

## 5 TRACKING GLOBAL PRECIPITATING SYSTEMS

In this chapter, our main objective is to evaluate the capability of the `pyForTraCC` algorithm to consistently track PSs in long-duration, global precipitation datasets from GSMap and IMERG for the 2015–2024 period. The analyses are structured to address the objectives defined in Section 1.1 of Chapter 1, and they are presented in the same order. We first examine the computational performance and processing time of the algorithm and describe the filters used to select PSs from the full tracking output. We then analyze the global and seasonal climatology of PSs, their regional patterns of genesis and dissipation, and their physical and morphological properties. Next, we investigate PS propagation, including the evaluation of vector-correction methods and trajectory characteristics across different regions. Finally, we characterize extreme PSs and describe the organization of the tracking output as a climatological dataset, emphasizing its relevance and availability to the scientific community.

### 5.1 Computational performance of tracking

The computational performance of the `pyForTraCC` algorithm is a critical aspect, especially given the data volume used in this thesis. The algorithm was designed to be efficient and scalable. However, its performance involves more than the ability to process large data volumes, it also depends on execution time and on the capacity to store and manipulate the results. The results demonstrate that `pyForTraCC` is efficient in terms of processing time and computational resource usage. They indicate that the algorithm can handle large data volumes within a reasonable time frame, without requiring advanced or specialized computational resources.

Before analyzing the results presented in this section, it is important to describe the computational environment used for processing. The experiments were performed on a workstation equipped with an AMD Ryzen Threadripper 3960X processor, with 24 cores and 48 threads, and 128 GB of RAM. The raw data were stored on a remote storage server, accessed through a 1 Gbps network connection. This environment was not exclusively dedicated to tracking tasks, because the computational resources were shared with other processes, which may have contributed to variations in processing time during the tracking stages.

As discussed in the The Algorithm Chapter 4, where the development of the tracking framework is divided into sequential modules, each responsible for a specific stage of the tracking process. The modules are: `FeaturesExtraction`, which extracts the

features of the PSs, **Spatial**, performs spatial operations and vector corrections. **Link**, which connects PSs over time, and **Concat**, which merges the results of all modules into a tracking table.

Each module is designed to run sequentially, and therefore, the total execution time is the sum of the execution times of each module. Table 5.1 presents the execution time and file size processed for each module of the algorithm, for both the GSMAp and IMERG products. The results are reported in terms of the number of files processed, processing time, and size of the generated files. Analysis of processing times in Table 5.1 reveals that, for both products, the **Spatial** module demands the longest processing time (17 hours and 6 minutes for GSMAp, and 63 hours, 5 minutes, and 36 seconds for IMERG), reflecting the complexity of the spatial operations and vector correction routines used during this stage of the tracking.

In terms of file size, the **Concat** module generated the largest data volume (406 GB for GSMAp and 1.1 TB for IMERG), as expected, since it combines the outputs of all previous modules into the files that compose the TT. The **FeaturesExtraction** module also produced files of considerable size (209 GB for GSMAp and 340 GB for IMERG), reflecting the amount of information extracted from each tracked PS. Regarding total execution time and data volume, the GSMAp tracking took approximately 28 hours and 20 minutes, generating a total of 709 GB of data. In contrast, the IMERG processing took around 101 hours and 39 minutes and produced a total of 1.8 TB of data.

Table 5.1 - Execution time and file size processed for each module of the algorithm.

<b>Module</b>	<b>Files</b>	<b>Proc. Time</b>	<b>Size</b>
GSMAP - FeaturesExtraction	95016	04:19:11	209GB
GSMAP - Spatial	-	17:06:00	77GB
GSMAP - Link	-	05:17:12	16GB
GSMAP - Concat	-	01:38:28	406GB
IMERG - FeaturesExtraction	175344	14:06:36	340GB
IMERG - Spatial	-	63:05:24	269GB
IMERG - Link	-	20:38:47	92GB
IMERG - Concat	-	03:48:54	1.1TB
GSMAP Total	<b>95016</b>	<b>28:20:51</b>	<b>709GB</b>
IMERG Total	<b>175344</b>	<b>101:39:41</b>	<b>1.8TB</b>

SOURCE: Author's production.

The computational performance results presented in Table 5.1 serve as a reference for the efficiency of the `pyForTraCC` algorithm when processing large data volumes. Execution time and data volume are key indicators for subsequent analyses, especially given the extensive outputs generated when tracking a 10-year period. Although the data volume is substantial, the output files are stored in tabular `parquet` format, which enables efficient and fast querying. This format facilitates the filtering of specific events, the separation of data subsets, and the grouping of events without requiring the entire dataset to be loaded into memory.

## 5.2 Filtering of events

Once the tracking of the PSs is consolidated and the results are concatenated into the TT, the number of detected events becomes very large. The sensitivity of the `pyForTraCC` algorithm, combined with the parameters defined in the `nameList`, such as the precipitation threshold of  $0.1 \text{ mm h}^{-1}$ , often leads to numerous detections. These detections include events that may be classified as noise or isolated cells. This sensitivity is desirable because it enables the tracking of a wide range of systems, from short-lived, small-scale events to large and long-lasting systems. However, analyzing all detected events becomes impractical when the objective is to examine more organized and larger-scale phenomena.

Another factor motivating the need for filtering is the evidence discussed in Chapter 2, which shows that many tracked events correspond to specific types of PSs with well-defined spatiotemporal characteristics. Examples include MCSs, which typically last more than 6 hours and cover areas larger than  $10,000 \text{ km}^2$  (HOUZE, 2004; MACHADO et al., 1998). In line with the objectives of this thesis, which seeks a more comprehensive analysis of PSs, a minimum duration filter of 3 hours (i.e., three consecutive GSMAp images) was applied. This filtering step isolates significant events and removes short-lived systems that may not represent organized meteorological phenomena or may be irrelevant to subsequent analyses.

Before advancing to more specific analyses, it is important to present preliminary statistics that summarize the global behavior of the tracked events in both GSMAp and IMERG databases. Since the outputs of `pyForTraCC` are stored in `parquet` format and systematically organized in the TT, it is possible to apply SQL queries using the `DuckDB` library and analyze the data directly on the structured tables. The SQL query procedure illustrated in Figure 5.1 shows how easily the distribution of tracked events can be extracted.

The query in Figure 5.1 retrieves all records that satisfy the precipitation threshold of 0.1 mm h<sup>-1</sup> and identifies the maximum duration of each unique PS (by uid). This duration (for example, a minimum of 3 hours based on the timestamp difference) is then associated with all corresponding records. In this way, statistics can be calculated by duration class, including the number of PSs, mean size, and size variability. In the last step of query, the total number of PSs is computed so that the results can also be expressed as percentages, providing a view of how events are distributed according to their duration.

Figure 5.1 - SQL query applied with DuckDB to the pyForTraCC Tracking Table (TT) in parquet format, filtering PS uids above the 0.1 mm h<sup>-1</sup> threshold and select by lifetime. For each duration class, the query returns the number and percentage of distinct PSs, as well as summary statistics of their size.

---

```

1   WITH b AS (
2       SELECT
3           uid,                                     -- unique user identifier
4           size,                                    -- data size
5           MAX(duration) OVER (PARTITION BY uid) AS duration -- max duration per uid
6       FROM read_parquet("/path_to_trackingtable/*.parquet", union_by_name=true)
7       WHERE threshold = 0.1                      -- filter records by threshold
8   )
9   SELECT
10      duration,                                -- duration bucket
11      COUNT(DISTINCT uid) AS uid_count,        -- number of unique uid's
12      ROUND(
13          100.0 * COUNT(DISTINCT uid)
14          / SUM(COUNT(DISTINCT uid)) OVER (), 2
15      ) AS percentage,                          -- percent of total uid's
16      ROUND(AVG(size), 2) AS mean_size,         -- average size
17      ROUND(STDDEV_SAMP(size), 2) AS std_size  -- size standard deviation
18  FROM b
19  GROUP BY duration                         -- use the common table expression
20  ORDER BY duration;                        -- aggregate by duration
                                            -- sort by duration

```

---

SOURCE: Author's production.

Based on query above, it is possible to obtain an initial view of the temporal properties of the tracked PSs. Figure 5.2 shows the distribution of all PSs tracked during the study period (2015–2024), organized as a function of their duration and mean area. Figure 5.2(a) presents the values for the GSMAp tracking, while Figure 5.2(b) shows the PS counts and mean sizes for the IMERG tracking. In both figures, the blue line represents the total number of PSs tracked (in logarithmic scale) as a function of duration (in hours). The gray triangles indicate the mean PS size for each duration class, and the error bars represent the standard deviation of the mean.

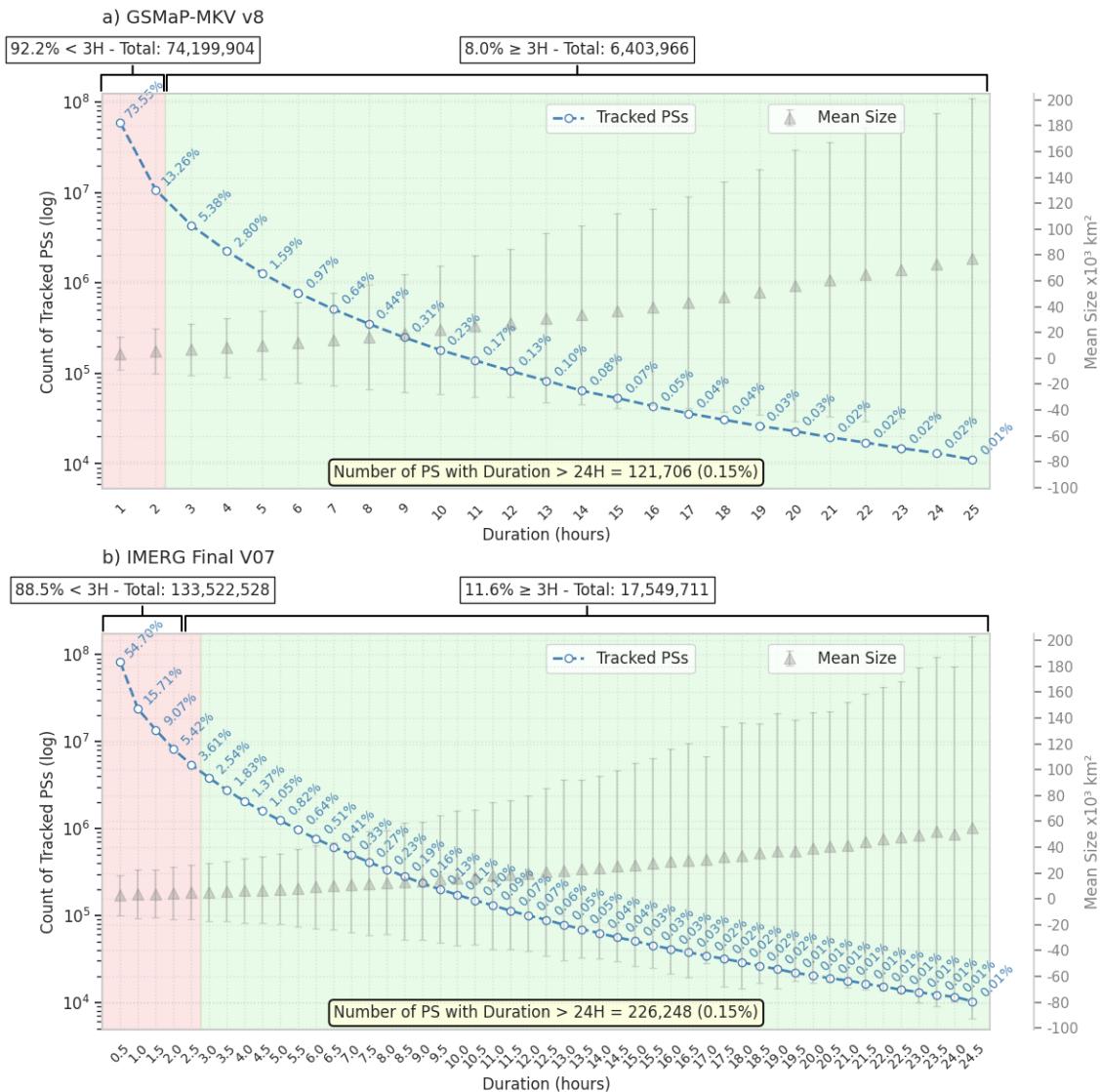
The red-shaded area highlights systems (Figure 5.2) with durations shorter than 3 hours, which account for the majority of tracked PSs (86.8% in GSMAp and 88.5% in IMERG). The green-shaded area corresponds to systems with durations equal to or longer than 3 hours, which are the ones considered in the subsequent characterization analyses and represent 13.3% and 11.6% of the total PSs for GSMAp and IMERG, respectively. The count and proportion of systems with a duration equal to 24 hours (last interval on the x-axis) are also highlighted, indicating the rarity of such events in the datasets used (approximately 0.15% for both products).

From the analysis of Figure 5.2, there is a clear tendency for the mean size of PSs to increase with longer durations. This relationship is expected, since longer-lived PSs tend to grow in scale, resulting in larger areas (BROWNING, 1986; WALLACE; HOBBS, 2006). The error bars of the mean PS sizes, which increase with the life cycle duration, corroborate this behavior. A gradual increase in mean size is observed with increasing duration, reaching average values close to 75,000 km<sup>2</sup> for 24-hour PSs in GSMAp and 60,000 km<sup>2</sup> for IMERG.

The standard deviation, represented by the error bars, suggests that the longer the PS life cycle, the greater the variability in system size. This indicates that, although there is a general trend of size increase with duration, the distribution of PS sizes is highly dispersed, and long-lived systems exhibit a wide range of sizes. This variability can be attributed to the complexity of the physical processes that govern the evolution of these systems (HOUZE, 1981).

However, it is important to emphasize that these global-scale patterns in Figure 5.2 provide only a general overview of the tracking datasets. That regionally focused analyses are more suitable for establishing robust premises about the variability and physical processes associated with the PSs.

Figure 5.2 - Distribution of tracked PSs as a function of life-cycle duration and mean size for (a) GSMAp-MKV v8 and (b) IMERG Final V07. The blue line shows PS counts (log scale), gray triangles with error bars show mean size and its standard deviation, and the red (green) band marks PSs with duration < 3 h ( $\geq 3$  h), with the 24 h bin highlighting long-lived events.



SOURCE: Author's production.

The high frequency of systems with durations shorter than 3 hours, as indicated by the red-shaded area in Figure 5.2, suggests that many of the events detected by the algorithm are short-lived and small in size. These systems may represent isolated cells or noise in the detection process, that is, data with some degree of

uncertainty. Applying the 3-hour filter is therefore a crucial step for refining the subsequent analyses, since it allows the focus to be placed on more organized and meteorologically relevant systems in terms of the variability of the tracked events (BROWNING, 1986; HOUZE, 2004; WHITE et al., 2017).

Even though events that pass the duration filter represent only 13.3% of GSMAp systems and 15.2% of IMERG systems, they still form a large sample (10,736,144 and 22,993,207 events, respectively), adequate for statistical and climatological analyses over the 10-year period. In this study, the 3-hour threshold was kept for IMERG solely to allow comparison with GSMAp. However, three consecutive IMERG images (1.5 hours) would likely be sufficient for tracking, given its higher temporal resolution.

### 5.3 Global census of precipitating systems

In this section, the global spatial-temporal distribution of the tracked PSs is analyzed, focusing on their occurrence and seasonal variability. The analyses are based on the filtered tracking results, considering only PSs with a duration equal to or greater than 3 hours, as described in Section 5.2. The objective is to provide a census of PSs on a global scale, identifying regions with higher occurrence and examining seasonal patterns that may influence their distribution.

#### 5.3.1 Global geographical distribution

After applying the minimum duration filter of 3 hours, as described in Section 5.2, the resulting was used to analyze the spatial distribution of significant PSs over the study period. Figure 5.3 presents the geographical distribution of the occurrence of tracked PSs, considering individually the spatial grouping of PSs detected based on their UIDs. The grouping that organizes the PSs spatially is based on the spatial coordinate values from the `array_x` and `array_y` columns of TT, which correspond to the coordinates of the clusters and grouped as a single PS occurrence. The counts represent the total number of events per pixel for the GSMAp product (panel a) and the IMERG product (panel b). Next to each map, the latitudinal profile of PS count is shown, allowing for the observation of aggregated meridional patterns.

