Km-Scale Simulations of Mesoscale Convective Systems Over South America-A Feature Tracker Intercomparison

Andreas F. Prein, Zhe Feng, Thomas Fiolleau, Zachary L. Moon, Kelly M. Nunez Ocasio, Julia Kukulies, Remy Roca, Adam C. Varble, Amanda Rehbein, Changhai Liu, et al.

JGR Atmospheres

RESEARCH ARTICLE
10.1029/2023JD040254

Km-Scale Simulations of Mesoscale Convective Systems Over South America—A Feature Tracker Intercomparison

Andreas F. Prein[1], Zhe Feng[2], Thomas Fiolleau[3], Zachary L. Moon[1,4], Kelly M. Núñez Ocasio[1], Julia Kukulies[1], Rémy Roca[3], Adam C. Varble[2], Amanda Rehbein[5], Changhai Liu[1], Kyoko Ikeda[1], Ye Mu[6], and Roy M. Rasmussen[1]

[1]NSF National Center for Atmospheric Research, Boulder, CO, USA, [2]Pacific Northwest National Laboratory, Richland, WA, USA, [3]LEGOS, CNES, CNRS, UPS, IRD, Université de Toulouse, Toulouse, France, [4]Earth Resources Technology (ERT), Inc., Laurel, MD, USA, [5]Department of Atmospheric Sciences, Institute of Astronomy, Geophysics and Atmospheric Sciences (IAG), University of Sao Paulo (USP), Sao Paulo, Brazil, [6]Department of Geography, University of California, Santa Barbara, CA, USA

Abstract Mesoscale convective systems (MCSs) are clusters of thunderstorms that are important in Earth's water and energy cycle. Additionally, they are responsible for extreme events such as large hail, strong winds, and extreme precipitation. Automated object-based analyses that track MCSs have become popular since they allow us to identify and follow MCSs over their entire life cycle in a Lagrangian framework. This rise in popularity was accompanied by an increasing number of MCS tracking algorithms, however, little is known about how sensitive analyses are concerning the MCS tracker formulation. Here, we assess differences between six MCS tracking algorithms on South American MCS characteristics and evaluate MCSs in kilometer-scale simulations with observational-based MCSs over 3 years. All trackers are run with a common set of MCS classification criteria to isolate tracker formulation differences. The tracker formulation substantially impacts MCS characteristics such as frequency, size, duration, and contribution to total precipitation. The evaluation of simulated MCS characteristics is less sensitive to the tracker formulation and all trackers agree that the model can capture MCS characteristics well across different South American climate zones. Dominant sources of uncertainty are the segmentation of cloud systems in space and time and the treatment of how MCSs are linked in time. Our results highlight that comparing MCS analyses that use different tracking algorithms is challenging. We provide general guidelines on how MCS characteristics compare between trackers to facilitate a more robust assessment of MCS statistics in future studies.

Plain Language Summary Large clusters of thunderstorms, called mesoscale convective systems (MCSs), are important in Earth's water and energy cycle including extreme weather events like large hail, strong winds, and heavy rainfall. To better understand MCSs, researchers have developed computer programs called MCS trackers that can identify and track MCSs throughout their lifespan. Different MCS tracking algorithms have been created and used for various purposes, but little is known about how sensitive the results are to the specific algorithm used. This study aims to address this knowledge gap by comparing six different MCS tracking algorithms and assessing their impact on the characteristics of MCSs in South America. We also analyze how sensitive high-resolution climate simulation evaluations are to the used tracking algorithm. The results show that the choice of tracking algorithm has a large influence on various characteristics of MCSs, such as their frequency, size, duration, and importance to the regional water cycle. However, when it comes to evaluating simulated MCS characteristics, the choice of tracker has less impact. Importantly, all trackers agree that the high-resolution climate model accurately represents MCS characteristics across different climate zones in South America.

Key Points:

- Mesoscale convective system (MCS) tracker formulation has a profound impact on MCS characteristics such as their frequencies, size, duration, and precipitation volume
- Tracker formulation uncertainties are smaller for evaluating modeled MCS characteristics but larger for MCS frequency statistics
- MCS tracking studies have to be compared cautiously, particularly when different tracking algorithms and MCS classifications are used

Supporting Information:

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

Correspondence to:

A. F. Prein,
prein@ucar.edu

Citation:

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*, e2023JD040254. https://doi.org/10.1029/2023JD040254

Received 19 OCT 2023
Accepted 28 MAR 2024

1. Introduction

Deep convective systems (DCSs) have lifetimes that span less than 1 hr to several days and spatial scales that span 10–1,000 km. They are an integral component of the global atmospheric circulation and water cycle (Cotton & Anthes, 1992). Mesoscale convective systems (MCSs) form from clusters of deep convective storms with horizontal scales over 100 km (Houze, 2014), and the "organization" of deep convective clusters is often categorized into types such as squall lines, bow echoes, line echo wave patterns, and mesoscale convective complexes (MCCs) (Markowski & Richardson, 2011) based on cloud and precipitation spatial patterns. MCSs play a crucial

role in regulating rainfall patterns and moisture distribution throughout the tropics and midlatitude regions downstream of mountain ranges, contributing up to 90% of the annual precipitation in these regions (Feng et al., 2021; Nesbitt et al., 2006; Nesbitt & Zipser, 2003; Schumacher & Rasmussen, 2020), though the value of this contribution varies substantially based on how MCSs are defined. MCSs also produce a majority of extreme precipitation events in many regions of the world (Prein, Mooney, & Done, 2023; Rasmussen et al., 2016; Roca & Fiolleau, 2020; Stevenson & Schumacher, 2014).

In this study, we focus on MCSs in South America since it is home to various climate zones that promote MCS development with a wide range of characteristics. Focusing on this region additionally allows us to leverage existing kilometer-scale climate simulations performed within the South America Affinity Group (SAAG) (Dominguez et al., 2024). Mesoscale and synoptic processes both have a role in the formation of MCSs but differ regionally, which leads to differences in MCS properties depending on geographical location. Tropical MCSs across the Amazon are organized synoptically by the seasonally migrating intertropical convergence zone (Rehbein et al., 2018) and equatorial waves (Anselmo et al., 2021; Serra et al., 2020) with key mesoscale circulation controls from the sea breeze (Cohen et al., 1995), low-level jet (Alcântara et al., 2011; Anselmo et al., 2020) and complex terrain including mountains, rivers, and vegetation (Rehbein et al., 2018; Silva Dias et al., 2002). Many MCSs in subtropical South America are related to the presence of the South American low-level jet (SALLJ) (Salio et al., 2007), which forms after the deflection of the northeasterly trade winds crossing the Amazon as they encounter the Andes mountains (Zhou & Lau, 1998). The SALLJ advects low-level moisture from the Amazon basin to subtropical South America (Jones, 2019; Marengo et al., 2004; Vera et al., 2006; Zhou & Lau, 1998) and supports the growth of MCSs through overnight hours much like the Great Plains low-level jet over the U.S (Velasco & Fritsch, 1987). During the monsoon period in South America south of the equator (October to April), upper-level large-scale circulations such as the Bolivian High, resultant from the latent heat release from the Amazon deep convective activity (Dias et al., 1983), can affect the SALLJ moisture flux and strength of the South Atlantic Convergence Zone (Carvalho et al., 2004), which increases convection and precipitation in subtropical regions.

The Andes intersect the westerly upper-level flow, leading to the formation of surface low-pressure regions in the lee of the Andes such as the Northwestern Argentinean Low and Chaco Low (Seluchi et al., 2003). These lows redirect the SALLJ south and even southwestward (Salio et al., 2002). As troughs pass over the Andes, northward propagating cold fronts in the lee are also produced, where they interact with the SALLJ and mountainous terrain to initiate deep convection (Marquis et al., 2021; Rasmussen & Houze, 2016) that grows into MCSs (Feng et al., 2022; Mulholland et al., 2018; Zhang et al., 2021). In mid-latitudes, subsidence in the lee of the Andes creates steep free tropospheric lapse rates with temperature inversions capping the low-level moisture, which helps build up high convective available potential energy and convective inhibition (Rasmussen & Houze, 2016; Ribeiro & Bosart, 2018). The combination of these thermodynamic conditions and multi-scale circulations as modulated by the complex terrain of South America produces some of the deepest and most intense storms (Nesbitt et al., 2021; Varble et al., 2021; Zipser et al., 2006) with the most prolific lightning (Cecil et al., 2015) and hail (Cecil & Blankenship, 2012; Kumjian et al., 2020) in the world. Tornadoes, on the other side, are more common over North America due to differences in low-level wind shear (Schumacher et al., 2021). These conditions also support MCCs (Maddox, 1980), which are the largest form of MCSs, that are larger and longer lived with greater rainfall volume over South America as compared to North America (Durkee & Mote, 2010; Velasco & Fritsch, 1987). How much larger and longer-lived South American MCCs are remains debated due to sensitivities to how MCCs are defined. Thus, a combination of many different multiscale circulations coupled with thermodynamic conditions affects the life cycle of MCSs.

The contribution of MCSs to total precipitation is potentially increasing globally as temperatures rise (Tan et al., 2015), though large uncertainties exists since global models struggle to simulate MCSs and their changes. In the U.S., the frequency and intensity of MCS precipitation have increased over the past three decades during the warm season and are projected to accelerate further under future warming (Feng et al., 2016; Hu et al., 2020; Prein et al., 2017). Over the Amazon basin, in contrast, MCSs have decreased from October to March and increased from June to August, with increased precipitation in both seasons (Rehbein & Ambrizzi, 2023b). Thus, regional sensitivities of MCSs to a changing climate vary, but such assessments remain uncertain. This is at least partly the case because there is no agreed-upon definition of an MCS but also because their representation in weather and climate models is imperfect (Prein et al., 2021; Zhang et al., 2021).

Detection and tracking of convective systems including MCSs is vital to better understanding the mechanisms that control their properties and their role in energy and water transport such that predictive models can be improved for informing critical societal decisions regarding water resources and other environmental issues. The increasing availability of sub-hourly, kilometer-scale satellite cloud and precipitation retrievals and the development of similar-scale models has improved the ability to monitor and track MCS life cycles. Automatic tracking algorithms are an indispensable tool for such large data sets because they enable us to understand the full lifecycle of MCSs and allow for a process-oriented model evaluation by focusing on dynamic features rather than atmospheric mean states.

Over the past 40 years, tracking algorithms have been developed to automatically and objectively detect and track convective systems from infrared (IR) geostationary satellites and more recently from high-resolution model data. The most common algorithms are based on a convective cluster detection step from the IR imagery, and on a tracking step linking cloud clusters identified from one time step to the next. An MCS is then defined as the succession of convective clusters in a time sequence of IR images. The detection step is generally based on the application of a single brightness temperature threshold on the IR images, to identify anvil clouds associated with convective clusters (Machado et al., 1998; Vila et al., 2008; Williams & Houze, 1987). Over the years, a number of evolutions have been implemented to better describe features of the convective clusters. Thus, by applying different brightness temperature thresholds at several levels between 213 and 253 K, we can access a volumetric analysis of convective systems, with their convective cores detected with a cold temperature threshold embedded in cloud anvils detected with warmer brightness temperature threshold (Mathon & Laurent, 2001; Núñez Ocasio et al., 2020a). To go further and based on the principle that brightness temperature increases from the convective core to the edges of the anvil, detect and spread techniques have been implemented and applied to the infrared imagery to decompose the high cold cloud shield into cloud clusters (Boer & Ramanathan, 1997; Feng et al., 2023; Heikenfeld et al., 2019; Roca & Ramanathan, 2000; Wilcox et al., 2023). The tracking step, for its part, is often based on an area-overlapping technique to link one cluster detected at one time step to another one at the next time step (Feng et al., 2023; Machado et al., 1998; Williams & Houze, 1987). Some studies have added cloud movement projection techniques to increase the area-overlapping accuracy (Núñez Ocasio et al., 2020a). Other methods use a search radius method and predict the position of the cluster's center of mass to match cloud clusters between two time steps (Heikenfeld et al., 2019; Sokolowsky et al., 2023). Another branch of algorithms considers that convective systems can be tracked only if they are contiguous in their space-time domain. These algorithms then work in a volume of IR images in three dimensions (longitude, latitude, time), and identify and track MCSs in a single step, by applying single brightness temperature thresholds (Prein, Mooney, & Done, 2023), or by applying more complex techniques derived from the detect and spread method (Fiolleau & Roca, 2013). The six MCS trackers participating in this study represent the wide variety of methodologies introduced above.

Merging and splitting of convective cells is common (Bluestein et al., 1990) making it an important process in MCS dynamics. MCSs grow upscale from multiple individual cells and can decay into multiple individual cloud systems as they weaken (Ćurić et al., 2009; Rotunno et al., 1988). At a larger scale, the high cold cloud shields can be shared between several MCSs, whose anvil clouds can split and merge with each other over time, which is a challenging process to depict in tracking schemes.

There are several approaches to verify if tracking algorithms correctly detect the convective systems of interest. For example, Machado et al. (1998) validated MCS tracking results through subjective examination of forecasters determining whether or not the detected outcome looks reasonable. This is, however, very time-consuming and requires a lot of expertise. Tracking results can also be evaluated based on our physical understanding of MCSs, that is, by validating if the MCS lifecycle exhibits a realistic evolution of precipitation formation or by setting upper bounds for the expected spatial and temporal extent given physical constraints such as the Rossby radius of deformation (Cotton et al., 1989). In addition to lifecycle characteristics, this paper also assesses tracker formulation uncertainty by comparing the results of different trackers given a common set of MCS criteria.

The goals of this paper are to understand the impact of feature identification and tracker formulation on the analysis of simulated MCSs in South America and to evaluate and compare model and observed MCSs in terms of size, duration, and intensity. The paper is organized as follows: Section 2 describes the satellite and observational data, the MCS tracker methodological developments and evaluation metrics; feature tracker comparison results and discussions are in Section 3; Section 4 provides the summary of the key findings and conclusions.

2. Data and Methods

We focus on data from 3 years that were selected based on different phases of El Ni no-Southern Oscillation (ENSO): (a) June 2010 to May 2011, which was a strong La Ni na event; (b) June 2015 to May 2016, which was a very strong El Ni no event, and; (c) June 2018 to May 2019, which was a weak El Ni no event. Active ENSO phases generate Rossby waves that affect South American deep convection intensity, phase, and seasonality (Rehbein & Ambrizzi, 2023a). We decided to focus our analyses on multi-year average statistics and leave assessments of ENSO impacts on MCS statistics for future studies to limit the amount of presented information.

2.1. Data

The GPM Integrated Multi-satellitE Retrievals (IMERG) precipitation data is a combined multi-satellite precipitation retrieval data set from a network of low-orbit passive microwave sensors (Huffman et al., 2015). A quasi-Lagrangian interpolation technique is applied to the passive microwave precipitation retrievals to fill in the gaps between microwave overpasses using motion vectors derived from numerical model-derived precipitable water (Tan et al., 2019). The precipitation retrievals used in GPM-IMERG differ depending on the surface types due to changes in emissivity and differences in land/ocean precipitation characteristics (Huffman et al., 2015). Derin et al. (2021) found that GPM-IMERG has better skill in detecting precipitation and representing its intensity over oceans, while it has higher false alarm rates over land.

The IMERG data used in the study is the Final Precipitation L3 Half Hourly 0.1° × 0.1° V06B data, which is corrected with monthly surface rain gauge measurements (Huffman et al., 2019). We average the half-hourly GPM IMERG precipitation data to hourly to match the model simulation output frequency. Despite the relatively fine spatiotemporal spacing of IMERG, its actual resolution is significantly coarser than its grid spacing (Guilloteau & Foufoula-Georgiou, 2020). It is worth noting that gridded precipitation data sets may not fully capture the most intense precipitation events measured by rain gauges (Rozante et al., 2018). However, Feng et al. (2021) demonstrate that when tracking MCSs across the United States, using IMERG precipitation yields comparable results to using radar-based precipitation estimates from hourly stage-IV data (Lin & Mitchell, 2005).

