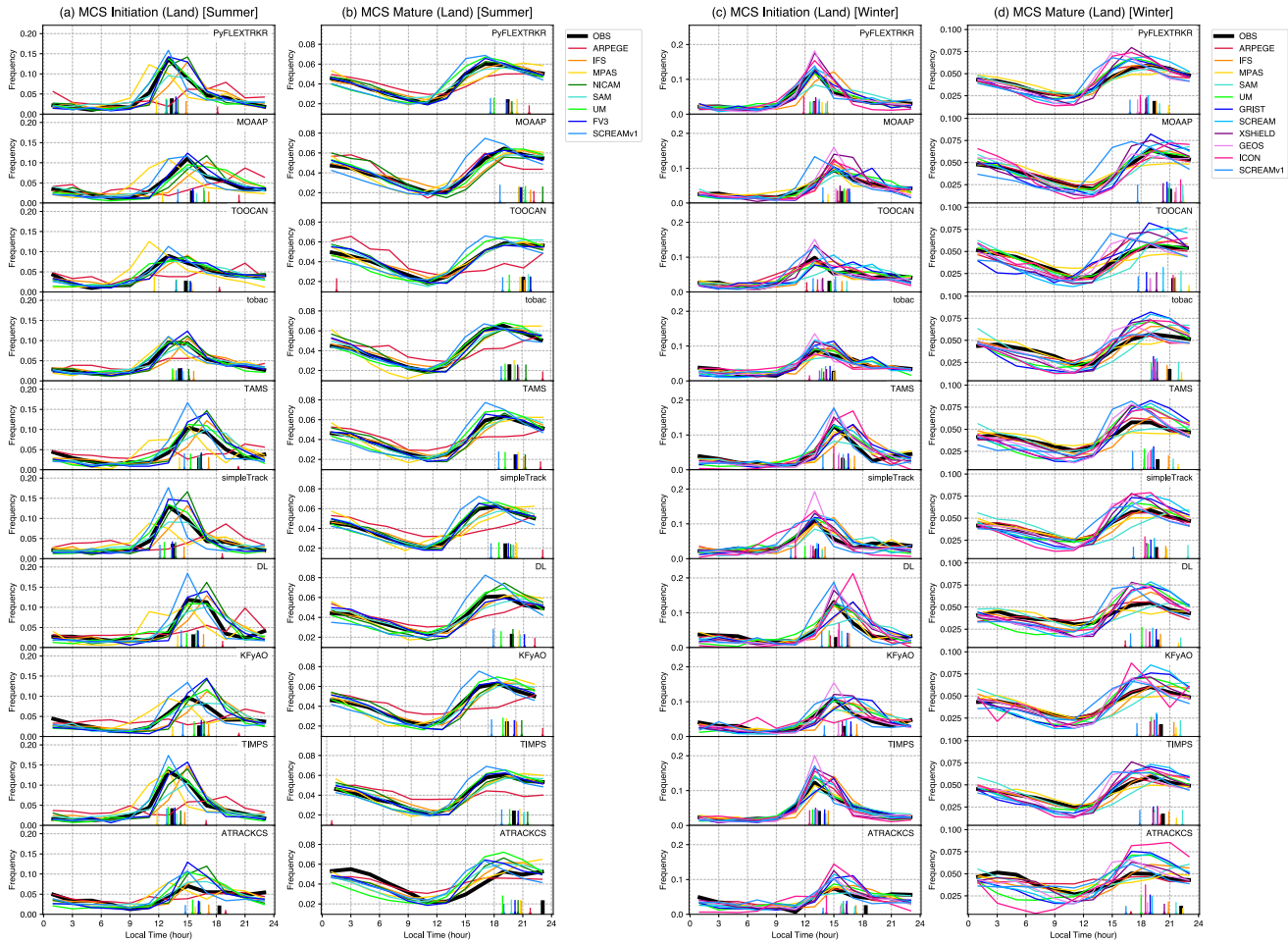

Figure 9. Diurnal cycle of tropical MCS frequency over land at initiation (a, c) [first hour an MCS is detected] and mature stages (b, d) [hours when the largest precipitation feature major axis length in the MCS > 100 km] for summer (a, b) and winter (c, d). Each row shows results from a tracker. The diurnal cycle amplitude and phase are shown as vertical bars in each panel, the height of the bars denotes the amplitude (taller bars have higher amplitude). Amplitude and phase are calculated using the first harmonic of the Fourier transform applied to the diurnal cycle frequency signal (Wallace, 1975). MCSs with lifetime-mean precipitation area fraction >70% over land are considered land MCS.

The model-observation differences in cloud shield properties are smallest in TOOCAN compared to other trackers. However, some models (e.g., MPAS, SCREAMv1, ICON) exhibit large differences in cloud shield area among trackers due to the spatial variability of simulated OLR/T_b. Models such as MPAS and ICON simulate extensive regions of cold clouds (T_b < 241 K), which cause simple threshold-based trackers (e.g., simpleTrack, KFyAO, TAMS) to identify large systems, while trackers employing more sophisticated approaches (e.g., TOOCAN, MOAPP, PyFLEXTRKR) segment these into smaller systems. These discrepancies reflect model biases in simulating spatial patterns of convective organization, as illustrated by tracking animations provided in the Movies S3–S6.

In contrast, the simulated MCS precipitation feature (PF, defined as the largest contiguous region with rain rate >2 mm hr^{-1}) area and mean rain rate show notably larger differences when compared with either version of IMERG (Figures 10c, 10d, 10h, and 10i). Following Feng, Leung, et al. (2021), this rain rate threshold was chosen because it results in closer agreement in PF area and mean rain rate with radar observations over the United States. Most models have much smaller PF area and much stronger PF mean rain rate (up to a factor of 2) throughout the MCS lifecycle, particularly during the mature phase for PF area and during the initiation phase for mean rain rate. These relative biases are consistent among the trackers, even though the absolute magnitudes differ. The underestimated PF area and overestimated mean rain rate compensate one another to produce a somewhat comparable simulated MCS rainfall volume to observed (Figures 10e and 10j).

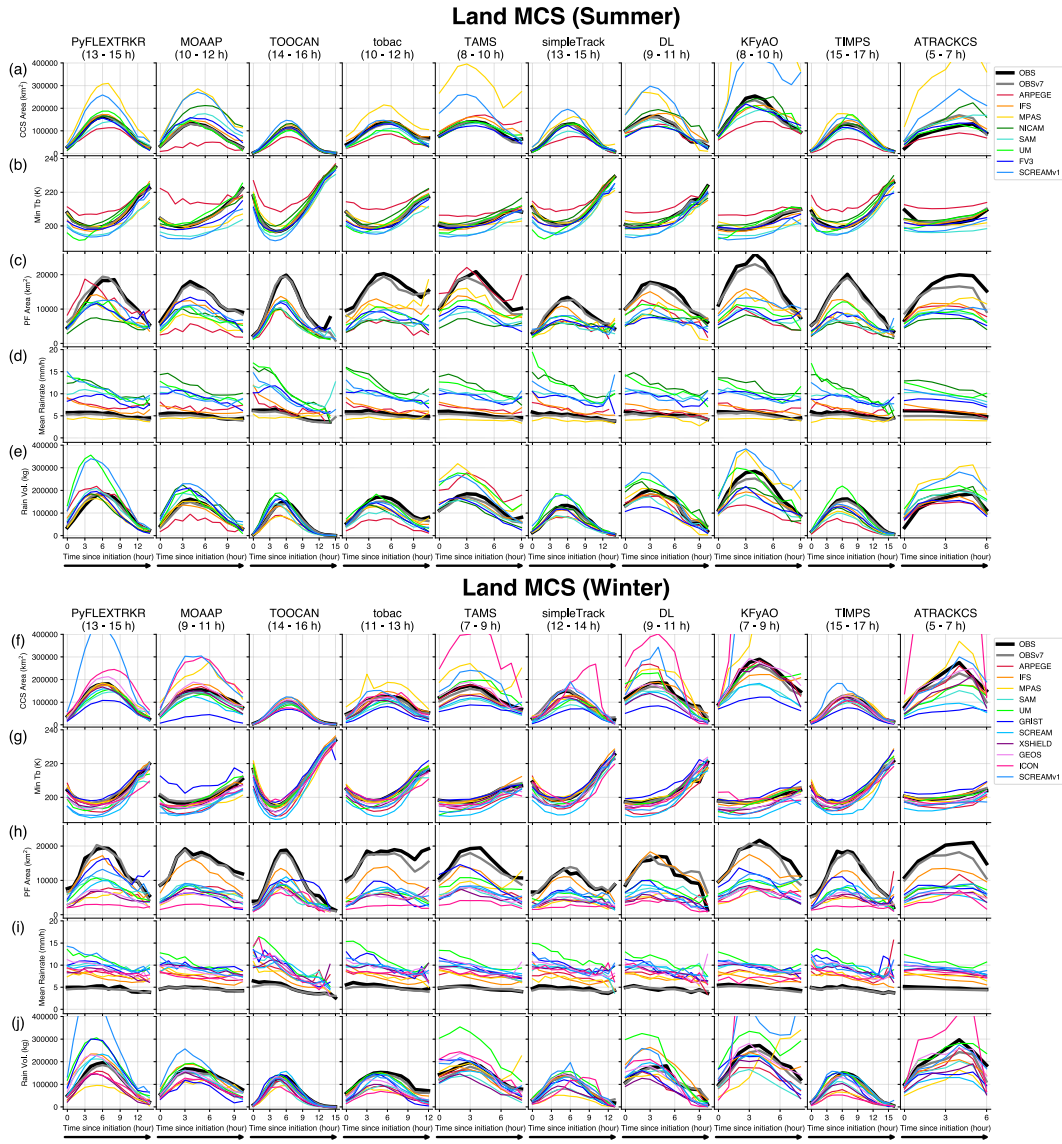

Figure 10. Composite lifecycle evolution for tropical land MCSs with median lifetimes: (a, f) cloud shield area, (b, g) minimum Tb, (c, h) largest PF area, (d, i) largest PF mean rain rate, and (e, j) total rain volume (includes all precipitation within MCS cloud shield). Each column shows results from a tracker. Thick black lines are observations with IMERG v6, gray lines are observations with IMERG v7. The median MCS lifetimes are shown above each column and were calculated among all simulations for each tracker. To exclude merge/split systems, MCSs with maximum cloud shield area occurring within the first or last 15% of the lifetime were removed.

