



Intercomparison of H SAF and IMERG heavy rainfall retrievals over a Mediterranean coastal region

Eric Peinó ^{a,*,} Marco Petracca ^{b,} Francesc Polls ^{a,} Mireia Udina ^{a,} Joan Bech ^{a,c}

^a Department of Applied Physics—Meteorology, Universitat de Barcelona, Barcelona, 08028, Spain

^b National Research Council of Italy, Institute of Atmospheric Sciences and Climate (CNR-ISAC), Rome, 00133, Italy

^c Water Research Institute, Universitat de Barcelona, Barcelona, 08028, Spain

ARTICLE INFO

Keywords:

Extreme precipitation
Satellite estimations
QPE
Cloud microphysics
IMERG
H SAF
Ground validation

ABSTRACT

Satellite-based precipitation products play a crucial role in providing global, continuous, and reliable estimates of rainfall, essential for understanding and managing Earth's water cycle. This study evaluates the accuracy of three H SAF satellite products (H61B, H64, and H68) and compares their performance with IMERG V06B (Early and Late Runs) products in detecting and estimating extreme precipitation events in the western Mediterranean region. The analysis is based on hourly and daily rainfall data collected from 186 rain gauges in Catalonia (NE Iberian Peninsula), using a point-to-pixel approach. The results show that satellite estimates tend to overestimate low precipitation accumulations (less than 2 mm and 10 mm in one hour and one day respectively), with this overestimation being more evident in the H68 product at the hourly scale and in IMERG Late at the daily scale. However, all products show a substantial decline in accuracy for higher precipitation amounts, particularly when they exceed 10 mm in one hour and 30 mm in one day (relative errors up to −60% and critical success index less than 33% in all cases). Despite its biases, IMERG Late product emerged as the most reliable for detecting substantial rainfall accumulations. Additionally, this analysis examined the relationship between the microphysical properties of the precipitating cloud top and estimated or observed surface precipitation. Accurate precipitation estimates from the satellite products were consistently associated with high values of Cloud Optical Thickness (COT) and Ice Water Path (IWP), while false alarms were often linked to low values of these variables. As expected, the results indicated a poorer performance in estimating precipitation associated with warm clouds. Additionally, these findings highlight the importance of integrating cloud optical and microphysical properties into retrieval algorithms to improve satellite precipitation estimates.

1. Introduction

Observing and measuring precipitation on a global scale is crucial to our understanding of the Earth system and has a substantial impact on society at multiple levels (Kidd et al., 2021; Kirschbaum et al., 2017; Skofronick-Jackson et al., 2017). Precipitation acts as a direct link between the global energy and water cycles, mediating energy exchange (Trenberth et al., 2009) and is the main factor controlling many natural hazards, such as droughts and floods (Vicente-Serrano et al., 2010). Accurate and reliable precipitation information is of great importance in water resources planning, hydrological simulation, environmental and ecological management, and irrigation management.

Currently, there are three main approaches to obtain precipitation observation and estimates: in situ measurements (mostly rain gauges), surface-based remote sensing (including weather radars) and airborne

or spaceborne remote sensing (Xie et al., 2022). Rain gauges are the most common and reliable way to directly measure precipitation on a point scale (Lanza and Stagi, 2008). However, the density and distribution of rain gauge networks vary significantly around the world, with few or no gauges in marine and mountainous areas making it difficult to provide spatially continuous precipitation information. Due to their ability to provide global and continuous coverage, satellite precipitation products are well-suited for overcoming the limitations of the spatially discontinuous distribution of rain gauges and have shown great potential in a wide range of applications (Serrat-Capdevila et al., 2014; Sun et al., 2018). However, these products are derived from indirect precipitation estimates, so errors in the products exist and vary across different regions, seasons, event types, and precipitation phases.

Since the launch of the first meteorological satellite in 1960, numerous satellites have been developed with sensors capable of providing

* Corresponding author.

E-mail address: epeino@meteo.ub.edu (E. Peinó).