The NASA Global Merged IR V1 infrared brightness temperature (T_b) data (Janowiak et al., 2017) is a merged data set combining all available operational geostationary meteorological satellite data. Viewing angle and parallax corrections have been applied to the data set. The Merged IR T_b product covers 60°S to 60°N and has a spatial resolution of 0.04° and temporal resolution of 30 min. The Merged IR product was regridded conservatively to match the IMERG 0.1°grid using conservative regridding in the Earth System Modeling Framework (ESMF) software (Collins et al., 2005). One of the 30-min T_b snapshots is used to represent convective clouds in an hour for tracking. Hourly data has been frequently used for MCS tracking (Feng et al., 2021; Kukulies et al., 2021; Núñez Ocasio et al., 2020a; Prein et al., 2020) and we follow this protocol mainly because of the availability of hourly modeled data. Future work will investigate the impact of using higher-frequency data on the presented conclusions.

We use the Weather Research and Forecasting (WRF) model version 4.1.5 (Powers et al., 2017; Skamarock & Klemp, 2008) to downscale hourly data from the fifth generation of the European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis (ERA5) (Hersbach et al., 2020) over the region shown in Figure 1. The simulation uses ∼4 km horizontal grid spacing with 1,471 × 2,028 grid cells in the horizontal and 61 stretched vertical levels. Each of the three simulations that focus on different ENSO states is initiated in May to allow for model spin-up. We use the Thompson microphysics scheme (Thompson et al., 2008), the Yonsei University Scheme (YSU) planetary boundary layer scheme (Hong et al., 2006), the RRTMG Shortwave and Longwave Schemes (Iacono et al., 2008), and the Noah-MP land surface model (Niu et al., 2011; Yang et al., 2011) including the Miguez-Macho and Fan groundwater scheme (Miguez-Macho & Fan, 2012). We use an empirical equation from Wu and Yan (2011) that estimates T_b from the modeled outgoing longwave radiation (OLR) at the top of the atmosphere. The estimated T_b as well as the hourly precipitation rates are conservatively regridded to the GPM IMERG grid by using the ESMF software (Collins et al., 2005). We analyze MCS characteristics in regions that are used in the IPCC sixth assessment report (Iturbide et al., 2020). Some of the selected regions are cropped to fit within the analysis domain (black polygon in Figure 1). The resulting analysis regions are displaced from the computational boundaries to allow for the spin-up of fine-scale structures from the coarser resolution ERA5 boundary conditions except for the Equatorial Atlantic Ocean (EAO) and Northwest South America (NWS) region.

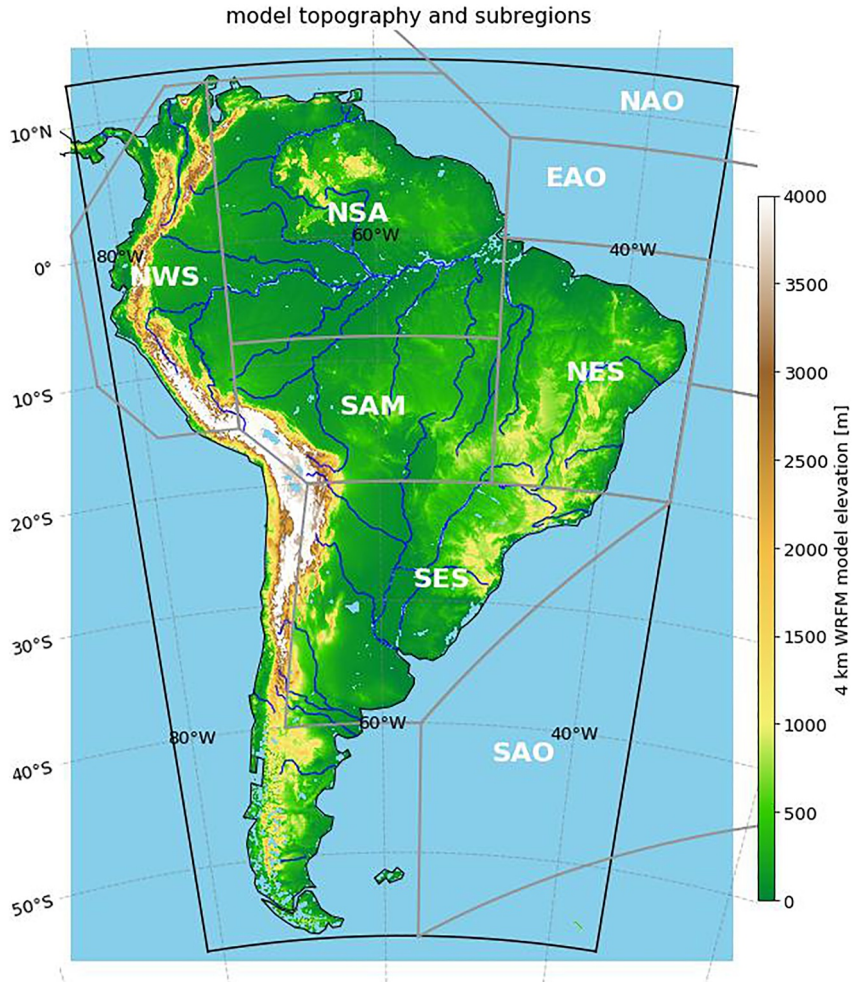

Figure 1. Simulation domain (outline of filled contours), model topography (colored contour), analysis region (black polygon), and outlines of sub-regions (gray lines). The regions are the same as in the IPCC sixth assessment report (Iturbide et al., 2020) and include Northwest South America (NWS), Northern South America (NSA), Equatorial Atlantic Ocean (EAO), South American Monsoon region (SAM), Northeast South America (NES), Southeast South America (SES), and the South Atlantic Ocean (SAO).

2.2. Methods

2.2.1. Tracking Thresholds

Several definitions using various cloud system parameters and thresholds have been used in the literature to provide an objective classification of MCS, preventing actual quantitative comparisons (e.g., recent reviews by Kukulies et al. (2023)). Houze (2004) broadly defined MCSs as "…a cumulonimbus cloud system that produces a contiguous precipitation area ∼100 km or more in at least one direction." We translate this definition into objective rules, which can be used in tracking algorithms, and can be applied to satellite-based observations. The following four criteria are used in all tracking algorithms that participate in this study:

1. The continuous $T_b \leq 241$ K area must be at least 40,000 km^2 for at least four continuous hours.
2. The maximum hourly precipitation underneath the ≤241 K T_b area must be larger than 10 mm hr^{-1} for at least 4 continuous hours.
3. The hourly precipitation volume must exceed 20,000 km^2 mm h^{-1} (e.g., 100 km × 100 km × 2 mm hr^{-1}) at least once in the lifetime of the MCS
4. The minimum T_b must be <225 K during the MCS lifetime to account for overshooting tops.

The minimum anvil cloud area (criterion 1) and the overshoot thresholds (criterion 4) are included to select clusters of cumulonimbus clouds that live longer than single-cell deep convective storms. The maximum hourly rainfall threshold (criterion 2) helps to select storms that produce heavy convective rainfall while the precipitation volume threshold (criterion 3) excludes storms whose precipitation footprint is too small or too weak. We have used similar criteria in previous work (e.g., Feng et al., 2021; Prein, Mooney, & Done, 2023) and found good agreement between objectively selected MCSs and MCSs selected by trained meteorologists. Classical definitions of MCSs, such as by Houze (2004, 2018) focus on the extent and duration of convective precipitation. Using precipitation as the primary classifier for MCSs is only feasible in regions with high-quality and high-resolution precipitation data sets that are typically based on radar observations. Such data sets are not available over South America since state-of-the-art satellite-based precipitation data sets such as GPM-IMERG have well-known biases (see Section 2.1) that prohibit their use for MCS classifications. We, therefore, translate Houze?s definition into rules that are objective and can be accounted for by using satellite-based observations. We decided to use T_b observations from geostationary satellites and track MCSs according to their cloud shield characteristics, which is common practice in MCS tracking studies (Feng et al., 2021; Hartman, 2021; Núñez Ocasio et al., 2020). All participating tracking schemes identify MCS candidates according to the Tb field and use GPM-IMERG precipitation in a second step to assess if the Tb objects qualify as MCSs.

2.3. MCS Feature Trackers

Python FLEXible object TRacKeR (PyFLEXTRKR)

PyFLEXTRKR (Feng et al., 2023) is an open-source Python package for tracking any 2D atmospheric features, with specific capabilities to track convective clouds from observations and model simulations. PyFLEXTRKR has a collection of multi-object identification algorithms, handles merging and splitting explicitly, and has been optimized for large data sets such as global kilometer-scale data. The package has a modular design that is easy to update and provides a suite of visualization, post-processing, and statistical analysis tools to facilitate scientific analysis of the tracking outputs.

The MCS tracking capability in PyFLEXTRKR jointly uses cloud top IR T_b and surface precipitation to identify and track convective systems and subsequently identifies MCSs. In this study, the detect-and-spread approach was used to identify individual deep convective systems: (a) A 10-grid (˜100 km) box filter smoothing was applied to the T_b field, and contiguous areas with $T_b < 225$ K larger than four grids (˜400 km^2) were labeled as cold cores (individual convective cloud object). (b) Each cold core was then spread outward to surrounding grid points until T_b reached 241 K. The grids with the closest distance to a nearby cold core were assigned the same label. Objects with area >800 km^2 were retained as candidate cloud systems. (c) Contiguous areas with smoothed precipitation (5-grid box filter) >3 mm hr^{-1} larger than 6 grids were defined as a precipitation feature (PF). Candidate cloud systems that share the same PF were combined to retain coherent PFs within a single convective system for tracking.

PyFLEXTRKR then tracks these convective systems based on their area overlap. Objects from two adjacent hours that have an overlap area fraction exceeding 0.5 were considered the same object. If more than one object exceeds the overlap fraction, the largest one was considered continuous and the smaller ones were labeled as merging or splitting. All convective systems exceeding 2-hr duration were tracked and saved. If a tracked system meets the MCS criteria (Section 2.2.1), the entire track is labeled as MCS, including convection initiation and upscale growth period and the decay period when the cloud shield area is below the minimum MCS area threshold. In addition, non-MCS cloud objects that merge with or split from an MCS are included as part of that MCS. The unique track numbers for each MCS were written to the pixel grid as masks, including the small merge/split cloud objects. Tracking was run continuously for each water year (from June to May) to obtain MCS tracks.

Tracking and Object-based Analysis of Clouds (tobac)

tobac (Heikenfeld et al., 2019; Sokolowsky et al., 2023) is a community-developed Python package for detecting, tracking, and analyzing clouds and other atmospheric phenomena. Due to its modular and flexible design, it can be used with user-defined tracking criteria on any atmospheric field (e.g., brightness temperatures or radar reflectivity) and on any gridded data set with two or three dimensions. In this study, we use *tobac version 1.4.2* to track MCSs based on the above-defined criteria.

The three main modules of *tobac* are feature detection, segmentation, and linking. In the feature detection, *tobac* identifies objects above or below a user-defined threshold over a minimum area. In this study, we used brightness temperature fields for the feature detection and required that a cloud object needs to be <241 K over at least 40,000 km^2. In addition, we require that at least one feature during the MCS lifetime contains a cold core of <225 K with no minimum area. For each of the identified feature, a center point is defined (in this study: the center of mass). In the segmentation procedure, these center points are used to identify all contiguous pixels around them below/above a specified threshold (here: 241 K). This done using watershedding, an image processing method that treats the input data as topographic maps and extends the area around a feature center point the same way water would flow until it meets a topographic barrier (the threshold). The segmented cloud features were co-located with the precipitation data to apply the additional precipitation-based criteria for MCS identification. It should be noted that the segmentation technique in *tobac* can result in time steps with detected cloud features that do not have an associated segmented area with their center location (see tobac documentation for details), which in turn influences the MCS lifetime when the latter is calculate based on the segmentation output.

The detected cloud features are linked over time using a search radius and their predicted propagation speed. In contrast to area-overlapping methods for the linking of features, *tobac* is based on particle tracking principles where the center points of features are assigned to a common track when they fall within the predicted radius of motion. The search radius for potential features can be adjusted by the user (we used a maximum propagation speed of 100 m s^{-1}) and if multiple features fall within the search radius, the feature with the path that is closest to the preceding motion direction is selected. While *tobac* has a postprocessing tool that identifies merges and splits based on the output from the feature detection, it has no explicit treatment of merging and splitting during the linking procedure.

Forecasting and Tracking the evolution of Cloud Clusters (ForTraCC)

ForTraCC's development started in the 1990s (Machado et al., 1998) making it one of the longest-standing cloud object tracking algorithms developed and still actively used (Vila et al., 2008). Currently, it is being used operationally for nowcasting at the Brazilian Center for Forecast and Climate Studies of the National Institute of Spatial Research (CPTEC/INPE; http://pindara.cptec.inpe.br/fortracc/). ForTraCC can work with radar reflectivity, precipitation, or OLR. However, ForTraCC is not able to meet all of the here defined meet MCS criteria within the model code requiring the development of a post-processing tool (see https://github.com/salvatirehbein/percolator).

In the current study, ForTraCC was set to identify and track all the objects with one or more contiguous pixels with brightness temperature equal to or above 241 K. The algorithm takes into account the potential occurrence of missing data or input failures (Vila et al., 2008). ForTraCC relies on the overlap (in our Case 5%) between consecutive images. The initiation can be (a) spontaneous; (b) merge, or (c) splits. In cases of merging, the larger system or the first one identified if they have the same size, will be tracked. If an MCS splits, the larger resultant system will continue to be tracked, while the smaller systems will become new individual systems.

ForTraCC tracks and stores all cold clouds, which increases its runtime (see Table 1). A post-processing program is used to filter out storms that do not meet our MCS criteria. First, we ensure that each cloud cluster defined by ForTraCC has a minimum area of 40,000 km^2 for at least four continuous hours, along with at least one pixel during the system's lifecycle with a minimum brightness temperature of 225 K. Next, the mask files generated by ForTraCC are overlapped with the corresponding precipitation field. This process verifies if the area and volume under the mask (i.e., cloud shield) meet the desired criteria.

TAMS

The Tracking Algorithm for Mesoscale Convective Systems (TAMS) is an open-source MCS tracking and classifying algorithm and Python package. One novelty of TAMS is its grid independence. Grid-independent tracking allows for the identification and tracking of both observed (satellite data) and simulated (model data) systems regardless of the type of grid and data resolution. The package includes a set of visualization and post-processing tools including functionality that allows matching a desired variable or atmospheric field to each MCS and calculating corresponding statistics. TAMS was initially developed to track and analyze tropical MCSs over Africa associated with African easterly waves (Núñez Ocasio et al., 2020a, 2020b) and a description of this initial version can be found in Núñez Ocasio et al. (2020a). The new in-development version of TAMS used in this study

Table 1
Key Characteristics and Computational Demands of Mesoscale Convective System (MCS) Tracking Algorithms

	Temporal linking method	Spatial segmentation	Program-ming language	Max. memory (GB)[a]	Runtime (s)[a]	Hardware
ForTraCC	Overlap	None	Fortran (R)	0.32 (32.25)	112 (153)	Intel 6240
PyFLEXTRKR	Overlap	Detect-and-spread	Python	8.98	72	AMD EPYC 7763, 1 node, 32 CPUs
TAMS	Overlap	Contour	Python	0.38 (2.1)	44 (84)	Intel 6240
tobac	Parcels, propagation speed	Water-shedding	Python	1.19	73	AMD EPYC 7763 Milan
MOAAP	Overlap	3D	Python	6.50	72	Intel 6240
TOOCAN	3D region growing (space + time)		C	8.6	69	AMD EPYC 7763 Milan

Note. For ForTraCC and TAMS two numbers are provided for the memory and runtime where the first number shows the demands of the tracking algorithm and the second number in brackets the demands of the post-processing program that creates an MCS mask file. The other trackers include the post-processing step in their algorithm. Temporal linking refers to the procedure of connecting MCSs in time while spatial segmentation refers to how a cloud field is segmented into individual MCSs. [a]Numbers correspond to tracking MCS to the completion of writing an MCS mask file. The common period is 1–7 November 2018 considering the entire domain (Figure 1) on the 0.1° GPM-IMERG grid.

follows the same main four steps as its predecessor: (a) Identify, (b) Track, (c) Classify, and (d) Assign variable(s).