To quantify model-observation differences, we show the relative biases of the median values of six key MCS characteristics for each tracker (Figure 11). Distributions of these MCS characteristics from select trackers are provided in Figures S10 and S11 in Supporting Information S1. In general, models simulate MCS lifetime and cloud shield (top rows in Figure 11) better than precipitation (bottom rows in Figure 11), as evident by the lower relative biases (lighter shading). Most trackers agree that most models have shorter-lived and smaller MCSs with lesser lifetime-total rainfall volume. All trackers show that models consistently produce too intense precipitation (Figures 11e and 11k) with overestimated heavy rainfall contribution to total rain volume (Figures 11f and 11l), particularly over land where most models are biased high by a factor of 2–3. MPAS is an outlier with much larger MCSs in both seasons and much weaker precipitation intensity in summer.

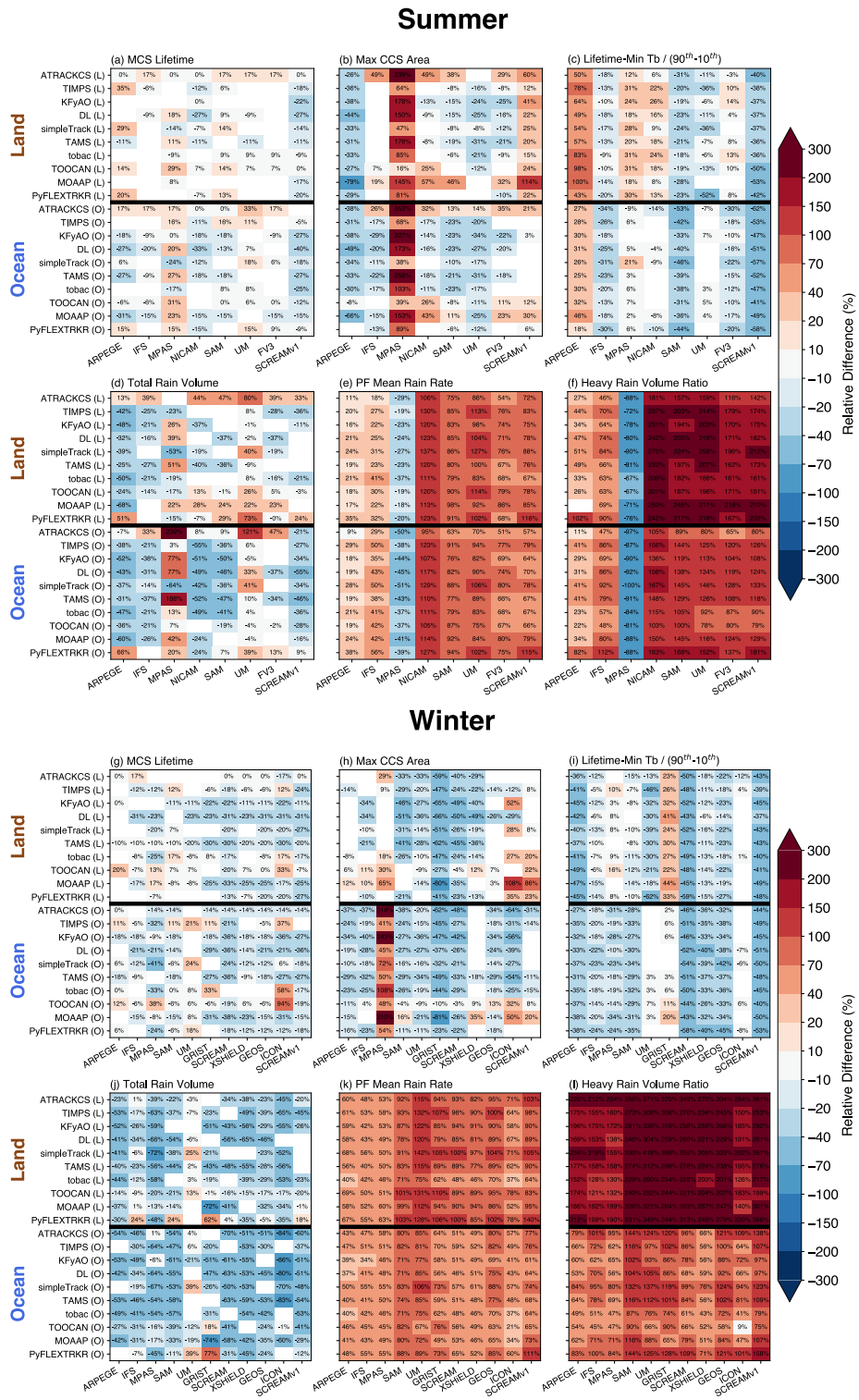

Figure 11. Relative difference in the median values of MCS properties between simulations and observations for each model (*x*-axis) and each tracker (*y*-axis). Top two rows are summer and bottom two rows are winter showing (a, g) MCS lifetime, (b, h) lifetime-maximum cold cloud shield area, (c, i) lifetime-minimum T_b, (d, j) lifetime-total rain volume, (e, k) PF mean rain rate, (f, l) heavy rain (>10 mm hr^{-1}) to total rain volume ratio. Land (ocean) MCS results are shown on the top (bottom) half of each panel. Relative difference is computed by (simulation–observation)/observation, except for T_b: (simulation–observation)/(90th–10th percentile in observation, typically 13–15 K). Only statistically significant differences (Mann-Whitney U rank test p value < 0.05) are shown.

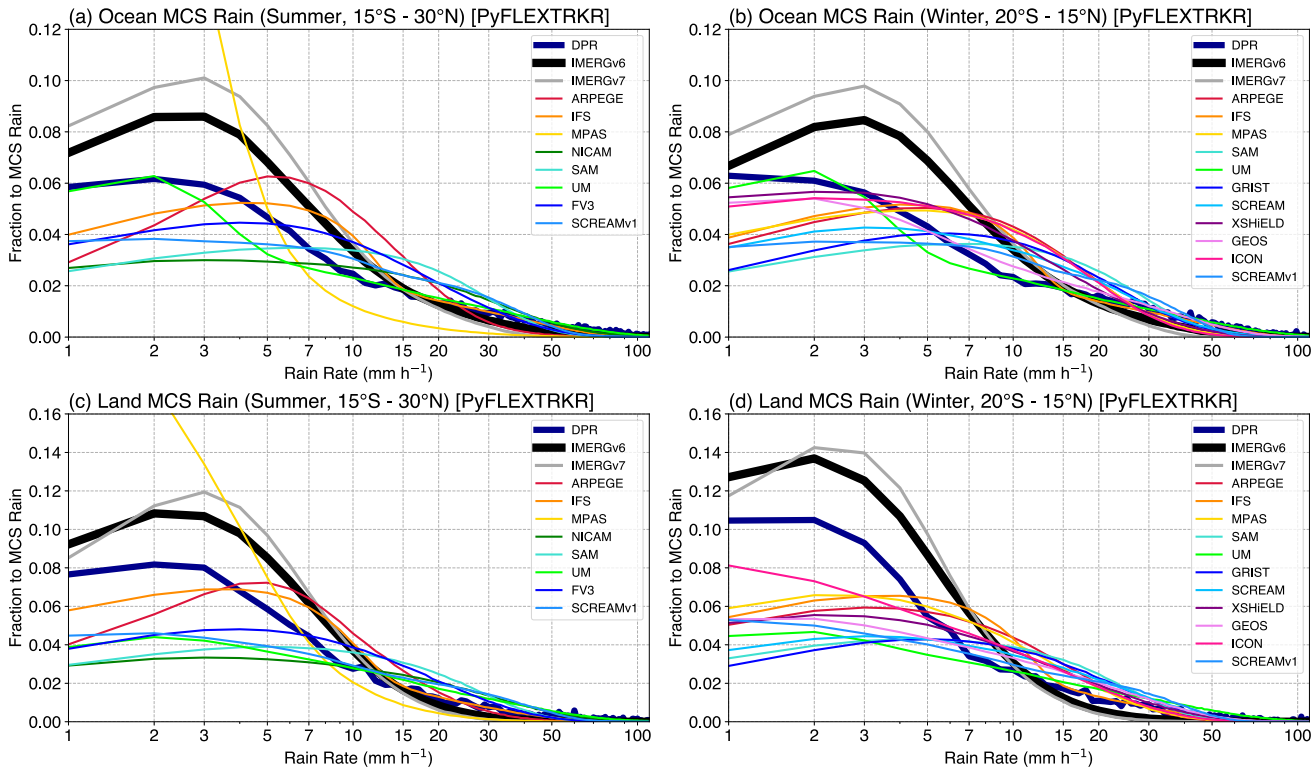

Figure 12. Contribution of hourly, 0.1° rain rates to tropical MCS rainfall amounts based on results from PyFLEXTRKR for (a, b) oceanic MCS and (c, d) land MCS. Contributions from DPR (thick dark blue lines) are calculated by collocating GPM DPR overpasses with MCS masks based on T_b + IMERG v6.

We further illustrate the model biases in simulating the contribution to MCS precipitation by hourly rain rate in Figure 12. The corresponding PDFs of MCS hourly rain rates are shown in Figure S12 in Supporting Information S1. Results show that weak rain rates (2–4 mm h^{-1}, typically associated with stratiform rain) contribute most to observed tropical MCS total rainfall volume both over ocean and land, with a sharp drop off toward more intense convective rain rates (>10–15 mm hr^{-1}). Compared to DPR, both IMERG versions exhibit a higher contribution from weak stratiform rain and a lesser contribution from intense convective rain. Despite the uncertainty in the IMERG product, most DYAMOND models show biases of the contribution toward heavier rain rates, with peak contributions between 5 and 10 mm hr^{-1}. The underestimation of weak rain rate contribution is consistently larger over land than over ocean, suggesting that stratiform rain biases are more pronounced over land. Results using other tracker outputs are qualitatively consistent with those in Figure 12 (not shown). Sensitivity tests involving coarsening the precipitation data to 0.25° × 0.25° (~28 × 28 km)—closer to the footprint of the microwave sensors used in the GPM IMERG data set—yield qualitatively consistent results (Figures S13 and S14 in Supporting Information S1).These findings, encompassing most available DYAMOND models, reaffirm the conclusions of Feng, Leung, et al. (2023) that models overestimate convective rainfall intensity and underestimate stratiform rainfall area and amount.

4.5. MCS Sensitivity to Moisture

We now turn toward examining the potential causes of the model biases in simulating MCS precipitation. Recall in Section 3 that most models have higher total precipitation-to-PW ratios than observations (Figure 3). In this subsection, we focus on the relationships between environmental moisture and MCSs, as previous studies have shown strong dependence of tropical MCS precipitation on moisture (Chen et al., 2017, 2022; Neelin et al., 2022; Schiro et al., 2020; Wolding et al., 2020).