<https://doi.org/10.1016/j.atmosres.2025.108311>

Received 9 December 2024; Received in revised form 19 June 2025; Accepted 20 June 2025

Available online 12 July 2025

0169-8095/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

observations from which surface precipitation can be derived (Kidd and Levizzani, 2019). However, it was not until the launch of the Tropical Rainfall Measuring Mission (TRMM) in December 1997 and, subsequently, the Global Precipitation Measurement (GPM) mission in February 2014, that missions dedicated specifically to satellite precipitation estimation were initiated. NASA's Integrated Multi-satellite Retrievals for GPM (IMERG) algorithm (Huffman et al., 2020) is one of the most reliable products for quantitative measurement of precipitation from satellites (Pradhan et al., 2022). The evaluation of IMERG performance has reflected a growing interest in its use, with studies stratifying the results according to time scales, topographic features, climatic conditions and precipitation accumulations, allowing a more specific description of IMERG behavior in different conditions and its most appropriate application (Lei et al., 2021; Moazami et al., 2022; Rojas et al., 2021; Yang et al., 2020; Zhang et al., 2022).

The Satellite Application Facility SAF on Support to Operational Hydrology and Water Management, (H SAF, or Hydrology SAF) is one of the current eight SAFs managed by European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT). H SAF is the main European concerted effort on satellite precipitation estimates, focusing on precipitation products from available geosynchronous and low-Earth-orbiting satellites using both passive-microwave (PMW) and infrared (IR) data (Mugnai et al., 2013). Since its inception, H SAF has adopted a validation strategy, which is now reaching the critical mass necessary to validate operational and climate products. Specifically, H SAF includes a product validation program and a hydrological validation program, involving several countries and instrumental for evaluating and improving H SAF precipitation products, as well as for assessing their usefulness in fighting floods, landslides, avalanches, and evaluating water resources (Mugnai et al., 2013; Puca et al., 2014).

Since there is no satellite product that performs best under all conditions, some of these studies have focused on comparing the performance of various satellite products in specific areas to enhance their application in diverse geographic and climatological contexts (Petracca et al., 2018; Baez-Villanueva et al., 2018; Li et al., 2013; Yu et al., 2020). In the Mediterranean regions several studies have assessed the performance of satellite precipitation products (SPPs) under complex topographic and climatic conditions, highlighting their potential and limitations. In southern Europe, multiple validations of IMERG have been conducted, evaluating its reliability across different spatial and temporal scales. In Greece, Italy, and Spain, studies have examined the influence of topography, land-sea transitions, and precipitation thresholds on IMERG estimates, revealing variations in accuracy depending on altitude, seasonality, and rainfall intensity (Kazamias et al., 2022; Caracciolo et al., 2018; Chiaravallotti et al., 2018; Tapiador et al., 2012; Navarro et al., 2020). Additionally, IMERG Late has been used to assess precipitation during extreme events such as Storm 'Filomena' (Tapiador et al., 2021).

Beyond IMERG, broader evaluations have analyzed multiple SPPs. In Turkey, a comparative study identified MSWXv100 (Hafizi and Sorman, 2023) as the most reliable product at the monthly scale, although its performance varied across different climatic regions. Kidd et al. (2018) found that the GPM constellation and its GPROF retrieval algorithm tend to overestimate light precipitation while underestimating heavy rainfall, posing challenges for hydrological applications in the region. Moreover, the integration of microwave-derived soil moisture products into hydrological models has demonstrated significant improvements in discharge predictions during extreme flow events (Cenci et al., 2016). However, during high-impact weather events such as Medicane Daniel in Greece, satellite products struggled to accurately capture localized intense convective rainfall, underscoring the need for ongoing algorithm enhancements (Katsanos et al., 2024).

According to the available literature, few studies evaluate H SAF precipitation products outside of the countries participating in the consortium's validation program (Kumah, 2023) and directly compare them with other products such as IMERG. To fill this gap, this

study aims to evaluate the accuracy of three H SAF products: P-IN-SEVIRI-PMW (H61B), P-AC-SM2RAIN (H64) and P-IN-PMW (H68), at hourly and daily temporal aggregations, and intercompares their performance with IMERG Early and Late products (hereafter Early and Late respectively) in the western Mediterranean. Given the nature of this objective, note that no specific research hypothesis is evaluated here. The validation includes assessing the ability of these products to detect and estimate extreme precipitation events. In addition, an analysis of the microphysical properties of the clouds responsible for precipitation events is conducted. This provides optimal thresholds for these variables related to precipitation mechanisms, which can be used as complementary information in quantitative precipitation estimates (QPE) algorithms.