The identification step consists of identifying regions of cloud top IR T_b within the 241 K region with overshoots (less than 225 K) larger than 4,000 km^2. (In the default version of TAMS 235 and 219 K are used as thresholds.) TAMS uses this criterion in addition to the four MCS criteria described in Section 2.2.1. These regions that are potential candidates to be MCSs are called Cloud Elements (CEs). Although these identification criteria may cause a late detection of initiation, they assure the system is an MCS and not a convective cell as well as assuring the targeting of raining clouds. The tracking is done on stored convex hull polygon shapes based on the CEs shapes using the overlapping method. For this study, the overlap threshold was 50% and the optional cloud projection or background flow was turned off. In the current simplified linking scheme, each CE at the current time step is matched with the maximum overlap "parent" CE from the previous time step, if the overlap condition is satisfied. This creates a list of "parents" and "kids" that then become one single family/MCS. Based on default criteria considering shape, size, and duration, each MCS can be classified into one of four possible categories: MCCs, Convective Cloud Clusters (CCCs), Disorganized Long-Lived, and Disorganized Short-Lived. However, TAMS was configured in this study to follow the set of criteria defined here. Tracks were filtered to remove MCSs that did not meet the criteria for this study. Parquet files were converted to gridded mask NetCDF files.

MOAAP

The Multi-Object Analysis of Atmospheric Phenomenon (MOAAP) algorithm (Prein, Mooney, & Done, 2023) is a Python-based MCS tracker previously used in Prein et al. (2021) and Poujol et al. (2020). The algorithm is similar to the Method for Object-Based Diagnostic Evaluation (MODE) Time Domain (MTD) (Clark et al., 2014; Davis et al., 2009; Prein et al., 2020). MOAAP is based on the connectedness of objects, meaning that objects must be adjacent in space and time (no minimum overlap criterion is used). It is designed to track multiple atmospheric features, such as cyclones, jet streaks, and atmospheric rivers but can also track single features such as MCSs. MOAAP operates through the following five steps to track MCSs.

1. The three-dimensional T_b field (time, latitude, longitude) is thresholded. This process produces a binary field where cells below the threshold are set to one (objects of interest), while all other cells are set to zero.
2. The binary field is passed to the Python label function of the multidimensional image processing tool (ndimage) from the SciPy package (Virtanen et al., 2020). This function identifies objects connected in space and time (grid cells that are co-located horizontally or diagonally in a 3-dimensional—latitude × longitude × time—matrix) and assigns them a unique label/index, resulting in a feature matrix.
3. MOAAP uses a merging and splitting function on the feature matrix. This function merges or breaks up objects connected in time but not in space. For instance, if two objects merge, the smaller object ends at the previous timestep and is assimilated into the larger object. Conversely, when an object splits into two, the larger object continues while the smaller one is treated as a new feature. The merging and splitting function incorporates a temporal threshold (we use 4 hr here) to ensure that only longer-lived merged and split objects are relabeled.

4. From the entire population of identified objects, we select a subset that satisfies specific criteria tailored to the atmospheric phenomena under consideration. This is the step where we account for the four MCS criteria defined in Section 2.2.1.
5. Once all objects qualifying as a specific phenomenon are identified, their characteristics are calculated.

TOOCAN

The TOOCAN algorithm (Tracking Of Organized Convection Algorithm through a 3-D segmentation) (Fiolleau & Roca, 2013) relies on a conceptual model of a convective system consisting of a 3D (longitude, latitude, time) cloud cluster made up of a convective core associated to its stratiform anvil evolving in the space-time domain. To identify such spatio-temporal cloud clusters, the algorithm works within a volume of IR images and applies a 3-D region growing technique to decompose the cold cloud shield, initially delineated by a 235 K threshold in the spatio-temporal domain into component MCSs. This technique consists of an iterative process of detection and dilatation of convective seeds in the spatiotemporal domain.

Convective seeds are first detected with a 190 K threshold. Note that this is a TOOCAN-specific setting that is stricter than the common criteria for overshoots of <225 K. Then, an intermediate cold cloud shield mask is identified in 3D at a 5 K warmer threshold. Only convective seeds with a minimum lifetime of three images and exceeding 625 km^2 per image are kept. The selected seeds are then spread in the spatio-temporal domain until they reach the edges of the intermediate cold cloud shield. This step consists of adding edge pixels belonging to the intermediate cold cloud shield to all already detected seeds. This iterative process of detection and dilation is repeated every 5 K from 190 to 235 K and is stopped when all the pixels below 235 K are associated with an MCS. Note that this means that TOOCAN continues to track clouds after they lose their overshoot. To fit with the MCS criteria defined in this study, the cold cloud shield boundaries have been set at 241 K. The multi-stage, multi-threshold technique allows an MCS identification independent of a single detection threshold. Also, the way the TOOCAN algorithm operates in 3D without the traditional detection and tracking steps allows the continuity of the tracking of the stratiform anvil associated with the MCS after its convective activity is ended. Isolated convective cells in the MCS initiation stages and scattered cirriform clouds in the MCS dissipation stages, disconnected on a single IR image, may be part of the same MCS allowing a coherent life cycle. With such methodology, the unphysical split and merge issues are resolved and all the MCS can be analyzed without filtering on merging and splitting. Finally, the method identifies the full spectrum of the convective organization, from small and short-lived systems to systems more organized lasting several days. For this study, we will focus on convective systems that meet the criteria defined previously.

It is important to highlight that although all trackers use the same four criteria to identify MCSs, each tracker has many additional assumptions incorporated into its source code that affect the classification. It is infeasible to homogenize all of these assumptions across the trackers and testing the effects of these assumptions on MCS classification is the prime motivation of this study. Table 1 summarizes high-level tracker characteristics and shows a comparison of memory needs and the speed of each tracker when identifying MCSs in a week-long period. All trackers, except for PyFLEXTRKR, used serial processing.

2.3.1. Evaluation Metrics

We ensure consistency in the evaluation of MCS characteristics by running the same analysis code on MCS mask files from each tracker (i.e., matrices with dimensions time, latitude, longitude that labels individual MCSs with a unique integer). All of the statistics are based on sampling over hourly MCS data except for the MCS duration, which integrates over the MCS lifetime. We introduce the evaluation metrics in the relevant locations in the results section to simplify the interpretation of results.

3. Results

3.1. Tracking of Idealized MCS Cases

We start with comparing the MCS trackers by applying them to four highly idealized test cases to more easily identify commonalities and differences among them (Figure 2). The first case (Figures 2a–2j) features three individual eastward-moving cloud objects with overshooting tops ($T_b \leq 221$ K). These clouds are growing and merging 8 hr after initialization ($t = 8$ hr). At this time, the northern and southern cells are losing their overshoots. The cells continue to move eastward until $t = 19$ hr when the northern cell splits off from the two southern cells.

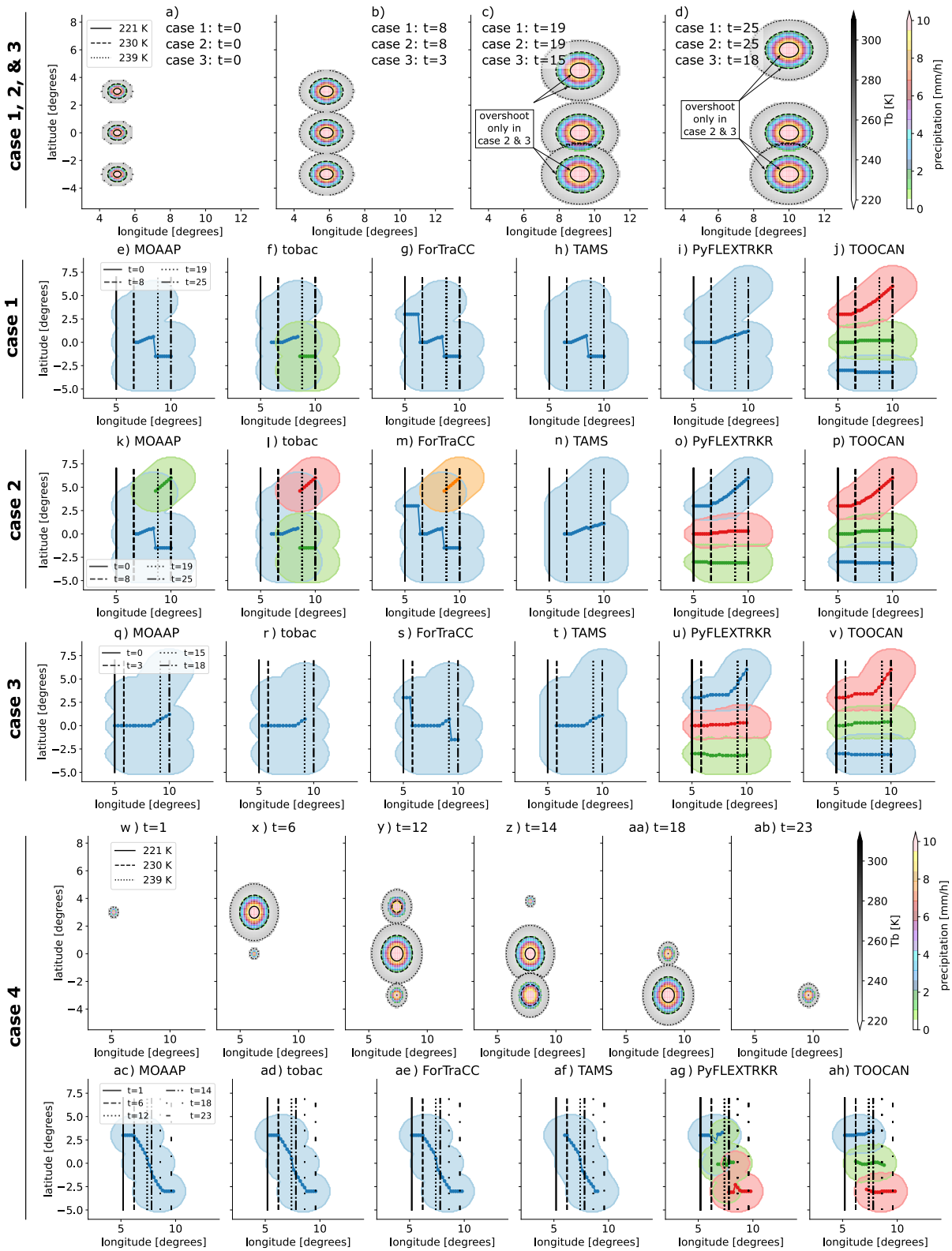

Figure 2.

MOAAP (Figure 2e), ForTraCC (Figure 2g), and TAMS (Figure 2h) identify a single MCS that starts at $t = 8$ hr ($t = 0$ hr in ForTraCC). These trackers exclude the northern cell from the MCS system once it splits off at $t = 19$ hr, which results in a sudden southward shift in the MCS track. ForTraCC has a discontinuity in its track also at $t = 8$ hr since it uses the northernmost cell as the initiation point after the systems have merged if there are multiple cells that merge at the same time. PyFLEXTRKR (Figure 2i) also identifies a single MCS but has an initiation at $t = 0$ hr and keeps the northern cell attached to the system after it splits off. Two MCSs are identified by tobac (Figure 2f). One system initiates at $t = 6$ hr and terminates at the time when the northern cell splits off. At this time, a new MCS is identified that consists of the central and southern cells. Finally, TOOCAN continues to track the three initial cells as separate systems and therefore identifies three MCSs that all start at $t = 0$ hr and terminate at $t = 25$ hr. It is important to mention that all of the found solutions satisfy our MCS criteria.

The second idealized case is identical to the first one but each of the three cells maintains its overshoot meaning the northern cell classifies as an MCS after it separates at $t = 19$ hr (Figures 2a–2d and 2k–2p). TAMS continues to identify one MCS with the northern cell remaining part of the system after it splits off. MOAAP, ForTraCC, and tobac identify the northern cell as a new MCS after it separates from the main system, but in contrast to most other trackers, tobac does not continue any of the previous track(s) after the split. This is most likely explained by the relatively large distance between the feature center points and the search radius that is used in tobac to connect the latter. PyFLEXTRKR identifies three MCSs instead of one due to the maintenance of the overshoots in this case, agreeing with TOOCAN whose classification is unchanged. Since PyFLEXTRKR uses the detect and spread method to segment cloud systems, the three separate overshoots that each satisfy the MCS criteria on their own result in three separately tracked MCSs (see Figure 8 and associated text in Feng et al. (2023)).

The third idealized case is similar to the first two except for only having 3 hr before the cells merge and 3 hr after the northern cell separates (Figures 2a–2d and 2q–2v). The 3-hr threshold is selected to test how the trackers deal with individual cells that are shorter-lived than the 4-hr minimum MCS lifetime. PyFLEXTRKR and TOOCAN find three systems similar to the second idealized case while MOAAP, tobac, ForTraCC, and TAMS each identify 1 MCS. However, there are differences between the start and end of the one identified MCS. MOOAP has a smooth track that initiates the MCS at $t = 0$ hr and follows it until the end of the simulation ($t = 18$ hr). tobac identifies the MCS at $t = 1$ hr and stops the system when the northern cell separates at $t = 15$ hr. This can be explained by a default smoothing procedure of the input data in tobac that can lead to some objects not being identified as an MCS cloud object if they have sizes close to the minimum area required. ForTraCC initiates the MCS at $t = 0$ hr but only follows the northern cell until the three cells merge at $t = 3$ hr, resulting in a discontinuity in the track. It also stops following the northern cell after it splits off at $t = 15$ hr, resulting in a second discontinuity in the track. TAMS starts identifying an MCS after the 3 cells merge at $t = 3$ hr and keeps all cells connected (similar to MOAAP) until the end of the simulation.

The fourth idealized case features an asynchronous development of the 3 cells with interactions (overlapping cloud shields) during their lifetimes (Figures 2w–2ah). MOAAP, tobac, and ForTraCC identify 1 MCS with identical tracks starting at $t = 0$ hr and ending at the end of the simulation at $t = 23$ hr. TAMS also identifies 1 system but with a shorter track (initiation happens at $t = 5$ hr and termination at $t = 20$ hr). TOOCAN keeps the 3 cells separated as individual MCSs during their entire lifetime with tracks predominantly moving eastward. PyFLEXTRKR also identifies three MCSs but features interactions between the individual cells when the northern cell terminates and the southern cell initiates, which results in small discontinuities in the MCS tracks.

Based on these idealized cases, we can expect that TOOCAN will identify more frequent and smaller MCSs followed by PyFLEXTRKR while TAMS and MOAAP might have the fewest and biggest systems with the other trackers being in between. TAMS likely produces larger MCSs due to its use of convex hulls when identifying

Figure 2. Mesoscale convective system (MCS) tracker intercomparison for four idealized cases. The first three cases initiate with three individual cells that all contain an overshooting top (a) and move eastward, grow, and merge (b). Afterward, the northern cell splits off (c) by moving toward the northeast (d). Case 1 differs from Case 2 because the northern and southern cells lose their overshoot after merging with the center cell. Case 3 differs from Case 2 by only having 3 hr before the cells merge and 3 hr after the cells split instead of 8 hr. The fourth case (w–ab) explores how trackers deal with 3 splitting and merging cells that develop asynchronously. The northern cell initiates at $t = 0$ hr, the central cell at $t = 6$ hr, and the southern cell at $t = 12$ hr. Each cell moves eastward at the same speed, growing for the first 7 hr and shrinking during the following 7 hr. The resulting MCS tracks (connected circles) and MCS footprints (contours) from the different trackers are shown in panels (e–v) for cases 1 to 3 and in panels (ac–ah) for Case 4. Individual MCSs have different colors. The vertical lines indicate the time of initiation (solid), merging of the cells (dashed), splitting off of the northern cell (dashed), and the end of the case study (dashed-dotted line).