Taking advantage of the MCS tracking data set from multiple trackers, we composite the evolution of PW (total column water vapor) as a function of the MCS lifecycle (Figure 13). Over ocean, a gradual moistening up to 24 hr prior to MCS initiation is seen from observations. In contrast, PW does not change significantly until ~6 hr prior

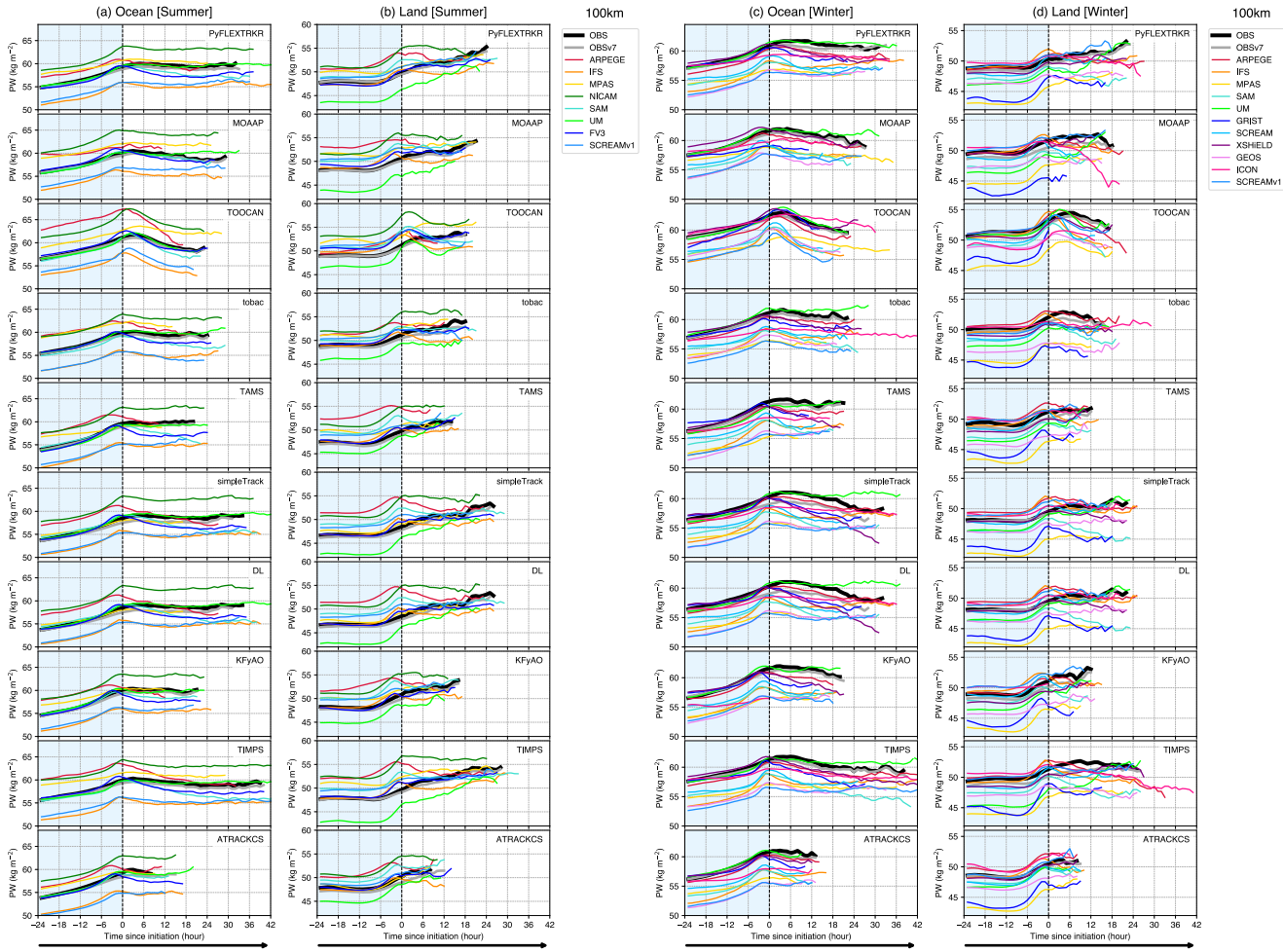

Figure 13. Composite lifecycle evolution of precipitable water (PW) associated with tropical MCSs over (a, c) ocean, and (b, d) land. Each row shows results from a tracker. Thick black lines are tracking using observations with IMERG v6, gray lines are tracking using observations with IMERG v7. Blue shaded periods are 24 hr before MCS initiation at the same location as initiation, while the white periods are following the MCS tracks. PWs are averaged within a 100 km circular radius centered at each MCS track centroid, and hours with less than 100 samples after initiation are removed. PW for observations are from ERA5.

to MCS initiation over land, followed by a rapid increase leading to initiation, possibly related to the strong diurnal cycle of surface heating and topography over land that is absent over ocean. Most DYAMOND models capture this moistening tendency prior to MCS initiation both over ocean and land, despite differences in the magnitude. The increases in PW before MCS initiation are largely consistent among trackers but diverge after initiation. Some trackers show relatively small change in PW over ocean during the MCS lifetime (e.g., PyFLEXTRKR, TAMS, simpleTrack, DL, KFyAO, tobac) while others show a gradual decrease (e.g., TOOCAN, MOAAP). In contrast, most trackers show an increase in observed PW after MCS initiation over land, but models tend to show flattening or decreasing PW. Such discrepancies among trackers during the MCS lifetime reflect the tracker formulation differences (Figure 10). The rise in PW after MCS initiation over land from ERA5 data is consistent with results from a long-term analysis (Muetzelfeldt et al., 2025) though the reason for the continued increase in environmental moisture following the MCS lifecycle remains unclear. Future studies should investigate this phenomenon further. Nevertheless, it is encouraging that global km-scale models are generally able to simulate the initiation of MCSs in response to environmental moistening rates that are similar to observations.

Lastly, we compare the relationships between MCS precipitation intensity and PW (Figure 14) from PyFLEXTRKR because results from other trackers are quite similar (Figures S15–S18 in Supporting Information S1). Over ocean, observed MCS precipitation remains weak for PW below 50 kg m^{-2} and increases sharply when PW exceeds 60 kg m^{-2} (Figures 14a and 14e) consistent with the critical PW value found in previous

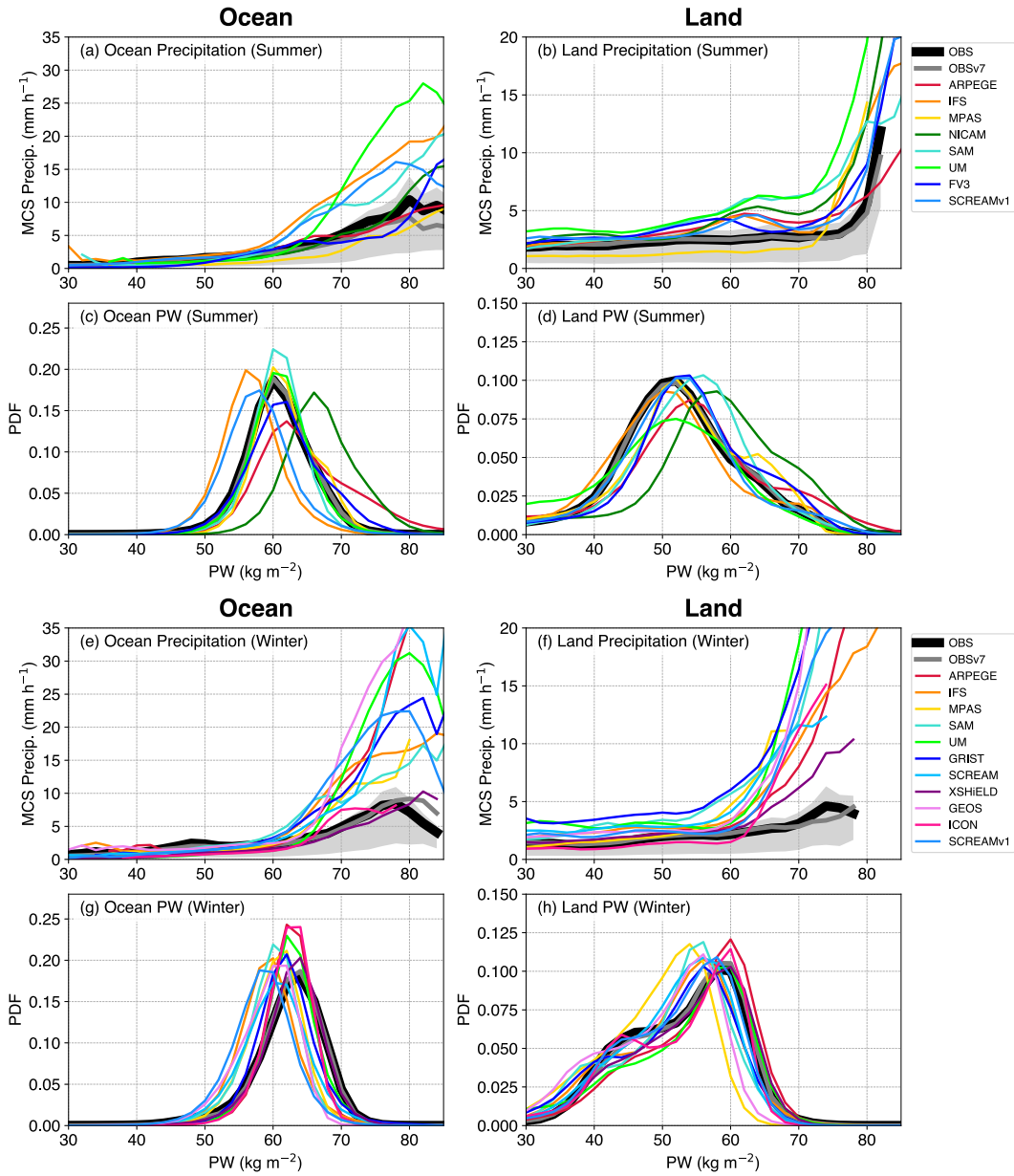

Figure 14. Mean tropical MCS hourly rain rate as a function of collocated precipitable water (PW) over (a, e) ocean and (b, f) land. PDFs of PW collocated with MCS over (c, g) ocean and (d, h) land. In (a, b, e, f), thick black lines are from IMERG v6, gray lines are from IMERG v7, light gray shadings show interquartile range from IMERG v6, and values with PW sample size <10 are excluded. In (c, d) and (g, h), both observations of PW are from ERA5. Both observed and simulated MCS precipitation were conservatively regridded to 0.25° × 0.25° to match ERA5. Only rain rates >0.1 mm hr^{-1} within MCS masks are included in the average. MCS tracking is from PyFLEXTRKR.

studies (e.g., Neelin et al., 2009). Over land, the dependence of observed MCS precipitation on PW is much weaker until PW reaches 70–75 kg m^{-2}, in contrast to some previous studies (Schiro et al., 2020), though it may be partly due to different sampling methods and spatiotemporal resolution for the precipitation data. DYAMOND models generally capture this rapid increase of MCS precipitation to PW, but the magnitudes vary widely. During summer, four out of eight models have much sharper oceanic MCS precipitation increase beyond the critical PW value while the rest are more comparable to observations. During winter, almost all models except XSHiELD show much stronger oceanic MCS precipitation intensity beyond the critical PW value. Over land, all models in both seasons simulate much steeper increases in MCS precipitation at lower PW values than observations (Figures 14b and 14f).