The visualization of tracking counts in Figure 5.3 provides an initial assessment of the algorithm's effectiveness in identifying and tracking PSs across the global domain. There is no degradation at the lateral edges of the maps, even near the  $180^{\circ}$  and  $0^{\circ}$  meridians, which shows that events crossing these longitudinal bound-

aries were correctly handled by pyForTraCC. This behavior is a direct result of the optimizations implemented in the algorithm (Section 4.1.3.2.4), which include procedures for handling spatial connectivity in global domains and prevent events from being fragmented when reaching the edges of the data field. The regularity and spatial continuity of the observed patterns also indicate that the tracking process was robust across all regions, with no evident gaps or counting artifacts that would suggest systematic errors. .

From this visual perspective, it is possible to proceed to an analysis of the spatial distribution of the occurrence of tracked PSs. Figure 5.3 reveals a global pattern of PS occurrence, with a clear predominance in tropical and subtropical regions (latitudes between 30°S and 30°N). This behavior is expected, since deep and organized convection is more frequent in these latitudes due to the greater availability of heat and moisture (MALHI; WRIGHT, 2004; MASUNAGA et al., 2005; TAKAHASHI et al., 2021).

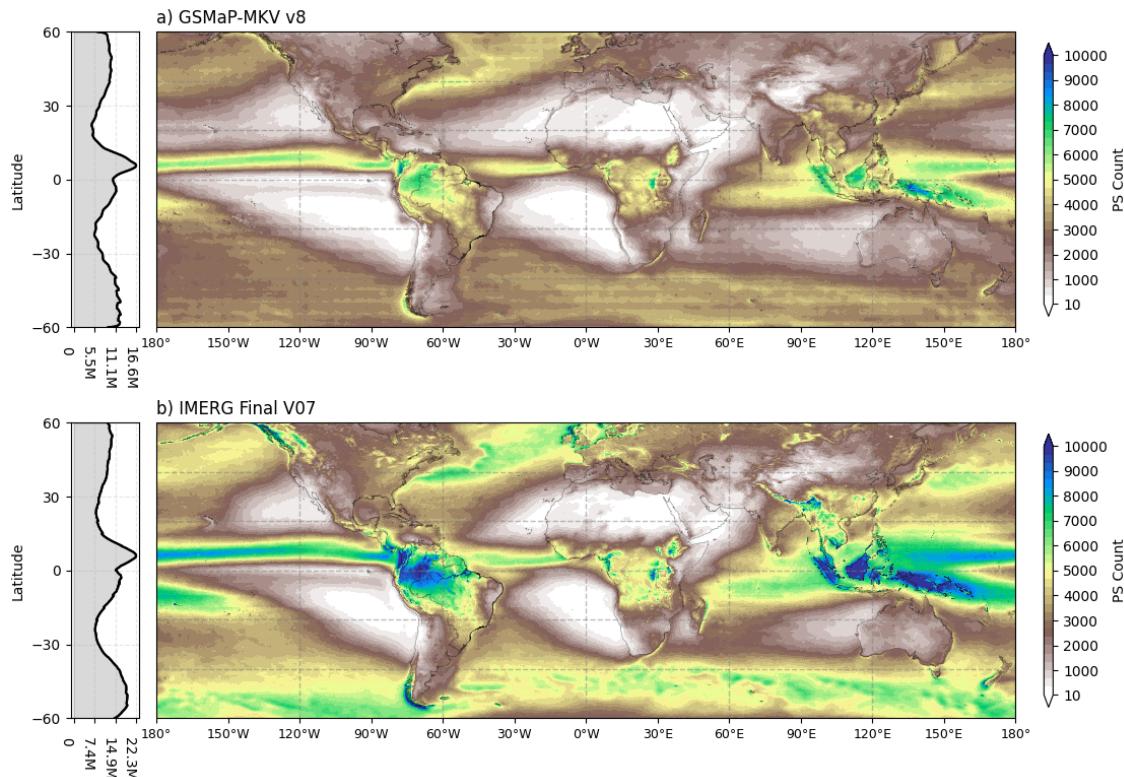
The maximum event counts are concentrated over tropical continental areas such as the Amazon Basin, Central Africa, Southeast Asia, and Pacific archipelagos, as well as over the intertropical convergence zones in the oceans, especially in the western tropical Pacific. These regions are known for their intense convective activity and are often associated with phenomena such as tropical cyclones and MCSs (MACHADO et al., 1998; MASUNAGA et al., 2005).

The latitudinal profiles (to the left of each map in Figure 5.3) reinforce this tropical predominance, with sharp peaks near the Equator. Both GSMAp and IMERG exhibit maximum PS count between approximately 5°S and 5°N, with a sharp decrease toward mid and high latitudes. This distribution is consistent with the climatological patterns of convective precipitation, as extensively described in the literature (MOORE et al., 1993; MONCRIEFF, 1992; HOUZE, 2004; MACHADO; LAURENT, 2004).

It is important to note that, although both products exhibit similar large-scale spatial patterns, the amplitudes are not the same. GSMAp presents fewer tracked events than IMERG, which reflects not only the intrinsic differences between the datasets but also their distinct temporal resolutions. These aspects must be considered when interpreting or comparing the trackings. As discussed in Section 3.3 of Chapter 3, the choice of precipitation product can significantly affect the detection and characterization of PSs, particularly in regions influenced by complex topography or limited calibration data (LEVIZZANI et al., 2020; HOUZE, 2012).

Therefore, Figure 5.3 not only confirms the ability of pyForTraCC to capture the main patterns of PS occurrence but also highlights the importance of accounting for the characteristics of each input product in spatial analyses. The differences between the two tracking outputs reinforce the need for caution when interpreting climatological results from satellite-derived datasets, particularly in studies of the global spatial and temporal variability of PSs.

Figure 5.3 - Global distribution of tracked PSs for the 2015–2024 period, after applying a minimum lifetime of 3 h. Shading indicates the number of PS occurrences per grid cell, obtained from the spatial grouping of clusters by uid in the Tracking Table, for (a) GSMAp–MVK v8 and (b) IMERG Final V07. The panels on the left show the corresponding meridional profiles of PS counts.



SOURCE: Author's production.

To consolidate the analysis of the spatial distribution of tracked PSs, Figure 5.4(a) presents the IMERG–GSMAp difference in PS count (IMERG minus GSMAp). Red shades indicate where IMERG detects more PSs, blue shades where GSMAp detects

more PSs, and white shades where both products agree. The latitudinal profile on the left summarizes the meridional structure of these differences.

A predominance of red tones over large portions of the globe indicates that IMERG generally reports higher PS frequency count than GSMAp, especially over tropical and extratropical oceans. Strong negative anomalies (dark blue) appear west of South America and over the Himalayas-regions dominated by complex orography, suggesting that GSMAp may be more sensitive to orographic precipitation and thus yields higher local PS count (HOUZE, 2012). Conversely, IMERG shows higher frequency across much of the tropics and the western Pacific, consistent with well-known convective climatologies (MASUNAGA et al., 2005).

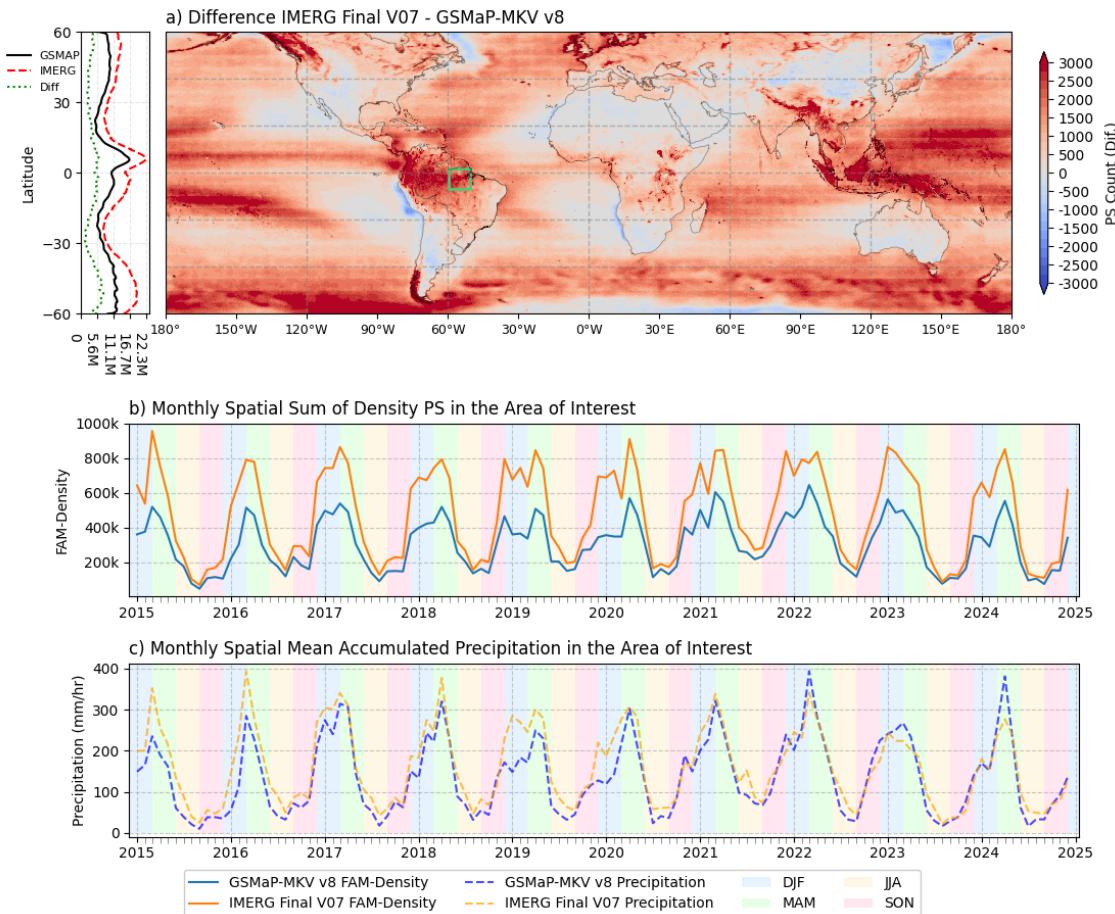
The latitudinal difference profile to the left of Figure 5.4(a) reinforces IMERG's dominance near the equator, with mostly positive values between about 10°S and 10°N. South of 30°S, the IMERG–GSMAp difference increases again, indicating higher PS frequency count in IMERG. These consistent signals demonstrate the algorithm's ability to capture and quantify product dependent discrepancies, which should be considered in subsequent PS characterizations.

To further examine these differences, a region of interest over the Amazon (green box in Figure 5.4(a)) was selected. The time series for this region (Figure 5.4b) shows that IMERG consistently detects more PSs than GSMAp, with this contrast becoming more pronounced during the peak rainfall season from early December to mid-May (LAURENT et al., 2002; WENDISCH et al., 2016). This seasonal increase in system counts, despite the two products presenting very similar accumulated rainfall totals (Figure 5.4c), suggests that the divergence arises from how the tracking algorithm, using the 0.1 mm/h threshold, interprets each rainfall field. The central explanation is that IMERG's higher temporal resolution (30 minutes, compared to GSMAp's 60 minutes) allows new events to start within each 30-minute interval, leading to a greater number of systems that exceed the detection threshold and are therefore counted as distinct PSs during this season.

The larger number of short-duration PSs identified in IMERG relative to GSMAp is consistent with differences in temporal sampling and with previous evidence that coarser temporal resolution tends to merge closely spaced convective pulses into fewer, longer-lived precipitation events, thereby reducing the number of distinct systems (WHITE et al., 2017; LIU, 2011). Independent evaluations also show that IMERG often provides competitive or improved performance in hydrological applications, especially in humid basins where bias-adjusted IMERG products tend to

achieve higher efficiencies and lower systematic bias than GSMAp (HUANG et al., 2025). IMERG additionally captures broader precipitating areas and higher extreme rain rates than GSMAp while maintaining comparable PS distributions (TAKAHASHI et al., 2021), supporting the interpretation that its higher PS counts largely reflect physically plausible systems rather than artefacts of the pyForTraCC tracking.

Figure 5.4 - Differences in tracked PS counts between IMERG and GSMAp. (a) Global IMERG–GSMAp PS count difference, with red (blue) indicating more PSs in IMERG (GSMAp), the zonal-mean latitudinal profile on the left, and the green box marking the region in Amazon. (b) Monthly spatial sum of PS density and (c) monthly spatial mean accumulated precipitation over the Amazon, with shaded bands indicating the DJF, MAM, JJA and SON seasons.



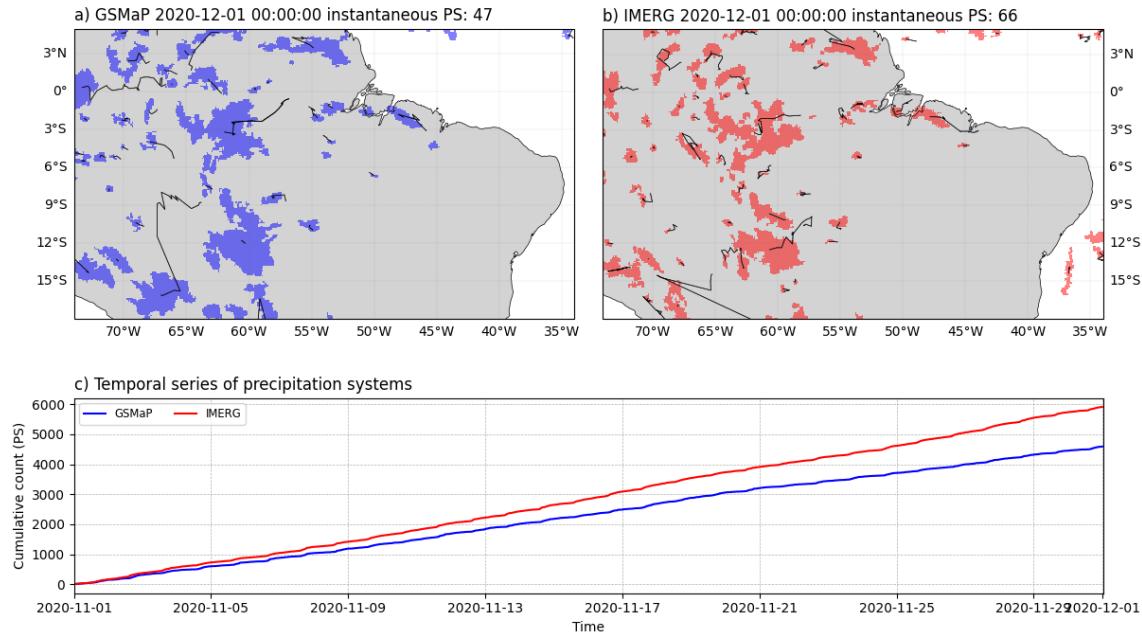
SOURCE: Author's production.

To further examine these differences, a targeted experiment over the Amazon Basin was conducted. The region of interest is a rectangle bounded by 74°W–34°W and 18°S–5°N, and all PSs whose trajectories intersect this box are included. The analysis period spans from 1 November to 1 December 2015, corresponding to one of the months with marked positive IMERG minus GSMAp differences in PS counts in Figure 5.4. At each time step within this month, only PSs that start inside the region are counted, and these counts are accumulated over the full period.

Figure 5.5 summarizes this experiment. Panels (a) and (b) show an instantaneous snapshot on 1 December at 00:00, with GSMAp clusters in blue and IMERG clusters in red. At this single time step, GSMAp presents 47 instantaneous precipitation clusters, whereas IMERG presents 66, already indicating a higher fragmentation of the precipitating field in the half hourly product. The lines emerging from the centroids represent the trajectories of each PS as tracked by pyForTraCC, and highlight that both products follow similar paths but assign different numbers of individual systems to those paths. Differences between GSMAp and IMERG trajectories are expected, as they reflect how long each UID is maintained by the tracking in each product and, consequently, the persistence of the systems in space and time.