The structure of this study is as follows. Section 2 provides a description of the study area, datasets used (H SAF and IMERG products and rain gauge data), the comparison methodology adopted in this study and the evaluation metrics employed. Section 3 shows results and discussions considering the validation of all products based on hourly and daily rainfall data for different accumulation thresholds according to the rain gauges and the relationship of microphysical variables at the cloud tops to the occurrence of surface precipitation. Finally, a summary and conclusion are provided in Section 4.

2. Materials and methods

2.1. Study area

Catalonia, located in the northeast of the Iberian Peninsula, is characterized by a complex topography that poses challenges for precipitation estimation using remote sensors, as well as for the estimation of the precipitation field using rain gauge stations (Peinó et al., 2022; Navarro et al., 2020; Trapero et al., 2009). Covering approximately 32107 km², Catalonia exhibits a wide range of climates due to its latitudinal position and the influence of the Mediterranean Sea. To the north, it is bordered by the Pyrenees, with elevations exceeding 3000 meters above sea level, while the inland area is predominantly flat with few orographic contrasts resulting from the erosion of the Ebro River and its tributaries (Fig. 1a). Additionally, the coastal and pre-coastal mountain ranges, oriented from northeast to southwest towards the Mediterranean Sea, enhance the development of flash floods from both a hydrological perspective (small catchments prone to flash floods) and a meteorological perspective (orographic forcing of Mediterranean air masses) (Jansa et al., 2014; Llasat et al., 2016). Consequently, this region experiences intense thunderstorms and subsequent flash floods annually (Llasat et al., 2014).

2.2. Datasets

2.2.1. Satellite precipitation products

Table 1 summarizes the key characteristics of the five satellite precipitation products analyzed in this study. The columns provide details on the original temporal and spatial resolutions of each product, the spatial coverage they offer, the temporal aggregation used for validation (1-hour or 24-hour accumulations), and the latency associated with each product. This framework ensures a systematic comparison across all products under a unified validation approach.

IMERG algorithm (Huffman et al., 2015) includes precipitation retrievals from PMW radiometers in low Earth orbits and from geostationary IR sensors to generate a global precipitation product. The algorithm intercalibrates, merges, and interpolates these estimates to provide a seamless dataset. It runs three times for each spatiotemporal estimate (Early, Late, and Final) where later runs incorporate more data and refined processing (Tan et al., 2019). This study uses data from the Early and Late V06B Level 3 products, which do not include gauge-based calibration, as this process is only applied in the Final run in this version. These products provide better results in retrieving heavy