Table 2
Characteristics of MCS Properties for Each Tracker Relative to Other Trackers

	Frequency	Rainfall contribution	Lifetime	Max cloud shield size	Min T_b	PF mean rain rate	Rain volume
PyFLEXTRKR	Average	High	Long	Average	Average	Average	High
DL + TempestExtremes	High	High	Average	Large	Average	Average	High
MOAAP	Average	Average	Average	Average	Average	Average	Average
TOOCAN	Low	Low	Average	Small	Cold	High	Low
tobac	Low	Low	Short	Average	Average	Average	Low
TIMPS	Average	Average	Long	Average	Average	Average	Average
simpleTrack	Average	Average	Long	Average	Average	Average	Average
KFyAO	Average	Average	Short	Large	Average	Average	High
TAMS	High	High	Short	Large	Average	Average	High
ATRACKCS	High	Average	Short	Small	Warm	High	Low

Note. Qualitative assessments are based on comparisons of tracker results applied to observations, as shown in Figure 5; Figures S5 and S20 in Supporting Information S1.

Simulated PW distributions also vary widely among the models, particularly during Summer (Figure 14c). Almost all models simulate lower PW than ERA5 over the ocean during winter (Figure 14g). Over land, many models simulate higher PW in summer and lower in winter. To mitigate the effects of model biases in PW on MCS precipitation, we rescaled the simulated PW PDFs to match that from the observations and the corresponding simulated MCS precipitation (Figure S19 in Supporting Information S1). Some of the overestimated oceanic MCS precipitation sensitivity to PW remains at high PW values, though the sharp increase in simulated MCS precipitation around the critical PW value compares better with observations after PW rescaling. However, this is not the case for land MCSs. This analysis suggests that many DYAMOND models simulated greater MCS precipitation sensitivity to environmental moisture than observations, though the biases in the simulated PW distribution may be partially contributing to the bias in MCS precipitation intensity. The cause of this higher sensitivity of MCS precipitation to environmental moisture in DYAMOND models is unclear but warrants future research. We discuss some possible reasons in Section 5.

5. Conclusions and Discussions

In this study, we comprehensively examined tropical MCS characteristics in DYAMOND models for both summer and winter phases by applying 10 different feature trackers to the simulations and observations. To provide a common ground for tracker intercomparisons, the same criteria were used for each tracker to identify MCS based on both infrared T_b and precipitation characteristics. We encompass a wide diversity of trackers that differ in their way of handling various elements of the identification and tracking of convective systems, such as the spatial segmentation procedure, temporal linking method, and treatment of merging/splitting. The ensemble of trackers results in a spread of a factor of 2–3 in the observed MCS frequency (Figure S5 in Supporting Information S1) and their contribution to total precipitation (Figure 5). We also compared a variety of MCS metrics and their relationships with environmental moisture, using data diagnosed by the trackers from both observations and simulations.

Table 2 summarizes the MCS characteristics derived from each tracker relative to the average characteristics across the tracker ensemble in this study. Rather than providing a ranking, the table illustrates how the behavior of each tracker compares to others, as no reference tracking data set on the global scale exists for the study period to quantitatively validate the derived MCS characteristics. Future research using any of the trackers examined in this study should carefully consider the behavior of their selected tracker, particularly when comparing results across studies, as differences between trackers may significantly influence outcomes. The choice of tracker should be guided by the specific scientific objectives of the study, since no single tracker is universally suited to all applications, as demonstrated by the variability in MCS characteristics shown in Table 2. Moreover, the performance of these trackers should be further assessed against reference data sets carefully curated by meteorologists and storm researchers (e.g., Cui et al., 2021; Machado et al., 1998).

Table 3
Summary of Model Biases in Simulated MCS Characteristics

	Frequency	Rain amount	Rainfall contribution	Diurnal cycle magnitude	Lifetime	Max cloud shield size	Min T_b	Mean PF rain rate	PW	Rain intensity-PW
Land										
Bias sign	X	–	–	X	–	–	–	+	X	+
Mean bias magnitude	Moderate	Moderate	Small	Moderate	Moderate	Moderate	Moderate	**Large**	Moderate	**Large**
Intermodel spread	**Large**	**Large**	Moderate	Moderate	Large	Large	**Large**	Moderate	**Large**	**Large**
Ocean										
Bias sign	–	–	–	X	–	–	–	+	X	+
Mean bias magnitude	Moderate	Moderate	Moderate	Small	Moderate	Moderate	Moderate	**Large**	Moderate	**Large**
Intermodel spread	**Large**	Moderate	Moderate	Moderate	Moderate	Moderate	**Large**	Moderate	**Large**	**Large**

Note. An "X" in the bias sign column indicates fewer than 50% models share the same bias sign. Color shading in each cell represents the confidence based on agreement across the tracker ensemble: high confidence (green), low confidence (orange), determined through visual inspection of tracker agreement regarding the sign or magnitude of the bias. Bold texts highlight instances of large mean bias or large intermodel spread accompanied by high confidence. "**PW**" refers to the lifecycle evolution of PW, as illustrated in Figure 13. "**Rain intensity-PW**" refers to the relationship between MCS rain rates and collocated PW associated with MCSs, as shown in Figure 14, Figures S12–S15 in Supporting Information S1. Separation of the results by season is provided in Table S2 in Supporting Information S1. Small, moderate, and large mean bias magnitudes are defined based on relative mean differences: *small* (0%–20%), *moderate* (20%–50%), and *large* (>50%). The categorization of small, moderate, and large intermodel spread for each tracker is determined as follows: 1. *Large spread*: If more than one model falls in each of the small, moderate, and large mean bias magnitude categories, or if at least 20% of the models exhibit opposite bias signs, classify the spread as *large*. 2. *Moderate spread*: If model bias magnitudes span two adjacent categories (small and moderate, or moderate and large), classify the spread as *moderate*. 3. *Small spread*: If model bias magnitudes fall into only one category, with no more than one outlier in other categories, classify the spread as *small*. Finally, the intermodel spread is determined by the most common spread classification across the tracker ensemble.

Despite the differences in MCS statistics among trackers, we illustrate various metrics that are robust among trackers to evaluate simulations. We summarize the main findings concerning the two leading questions about the sensitivity of simulated MCS characteristics to tracker formulations, and how the robust model biases depend on environmental moisture.

- *MCS frequency and precipitation amount*: Models are generally skillful in simulating tropical mean MCS frequency, with mean biases of −2%–8% over land and −8%–8% over ocean (summer vs. winter), though large variability exists among the models (Figure 6). Most models underestimate MCS precipitation amount (−14% over land, −28% to −37% over ocean) and their contribution to total precipitation (Figures 7 and 8), with smaller multi-model mean biases over land (−13%) than over ocean (−21%). The smaller bias over land may be related to the stronger diurnal cycle that drives MCS development compared to that over the ocean.
- *MCS characteristics*: MCS diurnal cycle and cloud shield characteristics are better simulated than precipitation (Figures 9 and 10). Most models overestimate MCS mean precipitation intensity (by a factor of 2–3) and underestimate stratiform rain contribution (up to a factor of 2), particularly over land (Figures 11 and 12).
- *MCS environments*: Models capture the moistening rate leading up to MCS initiation both over ocean and land, but a large inter-model spread in PW is found (Figure 13).
- *MCS precipitation intensity*: Most models simulate the exponential increase of MCS precipitation intensity beyond the critical PW value over ocean, although the MCS precipitation sensitivity to PW is overestimated (by a factor of 2–3) in most models (Figure 14).

This study demonstrated a range of metrics for evaluating MCS statistics in global km-scale simulations against observations, which remain robust across different tracker formulations. These metrics include grid-scale MCS frequency and precipitation, diurnal cycle, MCS cloud shield and precipitation feature characteristics, as well as relationships between PW and MCS precipitation intensity. Table 3 provides a qualitative assessment of multi-model mean biases, intermodel spread, and confidence across the tracker ensemble. These metrics offer valuable insights for future km-scale model development efforts, particularly by highlighting areas where model mean biases and intermodel spread are most pronounced (emphasized in bold in Table 3).

The model biases in overestimation of convective precipitation intensity and their contribution to total MCS precipitation, as well as underestimation of stratiform precipitation, are consistent with many previous studies

employing km-scale models. These seemingly common biases in models that have different dynamical cores, microphysics parameterizations and other components suggest that the biases are not specific to a particular model or a single scheme, but a broader issue for km-scale models. It may be related to model's insufficient resolution to properly resolve entrainment mixing induced dilution in convective updrafts (Lebo & Morrison, 2015; Morrison et al., 2020), leading to convective updrafts being too strong, which lead to further biases in microphysical processes such as excessive riming and precipitation efficiency (Fan et al., 2017; Kukulies et al., 2024). MCS updrafts and precipitation intensity over land are also sensitive to wind shear via modulations of entrainment dilution (Maybee et al., 2024). In addition, convective updraft widths are often too wide in km-scale models (Prein et al., 2021; Wang et al., 2020, 2022), leading to overly strong vertical transport of condensates that would otherwise sediment more easily in thinner updrafts (Varble et al., 2014a), and biases in momentum transport that could affect MCS evolution (Varble et al., 2020). Future studies should further investigate the cause of these biases to improve the skills in simulating MCS precipitation for km-scale models.