Panel (c) shows the cumulative number of newly born PSs as a function of time for the same region and period. Both curves increase almost linearly, but the IMERG curve stays above the GSMAp curve throughout. By the end of the month, GSMAp accumulates 4,591 PSs, whereas IMERG reaches 5,912 PSs, which is about thirty percent more events. This difference is consistent with the temporal sampling: with thirty minute images, IMERG can register systems that develop and cross the  $0.1 \text{ mm h}^{-1}$  threshold between two consecutive hourly fields. In GSMAp, several of these systems are either missed or merged with neighbouring precipitation features. This case study therefore indicates that the higher PS counts obtained with IMERG in the Amazon region are mainly a consequence of its finer temporal resolution. Additionally, product-dependent differences in the underlying precipitation fields may contribute to these discrepancies and will be examined later in the thesis.

Figure 5.5 - Case study to investigate discrepancy in screening counts between November and December 2015. Panels (a) and (b) show instantaneous PS fields on 1 December at 00:00 UTC for GSMAp (blue) and IMERG (red), with PS clusters and trajectories. Panel (c) presents the cumulative number of NEW PSs during November 2015.



SOURCE: Author's production.

The spatial analysis has validated the robustness of the pyForTraCC algorithm and revealed that the discrepancies between GSMAp and IMERG in terms of tracked PS count are systematic and significant. These differences are likely rooted in the distinct temporal resolutions and sensitivities of the two products, which influence how events are detected and tracked. This understanding is crucial for interpreting subsequent analyses and highlights the importance of considering product characteristics in climatological studies of PSs .

### 5.3.2 Seasonal precipitating systems occurrence

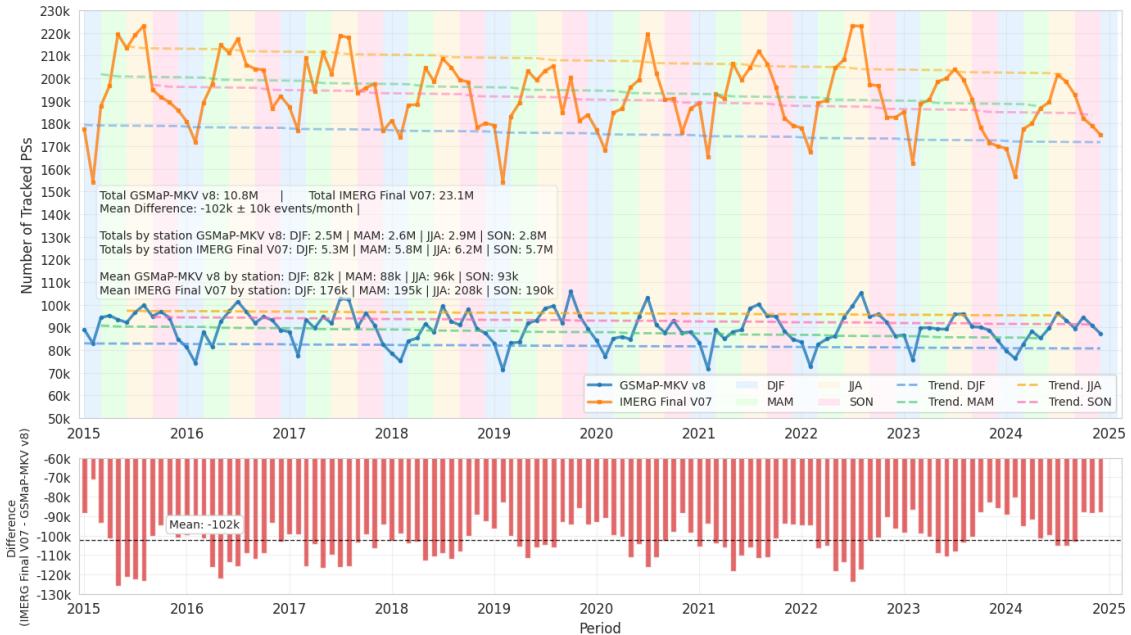
Identifying the seasonal occurrence of PSs is essential for understanding the climate regimes that modulate their distribution and frequency (WALSH; LAWLER, 1981). Although the tracked series spans only 10 years (2015–2024), which is shorter than the 30-year standard used for climatological characterization (ARGUEZ; VOSE, 2011), it is still sufficient to reveal clear and meaningful seasonal patterns in PS occurrence.

The Figure 5.6 presents the temporal series of the global tracked PS, illustrating the seasonal distribution of PSs counts. The upper panel shows the monthly mean time series for GSMAp (blue) and IMERG (orange), with background colors indicating the seasons (DJF, MAM, JJA, SON) and dashed lines denoting linear trends. The central box text summarizes the totals (10.8 million for GSMAp and 23.1 million for IMERG) and seasonal means, while the lower panel displays the difference (IMERG – GSMAp), with a mean offset of about –102k events per month.

The highest counts occur during boreal summer (JJA), with 2.9M PSs tracked by GSMAp and 6.2M by IMERG, while the lowest counts are in boreal winter (DJF), with 2.5M (GSMAp) and 5.3M (IMERG). Intermediate values are observed in spring (MAM: 2.6M GSMAp, 5.8M IMERG) and autumn (SON: 2.8M GSMAp, 5.7M IMERG). Seasonal trends highlight the predominance of JJA and the lower frequency in DJF, but also reveal differences in transitional seasons: for IMERG, MAM exceeds SON, whereas GSMAp shows the opposite.

As previously observed in the Amazon region (Section 5.3), global totals also reveal significant discrepancies between the two products, with IMERG consistently reporting higher PS counts than GSMAp. This pattern is consistent with the earlier hypothesis that IMERG’s finer temporal resolution allows for the detection of more events, leading to a higher overall count of tracked PSs. The mean monthly difference of approximately 102k events globally underscores the systematic nature of this discrepancy. It is crucial to consider this inherent difference in all subsequent analyses, as it may significantly influence the perceived frequency and distribution of PSs on a global scale. Understanding the nature of these discrepancies is essential for accurately interpreting the climatological characteristics of tracked PSs and for making informed comparisons between datasets.

Figure 5.6 - Global monthly time series and seasonal census of tracked PSs from GSMAp–MVK v8 and IMERG Final V07 for 2015–2024. The upper panel shows the monthly counts for each product, with colored background bands indicating the seasons (DJF, MAM, JJA, SON) and dashed lines representing seasonal mean values summarized in the central text box. The lower panel presents the monthly difference (IMERG–GSMAp), highlighting the excess of PSs detected by IMERG, with a mean offset of about  $-1.0 \times 10^5$  events per month.



SOURCE: Author's production.

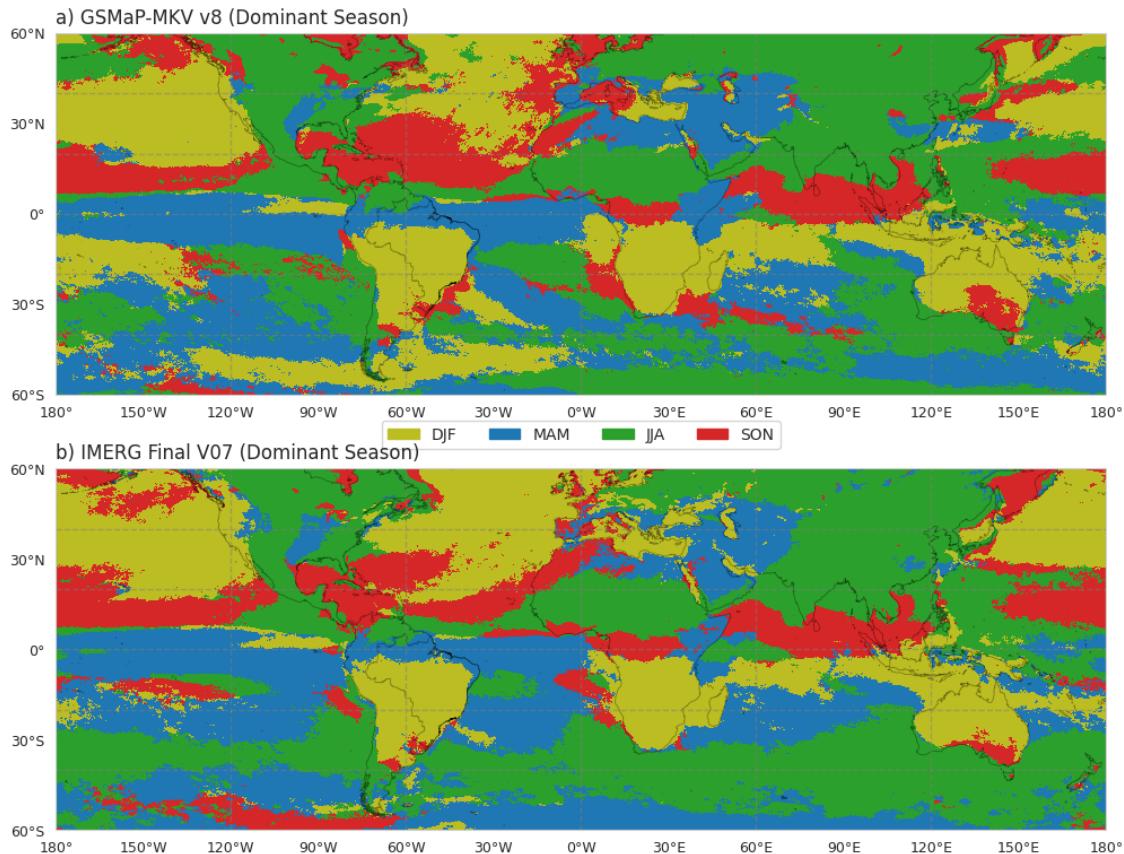
To complement the temporal analysis, which revealed the seasonal variability and differences between the GSMAp and IMERG products, Figure 5.7 offers a spatial view of the global distribution of the dominant season in PS occurrence. This map was generated by calculating the seasonal maximum of PS counts for each pixel, thus identifying the season with the highest frequency. The colors represent each season: DJF (dark yellow), MAM (blue), JJA (green), and SON (red). Complementary figures reporting the total number of PSs identified in each season are provided in Appendix A.5, offering a quantitative perspective to the spatial patterns shown here.

A comparative analysis of the panels in Figure 5.7 shows strong agreement between the GSMAp (a) and IMERG (b) datasets in identifying large-scale seasonal patterns. Over South America, both products consistently indicate DJF as the dominant

season across most of the continent, reflecting the concentration of PS activity during the austral summer. In the equatorial continental zone, both also identify MAM as the prevailing season, marking the transition period associated with a secondary peak in PS occurrence (MALHI; WRIGHT, 2004; MASUNAGA et al., 2005). Similar consistency appears in other continental regions, including central Africa and parts of Asia. This spatial agreement between two independent datasets reinforces the ability of the tracking methodology to capture intrinsic system characteristics, in line with previous findings such as those of Feng et al. (2021), who reported comparable seasonal patterns for the precipitation fraction of MCSs.

When extending the analysis to a broader scale in Figure 5.7, the overall spatial coherence between GSMAp and IMERG remains evident, but some differences become more noticeable, particularly over the oceans. GSMAp tends to exhibit a more fragmented pattern at the 30°S latitude, diverging from the more continuous distribution of IMERG for the MAM and SON seasons. In summary, the global map of dominant seasons highlights the general agreement between the two tracked datasets in identifying the prevailing seasons of PS occurrence. These differences, however, underscore the importance of considering the specific characteristics of each dataset when interpreting the spatial patterns of PS activity.

Figure 5.7 - Global map of the dominant season of PS occurrence for GSMAp–MVK v8 (a) and IMERG Final V07 (b), obtained by identifying, at each grid cell, the season (DJF, MAM, JJA, SON) with the highest number of tracked PSs during 2015–2024. The maps summarize the prevailing seasonal regimes of PS activity.



SOURCE: Author's production.

Another census perspective is provided by the Hovmöller diagrams in Figure 5.8, which illustrate the latitudinal distribution of PS counts as a function of month for both GSMAp (panel a) and IMERG (panel b). These diagrams offer a comprehensive view of the seasonal modulation of PS occurrence across different latitudes, highlighting the temporal evolution of PS activity throughout the year.

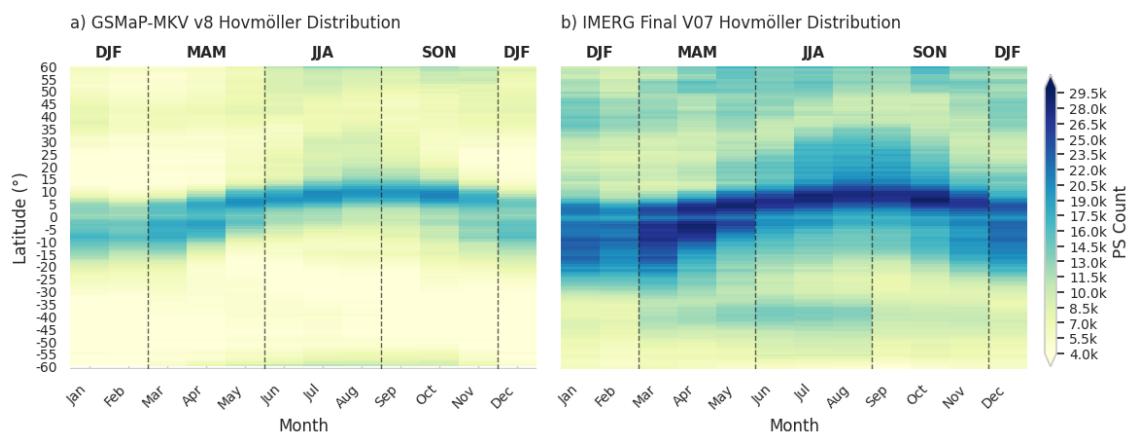
The diagrams show that tracked systems are overwhelmingly concentrated in the tropics. Both datasets exhibit a persistent equatorial core of activity, reaching over 29.5k events in IMERG and about 20k in GSMAp at seasonal maxima. This con-

centration is consistent with the greater availability of heat and moisture that fuels deep and organized convection in tropical regions (MALHI; WRIGHT, 2004; MASUNAGA et al., 2005; TAKAHASHI et al., 2021). This result confirms the dominance of tropical regions in the global totals, as also observed in the frequency counter maps in Figure 5.3.

Beyond this persistent maximum, the diagrams also capture the seasonal north-south migration of the main precipitation band, a pattern directly associated with the corresponding shift of the ITCZ (SCHNEIDER et al., 2014). The annual cycle is evident, with the largest counts occurring during the boreal summer (JJA), when the ITCZ is in its northernmost position, and the lowest during the boreal winter (DJF), when it shifts southward. These findings corroborate the time series analysis in Figure 5.6. Transitional seasons (MAM and SON) present intermediate values, marking the gradual shift between the peaks of activity.

A key feature evident in the diagrams is the offset between the products. IMERG consistently reports higher PS counts and a broader latitudinal span of active systems, whereas GSMAp shows lower totals and a more confined band of high occurrence. The presence of events between March and September at latitudes from 35°S to 50°S in IMERG also contributes to the higher counts in the southern extratropics, consistent with the patterns shown in Figure 5.3.

Figure 5.8 - Hovmöller diagrams of monthly PS counts (2015–2024) as a function of latitude, summarizing the seasonal variability in zonal-mean occurrence for (a) GSMAp–MVK v8 and (b) IMERG Final V07.



SOURCE: Author's production.

In summary, this section presented a global census of PS counts, and seasonality, revealing their spatial distribution while highlighting product differences between GSMAp and IMERG. A systematic bias is evident in the total number of tracked systems, with IMERG consistently reporting higher counts, largely attributable to its finer temporal resolution. However, both tracked datasets capture the similar large-scale spatial and seasonal patterns of PS occurrence, consistent with established climatological knowledge. Crucially, the results demonstrate the capability of pyForTraCC to robustly capture the census of PS occurrences over a global, wrap-around domain: it preserves spatial continuity at  $0^\circ/180^\circ$  (Figure 5.3) and recovers the large-scale agreement between GSMAp and IMERG while objectively quantifying where they diverge (Figures 5.2, 5.7). Therefore, the diagnosed GSMAp–IMERG discrepancies are attributable to dataset characteristics (temporal resolution, sensitivity of the retrieval precipitation algorithms) rather than tracking artifacts, reinforcing confidence in the methodology. A detailed characterization of PS properties (precipitation, area, lifetime, and their evolution) follows in Section 5.4.