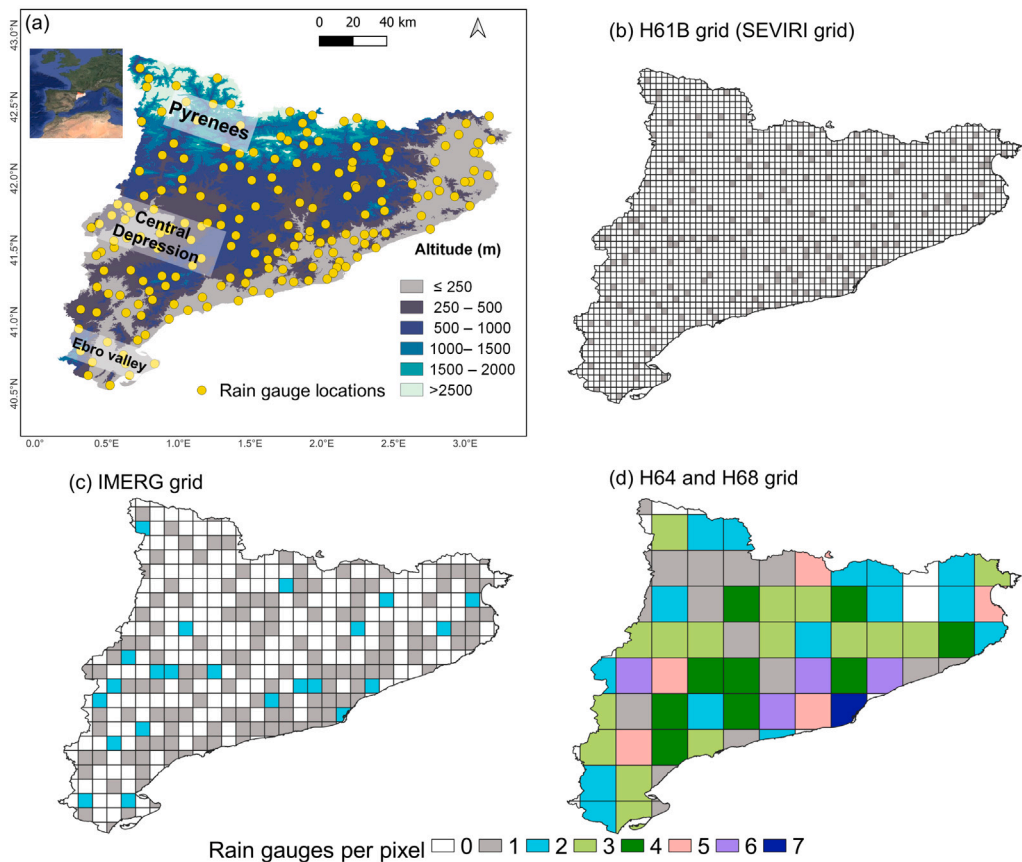


Fig. 1. (a) Digital elevation model of Catalonia and rain gauge network distribution (red symbols). (b) Number of rain gauges per pixel in H61B product considering the spatial resolution of each satellite grid cell in the Catalonia domain. Similarity panels (c) and (d) show the grid and rain gauges per grid cell for IMERG and for H64 and H68 products, respectively. Note that the H61B grid is not a regular grid, but has the geostationary observation geometry (SEVIRI).

Table 1
Summary of satellite precipitation product characteristics used. IMERG corresponds to version 06B. H SAF products (H61B, H64, H68) do not use explicit version numbers; instead, only the latest version that passes the internal review process is disseminated, replacing any previous versions through full retrospective reprocessing. The versions employed in this study were officially declared operational on June 9, 2022, and applied uniformly to the entire dataset analyzed. Further details on the validation process can be found in the respective Product Validation Reports (H SAF, 2022a,b,c).

Product	Temporal Resolution	Spatial Resolution	Coverage	Temporal aggregation	Latency
H61B	1h/24h	~ 3 km at the sub-satellite point	60° S - 75° N, 80° W - 80° E	1h/24h	~ 10 min
H68	30 min	0.25° x 0.25°	60° S - 75° N, 60° W - 60° E	1h	~ 4h
H64	24h	0.25° x 0.25°	60° S - 75° N, 60° W - 60° E	24h	~ 14h
IMERG (Early, Late)	30 min	0.1° x 0.1°	quasi-global coverage	1h/24h	Early ~ 4h, Late ~ 14h

precipitation compared to the IMERG Final product (Furl et al., 2018; Peinó et al., 2022, 2024). The output is a gridded product with a spatial resolution of $0.1^\circ \times 0.1^\circ$ and a temporal resolution of 30 min, with a latency of approximately 4 and 14 h, respectively, for Early and Late Runs.

The selection of H SAF products was based on their availability within the study period and their current operational status. Priority was given to products employing different retrieval techniques and, where possible, ensuring mutual independence, in order to provide a comprehensive assessment of satellite precipitation estimates across the 18 case studies analyzed. For this reason, the near-real-time product P-AC-SEVIRI-PMW (H61B) was selected, as it integrates instantaneous precipitation fields generated by the P-IN-SEVIRI-PMW (H60B) product. These fields are based on intercalibrated Level-2 passive microwave (PMW) instantaneous precipitation index estimates, combined with information from the $10.8 \mu\text{m}$ channel of the SEVIRI instrument (EUMETSAT, 2020). The integration of these data sources enhances the spatial and temporal continuity of precipitation estimates. The product provides hourly accumulated precipitation and 24-hour accumulated precipitation at 00, 06, 12, and 18 UTC. The final output has a spatial

resolution of ~ 3 km at the sub-satellite point, a temporal resolution of 1 h and 24 h, and a latency of approximately 10 min.