Our analysis showed a wide range of PW simulated by the DYAMOND models, pointing to the need to better understand how air-sea interactions are treated in global km-scale models. While the DYAMOND models are atmosphere-only, some global km-scale models can now be run coupled with ocean models (e.g., ICON-Sapphire, Hohenegger et al. (2023)). Deep convection, particularly MCSs, can significantly enhance surface gustiness through convective downdrafts and momentum transport, which alter the ocean surface conditions and the coupling with the atmosphere. These processes are more explicitly simulated in km-scale models and may play important roles in the feedback between ocean and convection. We also identified consistent model biases in MCS precipitation amount and intensity distribution. The cause of the higher sensitivity of MCS precipitation to PW should be further investigated in future studies. In addition, recent observational studies have highlighted other environmental factors, such as low-level ascent, deep-layer wind shear, divergence difference between upper and lower levels, and vertically integrated moisture flux convergence, which are correlated with MCS characteristics such as lifetime, size, and precipitation (Chen et al., 2023; Galarneau et al., 2023; Muetzelfeldt et al., 2025). Future research should explore these relationships in the DYAMOND models to gain a better understanding of the biases in MCS characteristics identified in this study.

Much of our MCS analysis focused on tropics-wide comparisons, but regional differences can be quite large and MCSs in the extratropics have not been examined. Future work should examine how well these models simulate MCS in varying geographic regions that have different drivers and forcing mechanisms. Lastly, the global km-scale models and the observations were processed at 1-hr resolution in this study. Most of the trackers can work with higher resolution that could affect identification and tracking of MCSs. Further, individual convective cells that aggregate to form MCSs can be tracked at temporal resolutions of ∼15 min or finer (e.g., Feng et al., 2022), which can greatly complement MCS studies to better isolate processes related to model biases. This warrants a higher resolution intercomparison exercise in the future.

Data Availability Statement

The NOAA CPC/NCEP global merged infrared brightness temperature (T_b) data can be downloaded from the NASA server (Janowiak et al., 2017). The GPM IMERG precipitation data can be accessed from NASA for V06B (Huffman et al., 2019) and V07B (Huffman et al., 2023). The HOAPS precipitable water data can be downloaded from Copernicus (Andersson et al., 2021). ERA-5 reanalysis data can be accessed from the Copernicus Climate Data Store (Hersbach et al., 2023). Model and observation data analyzed in this study can be accessed via Globus at: https://shorturl.at/A6YoA. A free-to-register Globus account is required to access the data. Description of the dataset is provided at: https://mcsmip.github.io/datasets/. Analysis codes and visualizations in this paper are available on GitHub (Feng, 2025). The PyFLEXTRKR code can be accessed on GitHub (Feng, Hardin, et al., 2024). The MOAAP code can be obtained on GitHub (Prein, 2024). The TAMS code can be accessed on GitHub (Moon & Núñez Ocasio, 2023). The tobac code can be obtained from GitHub (tobac, 2024). The simpleTrack code is available on GitHub (Stein, 2025). The DL code can be downloaded from GitHub (Molina et al., 2024). The KFyAO code is not open source but the original code is available at Huang (2017). Questions about KFyAO can be directed to Yanluan Lin (yanluan@tsinghua.edu.cn), Xiaomeng Huang (hxm@tsinghua.

Acknowledgments
MCSMIP is a grass-root community effort comprising international researchers from universities, laboratories, and agencies. The co-chairs and committee members include Zhe Feng, L. Ruby Leung, Andreas Prein, Julia Kukulies, Kelly Núñez Ocasio, Zachary Moon, Fengfei Song, Jinyan Song, Wenhao Dong, Manikandan Rajagopal, and John Mejia. Detailed information about the MCS catalogue developers is available on the MCSMIP website (https://mcsmip.github.io/). This effort is supported by the U.S. Department of Energy Office of Science, Biological and Environmental Research (BER) as part of the Regional and Global Model Analysis program area through the Water Cycle and Climate Extremes Modeling (WACCEM) scientific focus area. The authors thank Dr. Lucas Harris and two anonymous reviewers for their constructive comments. We also thank the SCREAM team for providing the simulation data from SCREAMv1, and Dr. Hassan Beydoun for the helpful discussions in interpreting the changes in the SCREAM model. Data analyses described in this study were performed using computational resources provided by the National Energy Research Scientific Computing Center (NERSC), a DOE Office of Science User Facility supported by the Office of Science of the US Department of Energy under contract DEAC02-05CH11231. We also like to acknowledge high-performance computing support from the Casper data analysis and visualization cluster (https://doi.org/10.5065/qx9a-pg09) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation. DYAMOND data management was provided by the German Climate Computing Center (DKRZ) and supported through the projects ESiWACE and ESiWACE2. The projects ESiWACE and ESiWACE2 have received funding from the European Union's Horizon 2020 research and innovation programme under Grants 675191 and 823988. This work used resources of the Deutsches Klimarechenzentrum (DKRZ) granted by its Scientific Steering Committee (WLA) under project IDs bk1040 and bb1153. PNNL is operated for DOE by Battelle Memorial Institute under contract DE-AC05-76RL01830. NCAR is sponsored by the National Science Foundation under Cooperative Agreement 1852977. JK and KMNO were sponsored by NCAR's Advanced Study Program. The work of WD was supported by NOAA's Science Collaboration Program and administered by UCAR's Cooperative Programs for the Advancement of Earth System Science (CPAESS) under awards NA16NWS4620043 and NA18NWS4620043B. The work of BM was funded by the National Environmental Research Council (NERC) funded LMCS

edu.cn), or Wenhao Dong (wenhao.dong@noaa.gov). The TOOCAN code is not open source. Questions about TOOCAN should be directed to Remy Roca (remy.roca@cnrs.fr) or Thomas Fiolleau (thomas.fiolleau@cnrs.fr). The Forward in Time (FiT) tracking program used in TIMPS is available on Zenodo (Skok, 2023), and the TIMPS code is available on Zenodo. The ATRACKCS code is available on Zenodo (Ramírez-Cardona et al., 2022).

References

Andersson, A., Graw, K., Schröder, M., Fennig, K., Liman, J., Bakan, S., et al. (2021). Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite Data—HOAPS4.0 (4.0). https://doi.org/10.24381/cds.92db7fef

Ashley, W. S., & Mote, T. L. (2005). Derecho hazards in the United States. *Bulletin of the American Meteorological Society*, *86*(11), 1577–1592. https://doi.org/10.1175/BAMS-86-11-1577

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

Bao, J., Stevens, B., Kluft, L., & Muller, C. (2024). Intensification of daily tropical precipitation extremes from more organized convection. *Science Advances*, *10*(8), eadj6801. https://doi.org/10.1126/sciadv.adj6801

Bentley, M. L., & Mote, T. L. (1998). A climatology of derecho-producing mesoscale convective systems in the central and eastern United States, 1986–95. Part I: Temporal and spatial distribution. *Bulletin of the American Meteorological Society*, *79*(11), 2527–2540. https://doi.org/10.1175/1520-0477(1998)079<2527:ACODPM>2.0.CO;2

Bony, S., Semie, A., Kramer, R. J., Soden, B., Tompkins, A. M., & Emanuel, K. A. (2020). Observed modulation of the tropical radiation budget by deep convective organization and lower-tropospheric stability. *AGU Advances*, *1*(3), e2019AV000155. https://doi.org/10.1029/2019AV000155

Brunner, M. I., & Dougherty, E. M. (2022). Varying importance of storm types and antecedent conditions for local and regional floods. *Water Resources Research*, *58*(12), e2022WR033249. https://doi.org/10.1029/2022WR033249

Chakraborty, S., Sullivan, S. C., & Feng, Z. (2023). An overview of mesoscale convective systems: Global climatology, satellite observations, and modeling strategies. In *Clouds and their Climatic Impacts* (pp. 195–221).

Chen, B., Liu, C., & Mapes, B. E. (2017). Relationships between large precipitating systems and atmospheric factors at a grid scale. *Journal of the Atmospheric Sciences*, *74*(2), 531–552. https://doi.org/10.1175/JAS-D-16-0049.1

Chen, S. S., Houze, R. A., & Mapes, B. E. (1996). Multiscale variability of deep convection in realation to large-scale circulation in TOGA COARE. *Journal of the Atmospheric Sciences*, *53*(10), 1380–1409. https://doi.org/10.1175/1520-0469(1996)053<1380:MVODCI>2.0.CO;2

Chen, X., Leung, L. R., Feng, Z., & Song, F. (2021). Crucial role of mesoscale convective systems in the vertical mass, water and energy transports of the south asian summer monsoon. *Journal of Climate*, *35*(1), 1–46. https://doi.org/10.1175/JCLI-D-21-0124.1

Chen, X., Leung, L. R., Feng, Z., & Yang, Q. (2022). Precipitation-moisture coupling over tropical oceans: Sequential roles of shallow, deep, and mesoscale convective systems. *Geophysical Research Letters*, *49*(7), e2022GL097836. https://doi.org/10.1029/2022GL097836

Chen, X., Leung, L. R., Feng, Z., & Yang, Q. (2023). Environmental controls on MCS lifetime rainfall over tropical oceans. *Geophysical Research Letters*, *50*(15), e2023GL103267. https://doi.org/10.1029/2023GL103267

Cheng, K.-Y., Harris, L., Bretherton, C., Merlis, T. M., Bolot, M., Zhou, L., et al. (2022). Impact of warmer sea surface temperature on the global pattern of intense convection: Insights from a global storm resolving model. *Geophysical Research Letters*, *49*(16), e2022GL099796. https://doi.org/10.1029/2022GL099796

Cheng, Y.-M., Dias, J., Kiladis, G., Feng, Z., & Leung, L. R. (2023). Mesoscale convective systems modulated by convectively coupled equatorial waves. *Geophysical Research Letters*, *50*(10), e2023GL103335. https://doi.org/10.1029/2023GL103335

Coniglio, M. C., Stensrud, D. J., & Wicker, L. J. (2006). Effects of upper-level shear on the structure and maintenance of strong quasi-linear mesoscale convective systems. *Journal of the Atmospheric Sciences*, *63*(4), 1231–1252. https://doi.org/10.1175/jas3681.1

Crook, J., Klein, C., Folwell, S., Taylor, C. M., Parker, D. J., Stratton, R., & Stein, T. (2019). Assessment of the representation of West African storm lifecycles in convection-permitting simulations. *Earth and Space Science*, *6*(5), 818–835. https://doi.org/10.1029/2018EA000491

Cui, W., Dong, X., Xi, B., & Feng, Z. (2021). Climatology of linear mesoscale convective system morphology in the United States based on random forests method. *Journal of Climate*, *34*(17), 7257–7276. https://doi.org/10.1175/JCLI-D-20-0862.1

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

Ding, T., Zhou, T., Guo, Z., Yang, Y., Zou, L., & Chen, X. (2024). Contribution of mesoscale convective systems to floods in the East Asian Summer Monsoon region. *Geophysical Research Letters*, *51*(13), e2023GL108125. https://doi.org/10.1029/2023GL108125