#### 5.4 Global properties of precipitating systems

The global properties of PSs (e.g. precipitation intensity, area, duration, and evolution) are fundamental for the characterization and understanding of rainfall events in meteorology (BROWNING, 1986; LIU; ZIPSER, 2015). These parameters describe the spatial and temporal behavior of PSs, regardless of their origin, providing valuable information about their structure, persistence, and development over time. In the context of nowcasting, they play a crucial role in improving short-term forecasts by helping to identify how rain systems evolve, expand, or dissipate (DIXON; WIENER, 1993; VILA et al., 2008). On a broader scale, the analysis of these properties supports the development of more robust and generalizable algorithms, enhances the representation of precipitation processes in climate studies, and contributes to a better understanding of global rainfall variability (MACHADO et al., 1998; MASUNAGA et al., 2005; HAYDEN et al., 2021).

This section characterizes the PSs tracked by pyForTraCC in the GSMAp and IMERG datasets (2015–2024), beginning with a detailed analysis of their precipitation properties. Initially, the global seasonal distribution of precipitation is examined to identify large-scale spatial patterns and their annual variability. Following this, an analysis of the regional distribution of annual precipitation is presented, segmented into Land, Ocean, Tropics, and Extratropics, to quantify key differences between these domains.

Subsequently, the analysis focuses on the morphological and dynamic properties of the systems, specifically their area and lifetime. We explore the mean area and mean duration for each product, their inter-product differences, and the seasonal latitudinal profiles. These diagnostics quantify the expected variation between size and duration, delineate regions and seasons where organized precipitation is most prevalent, as documented in the literature (HOUZE, 2004; MACHADO et al., 1998; MALHI; WRIGHT, 2004).

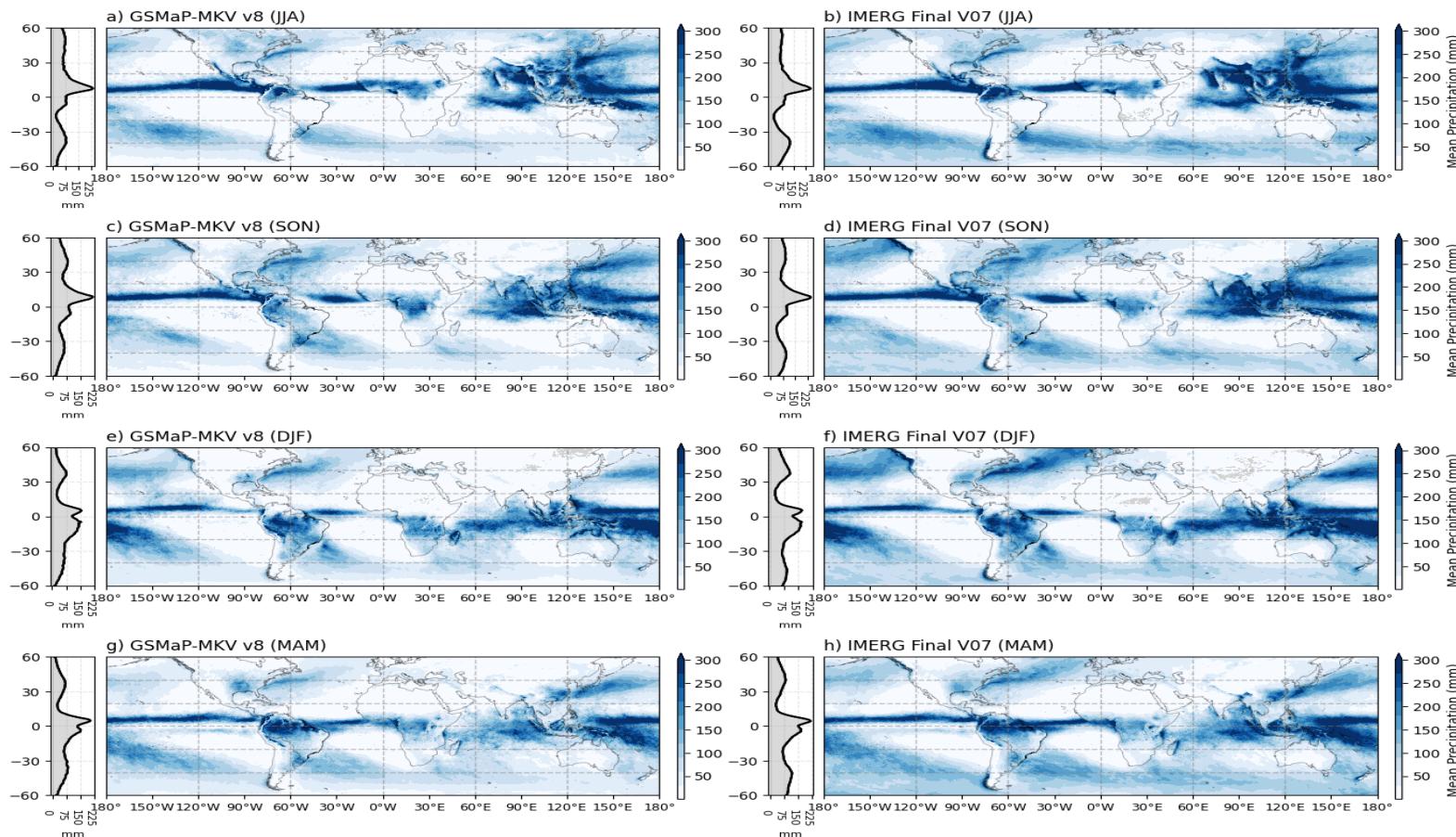
#### 5.4.1 Precipitation distribution

The analysis of precipitation distribution within the tracked PSs is crucial for understanding their contribution to the global hydrological cycle and their spatial-temporal variability. In many works, the focus is often on precipitation patterns overviewing (HOUZE, 1981; TRENBERTH et al., 2003; MASUNAGA et al., 2005; SUN et al., 2018). However, by concentrating on PSs identified through a tracking algorithm as done in this study, we can isolate the precipitation associated with organized systems, which are often responsible for significant weather events.

Figure 5.9 illustrates the global seasonal mean accumulated precipitation within the footprints of the tracked PSs for both GSMAp (left panels) and IMERG (right panels) over the 2015–2024 period. The figure reveals a strong correspondence between the two products in identifying the main global precipitation patterns. Both datasets highlight the tropical regions as the most active areas, with peak precipitation accumulations in the western Pacific, the Indian Ocean, and the tropical Atlantic, consistent with established climatological knowledge (TRENBERTH et al., 2003; MASUNAGA et al., 2005).

Both datasets clearly delineate the ITCZ as a persistent band of intense precipitation near the equator. This seasonal migration is explicitly quantified by the zonal mean precipitation profiles displayed to the left of each map of Figure 5.9, which summarize the meridional structure of precipitation. These profiles clearly show a northward displacement of the precipitation peak during the boreal summer (JJA, panels a-b) and a southward shift during the austral summer (DJF, panels e-f). This behavior is a well-documented feature of global atmospheric circulation and is successfully captured by the tracking-based analysis (WALLACE; HOBBS, 2006; SCHNEIDER et al., 2014; GILL, 2016).

Figure 5.9 - Global seasonal mean precipitation associated with tracked PSs for GSMAp–MVK v8 (a, c, e, g) and IMERG Final V07 (b, d, f, h) during 2015–2024. Shading shows the mean accumulated precipitation within PS footprints for JJA, SON, DJF, and MAM, and the panels to the left of each map present the corresponding zonal-mean precipitation profiles.



SOURCE: Author's production.

To further investigate the regional characteristics and product discrepancies, Figure 5.10 presents boxplots of the monthly mean precipitation stratified by four distinct domains: Land (a), Ocean (b), Tropics (c), and Extratropics (d). The tropics are defined as the region between  $-23.5^{\circ}$  and  $23.5^{\circ}$ , while the extratropics encompass all latitudes poleward of these bounds. The boxplots summarize the distribution of monthly mean precipitation rates within the footprints of tracked PSs for both GSMAp and IMERG for 10-year. This stratification allows for a detailed comparison of precipitation characteristics across different surface types and climatic zones.

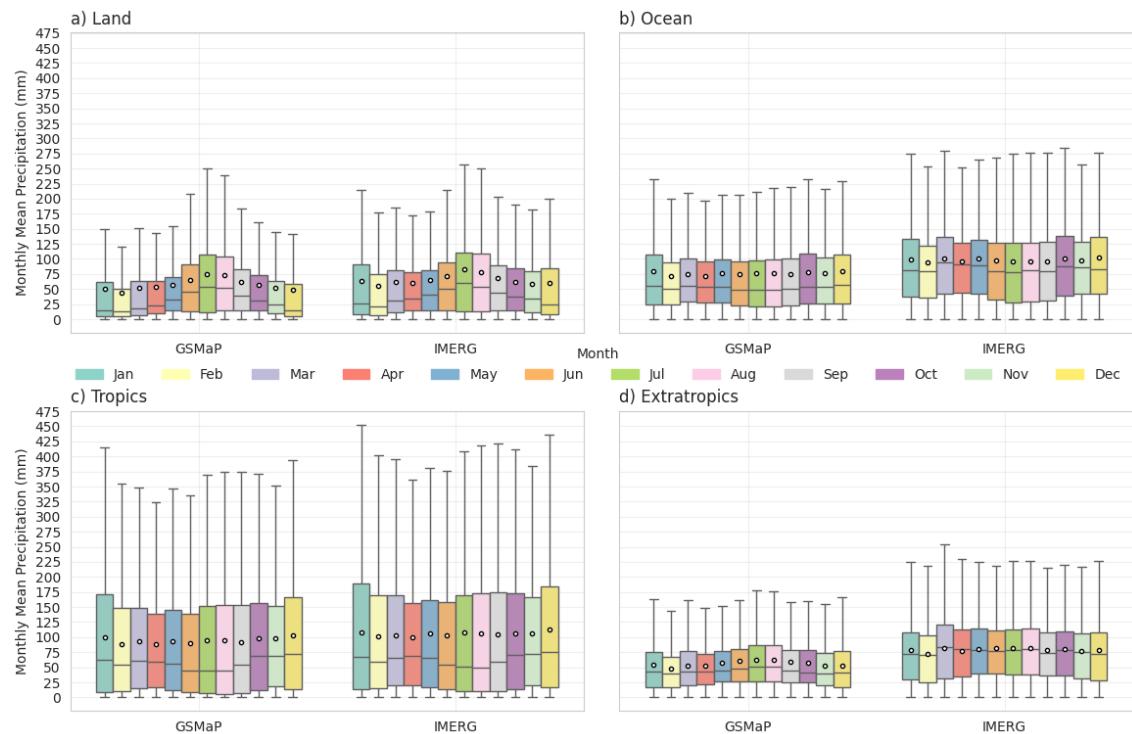
A primary observation across all domains in the Figure 5.10 is the difference in the magnitude of precipitation reported by the two products. IMERG consistently shows higher median values and a wider interquartile range compared to GSMAp, indicating that PSs tracked in IMERG are, on average, more intense and more variable. This contrast is particularly evident over the oceans (panel b) and in the tropics (panel c), where IMERG presents substantially higher median precipitation and a distribution skewed toward stronger events, as suggested by the longer upper whiskers and the larger number of outliers. These findings are consistent with the product characteristics discussed in Section 3.3 and with validation studies showing that IMERG is generally more sensitive to intense convective precipitation, especially in tropical and maritime environments (HUANG et al., 2025; YANG et al., 2024).

The land-ocean contrast is also clearly depicted. Over land (panel a), both products show a more pronounced seasonal cycle, with higher precipitation variability (larger boxes and longer whiskers) compared to the ocean. Conversely, oceanic precipitation (panel b) is more persistent throughout the year, although with the same systematic offset between the products. When comparing climatic zones, the tropics (panel c) exhibit significantly higher precipitation rates and greater month-to-month variability than the extratropics (panel d), a direct reflection of the dominant role of deep convection in the tropical climate system (MALHI; WRIGHT, 2004; MASUNAGA et al., 2005). In the extratropics, precipitation is weaker and less variable, consistent with the prevalence of synoptic-scale systems that produce more moderate and widespread rainfall.

Taken together, the spatial diagnostics from the maps (Figure 5.9) and the regional boxplot distributions (Figure 5.10) reveal a systematic positive offset in IMERG, especially within the tropical belt and over oceanic and monsoon regions. This behavior is consistent with known differences in sensor calibration, retrieval algorithms,

and temporal resolution between the products, as discussed in Chapter 3 (HUANG et al., 2025; YANG et al., 2024). Despite these quantitative discrepancies, the strong qualitative agreement in the land–ocean and tropics–extratropics contrasts, as well as in the seasonal cycle of PS-related precipitation, confirms that the pyForTraCC algorithm consistently captures the main climatological features of global precipitation associated with PSs in both datasets.

Figure 5.10 - Boxplots of monthly mean precipitation associated with tracked PSs for GSMAp–MVK v8 and IMERG Final V07 (2015–2024), stratified by four domains: Land (a), Ocean (b), Tropics (c), and Extratropics (d). Each panel summarizes the distribution of precipitation within tracked PSs.



SOURCE: Author's production.

In summary, the analyses in this subsection characterized the seasonal and regional patterns of precipitation associated with the tracked PSs. Although this investigation does not address in detail the microphysical or dynamical causes of these patterns, it provides a broad overview of the precipitation characteristics within the identified systems. This climatological characterization forms the basis for the subsequent analyses in this section, which will examine other morphological and dynamical properties of the PSs, such as their area, duration, and life-cycle characteristics.

### 5.4.2 Precipitating systems features

Some PSs features like area and duration are fundamental properties that reflect their degree of organization and persistence. As demonstrated by Machado et al. (1998), who investigated the life cycle variations of MCSs, there is a direct relationship between the average size of a system and its lifetime. Larger PSs tend to exhibit greater spatial organization and stratiform development, which allows them to maintain rainfall over extended periods. In contrast, smaller systems are usually short-lived and less organized (TAKAHASHI et al., 2021). Building on these characteristics, PS tracking becomes essential, and allows for the quantification of these properties over time and space, providing a more comprehensive understanding of the system's behavior.

Beyond their climatological relevance, these PS features are also highly valuable for nowcasting applications. Real-time monitoring of PS area, duration, and displacement provides objective diagnostics of the current degree of organization and of the short-term evolution of the systems. These diagnostics can be assimilated into operational nowcasting frameworks to anticipate the location, intensity, and persistence of rainfall on hourly time scales (WILSON et al., 1998; WMO, 2017).

Figure 5.11 presents the spatial distribution of mean PS area (panels a–c) and duration (panels d–f). The top and middle rows show the results for GSMAp (panels a and d) and IMERG (panels b and e), while the bottom row depicts their differences (panels c and f). These maps are derived from a post-processing step that groups systems by their unique identifier `uid`, from which the values of `duration` and `size` are obtained, `size` is converted to pixel area ( $\text{km}^2$ ) at  $0.1^\circ$  spatial resolution. The computed mean area and total duration are then assigned to all grid points covered by each PS.