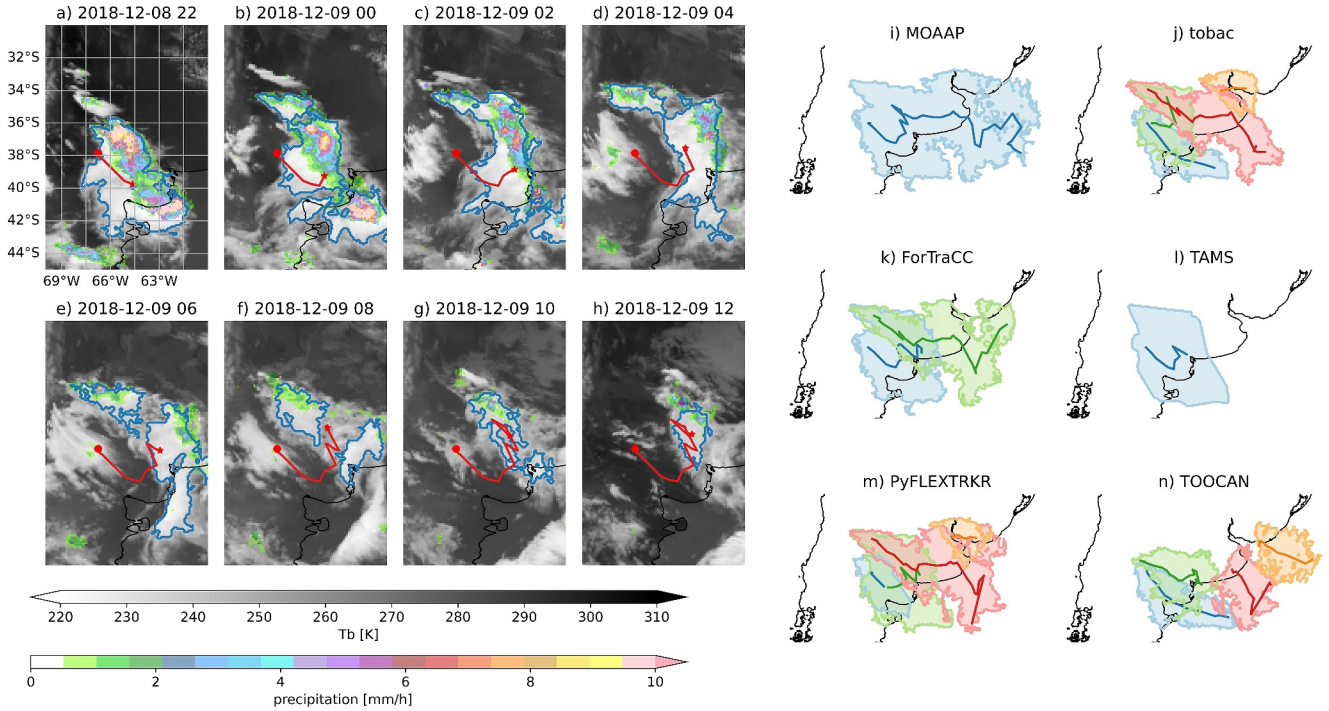

Figure 3. Similar to Figure 2 but for tracking an observed deep convective outbreak that occurred during December 2018 in southeastern South America. Panels (a–h) show the evolution of the cloud shield (gray shading), precipitation rates (colored shadings), the outline of the detected Mesoscale convective system (MCS) (blue contour), and the MCS track (red contour) based on results from MOAAP. The red circle shows the initiation point of the MCS. Panels (i–n) show the tracks and outline of the identified MCSs based on (i) MOAAP, (j) tobac, (k) ForTraCC, (l) TAMS, (m) PyFLEXTRKR, and (n) TOOCAN. Different colors indicate individual MCSs.

anvil clouds (see Figure 3 for an example). Additionally, ForTraCC, PyFLEXTRKR, and TOOCAN were able to follow the cells from $t = 0$ hr to the end of the simulations for all cases, while other trackers missed some of the early or late stages of MCS development for some cases. tobac tends to initiate new tracks instead of preserving one of the previous tracks when splitting occurs, which should result in higher initiation frequencies and shorter lifetimes. TAMS consistently detects MCSs hours after convection initiation. This late initiation detection is mainly due to its identification criteria of being a convective area with an embedded cold core area size threshold.

3.2. Tracking of MCSs During a Deep Convective Outbreak in Argentina

The idealized cases discussed in Section 3.1 capture some of the variability of MCS evolutions but certainly do not cover all possibilities that can occur in real cases. While it is impossible to analyze the thousands of MCSs that were identified during the 3-year analysis period, we want to highlight similarities and differences between the idealized cases and an observed deep convective outbreak that occurred during the Cloud, Aerosol, and Complex Terrain Interactions (Varble et al., 2021) and Remote sensing of Electrification, Lightning, And Mesoscale/microscale Processes with Adaptive Ground Observations (RELAMPAGO) (Nesbitt et al., 2021) field campaigns in Argentina in December 2018 (Figure 3).

Deep convection was triggered around −38°S and −67°W (red circle in Figures 3a–3h) on 8 December 2018 at 16:00 UTC and rapidly grew upscale moving toward the southeast during the subsequent hours. While the initial deep convection decayed after about 8 hr, new deep convection formed to the north and continued until 10 December 2018 16:00 UTC when it eventually decayed over the South Atlantic.

Similar to the idealized cases, MOAAP and TAMS have the fewest systems with one identified MCS, while tobac, PyFLEXTRKR, and TOOCAN identify the most systems with four MCSs each. MOAAP has the largest total MCS area extent (total area under the tracked cloud shield) while TAMS has the smallest total extent since it identifies only one fairly short-lived MCS. The four MCSs that were identified by tobac, PyFLEXTRKR, and TOOCAN have different extents and tracks, highlighting the complexity of cloud field decomposition including splitting and merging of systems within the MCS trackers.

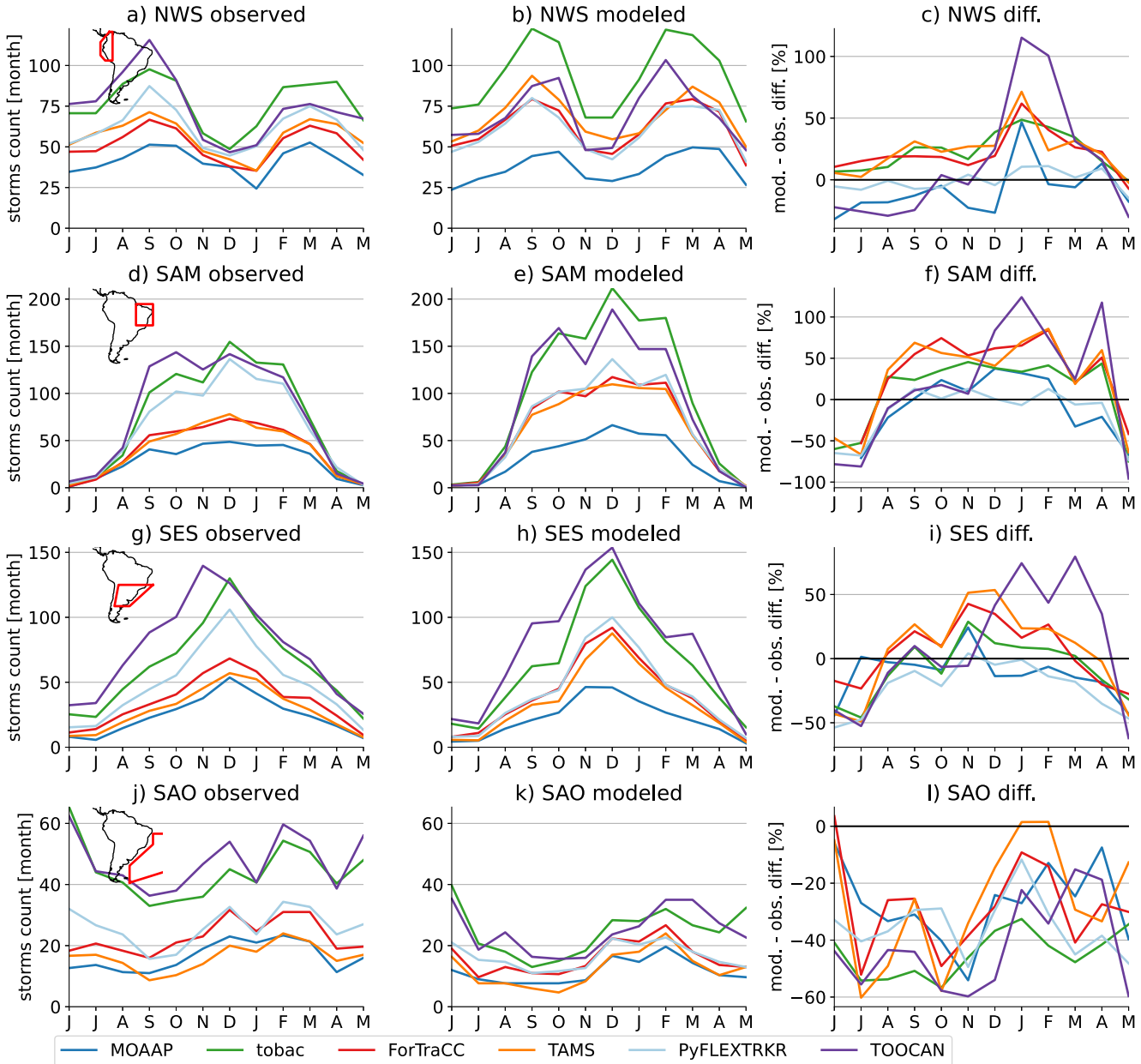

Figure 4. Monthly mean Mesoscale convective system (MCS) frequencies (averaged over the 3 years) for observed MCSs (first column), modeled MCSs (center column), and their relative differences (right column). Results are shown for the NWS, SAM, SES, and SAO regions (top-down).

3.3. Annual Cycle of Monthly MCS Frequencies

Moving on from investigating single cases, we now analyze the observed and simulated 3-year average monthly frequency of MCSs in different subregions. An MCS is assigned to a region if at least half of its track centroid (geometric center of the ≤241 K cloud shield) is within the region during the MCS's lifetime. The NWS region exhibits a double peak in MCS occurrence during September and March, which is captured by all trackers (Figure 4a). The main difference between the trackers is the average number of MCSs per year, which varies between 493 in MOAAP and 919 in tobac. These frequency differences between trackers are similar for simulated MCSs (Figure 4b). Comparing simulated to observed MCS frequencies shows that there are only 2 months where all trackers agree on the sign of the differences (i.e., the model has too many MCSs in January and too few in May (Figure 4c)).

The SAM region has a pronounced dry period during winter and a long period with high MCS activities from September to February (Figures 4d and 4e). All trackers agree that the simulations have too few MCSs during the dry season and too many from September to February. However, large differences exist about the magnitude of the overestimation ranging from close to zero in PyFLEXTRKR to more than 100% in January when using TOOCAN. Similar results are found for the SES region, while larger differences exist for the SAO region where all trackers agree on a low bias in simulated MCS frequencies. Results for all subregions are shown in Figure S1 of Supporting Information S1.

3.4. MCS Characteristics

Next, we investigate how MCS characteristics depend on the tracker formulation and how this uncertainty affects model evaluation. We show results from the NSA region as a representative example here but show other regions in Figures S2–S9 of Supporting Information S1 and discuss them further below.

Peak MCS cloud shield sizes during the MCS lifetime per definition have to be larger than 40,000 km^2, which is met by most trackers in the NSA region except for a few MCSs in PyFLEXTRKR, TOOCAN, and tobac (Figure 5a). In PyFLEXTRKR this might in part be related to the assumption of a constant grid cell area. The smallest MCSs are identified when using TOOCAN while the largest systems are found when using TAMS. ForTraCC, TAMS, and tobac suggest that the simulated MCSs are smaller than observed systems while the other trackers have similar observed and simulated size distributions. MCSs in NSA move slowest when using TOOCAN and are fastest when using tobac (Figure 5b). The smaller speed in TOOCAN is likely related to not having any mergers and splits in the tracking. All trackers agree that simulated systems move slightly slower than observed MCSs. Large model-observation differences occur for the MCS lifetime-maximum 95th-percentile (P95) hourly precipitation rate (Figure 5c) and mean precipitation rate (Figure 5d) with simulated rates being significantly higher than observed. This is at least partly related to deficiencies in accurately capturing precipitation frequencies and intensities in GPM-IMERG (Dominguez et al., 2024; Guilloteau & Foufoula-Georgiou, 2020; Rozante et al., 2018; Zhang et al., 2021), and simulated convective updrafts being too large and strong at 4-km grid spacing (Fan et al., 2017; Varble et al., 2020; Wang et al., 2020). Tracker-dependent model-observations differences are smaller for the mean and P95 precipitation rate characteristics than for other MCS characteristics, though TOOCAN produces higher P95 values compared to the other trackers. MCS lifetime-average precipitation volumes are smallest in TOOCAN and similar across the other trackers (Figure 5e), in agreement with MCS maximum size statistics. Interestingly, differences between modeled and observed precipitation volumes are much smaller than the differences in MCS precipitation rates. This is caused by GPM-IMERG having larger precipitation areas than simulated, which offsets lower precipitation intensities compared to the simulations(Dominguez et al., 2024; Zhang et al., 2021). Finally, there are large tracker formulation differences concerning MCS duration. MOAAP has the longest-lived systems with median values of 17 hr while median MCSs in tobac only live for ∼8 hr. Also, the interquartile range of the duration distribution varies from ∼10 hr in MOAAP to 3 hr in TOOCAN. Most trackers produce little differences between observed and simulated MCS lifetimes, though ForTraCC, TOOCAN, and tobac suggest MCSs may be longer lived in the observations. The duration of MCSs in tobac can be influenced by the applied segmentation technique that can result in time steps with detected cloud features that do not have an associated segmented area with their center location (see Section 2.3). The results for other regions are shown in Figures S2–S9 of Supporting Information S1.