The H64 product (EUMETSAT, 2022) derives precipitation estimates by merging PMW-based retrievals from H-AUX-23 and H67 with precipitation inferred from soil moisture variations using the SM2RAIN algorithm (Brocca et al., 2014; Ciabatta et al., 2018; Koster et al., 2016; Massari et al., 2017). Over oceanic regions, the algorithm relies exclusively on PMW-derived precipitation estimates. The product is provided at a spatial resolution of $0.25^\circ \times 0.25^\circ$ with a temporal resolution of 24 h, and a latency of approximately 14 h.

The H68 Level 3 product (EUMETSAT, 2021) combines instantaneous precipitation estimates from multiple PMW radiometers, including SSMIS, AMSR-2, GMI, ATMS, and AMSU/MHS. The algorithm intercalibrates these estimates to ensure consistency across sensors and integrates them into a common reference framework. The final dataset consists of 30 min precipitation estimates on a $0.25^\circ \times 0.25^\circ$ spatial grid, with a latency of approximately 4 h.

2.2.2. Rain gauge data

The network of automatic surface weather stations of the Meteorological Service of Catalonia (SMC for its acronym in Catalan) was used for the validation of satellite precipitation products as in previous studies (Peinó et al., 2022, 2024). It includes 186 stations distributed across the region (Fig. 1), with an average minimum distance between station of 8.7 km and a density of 1.13 stations per 100 km², which is six times the minimum threshold recommended by the WMO for inland flat and undulating areas (WMO, 1994). Semi-hourly precipitation records with a resolution of 0.1 mm were utilized. A quality control scheme was applied to the SMC rain gauge records based on comparisons with nearby stations and correlation analysis (Llabrés-Brustenga et al., 2019, 2020).

2.2.3. NWC SAF product

To explore the relationship between precipitation events observed by rain gauges and satellite estimates and their associated cloud-top microphysical properties, data from the Cloud Microphysics (CMIC) product were analyzed. This product was compiled within the framework of the EUMETSAT Nowcasting and Very Short Range Forecasting SAF. CMIC was developed to support nowcasting applications, enabling the characterization of rapidly developing storms (Lahuerta García, 2021). This product has parallax correction applied to the satellite observations of cloud-top properties. In this study, the following five variables were analyzed: (1) cloud-top phase, considering only time intervals with liquid, ice, or mixed-phase presence; (2) cloud droplet effective radius (R_{eff}), defined as the weighted mean of the droplet size distribution; (3) cloud optical thickness (COT); (4) cloud liquid water path (LWP) and (5) cloud ice water path (IWP), with the last two quantifying the vertically integrated amounts of liquid and frozen water droplets, respectively. Among these variables, only the 'cloud-top phase' is available during night and twilight conditions.

2.3. Methods

2.3.1. Selection of case studies

Eighteen case studies between 2021 and 2023 were considered, based on full days during which at least one station recorded precipitation accumulations exceeding 2 mm in one hour and 10 mm in one day, following the selection criteria applied by Peinó et al. (2024) over the same region of study. Most precipitation was of non-frontal convective origin; frontal passages occurred on only 3 of the 18 days analyzed, according to UK Met Office sea level pressure analysis charts (UK Met Office, 2025). Additionally, on these selected days, several stations reported extreme precipitation values, often surpassing the intense precipitation threshold established by the Meteorological Service of Catalonia (20 mm in 30 min) —see Appendix A for more details—. Table 2 provides a summary of the case studies and the number of stations that exceeded selected precipitation thresholds on an hourly and daily scale. This approach allowed us to construct a database representing extreme precipitation episodes, facilitating the assessment of the remote sensing systems' capabilities in such events.