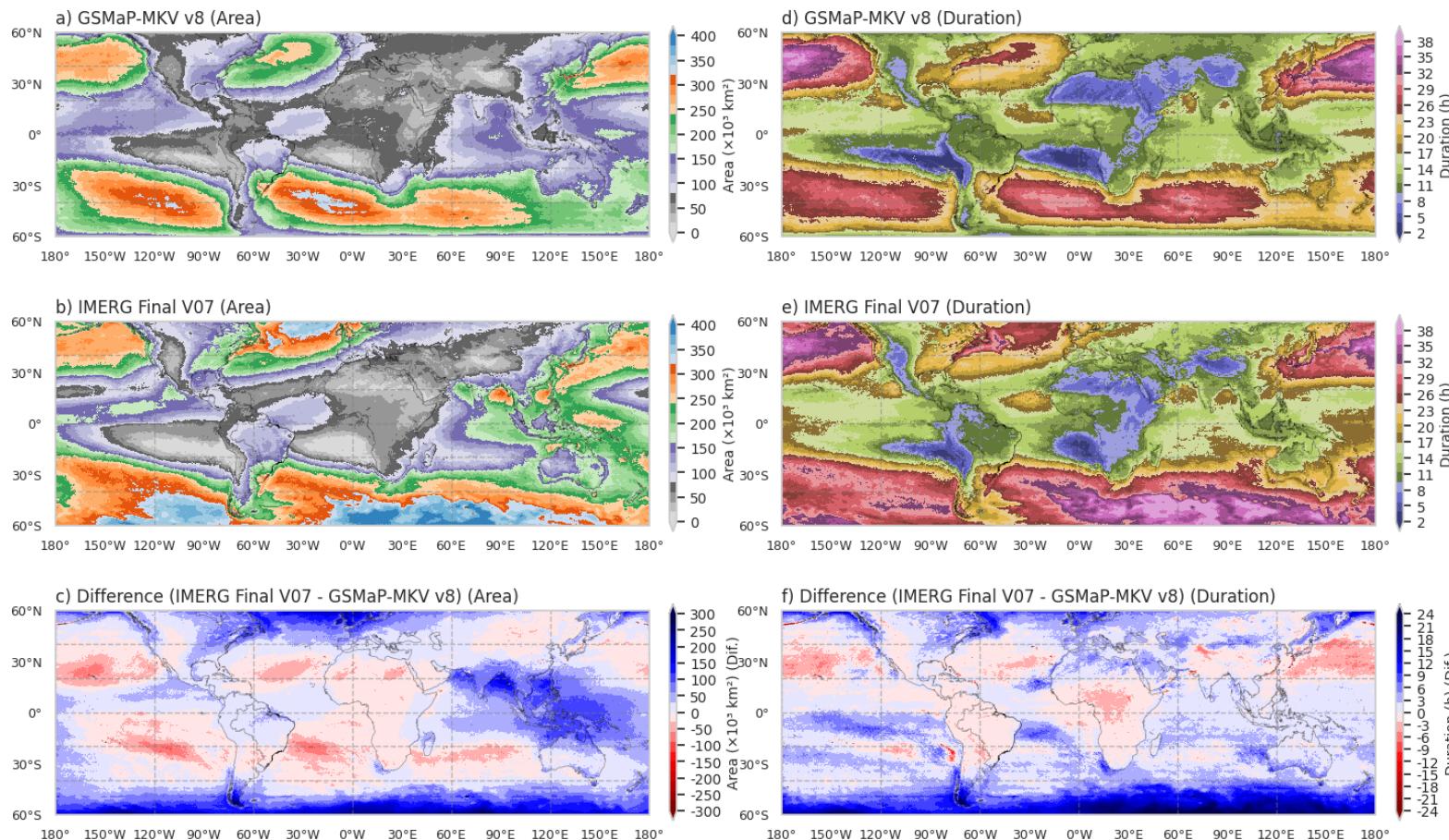
The spatial patterns of mean area (Figure 5.11 a–b) show that both datasets capture large-scale PSs predominantly over the tropical and subtropical oceans, especially the western Pacific, tropical Atlantic, and Indian Ocean. In IMERG, extensive regions exhibit values above 250,000  $\text{km}^2$  over these oceans in latitudes below 30°S, while GSMAp shows a similar pattern but with generally lower values, rarely exceeding 200,000  $\text{km}^2$ . Over continental regions, both products indicate smaller mean areas, with GSMAp often reporting values below 100,000  $\text{km}^2$  in regions such as the Amazon and Congo basins. A difference in substance is observed in South Asia and the Maritime Continent, where IMERG tends to report larger average areas.

The spatial patterns of mean duration (Figure 5.11 panels d–e) indicate broad agreement between GSMAp and IMERG, with both products identifying long-lived PSs over tropical oceans and south extratropics, particularly in the Southern Hemisphere. The most notable discrepancies appear over oceanic regions, where GSMAp tends to represent longer-lived systems, as well as over parts of Europe and the United States. More substantial differences are also evident in the Southern Hemisphere, particularly at latitudes below 40°S, where IMERG consistently reports shorter event lifetimes.

From a meteorological standpoint, the spatial patterns shown in Figure 5.11 align with established understanding of the organization and longevity of PSs across different regions of the globe (HOUZE, 2004; MASUNAGA et al., 2005). Subtropical and tropical oceanic regions, particularly the western Pacific, the tropical Atlantic, and the Indian Ocean, stand out as areas where PSs exhibit the largest mean areas and lifetimes, often exceeding 250,000 km<sup>2</sup> and mean durations greater than 26 hours. Comparable patterns were also reported by Takahashi et al. (2021) and Feng et al. (2021); however, the latter study focused on MCSs, a distinct class of atmospheric systems characterized by more intense associated precipitation (BROWNING, 1986; HOUZE, 1997).

The difference maps (Figure 5.11 panels c and f) further emphasize these divergences. GSMAp generally reports smaller mean areas and longer durations over many continental regions, while IMERG tends to capture larger areas mainly over the monsoonal regions of South Asia and the Maritime Continent, and longer durations over South Hemisphere extratropics. However, the differences are not pronounced, with most areas showing discrepancies low values (light shading), mainly over continental regions. This suggests that while the two products may differ in their absolute values, they generally agree on the spatial distribution of these properties.

Figure 5.11 - Spatial distribution of the mean area and duration of tracked PSs. (a and b) Maps of mean area ( $\text{km}^2$ ) and (c) is the difference (IMERG - GSMAp). (d and e) Maps of mean duration (hours) and (f) is the difference (IMERG - GSMAp).



SOURCE: Author's production.

As a complement to the previous analyses and to highlight both inter-product and regional differences, Figure 5.12 offers a quantitative view of PS characteristics. It presents the frequency distributions of PS size (panel a) and duration (panel b), stratified by the same regional domains used in Section 5.4.1: Land, Ocean, Tropics, and Extratropics. Each panel contains eight curves, with solid lines representing GSMAp and dashed lines representing IMERG. The x-axis shows the PS area ( $\text{km}^2$ ) in panel (a) and duration (hours) in panel (b), while the y-axis indicates the percentage of events, normalized by the total number of tracked systems in each region.

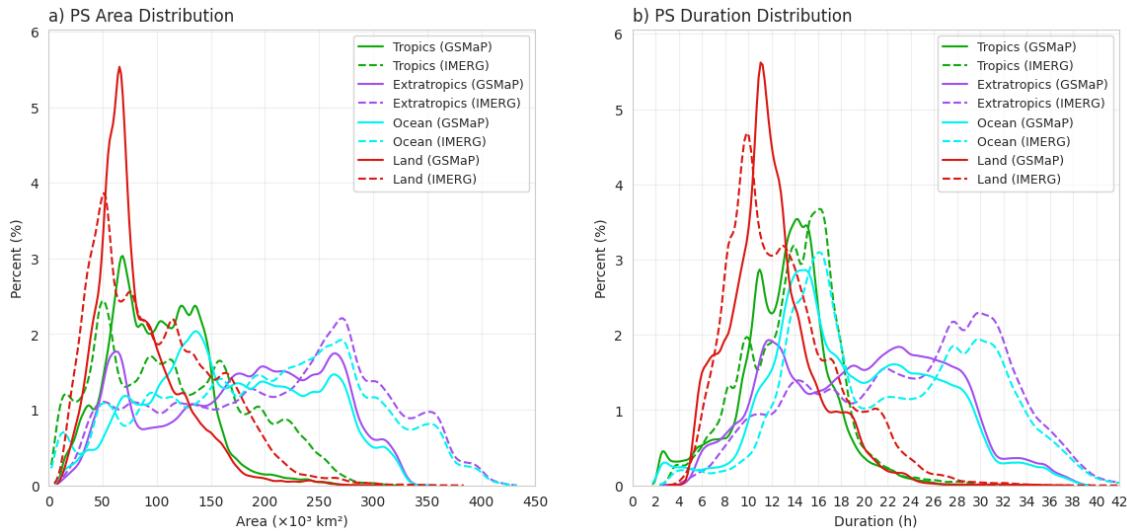
The area distributions in Figure 5.12(a) show distinct differences between the products and across regions. Over land (red lines), GSMAp displays a well-defined peak between 50,000  $\text{km}^2$  and 100,000  $\text{km}^2$ , indicating this as the modal size of systems detected over continental areas. IMERG, in contrast, peaks at smaller sizes near 50,000  $\text{km}^2$  and exhibits a more pronounced tail toward larger areas.

Over the ocean (blue lines), both distributions broaden and shift to larger sizes compared to land. GSMAp shows a peak around 140,000  $\text{km}^2$ , while IMERG peaks near 280,000  $\text{km}^2$ , indicating that larger systems are more common in oceanic environments. Tropical distributions (green lines) follow the same pattern, with IMERG skewed toward smaller sizes and GSMAp favoring larger systems.

In the extratropics (purple lines), both products present broader area distributions, but IMERG continues to show a higher frequency of large systems (greater than 260,000  $\text{km}^2$ ) compared to GSMAp.

The analysis of the duration distributions (Figure 5.12b) corroborates and deepens these observations. Over land (red lines), shows a peak at shorter durations (11 and 12 hours for IMERG and GSMAp, respectively), with IMERG exhibiting a sharper peak and a more pronounced tail toward shorter durations. This suggests that IMERG is more sensitive to short-lived systems, likely due to its finer temporal resolution. Over the ocean (blue lines), both products show a broader distribution, with IMERG peaking at longer durations (16 hours) compared to GSMAp (16 hours). The tropical distributions (green lines) show a similar pattern, with IMERG again skewed toward shorter durations and GSMAp favoring longer-lived systems. In the extratropics (purple lines), both products display a broader distribution, but IMERG shows a higher frequency of shorter-lived systems compared to GSMAp.

Figure 5.12 - Frequency distributions of PS area (panel a) and duration (panel b) for GSMAp–MVK v8 (solid lines) and IMERG Final V07 (dashed lines), stratified by four regions: Land, Ocean, Tropics, and Extratropics. Percentages are normalized by the total number of tracked PSs in each domain.



SOURCE: Author's production.

Synthesizing these results, the distributional analyses provide clear support for the central hypothesis discussed in earlier sections: the higher temporal resolution of IMERG enables the detection of short-lived events. This is most evident in the duration distributions (Figure 5.12b), where IMERG consistently shows a stronger skew toward shorter lifetimes, particularly over land, in the tropics, and in the extratropics. This pattern indicates that the 30-minute sampling of IMERG resolves individual precipitation pulses that the hourly GSMAp data may miss or merge into a single, longer-lasting system.

The area distributions (Figure 5.12a) reinforce this interpretation. The higher frequency of smaller systems detected by IMERG over land and in the tropics is consistent with the identification of short-lived, smaller-scale, and often more intense cells. Consequently, the larger population of short-duration systems in the IMERG dataset directly explains the higher total PS counts reported in Section 5.3.1, as these brief events substantially increase the overall census.

#### 5.4.3 Evolution of PS area over their life cycle

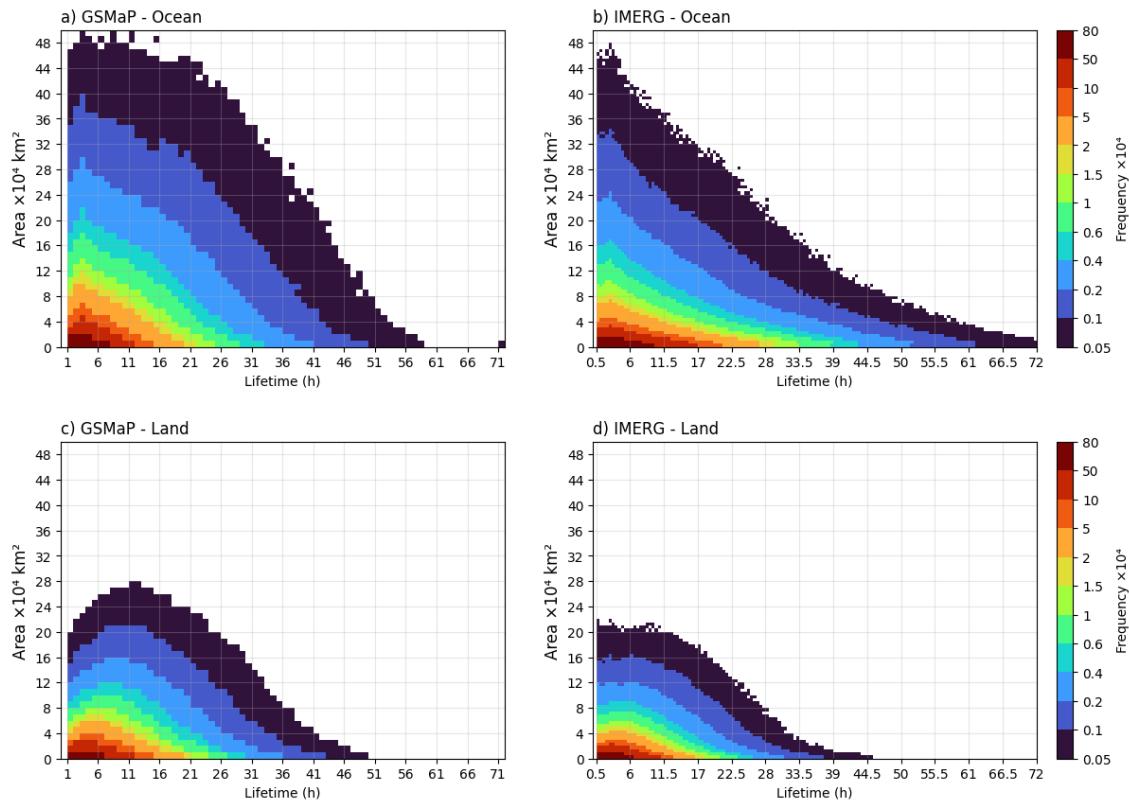
Building on the previous discussion of how GSMAp and IMERG differ in the size and duration of the systems they capture, especially between oceanic and continental regions, this subsection examines how PS area evolves throughout the life cycle. By linking spatial scale with duration, we combine the two pieces of information introduced so far and highlight how input-data characteristics influence the morphology of tracked systems. Figure 5.13 summarizes these life-cycle composites, showing the evolution of PS area from onset to decay, stratified by surface type (ocean in panels a–b and land in panels c–d). Area values from all clusters contributing to each PS were collected at every time step; frequencies were raised to the fourth power to emphasize larger systems, and bins with relative frequency below 0.05% were masked to improve visualization.

The comparison between panels (a) and (b) of Figure 5.13 over the ocean reveals distinct differences in the morphology of life cycle composites for GSMAp and IMERG. In IMERG, the region of highest frequency remains concentrated in relatively small to moderate areas across nearly the entire lifetime spectrum, with the distribution tail extending toward larger systems and longer lifetimes. This pattern indicates that long-duration PSs in IMERG are predominantly compact systems, suggesting the persistence of smaller-scale events, as previously discussed in Section 5.4.2. In contrast, GSMAp shifts the frequency distribution toward progressively larger areas as lifetime increases, such that longer-duration systems tend to be more spatially dispersed. Both products exhibit the typical pattern of initial area growth followed by decay; however, they differ in how size and persistence are coupled: IMERG favors the longevity of compact systems, while GSMAp emphasizes the development of larger, more dispersed systems with extended lifetimes.

Over land, Figure 5.13(c–d) shows that PSs tend to be smaller and shorter-lived than their oceanic counterparts. In both datasets, most events evolve with areas below 40,000 km<sup>2</sup> during the first 10 hours of their life cycle. GSMAp (panel c) exhibits a pronounced peak near 280,000 km<sup>2</sup> at around 12 hours, whereas IMERG (panel d) peaks near 200,000 km<sup>2</sup> at the same time. This indicates that, even over land, GSMAp tends to capture larger systems, while IMERG again emphasizes smaller-scale events. In GSMAp, the growth phase is more marked, with a clear increase in area up to about 12 hours, followed by a gradual decline. IMERG shows a more subdued growth phase and a broader area distribution throughout the life cycle, suggesting a more heterogeneous population of systems.

Although a direct comparison of the life cycle morphology area of PS between GSMAp and IMERG represents a novel contribution, these findings can be contextualized within the broader literature on PS characteristics. Previous studies like Takahashi et al. (2021), Feng et al. (2021), discussing the evolution PS area over the tropics, and create a divided view based on the lifetime range. They found a similar pattern of area evolution, with area curves for events with lifetimes between 21 and 25 hours peaking around  $400,000 \text{ km}^2$ , while shorter-lived events (6-10 hours) peaked closer to  $100,000 \text{ km}^2$ . Similar results can be compared with Machado et al. (1998) and Inoue et al. (2009), who analyzed the life cycle of MCSs.

Figure 5.13 - Life-cycle distributions of PS area for GSMAp–MVK v8 and IMERG Final V07, stratified by surface type. Panels (a)–(b) show oceanic systems and panels (c)–(d) show continental systems. Each diagram displays the joint frequency of area and lifetime, with low-frequency bins masked to emphasize the life-cycle patterns.