Figure 6 shows an overview of observed (x-axis) and modeled (y-axis) median MCS characteristics for different trackers (colors) and all subregions (symbols). For lifetime-maximum MCS size (Figure 6a), regional differences are similar between trackers with most trackers having the smallest MCSs in NAO and the largest MCSs in SAO, SES, and NES. MCS speeds are similar between observations and modeled systems with most data points lying close to the one-to-one line (Figure 6b). Most trackers simulate the fastest storms in SAO and the slowest in the EAO. There are large regional dependencies in the simulation of P95 precipitation rate with small differences in the SAO region and the largest differences in NWS, NES, and SES (Figure 6c). TOOCAN MCSs feature the heaviest P95 precipitation rates in most regions while systems identified by ForTraCC generally have the lowest rates (Figure 6c). Mean precipitation rates also agree better between simulated and observed storms in the SAO region (Figure 6d). In all other regions, simulated WRF MCSs have much higher mean precipitation rates than observed systems with small uncertainties due to tracker formulation. Simulated MCS average precipitation volumes are systematically smaller than observed volumes when tobac, ForTraCC, or TAMS are used, while the sign of differences is regionally dependent when the other trackers are used (Figure 6e). Finally, all trackers

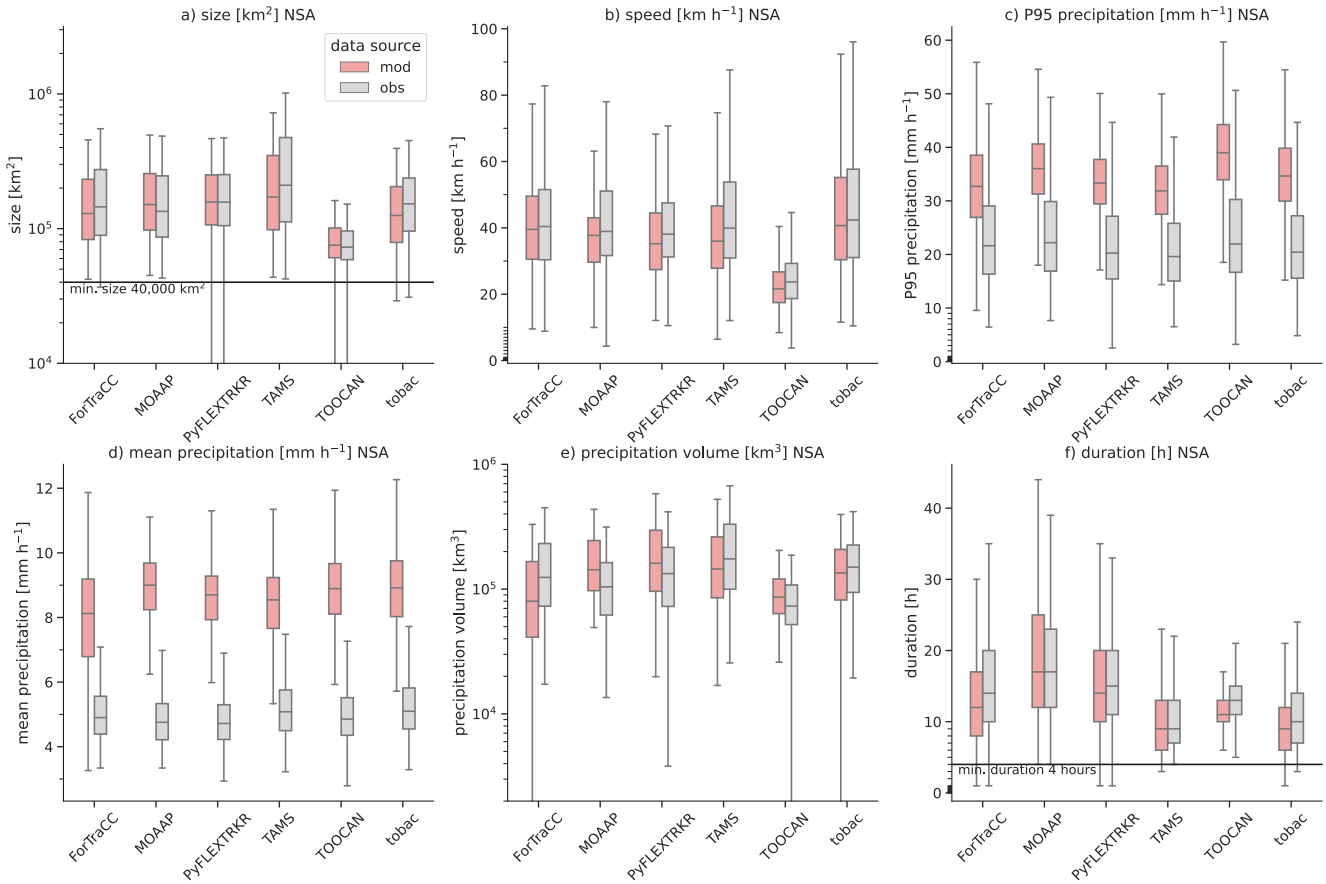

Figure 5. Observed (gray) and simulated (red) Mesoscale convective system (MCS) characteristics in the NSA region. Shown are MCS (a) peak size, (b) median speed, (c) lifetime-maximum 95th percentile precipitation rate (P95), (d) mean precipitation rate, (e) mean precipitation volume, and (f) duration distributions. Precipitation statistics only consider grid-scale precipitation rates larger than 2 mm hr^{-1}. The box width shows the interquantile range with the median indicated as a horizontal line within the box. The whiskers extend to the maximum or minimum data point or to 1.5 times the interquantile range dependent on which one is smaller.

except MOAAP and PyFLEXTRKR feature slightly shorter-lived modeled MCSs in all regions relative to observed (Figure 6f). MOAAP systematically detects the longest-lived MCSs while tobac has the shortest-lived systems as exemplified using the idealized case studies of Section 3.1.

3.5. MCS Life Cycles

MCSs are known to go through different life cycle stages that are frequently differentiated into (a) a growth stage, where individual storms deepen with merging anvils to create a larger CCC, (b) a mature stage in which the MCS reaches maximum size, convective regions are the most spatially connected, and stratiform precipitation is maximized, and (c) a decay stage where the MCS precipitation decreases and becomes the system becomes downdraft dominated (Machado et al., 1998).

Here we investigate how different trackers depict the MCS life cycle and what impact the tracker formulation has on the model evaluation. We focus the analysis on short-lived (duration ≤12 hr) and long-lived (duration between >16 hr and ≤20 hr) MCSs. We only consider MCSs that initiate with a ≤241 K T_b area of less than 40,000 km^2 to minimize the effect of MCSs that split off from an existing MCS. We chose the NSA region as a representative example (Figure 7) while results for other regions are shown in Figures S10–S16 of Supporting Information S1. All trackers show a rapid expansion of the anvil area after MCS initiation, followed by a stabilization of the area, and a decay of the anvil size (Figures 7a–7f). However, the shapes of these curves vary depending on the tracker, being close to bell-shaped in TOOCAN but skewed in tobac and TAMS (especially visible in short-lived MCSs). Note that the bend in the tail of some of the long-lived MCS distributions at hour 16 is artificial since we include MCSs that live between 16 and 20 hr in the statistics. The differences between the peak size in the short-lived and

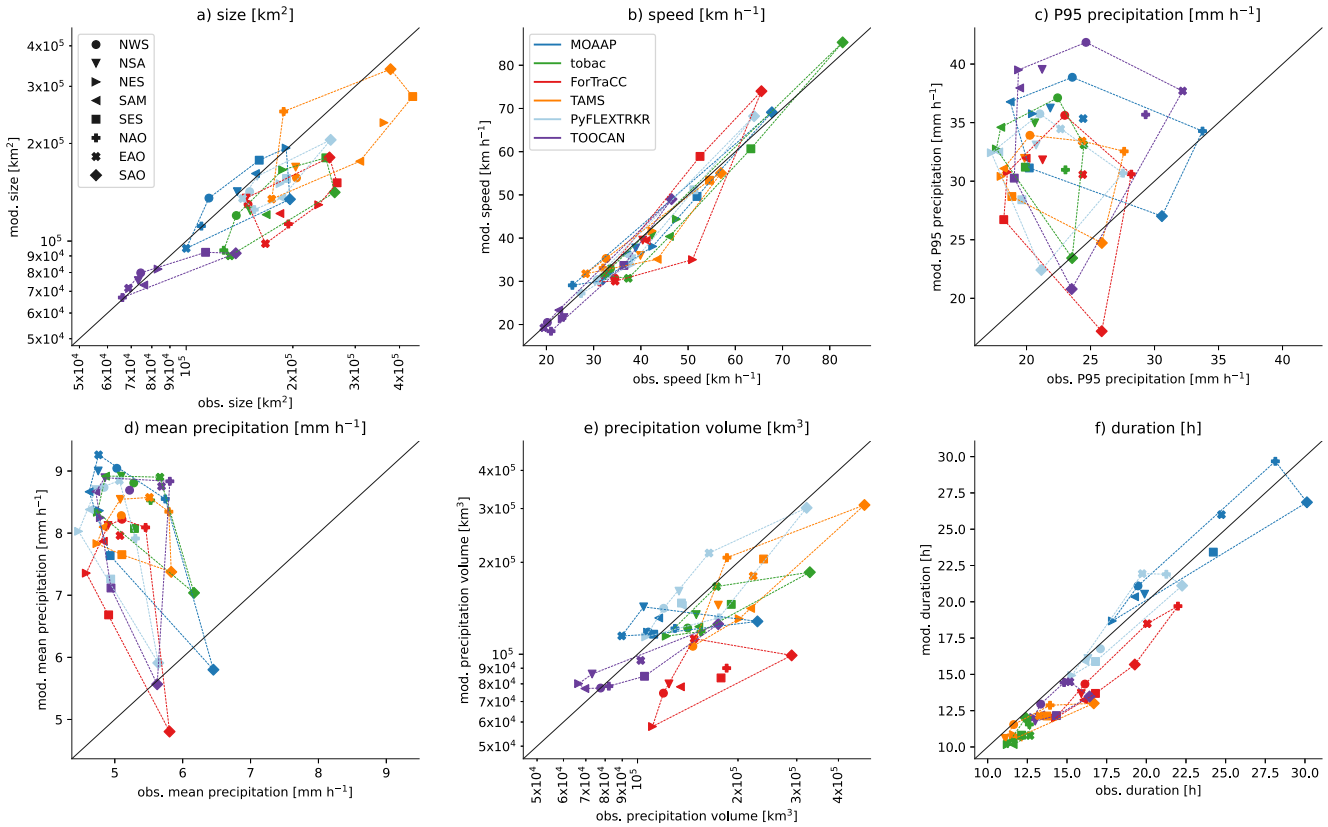

Figure 6. Observed (*x*-axis) and modeled (*y*-axis) median Mesoscale convective system (MCS) characteristics for (a) peak size, (b) median speed, (c) lifetime-maximum 95th percentile precipitation rate (P95), (d) mean precipitation rate, (e) mean precipitation volume, and (f) duration. Results from different trackers are shown in different colors (see legend in panel b). Results for different regions are shown with varying symbols (see legend in panel a). The dotted lines are convex hulls that incorporate all regional variations of MCS characteristics (polygons that incorporate all same colored points). The diagonal line represents a 1:1 relationship between observed and simulated MCS characteristics.

long-lived storms are also noteworthy. In MOAAP and ForTraCC, these two categories of storms reach similar peak sizes while TAMS has much smaller short-lived storms than long-lived ones. All trackers show similar MCS size evolution in the observations and the simulations.

Larger differences between observed and simulated MCS life cycles exist for P95 precipitation with modeled systems producing much heavier rainfall (Figures 7g–7l; similar to what is shown in Figures 5c and 6c). TOOCAN features P95 precipitation that rapidly intensifies within hours after initiation and that has a long decay period afterward. MOAAP, PyFLEXTRKR, and ForTraCC feature similar behavior but with a much less pronounced increase and decay, while tobac and TAMS do not show the initial intensification of P95 precipitation. There are likely two reasons for these differences. First, the results include MCSs that initiate by splitting from other systems. Second, TAMS, tobac, and MOAAP might miss the earliest few hours of the MCS life cycle as is shown in Section 3.1.

MCS life cycles of precipitation volume (Figures 7m–7r) are similar to those of MCS anvil size since these two properties are closely connected. The differences between tracking schemes are also similar. Most trackers show good agreement between simulated and observed precipitation volumes although peak volumes of short-lived systems can differ depending on the tracker.

For the long-lasting systems, three trackers (TOOCAN, PyFLEXTRKR and MOAAP) show similar bell-shaped life cycles of rain volume and further reveal a peak in the simulated precipitation volume that is ~2 hr earlier than that in the observations, possibly indicative of a systematic bias simulated life cycles. However, the other trackers do not exhibit a noticeable time lag between the modeled and observed precipitation volume peaks.

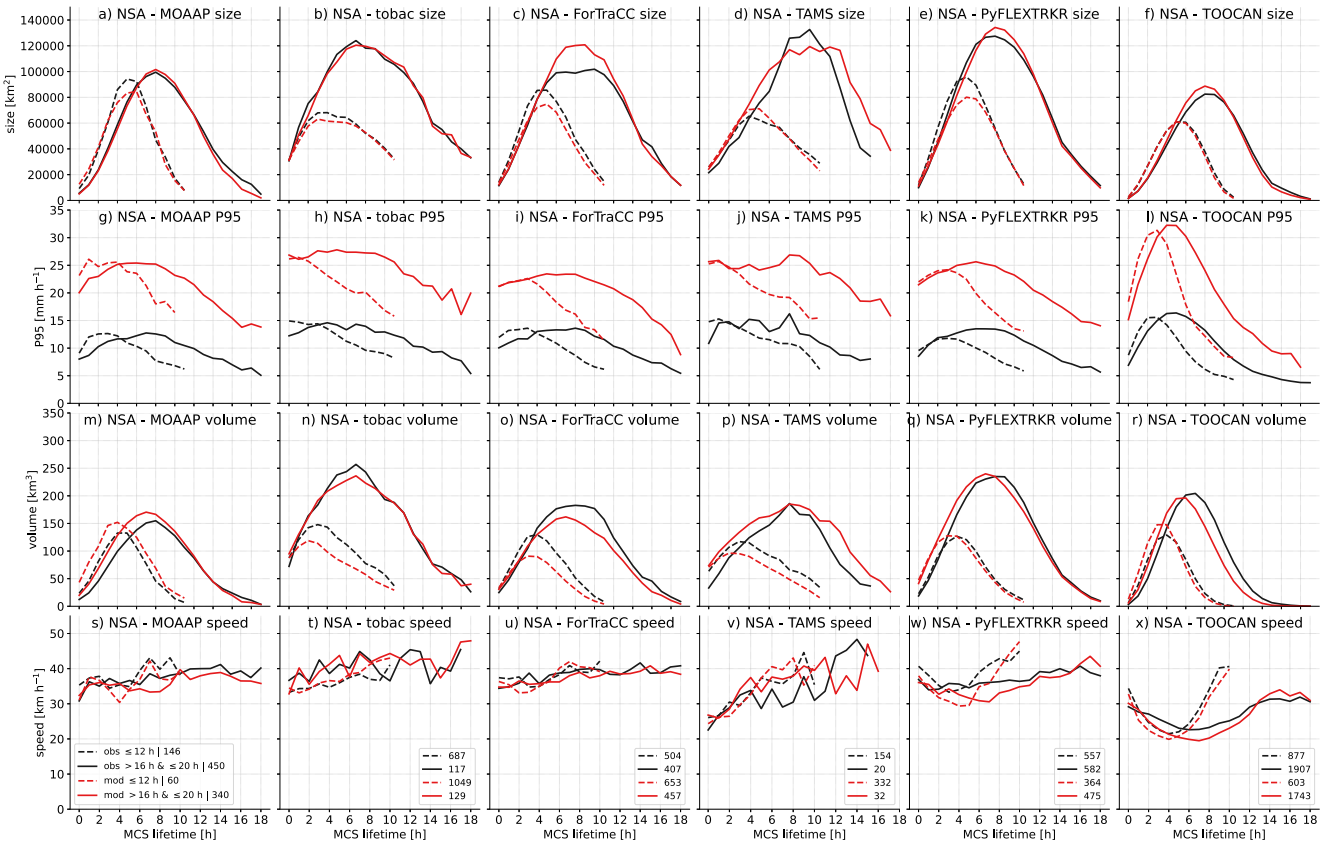

Figure 7. Evolution of short-lived (≥4 hr and ≤12 hr; dashed lines) and long-lived (≥16 hr and ≤20 hr) mean Mesoscale convective system (MCS) size (first row), 95th percentile precipitation rate (second row), precipitation volume (third row), and speed (bottom row) in the NSA region. Mean observed and simulated characteristics are shown with red and black lines, respectively. Results from different trackers are shown in columns. The number of MCSs in each analysis is shown in the legend. Only initiations that start with cloud shields smaller than 40,000 km^2 are incorporated to reduce the effect of MCS splits on the statistics.

Lastly, MCS speed slightly increases over time in most trackers (particularly in TAMS), though PyFLEXTRKR and TOOCAN exhibit a decrease followed by an increase (Figures 7s–7x). This is likely due to how the splitting and merging are handled in these two trackers and due to the usage of MCS cloud shield geometric center displacements to calculate movement speed. All trackers show that observed and simulated MCS movement speeds are in good agreement.