2.3.2. Pixel-based comparison method

Validation considered both hourly and daily temporal aggregations. For the IMERG and H68 products, aggregation was only performed when 100% of the required half-hourly intervals were available—two for hourly and 48 for daily values. In the case of H68, which relies solely on passive microwave (PMW) overpasses, the spatial and temporal coverage is intermittent. Although H68 can occasionally provide two consecutive half-hourly observations at a given location —making hourly comparisons possible— it never provides 48 consecutive half-hourly estimates at the same location. For this reason, daily-scale validation was not feasible for H68, and the sample size was significantly reduced compared to the other products.

A point-to-pixel (Gentilucci et al., 2022; Li et al., 2022, 2018; Xie et al., 2022) analysis was conducted in the study area to compare the time series of the selected rain gauge stations with the corresponding pixel values of the selected satellite products. The implicit assumption of this methodology is that the rain gauges are representative observations of the respective pixels of the products. Prior to the final choice of the method used here, we performed a sensitivity analysis by varying the method to spatially match the records of satellite products and those obtained through rain gauges. Although all are similar to the point-to-pixel method, three variants were evaluated: (1) In case there was more than one rain gauge in one satellite pixel, the independence of the precipitation records in each one was maintained for the comparison; (2) we compared the rain gauge information with the eight pixels neighboring the one containing it (Ebert, 2008) and selected for the comparison the one with the least bias and, (3) in case there was more than one rain gauge per pixel we took the arithmetic mean of the rain gauge records and compared it directly with the pixel containing them. These variants yielded results that were quite similar, so variant (2) was applied to H61B due to its high spatial resolution (Fig. 1) and variant (3) for the rest of the products.

Although the satellite products analyzed have different spatial resolutions, studies such as (Baez-Villanueva et al., 2018; Wang et al., 2019) suggest that there are no substantial improvements when applying a resampling method between them. In this case, we also believe that products such as H61B and even IMERG would be significantly more affected if this technique were applied to bring them to the resolution of H68 and H64.

The microphysical variables extracted from the CMIC product have a spatial resolution of ~ 3 km at sub-satellite point (SEVIRI grid) and a temporal resolution of 15 min. To account for spatial differences, the values at the closest grid point to each meteorological station were initially selected, and subsequently, the matching method explained earlier was applied for each product. Regarding temporal resolution, the four 15 min CMIC intervals corresponding to the hour of observation were aggregated into 1-hour resolution. The cloud-top phase variable was aggregated by taking the mode of the data, and in cases of ties, a new value was assigned, which was later treated as a mixed phase. For the remaining variables, the mean value was used.

Note that the point-to-pixel method, although it has advantages over other approaches (Yu et al., 2020), arises from the inherent challenges of comparing point-based precipitation measurements, such as those from rain gauges, with gridded satellite-derived products. The fundamental difference lies in spatial scale: rain gauges measure precipitation at a single point location, whereas satellite products provide area-averaged estimates over regions several orders of magnitude larger. This mismatch can lead to significant discrepancies, particularly given the high spatial variability of precipitation fields. These structural differences can influence the frequency of false alarms and impact the detection of extreme precipitation events at a local scale. Moreover, when multiple rain gauges are averaged within a satellite pixel, the representation of high rain rates may be attenuated, especially if the number of gauges is small. In regions like Catalonia, characterized by complex topography and high climatic variability, such discrepancies may introduce biases and affect the accuracy of comparisons. Therefore, it is essential to consider these limitations when interpreting the results, as understanding and mitigating these effects is key to conducting more accurate analyses and drawing meaningful conclusions in satellite precipitation validation studies.

2.4. Verification scores

Four continuous statistics were calculated to quantify the magnitude of differences between satellite estimates and rain gauges: modified Kling–Gupta Efficiency (KGE), Correlation Coefficient (CC), percent bias (Rbias), and Root Mean Square Error (RMSE). KGE describes overall accuracy as a multi-component metric integrating correlation

Table 2

Dates of selected cases and number of stations reaching or exceeding selected rainfall thresholds on an hourly and daily scale.