SOURCE: Author's production.

These results reinforce that PS morphology is strongly controlled by the spatial and temporal characteristics of each product and by the underlying surface type. To investigate how these morphological differences translate into broader statistical patterns based on duration, the subsequent analyses classify the systems according to their life cycle.

Displacement propensity is first assessed indirectly by stratifying events according to their total lifetime distribution. Let  $\tau_i$  denote the duration, in hours, of the  $i$ -th complete PS. The arithmetic mean and the sample standard deviation of the duration set are then computed as follows:

$$\mu_{\text{duration}} = \frac{1}{N} \sum_{i=1}^N \tau_i, \quad \sigma_{s,\text{duration}} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (\tau_i - \mu_{\text{duration}})^2} \quad (5.1)$$

where  $N$  is the total number of complete PSs,  $\mu_{\text{duration}}$  is the mean duration, and  $\sigma_{s,\text{duration}}$  is the sample standard deviation of the duration set. We then classify the PSs into three duration categories (Short, Medium, and Long) based on the following criteria:

$$\text{Class}(\tau_i) = \begin{cases} \text{Short}, & \tau_i < \mu_{\text{duration}}, \\ \text{Medium}, & \mu_{\text{duration}} \geq \tau_i < \mu_{\text{duration}} + \sigma_{s,\text{duration}}, \\ \text{Long}, & \tau_i \geq \mu_{\text{duration}} + \sigma_{s,\text{duration}} . \end{cases} \quad (5.2)$$

where the three duration categories are defined as:

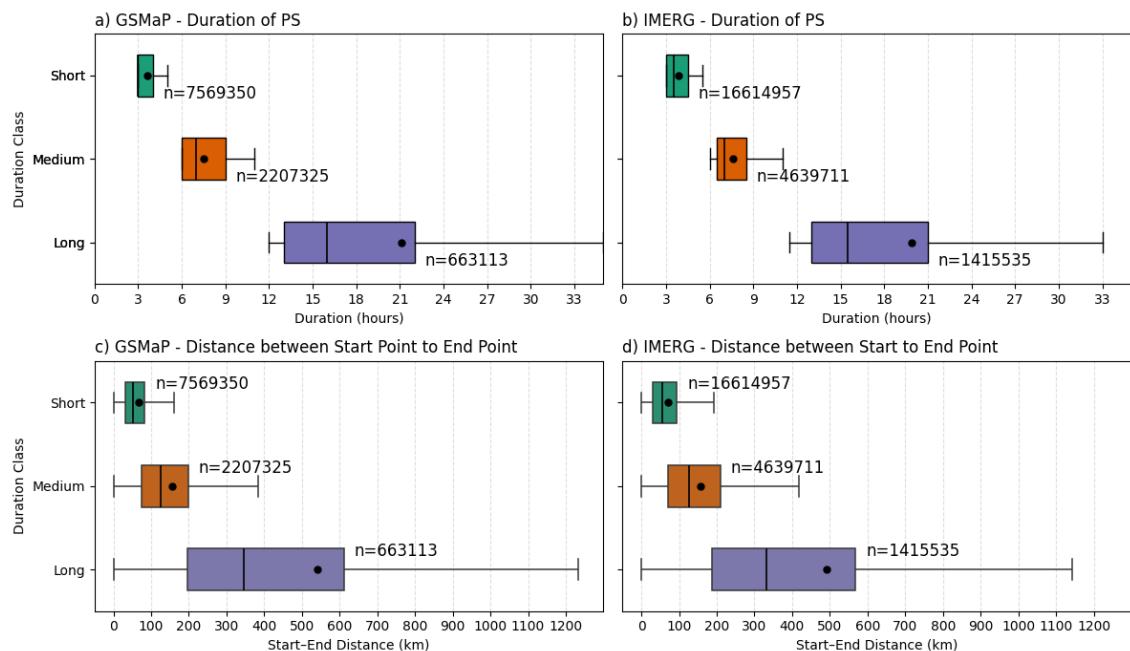
- **Short PSs:** Events with durations below the global mean ( $\mu_{\text{duration}}$ ).
- **Medium PSs:** Events with durations between the global mean and one standard deviation above it ( $\mu_{\text{duration}} + \sigma_{s,\text{duration}}$ ).
- **Long PSs:** Events with durations exceeding one standard deviation above the global mean.

Employing the duration classification from Eq. 5.2, we summarize the statistics of each category in Figure 5.14 for both GSMAp and IMERG track datasets. The figure displays the boxplot distribution of the duration classification groups (panel a-b) and the distance between start (genesis) to end points (dissipation) of tracked

PSs (panel c-d). In GSMAp, *Short* systems have a median of 2 h with IQR 2–3 h, *Medium* center at 6 h (IQR 5–8 h), and *Long* at 15 h (IQR 12–21 h). IMERG mirrors this pattern with medians of 2.5 h (IQR 2–3.5 h), 6 h (IQR 5.5–7.5 h), and 14.5 h (IQR 12–20 h), respectively.

Figure 5.14 Panels (c–d) present, for the same classes, the distributions of the distance between genesis and dissipation points. For Short events, both datasets indicate short displacements. GSMAp registered a median of 51.6 km (mean = 67.0 km; IQR = 31.0–83.2 km), while IMERG showed a median of 56.3 km (mean = 71.5 km; IQR = 29.5–94.9 km). For Medium events, the distributions shift to intermediate and closely matched values. GSMAp showed a median of 125.0 km (mean = 156.0 km; IQR = 74.6–198.3 km), very similar to IMERG's median of 127.8 km (mean = 158.2 km; IQR = 69.6–209.5 km). Finally, for Long events, GSMAp presented a median of 346.8 km (mean = 540.7 km; IQR = 196.3–611.4 km), and IMERG showed a median of 332.9 km (mean = 492.0 km; IQR = 186.5–568.7 km).

Figure 5.14 - Statistics of PS duration classification. (a and b) Boxplots of duration (hours) by class for GSMAp and IMERG, respectively. (c and d) Boxplots of the distance between genesis and dissipation points (km) by class for GSMAp and IMERG, respectively.



SOURCE: Author's production.

Given the duration-based classes defined in Eq. 5.2 and summarized in Figure 5.14, we now restrict the analysis to the *Long* class ( $\tau \geq 12$  h), with the aim of focusing on events with larger horizontal displacement and minimizing contamination from the more numerous short- and medium-lived, quasi-stationary systems. Figure 5.15 shows, for GSMAp and IMERG separately, the difference between dissipation and initiation counts (dissipation minus initiation) for this subset of events: red shading marks grid cells where long-duration PSs dissipate more often than they initiate, whereas blue shading highlights preferred initiation regions.

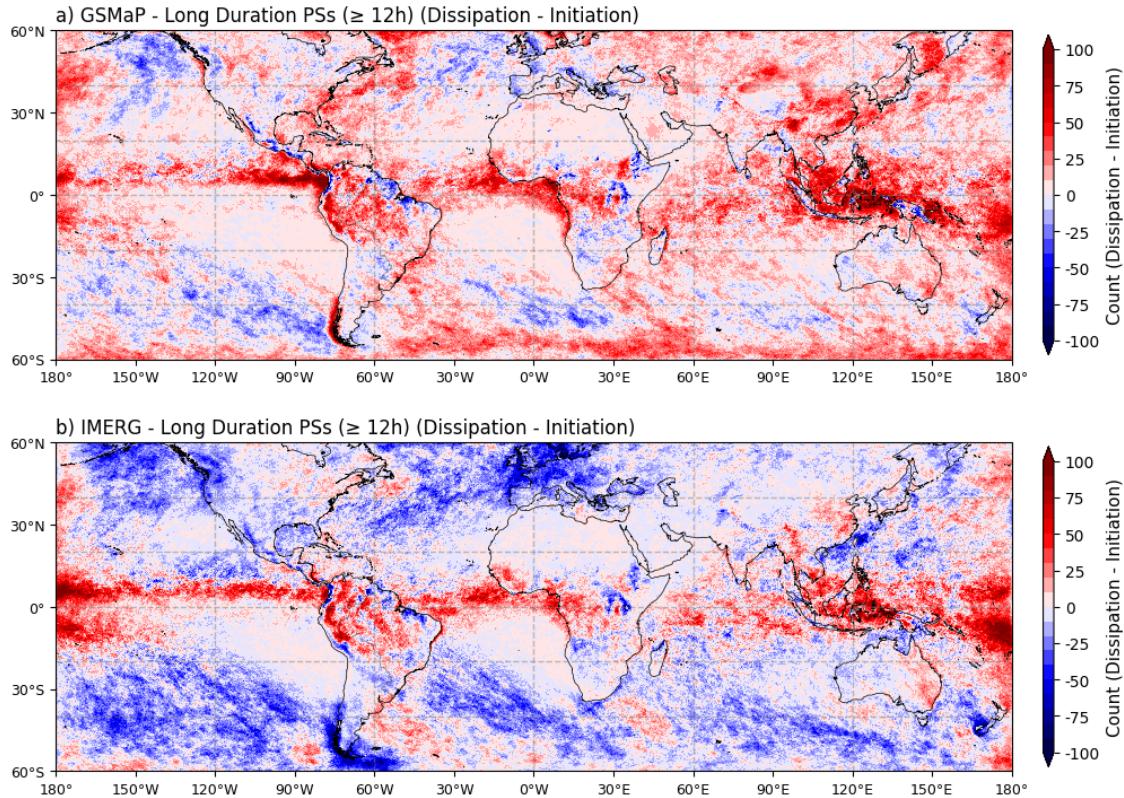
A first notable feature of Figure 5.15 is the marked contrast between GSMAp and IMERG in narrow bands near  $\pm 60^\circ$  of latitude, where large positive and negative values of dissipation minus initiation coincide with the edges of the analysis domain. These signals arise mainly from the truncation of IMERG to  $60^\circ$  S– $60^\circ$  N, adopted to match the native extent of GSMAp: long-lived PSs that form poleward of these bounds and later enter the domain, or that leave the domain before dissipating, are only partially sampled and are counted either as initiation or as dissipation events near the boundaries. The large anomalies at  $\pm 60^\circ$  are therefore best interpreted as artefacts of domain configuration, and the physical discussion is focused on tropical and mid-latitude bands, where both products provide more complete track coverage. Within the tropics, the main initiation and dissipation patterns are generally similar in GSMAp and IMERG, although some localized discrepancies remain; for example, along the eastern slopes of the southern–central Andes near  $\sim 20^\circ$  S, IMERG indicates a pronounced band of initiation that is weaker or of opposite sign in GSMAp. Such differences over complex terrain are consistent with previous evaluations that documented reduced skill and systematic biases of IMERG over the orographically enhanced precipitation regimes of the Chilean Andes (ROJAS et al., 2021).

The dominant global patterns in the tropical band between  $30^\circ$  S and  $30^\circ$  N are broadly consistent between GSMAp and IMERG and closely resemble those reported by (HAYDEN et al., 2021), who also restricted their analysis to storms with lifetimes of at least 12 h to investigate preferred regions of genesis and decay. In that study, tracks derived from IMERG precipitation features showed initiation hotspots over central Africa, the Maritime Continent and northern South America, with dissipation regions displaced downstream along the tropical and extratropical storm tracks. Our maps reproduce these main features: long-duration PSs preferentially initiate along tropical convergence and monsoon regions and over orographic lifting zones, while dissipation becomes more frequent downstream and poleward, especially along subtropical belts and in the lee of major mountain ranges such as the

Andes, Rockies and Himalayas. Methodologically, (HAYDEN et al., 2021) applied a simplified tracking algorithm to large IMERG precipitation features, whereas we use pyForTraCC to track PSs in both GSMAp and IMERG, explicitly handling events that skirt or cross the lateral boundaries of the data domain and thereby helping to resolve the edge effects affecting some of the long-lived tracks in (HAYDEN et al., 2021).

Over South America, our initiation–dissipation patterns for long-lived PSs are also consistent with the MCS initiation climatology of (PREIN et al., 2024). Using IMERG, geostationary IR data and kilometre-scale simulations, those authors diagnosed the initiation frequency of MCSs in  $2^\circ \times 2^\circ$  boxes, considering only initiations whose cloud shields were smaller than  $40\,000\text{ km}^2$  and defining initiation as the geometric centre of the MCS at its first detection. Their analysis combined six tracking algorithms (MOAAP, tobac, ForTraCC, TAMS, PyFLEXTRKR and TOOCAN) under a common MCS definition and revealed robust initiation hotspots along the northeastern Brazilian coast, over the southern Amazon Basin, across the Guiana Highlands and along the eastern slopes of the central and northern Andes. Our maps show enhanced initiation of long-lived PSs in these same regions, while additionally documenting that many of these systems dissipate over the lee of the Andes and the subtropical South Atlantic, underscoring their role in exporting convection and moisture from continental source regions toward downstream oceanic storm tracks.

Figure 5.15 - Hotspots of dissipation minus initiation for long-duration precipitating systems (PSs with lifetimes  $\geq 12$  h) tracked by pyForTraCC. Panels show (a) GSMAp and (b) IMERG. Red (blue) shading indicates grid cells where long-lived PSs dissipate more (less) frequently than they initiate.



SOURCE: Author's production.

Identifying where long-lived PSs preferentially initiate and where they dissipate is essential for characterizing regional hotspots of precipitation. The close agreement between GSMAp and IMERG, together with the consistency with previous tracking-based studies (HAYDEN et al., 2021; PREIN et al., 2024; LIU; ZIPSER, 2015; ROJAS et al., 2021), supports the use of pyForTraCC as a tool for comparing life-cycle statistics across satellite precipitation products with distinct retrieval characteristics. Having established the main genesis and dissipation hotspots of long-duration PSs, the next section examines their kinematics, investigating the overall displacement of these systems throughout their lifetimes.

## 5.5 Global displacement of precipitation systems

In this section delves into the analysis of the kinematics of PSs, investigating the trajectory patterns and propagation mechanisms that govern their displacement on a global scale. Using the tracked dataset, the objective is to characterize the preferred pathways, average speeds, and the seasonal and regional variability of PS movement. Understanding these patterns is crucial not only for the climatological characterization of the systems but also for enhancing the predictability of precipitation events through extrapolation techniques (WATSON et al., 1988; EICHHOLZ, 2017; LEAL et al., 2022).

Before analyzing the trajectories, it is necessary to evaluate the quality and robustness of the displacement vectors that support this analysis. As detailed in Section 4.1.3.2.5, the `pyForTraCC` algorithm applies vector correction methods (e.g., *MCor*, *SCor*, *ICor*, *ECor*, *OCor*) to reduce uncertainties associated with the non-rigid morphology of PSs. This section therefore begins with a quantitative examination of these events and of their impact on the performance of the correction methods. The goal is to establish the reliability of the kinematic data that will later be used to describe propagation patterns across different climatic and geographical domains.

### 5.5.1 "Continuous" and "Non-continuous" quantification

To clarify the performance of the vector-correction methods and to aid the interpretation of the spatial patterns presented in the next section, it is useful to examine the dynamical nature of the tracked PSs. Over the entire tracking period (2015–2024), the occurrence of each tracking-status class defined in Section 4.1.3.2.5 is quantified and, for each vector-correction method, the effectiveness with which these classes are corrected is evaluated. Because the activation of each correction scheme depends on specific tracking situations, these status-based diagnostics provide a direct measure of how much each method actually corrects in practice and of the residual uncertainties associated with their use. For example, the *SCor* method is typically applied when a PS splits into two or more parts, while the *MCor* method is used when two or more PSs merge into a single entity. In the analyses below, the tracking status is assessed in terms of absolute counts and relative frequencies, and the spatial distribution of these occurrences is provided in the supplementary material (Appendix A.8).

Table 5.2 summarizes the occurrences of each tracking status the GSMAp and IMERG datasets over entire tracking period (2015–2024). The analysis reveals that the vast majority of events are classified as continuity (CON), accounting for 63.87% for GSMAp and 75.32% for IMERG. This result indicates that, for most of their life-cycle, precipitating PSs maintain sufficient spatiotemporal coherence to be tracked without complex interactions. System genesis (NEW) is the second most frequent category, albeit with a notable difference between the products: 16.50% in GSMAp versus 8.40% in IMERG. This discrepancy may be attributed to different sensitivities in precipitation detection, and reinforce the earlier hypothesis that IMERG captures more short-lived events. The merging (MRG) and splitting (SPL) events are less common but still significant, with GSMAp showing 12.03% and 3.68%, respectively, while IMERG reports 8.96% and 4.11%. The combined events (MRG/SPL and NEW/SPL) are relatively rare, together accounting for less than 4% of all events in both datasets.