We emphasize that the model-observation differences of the composite MCS life cycle characteristics are more consistent among the trackers than the evolution of the composite values themselves. Except for simulating approximately twice as high P95 precipitation, the model is able to simulate the evolution of the observed MCS cloud size, rainfall volume, and movement speed well, regardless of which tracker is used. This supports these metrics as being robust for evaluating the performance of the simulations.

3.6. MCS Initiation by Location

Regional hotspots of MCS initiation are identified by most trackers over the northeast Brazilian coast, a few hundred kilometers inland from this coastline, over the Guiana Highlands, the western slopes of the Colombian Andes, and the eastern slopes of the Peruvian, Bolivian, and Argentinian Andes (Figure 8). However, the frequency of MCS initiation at these hotspots can vary by an order of magnitude. TAMS has the lowest MCS initiation frequency while TOOCAN has the highest, which is true for observed and simulated MCSs. The general spatial pattern of MCS initiation is similar between the observations and simulations but there are differences in initiation frequency (Figure 8 bottom row). These model-observation differences strongly depend on the tracker formulation with generally lesser absolute differences when MOAAP and PyFLEXTRKR are used (relative differences are largest in TAMS; not shown) and mostly positive differences (more modeled initiations) when

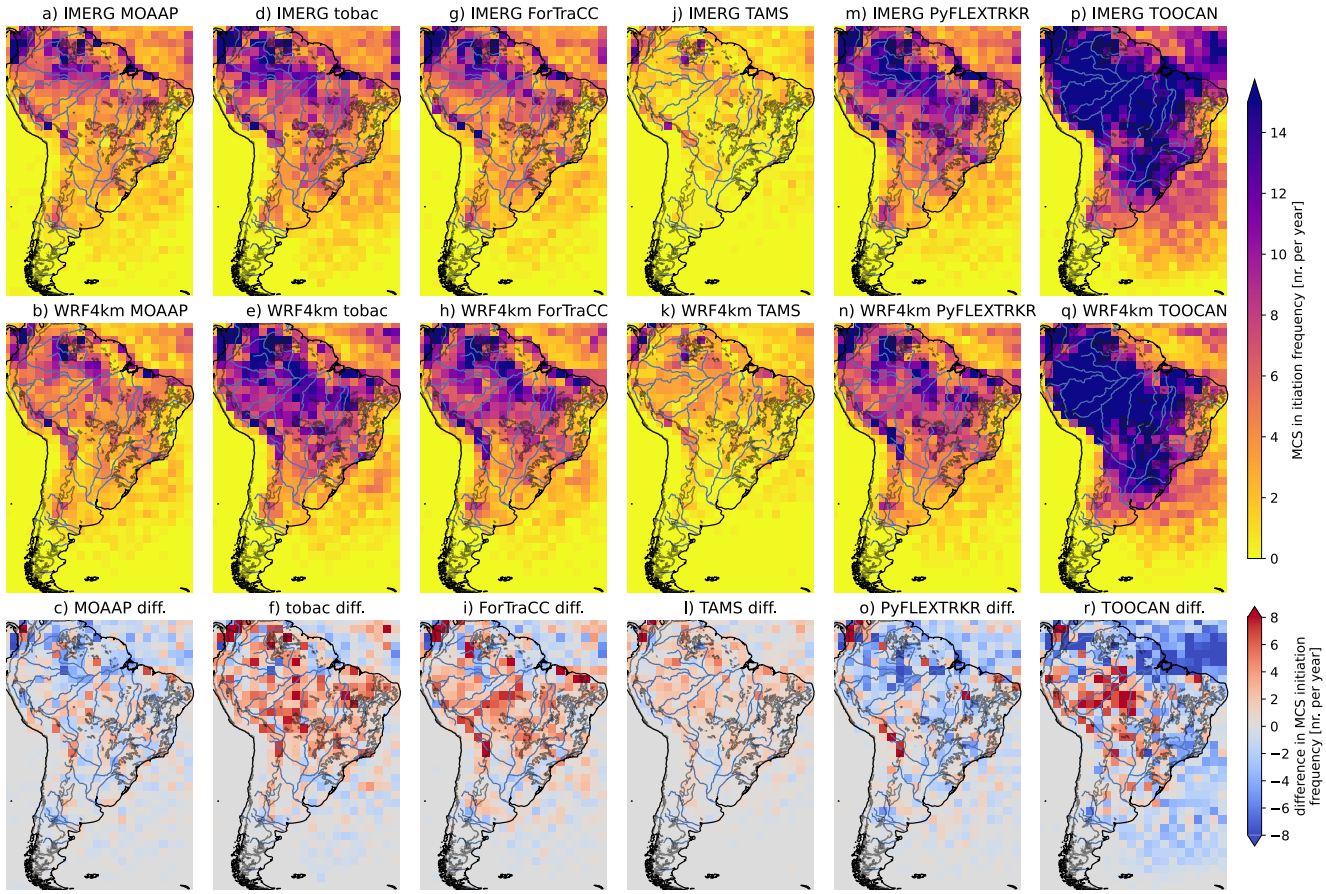

Figure 8. Initiation frequency of Mesoscale convective systems (MCSs) in 2° × 2° regions based on observations (top row), the simulation (middle row), and their difference (model minus observed; bottom row). Only initiations that start with cloud shields smaller than 40,000 km^2 are incorporated to reduce the effect of MCS splits on the statistics. We consider an initiation to be the geometric center of the MCS cloud shield at the time of its first detection.

using tobac, TAMS, and ForTraCC. There are only a few regions where all trackers agree on the sign of the difference. Systematically higher model frequencies are found along the eastern slopes of the Bolivian Andes, in northwestern Colombia, and over the southern Amazon Basin. Consistent model underestimation of MCS initiation frequency among trackers is rarer and only occurs off the coast of northeastern Brazil.

3.7. MCS Frequency by Location

The frequency of MCSs is less sensitive to the tracker formulation than the frequency of MCS initiation (compare Figure 9 with Figure 8). Initiation frequencies only consider the grid cell with the geometric center of the MCS cloud shield during its first detection, which is highly dependent on the tracker formulation as shown in Figures 4 and 8. In contrast, MCS frequencies consider the ≤241 K T_b footprint over the entire MCS life cycle where each MCS is only counted once in each grid cell eliminating the double counting of long-lived slow-moving systems. For instance, a tracker that produces many small and short-lived MCSs can result in the same MCS frequency as a tracker that produces few, large-scale, and long-lived MCSs. TOOCAN has the lowest MCS frequencies, likely because of the smaller systems that are identified, while TAMS has the highest frequencies, which is probably related to its use of convex hulls. The differences between modeled and observed frequencies are more similar between trackers than those for MCS initiation with all trackers agreeing on less simulated MCS frequencies over ocean regions, southern South America, and northeastern South America. Larger uncertainties exist in the Amazon basin.

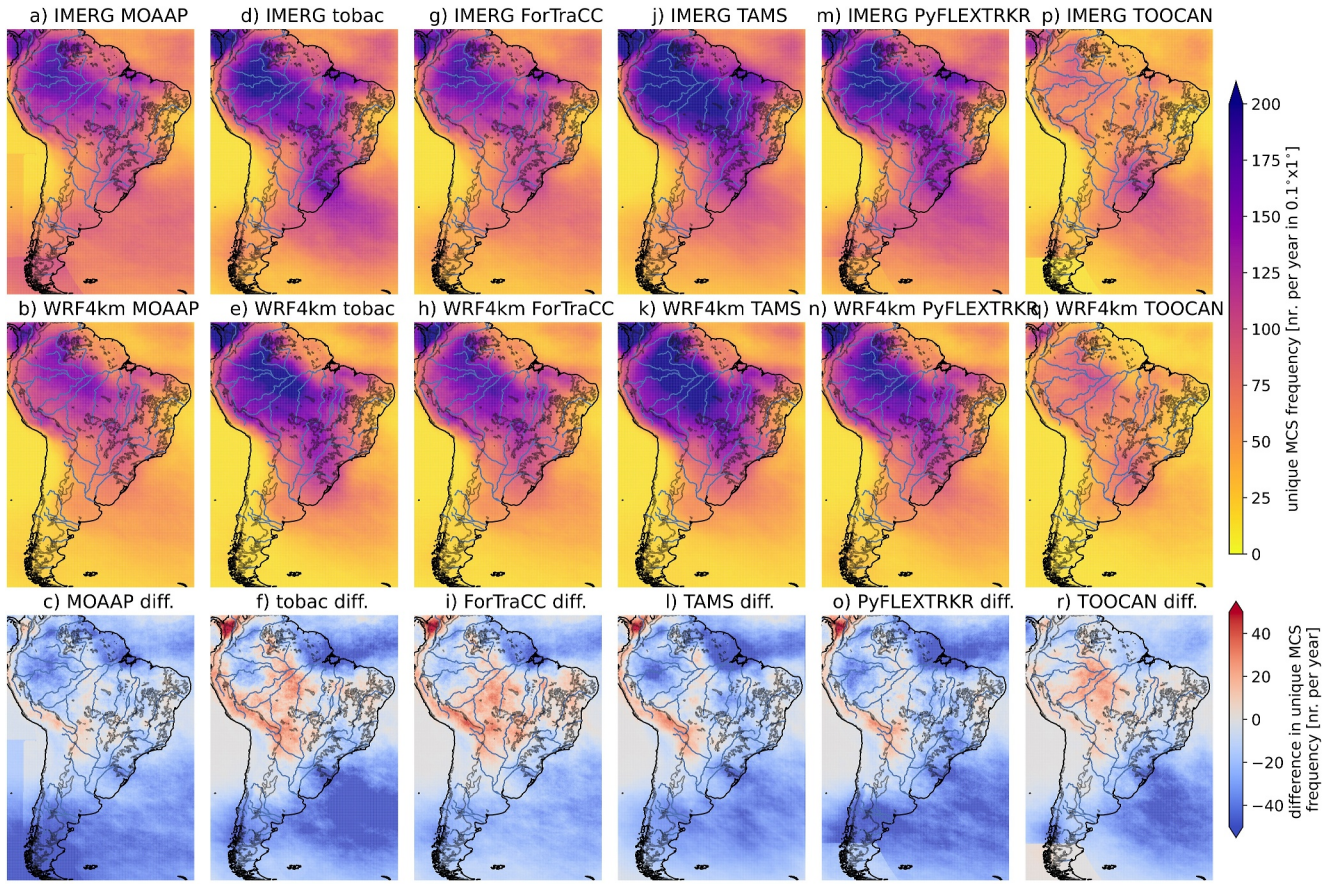

Figure 9. Frequency of Mesoscale convective systems (MCSs) in $0.1° \times 0.1°$ regions based on observations (top row), the simulation (middle row), and their difference (model minus observed; bottom row). We use the extent of the MCS cloud shield in this calculation and each MCS is only counted once in each cell (i.e., if a slow-moving MCS occupies a grid cell for 10 hr, it is only counted once).

3.8. MCS Contribution to Total Precipitation by Location

Lastly, we analyze the contribution of MCSs to total annual rainfall (Figure 10). This analysis is affected by a large range of MCS characteristics including frequency, size, longevity, and precipitation rates. Applying different trackers results in a wide range of MCS contributions to total precipitation with PyFLEXTRKR producing the highest contributions while TOOCAN and ForTraCC produce the lowest. There is agreement on the continental maximum of MCS contribution over the La Plata basin which varies between ~60% in ForTraCC and TOOCAN to more than 80% in PyFLEXTRKR. There is also agreement among trackers that the simulation produces a lesser fraction of precipitation from MCSs over large parts of the study region, particularly over Southern Argentina and Chile and over the equatorial and southern Atlantic. The differences over Patagonia are influenced by extratropical cyclones and atmospheric rivers that may produce erroneous MCS identifications in observations with the tracker definitions used since MCSs are not expected frequently in this region.

4. Conclusions

We compared the results of six MCS trackers to understand how sensitive MCS statistics are to the formulation of the tracking algorithm and what impact this has on the evaluation of km-scale regional climate model simulations over South America. We performed this analysis for three water years (June to May) over South America, each differing concerning their El Niño phase, but only focused on multi-year average statistics to limit the presented information.

Uncertainties in observed precipitation present difficulties in interpreting model-observation comparisons. There are documented high biases in the occurrence frequencies of light precipitation rates in GPM-IMERG over the

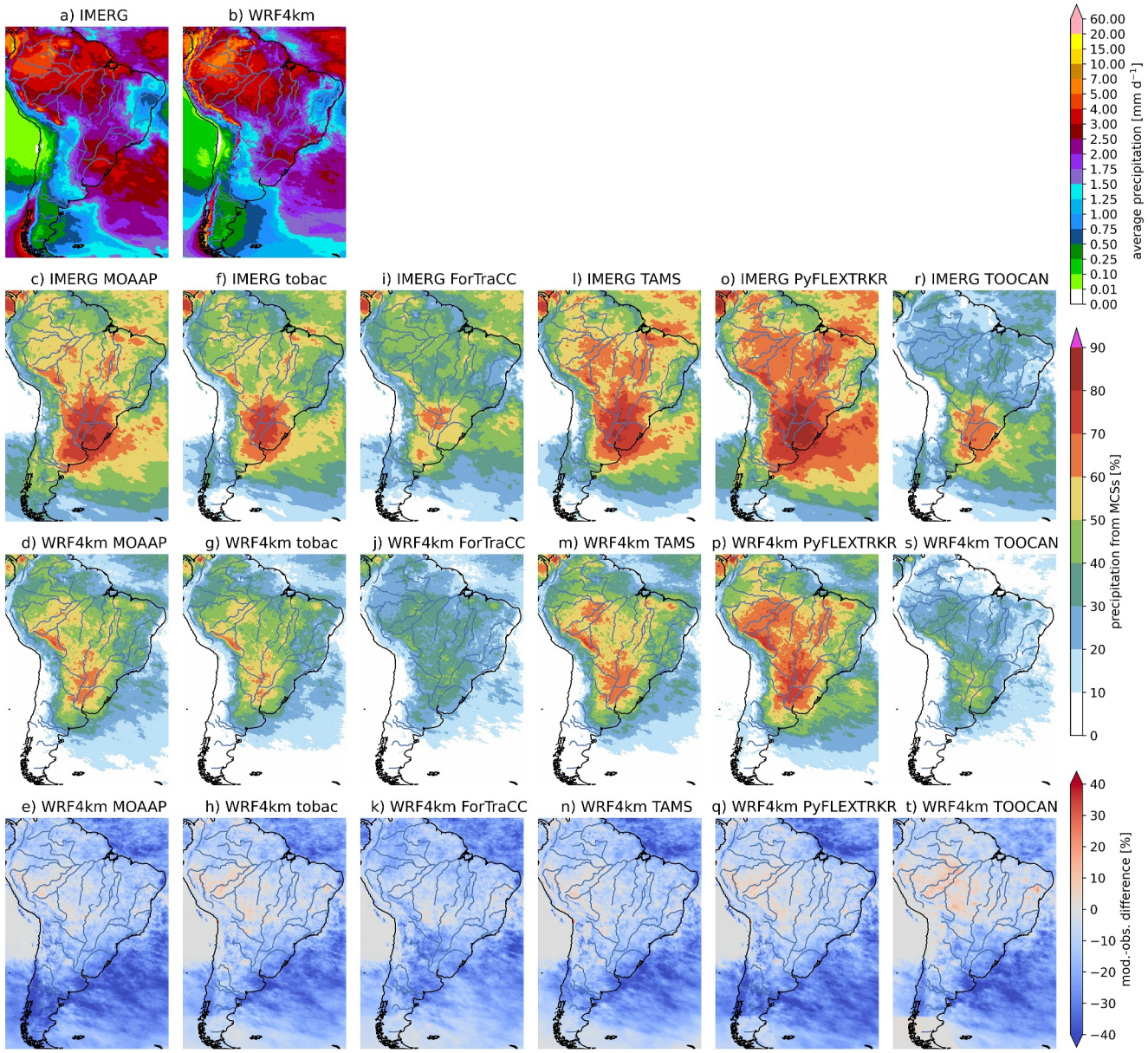

Figure 10. (a) Average observed and (b) simulated precipitation. Fraction of Mesoscale convective system to total precipitation based on observations (second row from top), simulations (third row from top) and their differences (model minus observed; bottom row).