Donahue, A. S., Caldwell, P. M., Bertagna, L., Beydoun, H., Bogenschutz, P. A., Bradley, A. M., et al. (2024). To exascale and beyond—The Simple Cloud-Resolving E3SM Atmosphere Model (SCREAM), a performance portable global atmosphere model for cloud-resolving scales. *Journal of Advances in Modeling Earth Systems*, *16*(7), e2024MS004314. https://doi.org/10.1029/2024MS004314

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*, 1–40. https://doi.org/10.1175/JCLI-D-20-0535.1

Dougherty, E., & Rasmussen, K. L. (2020). Changes in future flash flood–producing storms in the United States. *Journal of Hydrometeorology*, *21*(10), 2221–2236. https://doi.org/10.1175/JHM-D-20-0014.1

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. (2025). *MCSMIP DYAMOND Analysis (version 2025.01.0)*. Zenodo. https://doi.org/10.5281/zenodo.14664123

Feng, Z., Chen, X., & Leung, L. R. (2024). How might the May 2015 flood in the U.S. Southern Great Plains induced by clustered MCSs unfold in the future? *Journal of Geophysical Research: Atmospheres*, *129*(8), e2023JD039605. https://doi.org/10.1029/2023JD039605

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*(10), 2753–2776. 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. (2024). *PyFLEXTRKR: A flexible feature tracking Python software for convective cloud analysis (2024.03.0)*. Zenodo. https://doi.org/10.5281/zenodo.10044929

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

project (NE/W001888/1), and used the UK collaborative data analysis facility JASMIN. CK was supported by a NERC independent research fellowship (NE/X017419/1). ACV was supported by the DOE BER Atmospheric System Research program through the Integrated Cloud, Land-surface, and Aerosol System Study (ICLASS) science focus area. MJM was supported by the U.S. DOE, Office of Science, Office of Biological and Environmental Research (BER), Regional and Global Model Analysis (RGMA) component of the Earth and Environmental System Modeling Program under Award DE-SC0022070 and National Science Foundation IA 1947282. MR was supported by NASA through awards NNX17AG74G, and 80NSSC20K0896.

Feng, Z., Leung, L. R., Hardin, J., Terai, C. R., Song, F., & Caldwell, P. (2023). Mesoscale convective systems in DYAMOND global convection-permitting simulations. *Geophysical Research Letters*, *50*(4), e2022GL102603. https://doi.org/10.1029/2022GL102603

Feng, Z., Leung, L. R., Houze, R. A., Hagos, S., Hardin, J., Yang, Q., et al. (2018). Structure and evolution of mesoscale convective systems: Sensitivity to cloud microphysics in convection-permitting simulations over the United States. *Journal of Advances in Modeling Earth Systems*, *10*(7), 1470–1494. https://doi.org/10.1029/2018MS001305

Feng, Z., Leung, L. R., Liu, N., Wang, J., Houze, R. A., Li, J., et al. (2021). A global high-resolution mesoscale convective system database using satellite-derived cloud tops, surface precipitation, and tracking. *Journal of Geophysical Research: Atmospheres*, *126*(8). https://doi.org/10.1029/2020JD034202

Feng, Z., Song, F., Sakaguchi, K., & Leung, L. R. (2021). Evaluation of mesoscale convective systems in climate simulations: Methodological development and results from MPAS-CAM over the United States. *Journal of Climate*, *34*(7), 2611–2633. https://doi.org/10.1175/JCLI-D-20-0136.1

Feng, Z., Varble, A., Hardin, J., Marquis, J., Hunzinger, A., Zhang, Z., & Thieman, M. (2022). Deep convection initiation, growth, and environments in the complex terrain of central Argentina during CACTI. *Monthly Weather Review*, *150*(5), 1135–1155. https://doi.org/10.1175/MWR-D-21-0237.1

Fery, L., & Faranda, D. (2024). Analysing 23 years of warm-season derechos in France: A climatology and investigation of synoptic and environmental changes. *Weather and Climate Dynamics*, *5*(1), 439–461. https://doi.org/10.5194/wcd-5-439-2024

Fiolleau, T., & Roca, R. (2013). An algorithm for the detection and tracking of tropical mesoscale convective systems using infrared images from geostationary satellite. *IEEE Transactions on Geoscience and Remote Sensing*, *51*(7), 4302–4315. https://doi.org/10.1109/TGRS.2012.2227762

Fitzpatrick, R. G. J., Parker, D. J., Marsham, J. H., Rowell, D. P., Guichard, F. M., Taylor, C. M., et al. (2020). What drives the intensification of mesoscale convective systems over the West African Sahel under climate change? *Journal of Climate*, *33*(8), 3151–3172. https://doi.org/10.1175/JCLI-D-19-0380.1

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

González-Alemán, J. J., Insua-Costa, D., Bazile, E., González-Herrero, S., Marcello Miglietta, M., Groenemeijer, P., & Donat, M. G. (2023). Anthropogenic warming had a crucial role in triggering the historic and destructive Mediterranean derecho in summer 2022. *Bulletin of the American Meteorological Society*, *104*(8), E1526–E1532. https://doi.org/10.1175/BAMS-D-23-0119.1

Gutowski, W. J., Ullrich, P. A., Hall, A., Leung, L. R., O'Brien, T. A., Patricola, C. M., et al. (2020). The ongoing need for high-resolution regional climate models: Process understanding and stakeholder information. *Bulletin of the American Meteorological Society*, *101*(5), E664–E683. https://doi.org/10.1175/BAMS-D-19-0113.1

Haarsma, R. J., Roberts, M. J., Vidale, P. L., Senior, C. A., Bellucci, A., Bao, Q., et al. (2016). High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6. *Geoscientific Model Development*, *9*(11), 4185–4208. https://doi.org/10.5194/gmd-9-4185-2016

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

Hagos, S., Feng, Z., Burleyson, C. D., Lim, K. S. S., Long, C. N., Wu, D., & Thompson, G. (2014). Evaluation of convection-permitting model simulations of cloud populations associated with the Madden-Julian Oscillation using data collected during the AMIE/DYNAMO field campaign. *Journal of Geophysical Research-Atmospheres*, *119*(21), 12052–12068. https://doi.org/10.1002/2014JD022143

Han, B., Fan, J., Varble, A., Morrison, H., Williams, C. R., Chen, B., et al. (2019). Cloud-resolving model intercomparison of an MC3E squall line case: Part II. Stratiform precipitation properties. *Journal of Geophysical Research: Atmospheres*, *124*(2), 1090–1117. https://doi.org/10.1029/2018JD029596

Heikenfeld, M., Marinescu, P. J., Christensen, M., Watson-Parris, D., Senf, F., van den Heever, S. C., & Stier, P. (2019). tobac 1.2: towards a flexible framework for tracking and analysis of clouds in diverse datasets. *Geoscientific Model Development*, *12*(11), 4551–4570. https://doi.org/10.5194/gmd-12-4551-2019

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., et al. (2023). ERA5 hourly data on single levels from 1940 to present. https://doi.org/10.24381/cds.adbb2d47

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Hohenegger, C., Korn, P., Linardakis, L., Redler, R., Schnur, R., Adamidis, P., et al. (2023). ICON-Sapphire: Simulating the components of the Earth system and their interactions at kilometer and subkilometer scales. *Geoscientific Model Development*, *16*(2), 779–811. https://doi.org/10.5194/gmd-16-779-2023

Houze, R. A. (2014). *Cloud dynamics* (2nd ed.). Elsevier/Academic Press.

Houze, R. A. (2018). 100 years of research on mesoscale convective systems. *Meteorological Monographs*, *59*, 17.11–17.54. https://doi.org/10.1175/AMSMONOGRAPHS-D-18-0001.1

Hsu, W.-C., Kooperman, G. J., Hannah, W. M., Reed, K. A., Akinsanola, A. A., & Pendergrass, A. G. (2023). Evaluating mesoscale convective systems over the US in conventional and multiscale modeling framework configurations of E3SMv1. *Journal of Geophysical Research: Atmospheres*, *128*(23), e2023JD038740. https://doi.org/10.1029/2023JD038740

Hu, H. C., Feng, Z., & Leung, L. Y. R. (2021). Linking flood frequency with mesoscale convective systems in the US. *Geophysical Research Letters*, *48*(9). https://doi.org/10.1029/2021GL092546

Huang, X. (2017). A comprehensive Mesoscale Convective System (MSC) dataset, links to files in MatLab and plain text format. https://doi.org/10.1594/PANGAEA.877914

Huang, X., Hu, C., Huang, X., Chu, Y., Tseng, Y.-h., Zhang, G. J., & Lin, Y. (2018). A long-term tropical mesoscale convective systems dataset based on a novel objective automatic tracking algorithm. *Climate Dynamics*, *51*(7–8), 3145–3159. https://doi.org/10.1007/s00382-018-4071-0

Huffman, G. J., Stocker, E. F., Bolvin, D. T., Nelkin, E. J., & Tan, J. (2019). GPM IMERG final precipitation L3 half hourly 0.1 degree x 0.1 degree V06 (06). https://doi.org/10.5067/GPM/IMERG/3B-HH/06

Huffman, G. J., Stocker, E. F., Bolvin, D. T., Nelkin, E. J., & Tan, J. (2023). GPM IMERG final precipitation L3 half hourly 0.1 degree x 0.1 degree V07 (07). https://doi.org/10.5067/GPM/IMERG/3B-HH/07

Iguchi, T., & Meneghini, R. (2021). GPM DPR Ku precipitation profile 2A 1.5 hours 5 km V07 (07). https://doi.org/10.5067/GPM/DPR/Ku/2A/07

Janowiak, J., Joyce, B., & Xie, P. (2017). NCEP/CPC L3 half hourly 4km global (60S–60N) merged IR V1. https://doi.org/10.5067/P4HZB9N27EKU

Kim, H., Son, S.-W., Kim, H., Seo, K.-H., & Kang, M.-J. (2023). MJO influence on subseasonal-to-seasonal prediction in the Northern Hemisphere extratropics. *Journal of Climate*, *36*(22), 7943–7956. https://doi.org/10.1175/JCLI-D-23-0139.1

Kukulies, J., Chen, D., & Curio, J. (2021). The role of mesoscale convective systems in precipitation in the Tibetan Plateau region. *Journal of Geophysical Research: Atmospheres*, *126*(23), e2021JD035279. https://doi.org/10.1029/2021JD035279

Kukulies, J., Lai, H. W., Curio, J., Feng, Z., Lin, C. G., Li, P. X., et al. (2023). Mesoscale convective systems in the third pole region: Characteristics, mechanisms and impact on precipitation. *Frontiers in Earth Science*, *11*. Review. https://doi.org/10.3389/feart.2023.1143380