Table 5.2 - Quantification of tracking status occurrences for GSMAp and IMERG products during the 2015-2024 period.

Status	GSMAp Qty.	IMERG Qty.	GSMAp %	IMERG %
CON	38,173,855	178,029,630	63.87%	75.32%
NEW	9,861,853	19,862,433	16.50%	8.40%
MRG	7,189,581	21,167,947	12.03%	8.96%
SPL	2,201,901	9,725,880	3.68%	4.11%
MRG/SPL	1,314,431	4,101,835	2.20%	1.74%
NEW/SPL	1,022,977	3,469,825	1.71%	1.47%

SOURCE: Author's production.

To better understand the PS characteristics associated with the dynamic interactions represented by the tracking statuses, the tracked events were grouped into two categories based on their life-cycle evolution: Continuous and Non-continuous. A Continuous PS is defined as one that maintains the continuity status (CON) throughout its entire life cycle, from genesis to dissipation. In contrast, a Non-continuous PS is one that experiences at least one interaction event (MRG, SPL, or MRG/SPL) during its life cycle.

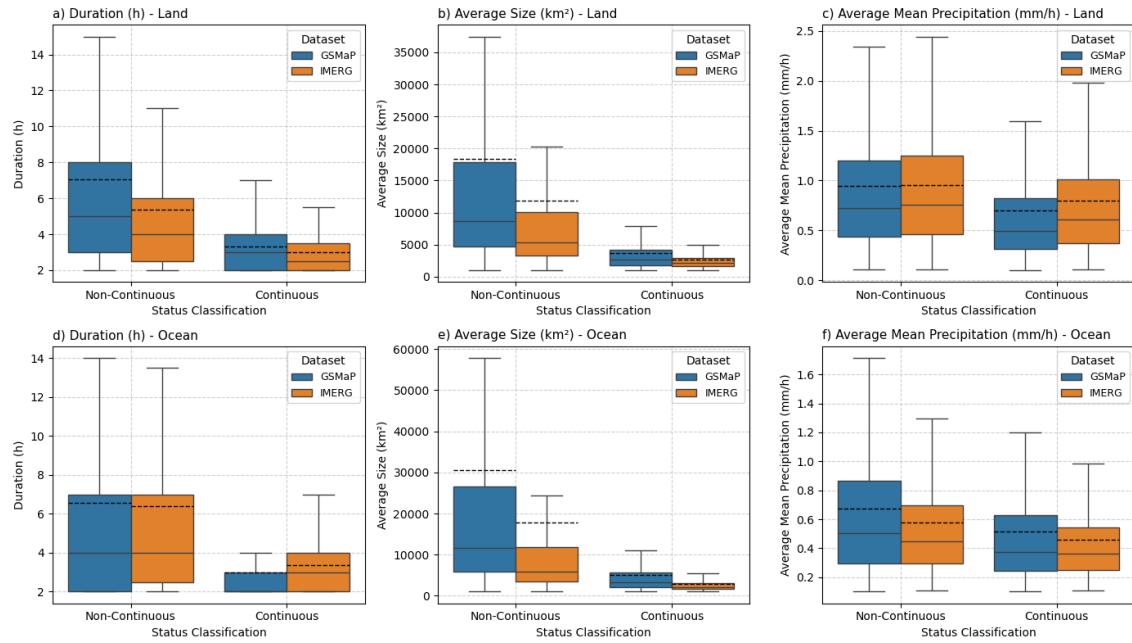
Figure 5.16 compares Continuous and Non-continuous PSs in terms of duration, average size and average mean precipitation for GSMAp and IMERG, separated

into land (panels a–c) and ocean (panels d–f). Over land, Non-continuous events are clearly longer lived and larger than Continuous ones, with the strongest contrast in average size (panel b), whose interquartile range is shifted toward much greater areas. Durations show the same behavior (panel a), with higher medians and means for Non-continuous systems, while average mean precipitation is only slightly higher (panel c). Across products, GSMAp systematically represents larger systems than IMERG, whereas IMERG tends to show marginally higher mean precipitation. Over the ocean (panels d–f), the same hierarchy between classes and products is maintained.

The land–ocean contrast is particularly clear in the oceanic regime. When mean size is considered, Non-continuous systems over the ocean remain substantially larger than Continuous ones. In GSMAp, the median area increases from about 4,000 km<sup>2</sup> for Continuous to 18,000 km<sup>2</sup> for Non-continuous events, with an interquartile range of roughly 10,000–27,000 km<sup>2</sup> (panel e). In IMERG, the median size similarly rises from about 2,500–3,000 km<sup>2</sup> to 8,000 km<sup>2</sup>. Mean precipitation also shows a stronger separation between classes over the ocean than over land. Median values increase from approximately 0.50 mm h<sup>-1</sup> for Continuous systems to about 0.70 mm h<sup>-1</sup> in GSMAp and 0.60 mm h<sup>-1</sup> in IMERG for Non-continuous events (panel f), whereas over land the difference between classes is on the order of 0.2–0.3 mm h<sup>-1</sup> (panel c). Across products, GSMAp systematically exhibits higher median rain rates and a broader upper tail than IMERG in the oceanic domain (panels d–f). IMERG, in turn, tends to report slightly larger median sizes for its Non-continuous events relative to its Continuous population (for example, 8,000 versus 2,500–3,000 km<sup>2</sup> in panel e).

Overall, these results show that Non-continuous PSs are systematically larger, longer lived, and more intense than Continuous systems, particularly over the oceans. This pattern is consistent with MCS climatologies that attribute a disproportionate share of regional precipitation to the largest and longest-lived systems (LAING; FRITSCH, 1997; FUTYAN; GENIO, 2007), with IMERG evidence that storm-total rainfall increases with lifetime and characteristic rain rate (TAKAHASHI et al., 2021), and with recent findings that discontinuous MCSs produce stronger precipitation than continuous ones (ZHANG; CHEN, 2025).

Figure 5.16 - Boxplots comparing the characteristics distinct dynamic interactions of PSs classified as "Continuous" and "Non-continuous" over Land (upper panels) and Ocean (lower panels) for both GSMAp and IMERG datasets. The characteristics analyzed are: (a and d) Duration (hours), (b and e) Average Size ( $\text{km}^2$ ), and (c and f) Average Mean Precipitation ( $\text{mm/h}$ ).



SOURCE: Author's production.

In summary, the contrasted behavior of Continuous and Non-continuous PSs highlights that the frequency of dynamic interaction events such as splits and merges in aspect of PS kinematics, not just a detail of their morphology. Since some of the displacement-vector correction methods in pyForTraCC are explicitly designed to handle these interaction regimes, it is essential to verify how they perform and how much they reduce kinematic uncertainties. In the next section, we present the results on how the displacement-vector corrections affect the propagation of PSs as a function of the classes introduced above.

### 5.5.2 Evaluation of displacement vector correction

The kinematics of PSs are fundamentally represented by their displacement vectors, which describe the movement of the systems. However, the morphological changes and the centroid-based tracking approach can introduce significant uncertainties in these vectors. The non-rigid nature of PSs means that their shapes and sizes can

change rapidly, leading to potential misrepresentations of their true movement if not properly accounted for (LEAL et al., 2022).

To reduce kinematic uncertainties, `pyForTraCC` employs several vector-correction methods (Section 4.1.3.2.5). For each consecutive image pair  $(t - 1, t)$ , an internal validation routine tests all available schemes and selects the one that yields the lowest FAR (Eq. 4.4). The counts reported in Table 5.3 therefore represent the number of times each method was the most effective option, that is, when it produced the smallest displacement error.

The *Optical Flow Method* (OCor) clearly dominates the statistics, accounting for 79.94% of best-performing corrections in GSMAp and 83.70% in IMERG. This strong prevalence underscores the central role of motion-estimation techniques in stabilizing displacement vectors when pure overlap information is not sufficient. The *Inner Cores Method* (ICor) is the second most frequently selected scheme, with 16.31% of activations in GSMAp and 14.51% in IMERG, indicating that internal intensity patterns often provide robust guidance when outer system boundaries are noisy or highly deformable.

In contrast, interaction-based methods such as *Merge* (MCor) and *Split* (SCor) are seldom the best choice, with activations below 4% in both products, reflecting the smaller fraction of time steps affected by mergers and splits. The *Ellipse Method* (ECor) is the least selected, with fewer than 0.1% of occurrences, which suggests that shape-based corrections are advantageous only in very specific configurations. Taken together, these results indicate a clear predominance of OCor and ICor, showing that most displacement uncertainties are controlled by translation and internal structure rather than by gross morphological changes.

Table 5.3 - Quantification of the application of displacement-vector correction methods selected by the automatic mode for GSMAp and IMERG products during the 2015–2024 period, in comparison to the non-correction approach.

Method	GSMAp Qty.	IMERG Qty.	GSMAp (%)	IMERG (%)
OCor	19,286,841	87,820,297	79.94%	83.70%
ICor	3,935,670	15,228,427	16.31%	14.51%
MCor	837,529	1,719,546	3.47%	1.64%
SCor	52,045	108,923	0.22%	0.10%
ECor	15,090	48,102	0.06%	0.05%

SOURCE: Author's production.

Under the tracking configuration used in this thesis, no single method emerges as a universal winner. Instead, the combination of correction techniques, chosen according to the prevailing situation and feature set (for example, occurrence of merges or splits), most effectively reduces displacement-vector uncertainty and mitigates artifacts associated with abrupt morphological changes. Under highly dynamic evolution, optical-flow vectors tend to perform best once systems grow in size or duration. In contrast, events without correction remains the safest option for small, short-lived systems. Both sets of results, IMERG systematically attains lower FAR than GSMAp within the winning methods and strata. Detailed analyses supporting these findings are presented in Appendix A.1.

Figure 5.17 provides a global assessment of FAR values. Panels (a–b) show the results obtained with the automatic selection of correction methods, while panels (c–d) present a map with no correction. Panels (e–f) display the gain Skill Score, which measures the improvement of the automatic selection relative to the baseline. The left column corresponds to GSMAp, and the right column to IMERG.

In GSMAp, the corrected field (Figure 5.17a) clusters mainly within  $0.20 \leq \text{FAR} \leq 0.35$ . Values near 0.20–0.25 dominate the equatorial Atlantic and Pacific, the tropical Indian Ocean, the Maritime Continent, and much of equatorial Africa and the Amazon, while 0.30–0.35 becomes more common toward the subtropics. The baseline field (Figure 5.17c) shows FAR frequently above 0.45 over large areas, indicating substantially poorer performance. The few areas that remain above  $\text{FAR} > 0.40$  under correction correspond mainly to two red belts: the southeastern Pacific and the South Atlantic subtropics (roughly  $20^\circ\text{--}35^\circ\text{S}$ ), where the number of tracked PS events is low as mentioned in Section 5.3.1, and the complex orography of the central Andes (LIU, 2011; HOUZE, 2012). The Skill Score (Figure 5.17e) quantifies these improvements, showing positive gains almost everywhere, with the largest enhancements (up to 0.50) in the deep tropics, especially along the Pacific (latitudes  $0^\circ\text{S}\text{--}10^\circ\text{N}$ ) and the Southern Hemisphere Zones of Convergence (latitudes  $10^\circ\text{S}\text{--}20^\circ\text{S}$ ).

For IMERG (panels b, d, f of Figure 5.17), the automatic mode concentrates *very low* FAR in pockets with  $\text{FAR} < 0.15$  distributed over different zones of globe, especially over regions where the number of tracked PS events is high (Section 5.3.1) and areas of intense precipitation (Section 5.4.1), such as the tropical oceans, the Maritime Continent, and equatorial Africa. The baseline without correction (Figure 5.17d) shows a more heterogeneous pattern, with FAR values frequently exceeding 0.35

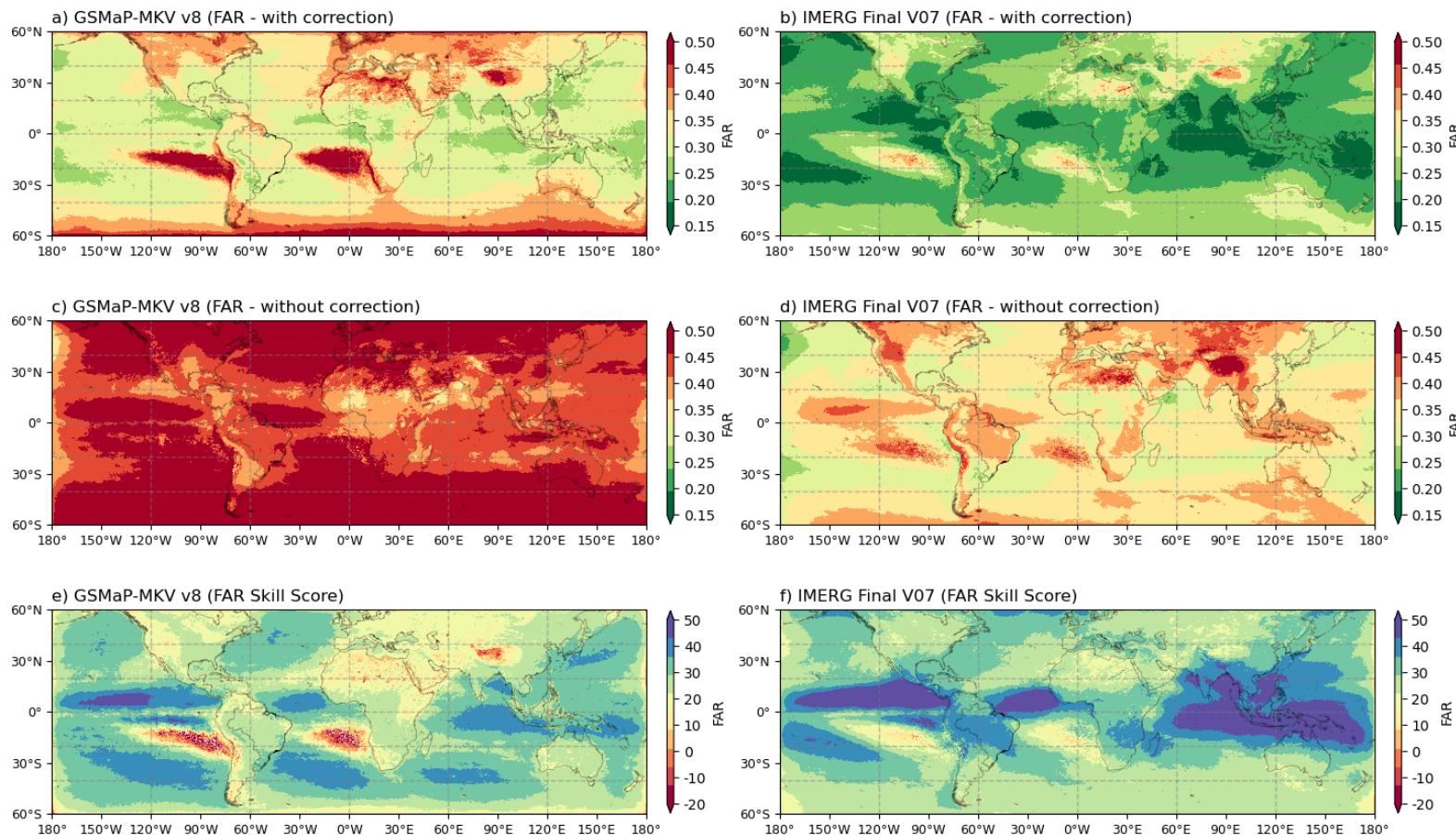
in many regions, particularly over continental areas. The Skill Score (Figure 5.17f) indicates substantial improvements with the automatic selection, with gains often exceeding 0.50 in the tropics and especially over the monsoonal regions of South Asia and the Maritime Continent.

In line with the distributions discussed in Section 5.3.1, regional contrasts in *FAR* are consistent with well-known oceanic–continental and orographic controls that modulate the morphology and continuity of PSs (HOUZE, 1997; HOUZE, 2004). Over oceans, where PSs are typically more organized on large scales and more persistent, *FAR* values are generally lower, indicating that the tracking and vector-correction methods capture their movement more effectively. Over continents, particularly in regions with complex terrain such as the Andes and the Himalayas, higher *FAR* values are found, reflecting the greater difficulty in tracking short-lived and fragmented systems.