Amazon basin (Dominguez et al., 2024; Rozante et al., 2018) and La Plata basin (Zhang et al., 2021), though heavy precipitation rates have also been shown to be biased high in a kilometer-scale season-long WRF simulation over the La Plata basin (Zhang et al., 2021). Particularly striking are the large land-ocean contrasts in simulated versus observed MCS statistics, which are partly related to differences in the GPM-IMERG precipitation retrieval over ocean and land.

The following points summarize our findings concerning the two leading questions about how MCS tracker formulation affects MCS statistics and how this uncertainty impacts the evaluation of km-scale climate models.

- MCS frequencies, as well as certain MCS characteristics, strongly depend on the tracker formulation, even when using the same MCS criteria. This means that statistics on the frequency, size, duration, or contribution to total precipitation of MCSs are susceptible to the tracker algorithm in use and should be interpreted

Table 2
Characteristics of Mesoscale Convective System (MCS) Properties for Each Tracker Relative to Other Trackers

	Frequency	Peak size	Duration	PR P95	Initiation detection	Termi-nation detection	Initiation frequency	Rainfall contribution
ForTraCC	Average	Average	Average	Low	At start	At end	Average	Low
MOAAP	Low	Average	Long	Above average	Sometimes late	At end	Below average	Average
PyFLEXTRKR	Average	Average	Average	Average	At start	At end	Average	High
TAMS	Average	Large	Short	Average	Late	Sometimes early	Low	Average
tobac	High	Average	Short	Average	Sometimes late	Sometimes early	High	Average
TOOCAN	High	Small	Short	High	At start	At end	Average	Low

Note. MCS occurrence frequency, peak size, duration, 95th percentile precipitation rate (PR P95), initiation detection timing, termination detection, initiation frequency, and contribution to total rainfall are shown from left to right. Note that relative MCS characteristics can vary by region and input data set (i.e., modeled vs. observed).

accordingly. A main source of uncertainty is the treatment of cloud system segmentation including splitting and merging in different tracking algorithms, in agreement with previous findings (Müller et al., 2022).

- The dependence of MCS characteristics on the tracker formulation is fairly systematic across geographical locations although some regional differences exist. Table 2 provides an overview of tracker MCS characteristics relative to average characteristics across the tracker ensemble. This should not be interpreted as a ranking of tracking schemes since no reference data set could be used to infer a quantitative assessment of derived MCS characteristics.
- The tracker formulation can affect the evaluation of model performance in profound ways. Agreements amongst the tracking schemes on the sign of model-observational differences are typically the exception, which is in part caused by the good performance (i.e., small differences) of the 4 km simulation in capturing many observed MCS characteristics. Statistics that are highly sensitive to the tracker formulation are the MCS frequency including the initiation frequency, the ratio of MCS to total precipitation, and MCS size and duration.
- Comparisons of observed and modeled MCS lifecycle characteristics (e.g., the development of cloud shield size, movement speed, and rain volume) are more robust and less dependent on the tracker used. Comparisons of MCS frequency differences by location and differences in MCS contributions to total precipitation are generally more robust, though disagreement exists for some locations such as the southern Amazon basin.

This study only focused on MCS tracker formulation uncertainty, neglecting uncertainties that stem from differences in how MCSs are defined. We use an arbitrary definition of MCSs that results in similar statistics as in published literature (Feng et al., 2021), but a modified definition could be warranted depending on the research question being asked and should be the focus of future assessments. Additionally, users of a particular tracker should configure it to fit the purpose of their work.

It is important to mention that while all trackers use a common definition of MCSs, there are settings in the tracking algorithm that have a substantial impact on the presented results. The segmentation treatment that treads the splitting and merging of MCSs is one of the most important procedures in MCS tracking and has a significant impact on MCS statistics such as frequency, size, and duration. Additionally, the method used to track MCSs in time (point-tracking: tobac; 2D overlap: ForTraCC, PyFLEXTRKR, TAMS, MOAAP; 3D segmentation: TOOCAN) can have also a substantial effect. Importantly, we find that the model versus observation differences are more consistent among the trackers than the MCS metrics meaning that model evaluation studies are less prone to MCS tracker formulation uncertainties.

Future work could also expand the present study to global scales to improve our understanding of tracker formulation uncertainties in regions that have different atmospheric conditions than those found in South America. Additionally, a better understanding of tracker formulation impacts on inter-annual variability and long-term trends in MCS statistics would be valuable.

Performing the analysis over a regional domain introduces uncertainties about the impact of the domain boundary on the MCS statistics. Most analysis regions are far away from the lateral boundaries and should not be affected except for EAO, NWS, and SAO. Particularly, the low difference of simulated MCS frequencies in the EAO

region might be related to a lack of MCSs that enter this region through the eastern boundary. Additionally, boundary effects on our results should be small since most MCSs initiate over land regions and the vast majority of those MCSs do not live long enough to reach the boundary.

While we do not recommend that all MCS tracking analyses need to use multiple tracking schemes because of the complexity this would introduce, we stress that studies using a single scheme in isolation, or in comparison with another study with a different tracker, have to be interpreted with caution. Context from many complementary methods can improve understanding of any single method's strengths and weaknesses for a specific application, providing robust support for generalized scientific conclusions.

Data Availability Statement

ERA-5 reanalysis data can be accessed from the Copernicus Climate Data Store (Copernicus, 2023). The GPM_MERGIR brightness temperature observations can be downloaded from the NASA server (GPM-MERGIR, 2023) and GPM-IMERG precipitation data can also be accessed from NASA (GPM-IMERG, 2023). The 4 km grid spacing WRF simulation data can be downloaded via Globus file transfer (SAAG, 2023). The MCS mask files from each tracker can also be accessed via Globus (Prein, Feng, et al., 2023). The MOAAP code can be downloaded from GitHub (Prein, 2023b). PyFLEXTRKR can also be accessed on GitHub (Zhe Fent, 2023). The TAMS code can be downloaded via GitHub (Núñez Ocasio, K. M. and Moon, Z. L., 2023). The tobac code can be downloaded from GitHub (Heikenfeld et al., 2023). The TOOCAN and ForTTraCC codes are not open source. Questions about TOOCAN should be directed to Rémy Roca (remy.roca@cnrs.fr) or Thomas Fiolleau (thomas.fiolleau@legos.obs-mip.fr) and questions about ForTTraCC to Amanda Rehbein (amanda.rehbein@usp.br). The code that was used for the analyses and visualizations in this paper can also be accessed from GitHub (Prein, 2023a).

Acknowledgments
The NSF National Center for Atmospheric Research (NCAR) is a major facility sponsored by the National Science Foundation (NSF) under Cooperative Agreement #1852977. We would like to acknowledge high-performance computing support from Cheyenne (https://doi.org/10.5065/D6RX99HX) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation, and the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy (DOE) Office of Science User Facility supported by the Office of Science of the US DOE under contract DEAC02-05CH11231. AR is also thankful for the supports from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) Grants 2021/07992-5 and 2022/05622-9. ZF and AV were supported by the U.S. DOE Office of Science Biological and Environmental Research (BER) as part of the Regional and Global Modeling Analysis (RGMA) and Atmospheric System Research (ASR) programs. Pacific Northwest National Laboratory is operated by Battelle for the US DOE under contract DEAC05-76RLO1830. AFP was supported by the FRONTIER project funded by the Research Council of Norway (project number 301777). YM was supported by the National Science Foundation Grants (AGS 1937899) and the NASA Future Investigators in NASA Earth and Space Science and Technology program (award number: 80NSSC24K0012).

References

Alcântara, C. R., Dias, M. A. S., Souza, E. P., & Cohen, J. C. (2011). Verification of the role of the low level jets in Amazon squall lines. *Atmospheric Research*, *100*(1), 36–44. https://doi.org/10.1016/j.atmosres.2010.12.023

Anselmo, E. M., Machado, L. A., Schumacher, C., & Kiladis, G. N. (2021). Amazonian mesoscale convective systems: Life cycle and propagation characteristics. *International Journal of Climatology*, *41*(7), 3968–3981. https://doi.org/10.1002/joc.7053

Anselmo, E. M., Schumacher, C., & Machado, L. A. (2020). The Amazonian low-level jet and its connection to convective cloud propagation and evolution. *Monthly Weather Review*, *148*(10), 4083–4099. https://doi.org/10.1175/mwr-d-19-0414.1

Bluestein, H. B., McCaul, E. W., Jr., Byrd, G. P., Walko, R. L., & Davies-Jones, R. (1990). An observational study of splitting convective clouds. *Monthly Weather Review*, *118*(6), 1359–1370. https://doi.org/10.1175/1520-0493(1990)118<1359:aososc>2.0.co;2

Boer, E., & Ramanathan, V. (1997). Lagrangian approach for deriving cloud characteristics from satellite observations and its implications to cloud parameterization. *Journal of Geophysical Research*, *102*(D17), 21383–21399. https://doi.org/10.1029/97jd00930

Carvalho, L., Jones, C., & Liebmann, B. (2004). The South Atlantic Convergence Zone: Persistence, intensity, form, extreme precipitation and relationships with intraseasonal activity. *Journal of Climate*, *17*(1), 88–108. https://doi.org/10.1175/1520-0442(2004)017<0088:tsaczi>2.0.co;2

Cecil, D. J., & Blankenship, C. B. (2012). Toward a global climatology of severe hailstorms as estimated by satellite passive microwave imagers. *Journal of Climate*, *25*(2), 687–703. https://doi.org/10.1175/jcli-d-11-00130.1

Cecil, D. J., Buechler, D. E., & Blakeslee, R. J. (2015). TRMM LIS climatology of thunderstorm occurrence and conditional lightning flash rates. *Journal of Climate*, *28*(16), 6536–6547. https://doi.org/10.1175/jcli-d-15-0124.1

Clark, A. J., Bullock, R. G., Jensen, T. L., Xue, M., & Kong, F. (2014). Application of object-based time-domain diagnostics for tracking precipitation systems in convection-allowing models. *Weather and Forecasting*, *29*(3), 517–542. https://doi.org/10.1175/waf-d-13-00098.1

Cohen, J. C., Silva Dias, M. A., & Nobre, C. A. (1995). Environmental conditions associated with Amazonian squall lines: A case study. *Monthly Weather Review*, *123*(11), 3163–3174. https://doi.org/10.1175/1520-0493(1995)123<3163:ecawas>2.0.co;2

Collins, N., Theurich, G., DeLuca, C., Suarez, M., Trayanov, A., Balaji, V., et al. (2005). Design and implementation of components in the Earth System Modeling Framework. *The International Journal of High Performance Computing Applications*, *19*(3), 341–350. https://doi.org/10.1177/1094342005056120

Copernicus. (2023). ERA5 hourly data on single levels from 1979 to present [Dataset]. *Copernicus Climate Data Store*. https://doi.org/10.24381/cds.adbb2d47

Cotton, W. R., & Anthes, R. A. (1992). *Storm and cloud dynamics*. Academic Press.

Cotton, W. R., Lin, M.-S., McAnelly, R. L., & Tremback, C. J. (1989). A composite model of mesoscale convective complexes. *Monthly Weather Review*, *117*(4), 765–783. https://doi.org/10.1175/1520-0493(1989)117<0765:acmomc>2.0.co;2

Ćurić, M., Janc, D., & Vučković, V. (2009). The influence of merging and individual storm splitting on mesoscale convective system formation. *Atmospheric Research*, *93*(1–3), 21–29. https://doi.org/10.1016/j.atmosres.2008.10.018

Davis, C. A., Brown, B. G., Bullock, R., & Halley-Gotway, J. (2009). The method for object-based diagnostic evaluation (MODE) applied to numerical forecasts from the 2005 NSSL/SPC Spring Program. *Weather and Forecasting*, *24*(5), 1252–1267. https://doi.org/10.1175/2009waf2222241.1

Derin, Y., Kirstetter, P.-E., & Gourley, J. J. (2021). Evaluation of IMERG satellite precipitation over the land–coast–ocean continuum. Part I: Detection. *Journal of Hydrometeorology*, *22*(11), 2843–2859. https://doi.org/10.1175/jhm-d-21-0058.1

Dias, P. S., Schubert, W. H., & DeMaria, M. (1983). Large-scale response of the tropical atmosphere to transient convection. *Journal of the Atmospheric Sciences*, *40*(11), 2689–2707. https://doi.org/10.1175/1520-0469(1983)040<2689:lsrott>2.0.co;2

Dominguez, F., Rasmussen, R., Liu, C., Ikeda, K., Prein, A., Varble, A., et al. (2024). Advancing South American Water and Climate Science through multidecadal convection-permitting modeling. *Bulletin of the American Meteorological Society*, *105*(1), E32–E44. https://doi.org/10.1175/BAMS-D-22-0226.1

Durkee, J. D., & Mote, T. L. (2010). A climatology of warm-season mesoscale convective complexes in subtropical South America. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *30*(3), 418–431. https://doi.org/10.1002/joc.1893

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., 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., Leung, L. R., Hagos, S., Houze, R. A., Burleyson, C. D., & Balaguru, K. (2016). More frequent intense and long-lived storms dominate the springtime trend in central us rainfall. *Nature Communications*, *7*(1), 13429. https://doi.org/10.1038/ncomms13429

Feng, Z., Leung, L. R., Liu, N., Wang, J., Houze, R. A., Jr., 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), e2020JD034202. https://doi.org/10.1029/2020jd034202

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

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

GPM-IMERG. (2023). GPM-IMERG half-hourly precipitation retrieval [Dataset]. *NASA*. Retrieved from https://gpm.nasa.gov/data/imerg

GPM-MERGIR. (2023). GPM-IMERG half-hourly brightness temperature observations [Dataset]. *NASA*. Retrieved from https://disc.gsfc.nasa.gov/datasets/GPM_MERGIR_1/summary

Guilloteau, C., & Foufoula-Georgiou, E. (2020). Multiscale evaluation of satellite precipitation products: Effective resolution of IMERG. In *Satellite precipitation measurement* (pp. 533–558). Springer.

Hartman, A. T. (2021). Tracking mesoscale convective systems in central equatorial Africa. *International Journal of Climatology*, *41*(1), 469–482. https://doi.org/10.1002/joc.6632

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

Heikenfeld, M., Marinescu, P. J., Freeman, S. W., Kukulies, J., Jones, W. K., Senf, F., et al. (2023). Tracking and object-based analysis of clouds (tobac) [Software]. *Zenodo*. https://doi.org/10.5281/zenodo.7335133

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

Hong, S.-Y., Noh, Y., & Dudhia, J. (2006). A new vertical diffusion package with an explicit treatment of entrainment processes. *Monthly Weather Review*, *134*(9), 2318–2341. https://doi.org/10.1175/mwr3199.1

Houze, R. A., Jr. (2004). Mesoscale convective systems. *Reviews of Geophysics*, *42*(4), RG4003. https://doi.org/10.1029/2004rg000150

Houze, R. A., Jr. (2014). *Cloud dynamics*. Academic Press.