Kukulies, J., Prein, A. F., & Morrison, H. (2024). Simulating precipitation efficiency across the deep convective gray zone. *Journal of Geophysical Research: Atmospheres*, *129*(24), e2024JD041924. https://doi.org/10.1029/2024JD041924

Laing, A. G., & Fritsch, J. M. (1997). The global population of mesoscale convective complexes. *Quarterly Journal of the Royal Meteorological Society*, *123*(538), 389–405. https://doi.org/10.1002/qj.49712353807

Lasher-Trapp, S., Orendorf, S. A., & Trapp, R. J. (2023). Investigating a derecho in a future warmer climate. *Bulletin of the American Meteorological Society*, *104*(10), E1831–E1852. https://doi.org/10.1175/BAMS-D-22-0173.1

Lebo, Z. J., & Morrison, H. (2015). Effects of horizontal and vertical grid spacing on mixing in simulated squall lines and implications for convective strength and structure. *Monthly Weather Review*, *143*(11), 4355–4375. https://doi.org/10.1175/MWR-D-15-0154.1

Leung, L. R., Boos, W. R., Catto, J. L., DeMott, C., Martin, G. M., Neelin, J. D., et al. (2022). Exploratory precipitation metrics: Spatiotemporal characteristics, process-oriented, and phenomena-based. *Journal of Climate*, *35*(12), 1–55. https://doi.org/10.1175/JCLI-D-21-0590.1

Li, J., Qian, Y., Leung, L. R., Chen, X., Yang, Z., & Feng, Z. (2023). Potential weakening of the June 2012 North American derecho under future warming conditions. *Journal of Geophysical Research: Atmospheres*, *128*(2), e2022JD037494. https://doi.org/10.1029/2022JD037494

Lin, G., Jones, C. R., Leung, L. R., Feng, Z., & Ovchinnikov, M. (2022). Mesoscale convective systems in a super parameterized E3SM simulation at high resolution. *Journal of Advances in Modeling Earth Systems*, *14*(1), 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

Ma, H.-Y., Klein, S. A., Xie, S., Zhang, C., Tang, S., Tang, Q., et al. (2018). CAUSES: On the role of surface energy budget errors to the warm surface air temperature error over the central United States. *Journal of Geophysical Research: Atmospheres*, *123*(5), 2888–2909. https://doi.org/10.1002/2017JD027194

Machado, L. A. T., Rossow, W. B., Guedes, R. L., & Walker, A. W. (1998). Life cycle variations of mesoscale convective systems over the Americas. *Monthly Weather Review*, *126*(6), 1630–1654. https://doi.org/10.1175/1520-0493(1998)126%3C1630:LCVOMC%3E2.0.CO;2

Madden, R. A., & Julian, P. R. (1971). Detection of a 40–50 day oscillation in the zonal wind in the Tropical Pacific. *Journal of the Atmospheric Sciences*, *28*(5), 702–708. https://doi.org/10.1175/1520-0469(1971)028<0702:DOADOI>2.0.CO;2

Madden, R. A., & Julian, P. R. (1972). Description of global-scale circulation cells in the tropics with a 40–50 day period. *Journal of the Atmospheric Sciences*, *29*(6), 1109–1123. https://doi.org/10.1175/1520-0469(1972)029<1109:DOGSCC>2.0.CO;2

Mapes, B., Tulich, S., Lin, J., & Zuidema, P. (2006). The mesoscale convection life cycle: Building block or prototype for large-scale tropical waves? *Dynamics of Atmospheres and Oceans*, *42*(1–4), 3–29. https://doi.org/10.1016/j.dynatmoce.2006.03.003

Maybee, B., Marsham, J., Klein, C., Parker, D. J., Barton, E., Taylor, C. M., et al. (2024). Wind shear effects on entrainment in convection-permitting models influence convective storm rainfall and forcing of tropical circulation. https://doi.org/10.22541/essoar.171536360.00696271/v1

Molina, M., Albright, M. G., & Feng, R. (2024). *Deep learning-based detection of mesoscale convective systems*. Zenodo. https://doi.org/10.5281/zenodo.13248327

Moon, Z. L., & Núñez Ocasio, K. M. (2023). *knubez/TAMS: v0.1.2*. Zenodo. https://doi.org/10.5281/zenodo.8393890

Morcrette, C. J., Van Weverberg, K., Ma, H. Y., Ahlgrimm, M., Bazile, E., Berg, L. K., et al. (2018). Introduction to CAUSES: Description of weather and climate models and their near-surface temperature errors in 5 day hindcasts near the southern Great Plains. *Journal of Geophysical Research: Atmospheres*, *123*(5), 2655–2683. https://doi.org/10.1002/2017JD027199

Morrison, H., Peters, J. M., Varble, A. C., Hannah, W. M., & Giangrande, S. E. (2020). Thermal chains and entrainment in cumulus updrafts. Part I: Theoretical description. *Journal of the Atmospheric Sciences*, *77*(11), 3637–3660. https://doi.org/10.1175/JAS-D-19-0243.1

Muetzelfeldt, M. R., Plant, R. S., Christensen, H. M., Zhang, Z., Woollings, T., Feng, Z., & Li, P. (2025). Environmental conditions affecting global mesoscale convective system occurrence. *Journal of the Atmospheric Sciences*, *82*(2), 391–407. https://doi.org/10.1175/JAS-D-24-0058.1

Müller, S. K., Caillaud, C., Chan, S., de Vries, H., Bastin, S., Berthou, S., et al. (2023). Evaluation of Alpine-Mediterranean precipitation events in convection-permitting regional climate models using a set of tracking algorithms. *Climate Dynamics*, *61*(1), 939–957. https://doi.org/10.1007/s00382-022-06555-z

Neelin, J. D., Martinez-Villalobos, C., Stechmann, S. N., Ahmed, F., Chen, G., Norris, J. M., et al. (2022). Precipitation extremes and water vapor. *Current Climate Change Reports*, *8*(1), 17–33. https://doi.org/10.1007/s40641-021-00177-z

Neelin, J. D., Peters, O., & Hales, K. (2009). The transition to strong convection. *Journal of the Atmospheric Sciences*, *66*(8), 2367–2384. https://doi.org/10.1175/2009JAS2962.1

Nesbitt, S. W., Cifelli, R., & Rutledge, S. A. (2006). Storm morphology and rainfall characteristics of TRMM precipitation features. *Monthly Weather Review*, *134*(10), 2702–2721. https://doi.org/10.1175/MWR3200.1

Núñez Ocasio, K. M., Davis, C. A., Moon, Z. L., & Lawton, Q. A. (2024). Moisture dependence of an African easterly wave within the West African monsoon system. *Journal of Advances in Modeling Earth Systems*, *16*(6), e2023MS004070. https://doi.org/10.1029/2023MS004070

Núñez Ocasio, K. M., Evans, J. L., & Young, G. S. (2020a). Tracking mesoscale convective systems that are potential candidates for tropical cyclogenesis. *Monthly Weather Review*, *148*(2), 655–669. https://doi.org/10.1175/MWR-D-19-0070.1

Núñez Ocasio, K. M., Evans, J. L., & Young, G. S. (2020b). A wave-relative framework analysis of AEW–MCS interactions leading to tropical cyclogenesis. *Monthly Weather Review*, *148*(11), 4657–4671. https://doi.org/10.1175/MWR-D-20-0152.1

Núñez Ocasio, K. M., & Moon, Z. L. (2024). TAMS: A tracking, classifying, and variable-assigning algorithm for mesoscale convective systems in simulated and satellite-derived datasets. *Geoscientific Model Development*, *17*(15), 6035–6049. https://doi.org/10.5194/gmd-17-6035-2024

Prein, A. (2024). *AndreasPrein/MOAAP: V1.1.1 (V1.1.1)*. Zenodo. https://doi.org/10.5281/zenodo.7561615

Prein, A., Liu, C., Ikeda, K., Bullock, R., Rasmussen, R., Holland, G., & Clark, M. (2017). Simulating North American mesoscale convective systems with a convection-permitting climate model. *Climate Dynamics*, *55*(1–2), 95–110. https://doi.org/10.1007/s00382-017-3993-2

Prein, A. F. (2023). Thunderstorm straight line winds intensify with climate change. *Nature Climate Change*, *13*(12), 1353–1359. https://doi.org/10.1038/s41558-023-01852-9

Prein, A. F., Feng, Z., Fiolleau, T., Moon, Z. L., Núñez Ocasio, K. M., Kukulies, J., et al. (2024). Km-scale simulations of mesoscale convective systems over South America—A feature tracker intercomparison. *Journal of Geophysical Research: Atmospheres*, *129*(8), e2023JD040254. https://doi.org/10.1029/2023JD040254

Prein, A. F., Liu, C. H., Ikeda, K., Trier, S. B., Rasmussen, R. M., Holland, G. J., & Clark, M. P. (2017). Increased rainfall volume from future convective storms in the US. *Nature Climate Change*, *7*(12), 880–884. https://doi.org/10.1038/s41558-017-0007-7

Prein, A. F., Mooney, P. A., & Done, J. M. (2023). The multi-scale interactions of atmospheric phenomenon in mean and extreme precipitation. *Earth's Future*, *11*(11), e2023EF003534. https://doi.org/10.1029/2023EF003534

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

Rajagopal, M., Russell, J., Skok, G., & Zipser, E. (2023). Tracking mesoscale convective systems in IMERG and regional variability of their properties in the tropics. *Journal of Geophysical Research: Atmospheres*, *128*(24), e2023JD038563. https://doi.org/10.1029/2023JD038563

Rajagopal, M., Zipser, E., Huffman, G., Russell, J., & Tan, J. (2021). Comparisons of IMERG version 06 precipitation at and between passive microwave overpasses in the tropics. *Journal of Hydrometeorology*, *22*(8), 2117–2130. https://doi.org/10.1175/JHM-D-20-0226.1

Rajasree, V. P. M., Cao, X., Ramsay, H., Núñez Ocasio, K. M., Kilroy, G., Alvey, G. R., et al. (2023). Tropical cyclogenesis: Controlling factors and physical mechanisms. *Tropical Cyclone Research and Review*, *12*(3), 165–181. https://doi.org/10.1016/j.tcrr.2023.09.004

Ramírez-Cardona, Á., Robledo, V., Rendón, Á. M., Henao, J. J., Hernández, K. S., Gómez-Ríos, S., et al. (2022). Algorithm for tracking convective systems (ATRACKCS) (version 1.0). https://doi.org/10.5281/zenodo.7709959

Robledo, V., Henao, J. J., Mejía, J. F., Ramírez-Cardona, Á., Hernández, K. S., Gómez-Ríos, S., & Rendón, Á. M. (2024). Climatological tracking and lifecycle characteristics of mesoscale convective systems in northwestern South America. *Journal of Geophysical Research: Atmospheres*, *129*(19), e2024JD041159. https://doi.org/10.1029/2024JD041159

Roca, R., & Fiolleau, T. (2020). Extreme precipitation in the tropics is closely associated with long-lived convective systems. *Communications Earth & Environment*, *1*(1), 18. https://doi.org/10.1038/s43247-020-00015-4

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation.