Crucially, the inter-product differences in *FAR* stem less from spatial sampling and more from how the vector-correction methods interact with the distinct precipitation characteristics of each dataset. IMERG’s higher temporal resolution allows it to capture short-lived events and to sample the evolution of PSs more smoothly, increasing the number of tracked systems while providing a more refined depiction of their life cycles. Rather than introducing additional morphological variability and inflating the initial *FAR* in the baseline scenario, the larger number of frames tends to reduce timing and displacement uncertainties when compared to GSMap, thereby decreasing the need for subsequent vector corrections.

Having established the skill score and limitations of the correction methods, the next section examines the propagation of PSs using the displacement vectors selected by the automatic mode, characterizing their main kinematic patterns across different regions and regimes.

Figure 5.17 - Spatial distribution of the False Alarm Ratio (FAR) for GSMAp (left) and IMERG (right). (a and b) FAR with automatic selection of correction methods. (c and d) FAR without any correction. (e and f) Skill Score representing the improvement achieved by the automatic selection relative to the baseline.



SOURCE: Author's production.

### 5.5.3 Propagation and displacement dynamics of precipitating systems

A central aspect of the kinematics of PSs is their displacement patterns, which can be synthesized through the zonal ( $u$ ) and meridional ( $v$ ) vector components derived from the optimized displacement vectors. For each cluster of PS the components are representative motion vector of the centroid displacement and distributed spatially according to the system's coordinate at each time step. Figure 5.18 presents the global composites of these average displacement vectors for GSMAp (left) and IMERG (right). The vectors are plotted on a regular grid, with their lengths scaled to represent the average speed of displacement (in  $\text{m s}^{-1}$ ), and their orientations indicating the prevailing direction of motion. Panels (c)–(d) of Figure 5.18 express the inter-dataset differences on the same grid: (c) the signed difference in mean displacement speed ( $\text{IMERG} - \text{GSMAp}$ ,  $\text{m s}^{-1}$ ); and (d) the signed difference in mean displacement *angle* (degrees), obtained from the bearings of the mean vectors in (a)–(b) and shown on  $[-180^\circ, 180^\circ]$ .

The mean displacement vectors in Figure 5.18a–b reveal a latitude-structured kinematic regime for clusters that is consistent with canonical wind climatology (HANLEY et al., 2010). In equatorial regions ( $\pm 10^\circ$  latitude), clusters predominantly exhibit westward motion with a slight poleward component and modest speeds, typically on the order of  $3\text{--}6 \text{ m s}^{-1}$ , with locally weaker means ( $\lesssim 3 \text{ m s}^{-1}$ ) over continental interiors. Moving into the south subtropics ( $-10^\circ\text{--}30^\circ$  latitude), vectors veer eastward and accelerate, particularly across the extratropical oceans ( $30^\circ\text{--}60^\circ$  latitude), where speeds peak at  $15\text{--}18 \text{ m s}^{-1}$ . Over land, propagation remains patchier and slower ( $6\text{--}10 \text{ m s}^{-1}$ ). This pattern aligns with the influence of the westerlies, preserving hemispheric symmetry and marking a clear transition from tropical easterlies to extratropical eastward motion (SALIO et al., 2007).

Quantitatively, the centroid velocities diagnosed by pyForTraCC are consistent with the ranges reported by previous global tracking studies, which document predominantly westward motions of about  $4\text{--}8 \text{ m s}^{-1}$  in the tropics and faster, mainly eastward propagation of about  $8\text{--}15 \text{ m s}^{-1}$  in the midlatitudes (FENG et al., 2021; HAYDEN et al., 2021). Similar magnitudes are reported for MCSs over southeastern South America and for precipitation and cloud structures embedded in extratropical cyclones, indicating that the PS displacements retrieved here fall within the expected range for organized systems (SALIO et al., 2007; FIELD; WOOD, 2007).

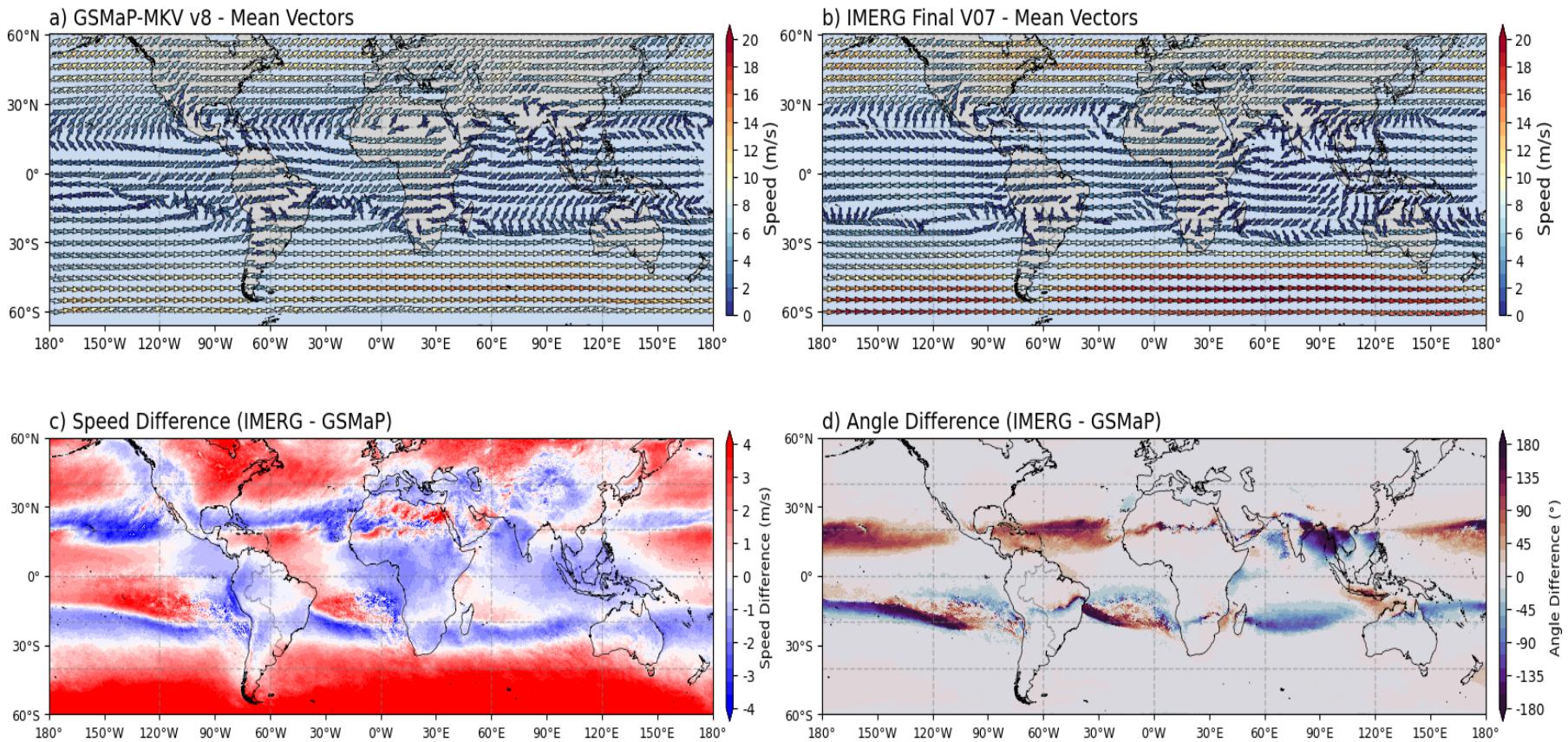
An important caveat is that the vector components in Figure 5.18(a–b) represent the vector-mean displacement of each PS cluster obtained from tracking, not direct

measurements of the ambient wind field. They summarize the net motion of the system as a whole between consecutive time steps and should be interpreted as a bulk kinematic descriptor of PS motion, rather than a detailed representation of the underlying flow. Other factors, such as the vertical and horizontal wind structure and the internal growth and decay of precipitation within the system, can also influence these displacements (CORFIDI, 2003) but are not explicitly resolved in this analysis. Differences between PS displacement vectors and standard wind climatologies or station observations are therefore expected and should not be interpreted as inconsistencies.

In the inter-dataset diagnostics (Figure. 5.18c–d), speed differences (IMERG – GSMAp) are modest in the deep tropics, typically within  $\sim 0.5\text{--}1 \text{ m s}^{-1}$ , but grow poleward of  $\sim 30^\circ$  in both hemispheres, forming zonal bands of  $\pm 2\text{--}3 \text{ m s}^{-1}$  with localized peaks near  $\sim 4 \text{ m s}^{-1}$  over midlatitude oceans. These contrasts likely reflect distinct precipitation retrieval behavior in cooler, more stratiform or frozen-hydrometeor regimes and edge effects from the analysis window ( $-60^\circ$  to  $60^\circ$ ) versus IMERG's native coverage to  $\pm 90^\circ$ , which can truncate tracks and bias speeds near the domain limits and high-elevation rings.

Angular differences (panel d) cluster along the ITCZ/SPCZ corridors and some coastal/orographic sectors (SCHNEIDER et al., 2014; GILL, 2016), with frequent offsets of  $30^\circ\text{--}90^\circ$  and occasional near-opposite headings ( $\pm 180^\circ$ ) where steering is complex or track splitting/merging occurs which was supplemented in the results in Appendix A.8. Overall, both products depict the same large-scale displacement regime, with IMERG generally yielding slightly faster and more zonally aligned motions in midlatitudes, while GSMAp shows more variability in angular orientation in convectively active tropical regions.

Figure 5.18 - Global composites of average displacement vectors for precipitation clusters from GSMAp (a) and IMERG (b), with vector lengths scaled to represent average speed ( $\text{m s}^{-1}$ ). Panels (c) and (d) show the signed differences between IMERG and GSMAp in mean displacement speed ( $\text{m s}^{-1}$ ) and angle (degrees), respectively.

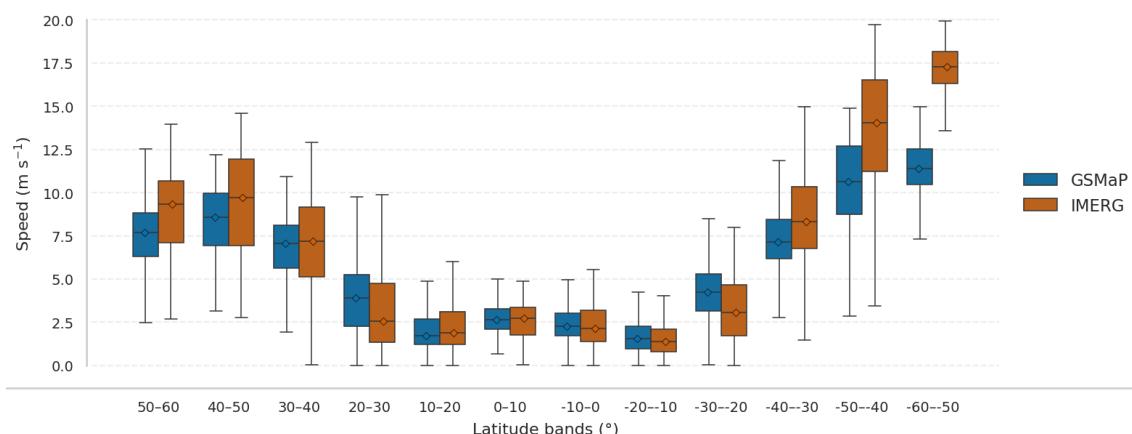


SOURCE: Author's production.

Latitude–binned displacement speeds ( $10^{\circ}$  bands) are synthesized in Figure 5.19 through boxplots that summarize the distribution of PS speeds within each band for both GSMAp and IMERG. The medians delineate a clear latitudinal structure: slowest speeds (medians  $\sim 1.5$ – $2.5 \text{ m s}^{-1}$ ) occur near the equator ( $10^{\circ}$ – $10^{\circ}$ ), increasing through the subtropics (medians  $\sim 3$ – $5 \text{ m s}^{-1}$  at  $20^{\circ}$ – $30^{\circ}$ ) and peaking in the midlatitudes (medians  $\sim 7$ – $9 \text{ m s}^{-1}$  at  $40^{\circ}$ – $50^{\circ}$ , reaching  $\sim 11$ – $13 \text{ m s}^{-1}$  at  $50^{\circ}$ – $60^{\circ}$ ). IMERG systematically yields higher speeds than GSMAp across most bands, with median offsets typically  $\approx 0.5$ – $1.5 \text{ m s}^{-1}$  in the tropics and up to  $\sim 2$ – $3 \text{ m s}^{-1}$  poleward of  $30^{\circ}$ . This progression is dynamically consistent with the climatological transition from weak equatorial steering to stronger trade-wind regimes and, poleward, to baroclinic westerlies and jet-stream influences (STENDEL et al., 2021).

In the Southern Hemisphere mid–high latitudes ( $-30^{\circ}$  to  $-60^{\circ}$ ) the largest interproduct differences emerge, with IMERG medians rising from  $\sim 8$ – $9 \text{ m s}^{-1}$  at  $-40^{\circ}$ – $-30^{\circ}$  to  $\sim 14 \text{ m s}^{-1}$  at  $-50^{\circ}$ – $-40^{\circ}$  and  $\sim 17$ – $18 \text{ m s}^{-1}$  at  $-60^{\circ}$ – $-50^{\circ}$ , whereas GSMAp medians are  $\sim 7.5$ – $8$ ,  $\sim 11$ , and  $\sim 12$ – $13 \text{ m s}^{-1}$ , respectively. These contrasts are consistent with the displacement patterns in Fig. 5.18 and likely reflect differences in precipitation representation near the tracking threshold ( $0.1 \text{ mm h}^{-1}$ ), aliasing linked to native temporal resolution (IMERG 15 min vs. GSMAp hourly), the more variable extratropical background flow, and the generally faster motion of oceanic systems (LUIZ; FIEDLER, 2024), and seen in previous Figure 5.18 (a–b).

Figure 5.19 - Latitudinal distribution of PS displacement speeds ( $\text{m s}^{-1}$ ) for GSMAp (blue) and IMERG (orange) across  $10^{\circ}$  latitude bands.



SOURCE: Author's production.

In summary, the tracking diagnostics based on the extracted vector components from datasets consolidate a coherent picture of PS propagation and displacement dynamics. Both GSMAp and IMERG capture a latitude-structured regime of motion that aligns with large-scale atmospheric circulation patterns, with IMERG generally yielding slightly faster and more zonally aligned displacements, particularly in mid-latitudes. Tropical PSs preferentially translate slowly westward with slight cross-equatorial components, while subtropical and mid-latitude systems accelerate and veer to predominantly eastward headings, reflecting the growing influence of westerly steering flows. These findings corroborate prior observational and modeling studies of convective system kinematics (FENG et al., 2021; HAYDEN et al., 2021), while also highlighting the sensitivity of displacement estimates to the underlying precipitation product characteristics and tracking configurations employed.

By the end of this section, we conclude that the tracking algorithm successfully identified features that are consistent with previous studies. The results show that the `pyForTraCC` vector methods remains robust, even with differences among the input datasets. This robustness indicates that the algorithm is suitable for use in multiple areas of meteorology.

## 5.6 Global decadal trends for precipitating systems

A common application in climate science is decadal trend analysis. In this context, we use the tracking datasets produced by `pyForTraCC` to evaluate trends in both the instantaneous behavior of precipitation clusters and the mean properties of complete PSs. We first analyse cluster diagnostics, focusing on accumulated precipitation and cluster area. And for the complete life cycle of PSs, we then examine trends in mean precipitation, mean duration, and mean area. These properties are computed for each tracked system over 2015–2024, mapped onto the regular tracking grid, and aggregated into monthly means by averaging the values of all PSs present in each grid cell.

The trend analysis adopts the non-parametric Mann–Kendall (MK) test (MANN, 1945; KENDALL, 1938) to evaluate the presence of monotonic trends in the monthly spatial time series of each PS property at every grid. This choice is well suited to precipitation-related fields, which often exhibit skewness, outliers, and mild serial/spatial dependence. For trend magnitude, the Theil-Sen estimator (THEIL, 1950; SEN, 1968) is then used to quantify the slope of significant trends. Statistical significance is assessed at the 5% level, and the results are masked to show those grid cells where trends are significant. The following figures summarize the