Houze, R. A., Jr. (2018). 100 years of research on mesoscale convective systems. *Meteorological Monographs*, *59*, 17.1–17.54. https://doi.org/10.1175/amsmonographs-d-18-0001.1

Hu, H., Leung, L. R., & Feng, Z. (2020). Observed warm-season characteristics of MCS and non-MCS rainfall and their recent changes in the central United States. *Geophysical Research Letters*, *47*(6), e2019GL086783. https://doi.org/10.1029/2019gl086783

Huffman, G., Bolvin, D., Braithwaite, D., Hsu, K., Joyce, R., Xie, P., & Yoo, S. (2019). *Algorithm theoretical basis document (ATBD) version 06. NASA Global Precipitation Measurement (GPM) Integrated Multi-Satellite Retrievals for GPM (IMERG)*. NASA. Retrieved from https://pmm.nasa.gov/data-access/downloads/gpm

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., Xie, P., & Yoo, S.-H. (2015). NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG). *Algorithm theoretical basis document (ATBD) version*, *4*(26), 30.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., & Collins, W. D. (2008). Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *Journal of Geophysical Research*, *113*(D13), D13103. https://doi.org/10.1029/2008jd009944

Iturbide, M., Gutiérrez, J. M., Alves, L. M., Bedia, J., Cimadevilla, E., Cofiño, A. S., et al. (2020). An update of IPCC climate reference regions for subcontinental analysis of climate model data: Definition and aggregated datasets. *Earth System Science Data Discussions*, *2020*(4), 1–16. https://doi.org/10.5194/essd-12-2959-2020

Janowiak, J., Joyce, B., & Xie, P. (2017). NCEP/CPC L3 half hourly 4 km Global (60°S–60°N) Merged IR V1, Edited by Andrey Savtchenko [Dataset]. *Goddard Earth Sciences Data and Information Services Center (GES DISC)*. https://doi.org/10.5067/P4HZB9N27EKU

Jones, C. (2019). Recent changes in the South America low-level jet. *Npj Climate and Atmospheric Science*, *2*(1), 20. https://doi.org/10.1038/s41612-019-0077-5

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., Li, P., et al. (2023). Mesoscale convective systems in the third pole region: Characteristics, mechanisms and impact on precipitation. *Frontiers in Earth Science*, *11*, 1143380. https://doi.org/10.3389/feart.2023.1143380

Kumjian, M. R., Gutierrez, R., Soderholm, J. S., Nesbitt, S. W., Maldonado, P., Luna, L. M., et al. (2020). Gargantuan hail in Argentina. *Bulletin of the American Meteorological Society*, *101*(8), E1241–E1258. https://doi.org/10.1175/bams-d-19-0012.1

Lin, Y., & Mitchell, K. E. (2005). 1.2 the NCEP stage II/IV hourly precipitation analyses: Development and applications. In *Proceedings of the 19th Conference Hydrology, American Meteorological Society, San Diego, CA, USA* (Vol. 10).

Machado, L., Rossow, W., Guedes, R., & Walker, A. (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<1630:lcvomc>2.0.co;2

Maddox, R. A. (1980). Mesoscale convective complexes. *Bulletin of the American Meteorological Society*, *61*(11), 1374–1387. https://doi.org/10.1175/1520-0477(1980)061<1374:mcc>2.0.co;2

Marengo, J. A., Soares, W. R., Saulo, C., & Nicolini, M. (2004). Climatology of the low-level jet east of the Andes as derived from the NCEP–NCAR reanalyses: Characteristics and temporal variability. *Journal of Climate*, *17*(12), 2261–2280. https://doi.org/10.1175/1520-0442(2004)017<2261:cotlje>2.0.co;2

Markowski, P., & Richardson, Y. (2011). *Mesoscale meteorology in midlatitudes*. John Wiley & Sons.

Marquis, J. N., Varble, A. C., Robinson, P., Nelson, T. C., & Friedrich, K. (2021). Low-level mesoscale and cloud-scale interactions promoting deep convection initiation. *Monthly Weather Review*, *149*(8), 2473–2495. https://doi.org/10.1175/mwr-d-20-0391.1

Mathon, V., & Laurent, H. (2001). Life cycle of Sahelian mesoscale convective cloud systems. *Quarterly Journal of the Royal Meteorological Society*, *127*(572), 377–406. https://doi.org/10.1002/qj.49712757208

Miguez-Macho, G., & Fan, Y. (2012). The role of groundwater in the Amazon water cycle: 1. Influence on seasonal streamflow, flooding and wetlands. *Journal of Geophysical Research*, *117*(D15), D15113. https://doi.org/10.1029/2012jd017539

Mulholland, J. P., Nesbitt, S. W., Trapp, R. J., Rasmussen, K. L., & Salio, P. V. (2018). Convective storm life cycle and environments near the Sierras de Córdoba, Argentina. *Monthly Weather Review*, *146*(8), 2541–2557. https://doi.org/10.1175/mwr-d-18-0081.1

Müller, S. K., Caillaud, C., Chan, S., De Vries, H., Bastin, S., Berthou, S., et al. (2022). Evaluation of Alpine-Mediterranean precipitation events in convection-permitting regional climate models using a set of tracking algorithms. *Climate Dynamics*, *61*(1–2), 1–19. https://doi.org/10.1007/s00382-022-06555-z

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

Nesbitt, S. W., Salio, P. V., Ávila, E., Bitzer, P., Carey, L., Chandrasekar, V., et al. (2021). A storm safari in subtropical South America: Proyecto RELAMPAGO. *Bulletin of the American Meteorological Society*, *102*(8), E1621–E1644. https://doi.org/10.1175/bams-d-20-0029.1

Nesbitt, S. W., & Zipser, E. J. (2003). The diurnal cycle of rainfall and convective intensity according to three years of TRMM measurements. *Journal of Climate*, *16*(10), 1456–1475. https://doi.org/10.1175/1520-0442-16.10.1456

Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., et al. (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *Journal of Geophysical Research*, *116*(D12), D12109. https://doi.org/10.1029/2010jd015139

Núñez Ocasio, K. M., Evans, J. L., & Young, G. S. (2020). 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. (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. (2023). Tracking algorithm for mesoscale convective systems (TAMS) [software]. *Zenodo*. https://doi.org/10.5281/zenodo.8393891

Poujol, B., Prein, A. F., & Newman, A. J. (2020). Kilometer-scale modeling projects a tripling of Alaskan convective storms in future climate. *Climate Dynamics*, *55*(11–12), 3543–3564. https://doi.org/10.1007/s00382-020-05466-1

Powers, J. G., Klemp, J. B., Skamarock, W. C., Davis, C. A., Dudhia, J., Gill, D. O., et al. (2017). The weather research and forecasting model: Overview, system efforts, and future directions. *Bulletin of the American Meteorological Society*, *98*(8), 1717–1737. https://doi.org/10.1175/bams-d-15-00308.1

Prein, A. F. (2023a). MCS tracker intercomparison analysis code. [Software]. *Zenodo*. https://doi.org/10.5281/zenodo.10022715

Prein, A. F. (2023b). Multi object analysis of atmospheric phenomenon (MOAAP) [Software]. *Zenodo*. https://doi.org/10.5281/zenodo.7561616

Prein, A. F., Feng, Z., Fiolleau, T., Moon, Z. L., Núñez Ocasio, K. M., Kukulies, J., et al. (2023). MCS mask files for observational and model data from multiple tracking schemes. [Dataset]. https://shorturl.at/rsGNT

Prein, A. F., Liu, C., Ikeda, K., Bullock, R., Rasmussen, R. M., Holland, G. J., & Clark, M. (2020). Simulating North American mesoscale convective systems with a convection-permitting climate model. *Climate Dynamics*, *55*(1), 95–110. https://doi.org/10.1007/s00382-017-3993-2

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., Ikeda, K., Liu, C., Clark, M. P., & Holland, G. J. (2017). The future intensification of hourly precipitation extremes. *Nature Climate Change*, *7*(1), 48–52. https://doi.org/10.1038/nclimate3168

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

Rasmussen, K. L., Chaplin, M., Zuluaga, M., & Houze, R. (2016). Contribution of extreme convective storms to rainfall in South America. *Journal of Hydrometeorology*, *17*(1), 353–367. https://doi.org/10.1175/jhm-d-15-0067.1

Rasmussen, K. L., & Houze, R., Jr. (2016). Convective initiation near the Andes in subtropical South America. *Monthly Weather Review*, *144*(6), 2351–2374. https://doi.org/10.1175/mwr-d-15-0058.1

Rehbein, A., & Ambrizzi, T. (2023a). ENSO teleconnections pathways in South America. *Climate Dynamics*, *61*(3–4), 1277–1292. https://doi.org/10.1007/s00382-022-06624-3

Rehbein, A., & Ambrizzi, T. (2023b). Mesoscale convective systems over the Amazon basin in a changing climate under global warming. *Climate Dynamics*, *61*(3–4), 1–13. https://doi.org/10.1007/s00382-022-06657-8

Rehbein, A., Ambrizzi, T., & Mechoso, C. R. (2018). Mesoscale convective systems over the Amazon basin. Part I: Climatological aspects. *International Journal of Climatology*, *38*(1), 215–229. https://doi.org/10.1002/joc.5171

Ribeiro, B. Z., & Bosart, L. F. (2018). Elevated mixed layers and associated severe thunderstorm environments in south and north America. *Monthly Weather Review*, *146*(1), 3–28. https://doi.org/10.1175/mwr-d-17-0121.1

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

Roca, R., & Ramanathan, V. (2000). Scale dependence of monsoonal convective systems over the Indian Ocean. *Journal of Climate*, *13*(7), 1286–1298. https://doi.org/10.1175/1520-0442(2000)013<1286:sdomcs>2.0.co;2

Rotunno, R., Klemp, J. B., & Weisman, M. L. (1988). A theory for strong, long-lived squall lines. *Journal of the Atmospheric Sciences*, *45*(3), 463–485. https://doi.org/10.1175/1520-0469(1988)045<0463:atfsll>2.0.co;2

Rozante, J. R., Vila, D. A., Barboza Chiquetto, J., Fernandes, A. d. A., & Souza Alvim, D. (2018). Evaluation of TRMM/GPM blended daily products over Brazil. *Remote Sensing*, *10*(6), 882. https://doi.org/10.3390/rs10060882

SAAG. (2023). Year-long simulations of different ENSO states over South America with the 4 km WRF model downscaling ERA5 [Dataset]. https://app.globus.org/file-manager?origin_id=0e88ae7f-297e-11ed-8dd0-9f359c660fbd&origin_path=%2F

Salio, P., Nicolini, M., & Saulo, A. C. (2002). Chaco low-level jet events characterization during the austral summer season. *Journal of Geophysical Research*, *107*(D24), ACL 32-1–ACL 32-17. https://doi.org/10.1029/2001jd001315

Salio, P., Nicolini, M., & Zipser, E. J. (2007). Mesoscale convective systems over southeastern South America and their relationship with the South American low-level jet. *Monthly Weather Review*, *135*(4), 1290–1309. https://doi.org/10.1175/mwr3305.1

Schumacher, R. S., Hence, D. A., Nesbitt, S. W., Trapp, R. J., Kosiba, K. A., Wurman, J., et al. (2021). Convective-storm environments in subtropical South America from high-frequency soundings during RELAMPAGO-CACTI. *Monthly Weather Review*, *149*(5), 1439–1458. https://doi.org/10.1175/mwr-d-20-0293.1

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/s43017-020-0057-7

Seluchi, M. E., Saulo, A. C., Nicolini, M., & Satyamurty, P. (2003). The northwestern Argentinean low: A study of two typical events. *Monthly Weather Review*, *131*(10), 2361–2378. https://doi.org/10.1175/1520-0493(2003)131<2361:tnalas>2.0.co;2

Serra, Y. L., Rowe, A., Adams, D. K., & Kiladis, G. N. (2020). Kelvin waves during GOAmazon and their relationship to deep convection. *Journal of the Atmospheric Sciences*, *77*(10), 3533–3550. https://doi.org/10.1175/jas-d-20-0008.1

Silva Dias, M. A. F. D., Petersen, W., Silva Dias, P. L. D., Cifelli, R., Betts, A., Longo, M., et al. (2002). A case study of convective organization into precipitating lines in the Southwest Amazon during the WETAMC and TRMM-LBA. *Journal of Geophysical Research*, *107*(D20), LBA 46-1–LBA 46-23. https://doi.org/10.1029/2001jd000375

Skamarock, W. C., & Klemp, J. B. (2008). A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. *Journal of Computational Physics*, *227*(7), 3465–3485. https://doi.org/10.1016/j.jcp.2007.01.037

Sokolowsky, G. A., Freeman, S. W., Jones, W. K., Kukulies, J., Senf, F., Marinescu, P. J., et al. (2023). *tobac* v1.5: Introducing fast 3D tracking, splits and mergers, and other enhancements for identifying and analysing meteorological phenomena. *EGUsphere*, *2023*, 1–37. https://doi.org/10.5194/egusphere-2023-1722

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

Tan, J., Huffman, G. J., Bolvin, D. T., & Nelkin, E. J. (2019). Imerg v06: Changes to the morphing algorithm. *Journal of Atmospheric and Oceanic Technology*, *36*(12), 2471–2482. https://doi.org/10.1175/jtech-d-19-0114.1

Tan, J., Jakob, C., Rossow, W. B., & Tselioudis, G. (2015). Increases in tropical rainfall driven by changes in frequency of organized deep convection. *Nature*, *519*(7544), 451–454. https://doi.org/10.1038/nature14339

Thompson, G., Field, P. R., Rasmussen, R. M., & Hall, W. D. (2008). Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Monthly Weather Review*, *136*(12), 5095–5115. https://doi.org/10.1175/2008mwr2387.1

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. C., Nesbitt, S. W., Salio, P., Hardin, J. C., Bharadwaj, N., Borque, P., et al. (2021). Utilizing a storm-generating hotspot to study convective cloud transitions: The CACTI experiment. *Bulletin of the American Meteorological Society*, *102*(8), E1597–E1620. https://doi.org/10.1175/bams-d-20-0030.1

Velasco, I., & Fritsch, J. M. (1987). Mesoscale convective complexes in the Americas. *Journal of Geophysical Research*, *92*(D8), 9591–9613. https://doi.org/10.1029/jd092id08p09591

Vera, C., Baez, J., Douglas, M., Emmanuel, C., Marengo, J., Meitin, J., et al. (2006). The South American low-level jet experiment. *Bulletin of the American Meteorological Society*, *87*(1), 63–78. https://doi.org/10.1175/bams-87-1-63

Vila, D. A., Machado, L. A. T., Laurent, H., & Velasco, I. (2008). Forecast and Tracking the Evolution of Cloud Clusters (ForTraCC) using satellite infrared imagery: Methodology and validation. *Weather and Forecasting*, *23*(2), 233–245. https://doi.org/10.1175/2007waf2006121.1

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., et al. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), 261–272. https://doi.org/10.1038/s41592-019-0686-2

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

Wilcox, E. M., Yuan, T., & Song, H. (2023). Deep convective cloud system size and structure across the global tropics and subtropics. *Atmospheric Measurement Techniques Discussions*, *2023*(21), 1–23. https://doi.org/10.5194/amt-16-5387-2023

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

Wu, X., & Yan, J. (2011). Estimating the outgoing longwave radiation from the FY-3B satellite visible infrared radiometer Channel 5 radiance observations. *Chinese Science Bulletin*, *56*(32), 3480–3485. https://doi.org/10.1007/s11434-011-4686-6

Yang, Z.-L., Niu, G.-Y., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., et al. (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins. *Journal of Geophysical Research*, *116*(D12), D12110. https://doi.org/10.1029/2010jd015140

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

Zhe Fent. (2023). PyFLEXTRKR [Software]. *Zenodo*. https://doi.org/10.5281/zenodo.7236445

Zhou, J., & Lau, K. (1998). Does a monsoon climate exist over South America? *Journal of Climate*, *11*(5), 1020–1040. https://doi.org/10.1175/1520-0442(1998)011<1020:damceo>2.0.co;2

Zipser, E. J., Cecil, D. J., Liu, C., Nesbitt, S. W., & Yorty, D. P. (2006). Where are the most intense thunderstorms on Earth? *Bulletin of the American Meteorological Society*, *87*(8), 1057–1072. https://doi.org/10.1175/bams-87-8-1057