Russell, J., Rajagopal, M., Veals, P., Skok, G., Zipser, E., & Tinoco-Morales, M. (2024). A dataset of tracked mesoscale precipitation systems in the tropics. *Geoscience Data Journal*, *12*(2). https://doi.org/10.1002/gdj3.275

Satoh, M., Stevens, B., Judt, F., Khairoutdinov, M., Lin, S.-J., Putman, W. M., & Düben, P. (2019). Global cloud-resolving models. *Current Climate Change Reports*, *5*(3), 172–184. https://doi.org/10.1007/s40641-019-00131-0

Schiro, K. A., Sullivan, S. C., Kuo, Y.-H., Su, H., Gentine, P., Elsaesser, G. S., et al. (2020). Environmental controls on tropical mesoscale convective system precipitation intensity. *Journal of the Atmospheric Sciences*, *77*(12), 1–48. https://doi.org/10.1175/JAS-D-20-0111.1

Schumacher, R. S., & Rasmussen, K. L. (2020). The formation, character and changing nature of mesoscale convective systems. *Nature Reviews Earth & Environment*, *1*(6), 300–314. https://doi.org/10.1038/s43017-020-00057-7

Skok, G. (2023). *Forward in Time (FiT) object analysis and tracking package (Version v2023_01_27)*. Zenodo. https://doi.org/10.5281/zenodo.8158841

Sokolowsky, G. A., Freeman, S. W., Jones, W. K., Kukulies, J., Senf, F., Marinescu, P. J., et al. (2024). tobac v1.5: Introducing fast 3D tracking, splits and mergers, and other enhancements for identifying and analysing meteorological phenomena. *Geoscientific Model Development*, *17*(13), 5309–5330. https://doi.org/10.5194/gmd-17-5309-2024

Song, J., Song, F., Feng, Z., Leung, L. R., Li, C., & Wu, L. (2024). Realistic precipitation diurnal cycle in global convection-permitting models by resolving mesoscale convective systems. *Geophysical Research Letters*, *51*(13), e2024GL109945. https://doi.org/10.1029/2024GL109945

Stein, T. (2025). *thmstein/simple-track: Simple track version 1.0.0 (version 1.0.0)*. Zenodo. https://doi.org/10.5281/zenodo.14679488

Stein, T. H. M., Hogan, R. J., Clark, P. A., Halliwell, C. E., Hanley, K. E., Lean, H. W., et al. (2015). The DYMECS project: A statistical approach for the evaluation of convective storms in high-resolution NWP models. *Bulletin of the American Meteorological Society*, *96*(6), 939–951. https://doi.org/10.1175/BAMS-D-13-00279.1

Stephens, G., Polcher, J., Zeng, X., van Oevelen, P., Poveda, G., Bosilovich, M., et al. (2023). The first 30 years of GEWEX. *Bulletin of the American Meteorological Society*, *104*(1), E126–E157. https://doi.org/10.1175/BAMS-D-22-0061.1

Stevens, B., Satoh, M., Auger, L., Biercamp, J., Bretherton, C. S., Chen, X., et al. (2019). DYAMOND: The DYnamics of the Atmospheric general circulation modeled on non-hydrostatic domains. *Progress in Earth and Planetary Science*, *6*(1), 61. https://doi.org/10.1186/s40645-019-0304-z

Stevenson, S. N., & Schumacher, R. S. (2014). A 10-year survey of extreme rainfall events in the central and eastern United States using gridded multisensor precipitation analyses. *Monthly Weather Review*, *142*(9), 3147–3162. https://doi.org/10.1175/MWR-D-13-00345.1

Surowiecki, A., & Taszarek, M. (2020). A 10-year radar-based climatology of mesoscale convective system archetypes and derechos in Poland. *Monthly Weather Review*, *148*(8), 3471–3488. https://doi.org/10.1175/MWR-D-19-0412.1

Thompson, R. L., Smith, B. T., Grams, J. S., Dean, A. R., & Broyles, C. (2012). Convective modes for significant severe thunderstorms in the contiguous United States. Part II: Supercell and QLCS tornado environments. *Weather and Forecasting*, *27*(5), 1136–1154. https://doi.org/10.1175/WAF-D-11-00116.1

tobac. (2024). *tobac—Tracking and object-based analysis of clouds (v1.5.3)*. Zenodo. https://doi.org/10.5281/zenodo.2577046

Ullrich, P. A., Zarzycki, C. M., McClenny, E. E., Pinheiro, M. C., Stansfield, A. M., & Reed, K. A. (2021). TempestExtremes v2.1: A community framework for feature detection, tracking, and analysis in large datasets. *Geoscientific Model Development*, *14*(8), 5023–5048. https://doi.org/10.5194/gmd-14-5023-2021

Varble, A., Morrison, H., & Zipser, E. (2020). Effects of under-resolved convective dynamics on the evolution of a squall line. *Monthly Weather Review*, *148*(1), 289–311. https://doi.org/10.1175/MWR-D-19-0187.1

Varble, A., Zipser, E. J., Fridlind, A. M., Zhu, P., Ackerman, A. S., Chaboureau, J.-P., et al. (2014a). Evaluation of cloud-resolving and limited area model intercomparison simulations using TWP-ICE observations: 1. Deep convective updraft properties. *Journal of Geophysical Research: Atmospheres*, *119*(24), 2013JD021371. https://doi.org/10.1002/2013JD021371

Varble, A., Zipser, E. J., Fridlind, A. M., Zhu, P., Ackerman, A. S., Chaboureau, J.-P., et al. (2014b). Evaluation of cloud-resolving and limited area model intercomparison simulations using TWP-ICE observations: 2. Precipitation microphysics. *Journal of Geophysical Research: Atmospheres*, *119*(24), 2013JD021372. https://doi.org/10.1002/2013JD021372

Wallace, J. M. (1975). Diurnal variations in precipitation and thunderstorm frequency over the conterminous United States. *Monthly Weather Review*, *103*(5), 406–419. https://doi.org/10.1175/1520-0493%281975%29103%3C0406%3ADVIPAT%3E2.0.CO%3B2

Wang, D., Giangrande, S. E., Feng, Z., Hardin, J. C., & Prein, A. F. (2020). Updraft and downdraft core size and intensity as revealed by radar wind profilers: MCS observations and idealized model comparisons. *Journal of Geophysical Research: Atmospheres*, *125*(11), e2019JD031774. https://doi.org/10.1029/2019JD031774

Wang, D., Prein, A. F., Giangrande, S. E., Ramos-Valle, A., Ge, M., & Jensen, M. P. (2022). Convective updraft and downdraft characteristics of continental mesoscale convective systems in the model gray zone. *Journal of Geophysical Research: Atmospheres*, *127*(16), e2022JD036746. https://doi.org/10.1029/2022JD036746

Wang, J., Fan, J., & Feng, Z. (2023). Climatological occurrences of hail and tornadoes associated with mesoscale convective systems in the United States. *Natural Hazards and Earth System Sciences*, *23*(12), 3823–3838. https://doi.org/10.5194/nhess-23-3823-2023

Williams, M., & Houze, R. A. (1987). Satellite-observed characteristics of winter monsoon cloud clusters. *Monthly Weather Review*, *115*(2), 505–519. https://doi.org/10.1175/1520-0493(1987)115<0505:SOCOWM>2.0.CO;2

Wolding, B., Dias, J., Kiladis, G., Ahmed, F., Powell, S. W., Maloney, E., & Branson, M. (2020). Interactions between moisture and tropical convection. Part I: The coevolution of moisture and convection. *Journal of the Atmospheric Sciences*, *77*(5), 1783–1799. https://doi.org/10.1175/JAS-D-19-0225.1

Wu, W.-Y., Ma, H.-Y., Lafferty, D. C., Feng, Z., Ullrich, P., Tang, Q., et al. (2024). Assessment of storm-associated precipitation and its extremes using observational data sets and climate model short-range hindcasts. *Journal of Geophysical Research: Atmospheres*, *129*(9), e2023JD039697. https://doi.org/10.1029/2023JD039697

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

Yang, Q., Majda, A. J., & Moncrieff, M. W. (2019). Upscale impact of mesoscale convective systems and its parameterization in an idealized GCM for an MJO analog above the equator. *Journal of the Atmospheric Sciences*, *76*(3), 865–892. https://doi.org/10.1175/JAS-D-18-0260.1

Yuan, J., & Houze, R. A. (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, M., Xie, S., Feng, Z., Terai, C. R., Lin, W., Tao, C., et al. (2024). Mesoscale convective systems represented in high resolution E3SMv2 and impact of new cloud and convection parameterizations. *Journal of Geophysical Research: Atmospheres*, *129*(18), e2024JD040828. https://doi.org/10.1029/2024JD040828

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

References From the Supporting Information

Crook, J., Morris, F., Fitzpatrick, R. G. J., Peatman, S. C., Schwendike, J., Stein, T. H., et al. (2024). Impact of the Madden–Julian oscillation and equatorial waves on tracked mesoscale convective systems over Southeast Asia. *Quarterly Journal of the Royal Meteorological Society*, *150*(760), 1724–1751. https://doi.org/10.1002/qj.4667

Dong, W. H., Lin, Y. L., Zhang, M. H., & Huang, X. M. (2020). Footprint of tropical mesoscale convective system variability on stratospheric water vapor. *Geophysical Research Letters*, *47*(5), e2019GL086320. https://doi.org/10.1029/2019GL086320

Fiolleau, T., & Roca, R. (2024). A deep convective systems database derived from the intercalibrated meteorological geostationary satellite fleet and the TOOCAN algorithm (2012-2020). *Earth System Science Data Discussions*, 1–42. https://doi.org/10.5194/essd-2024-36

Skok, G., Bacmeister, J., & Tribbia, J. (2013). Analysis of tropical cyclone precipitation using an object-based algorithm. *Journal of Climate*, *26*(8), 2563–2579. https://doi.org/10.1175/JCLI-D-12-00135.1

Skok, G., Tribbia, J., Rakovec, J., & Brown, B. (2009). Object-based analysis of satellite-derived precipitation systems over the low- and midlatitude Pacific Ocean. *Monthly Weather Review*, *137*(10), 3196–3218. https://doi.org/10.1175/2009MWR2900.1