Dates dd/mm/yyyy	Hourly thresholds (mm)				Daily thresholds (mm)			
	0.5	2	15	30	1	10	50	100
09/05/2021	49	15	0	0	100	14	1	0
17/06/2021	81	48	13	3	141	93	10	0
31/07/2021	98	75	13	1	183	96	4	0
25/08/2021	25	14	8	2	88	34	3	0
01/09/2021	69	31	5	2	114	39	7	2
09/09/2021	31	15	3	2	97	28	0	0
18/09/2021	55	25	3	1	125	57	0	0
23/11/2021	121	88	6	3	184	180	37	9
12/03/2022	181	109	2	0	187	175	24	3
20/03/2022	66	27	2	0	132	52	3	2
17/08/2022	32	23	9	4	94	35	4	0
24/08/2022	45	19	3	0	72	18	1	0
25/08/2022	80	45	7	1	123	62	2	0
23/09/2022	99	76	11	2	167	111	8	2
29/06/2023	113	79	16	3	170	109	6	1
27/07/2023	89	71	16	0	139	73	0	0
03/09/2023	25	21	8	5	44	25	11	7
15/09/2023	139	112	10	5	185	161	16	1

Table 3

Summary of the continuous verification metrics used to evaluate satellite products.

Name	Formula	Unit	Perfect score
Kling–Gupta Efficiency (KGE)	$KGE = 1 - \sqrt{(r-1)^2 + \left(\frac{\sigma_s}{\sigma_o} - 1\right)^2} + (\mu_s - \mu_o)^2$	–	1
Correlation Coefficient (CC)	$CC = \frac{\sum_{i=1}^n (S_i - \bar{S})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^n (S_i - \bar{S})^2} \sqrt{\sum_{i=1}^n (O_i - \bar{O})^2}}$	–	1
Relative Bias (Rbias)	$Rbias = \left(\frac{\sum_{i=1}^n (S_i - O_i)}{\sum_{i=1}^n O_i} \right) \times 100\%$	%	0
Root-Mean Square Error (RMSE)	$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (S_i - O_i)^2}$	mm	0

S_i - value of satellite precipitation estimates for the i th event, O_i - value of rain-gauge observation for the i th event, $\bar{S}$ and $\bar{O}$ - corresponding mean values, n - number of observed records, $\beta = \bar{S}/\bar{O}$, $\gamma = (\bar{S}/\sigma_{S_i})/(\bar{O}/\sigma_{O_i})$, σ_{S_i} and σ_{O_i} - corresponding standard deviations.

Table 4

Contingency table for comparing rainfall observed by rain gauges and estimated by satellite products for a given threshold (th).

Estimated rainfall	Observed rainfall	
	Gauge rainfall $\geq$ th	Gauge rainfall $<$ th
Satellite rainfall $\geq$ th	Hits (H)	False alarms (F)
Satellite rainfall $<$ th	Misses (M)	Correct Negatives (CN)

coefficient (CC), bias ratio (β), and variability ratio (γ), proposed by Gupta et al. (2009) and revised by Kling et al. (2012). The range of KGE is $(-\infty, 1]$, with a perfect value of 1. Higher values of KGE indicate better accuracy. Rbias quantifies the deviation of satellite products from gauge data, ranging from $-\infty$ to ∞ , with 0 representing a perfect value. The sign of Rbias (positive or negative) indicates the direction of deviation (overestimation or underestimation) and absolute value reflects the magnitude of deviation. RMSE provides a measure of the error that gives more weight to large errors without showing the direction of deviation between satellite products and rain gauges. Table 3 shows details of the calculation of these verification scores.

Additionally, two categorical indices based on contingency tables (Table 4) were used. The Hansen and Kuipers (HK) score (Table 5) is the difference between the Probability of Detection (POD) and the Probability of False Detection (POFD) (Jolliffe and Stephenson, 2012; Trappero et al., 2013). The HK statistic measures the forecasting system's ability to distinguish observed “yes” and “no” cases. The second score, Critical Success Index (CSI) quantifies the proportion of correct predictions relative to the total number of event and non-event predictions, focusing on the model's ability to detect events (Ebert, 2008).

For statistical analysis, only precipitation records determined by the following thresholds were considered: 0.5 mm one hour and 1 mm in one day. Additionally, in the second part of the study, results were stratified based on predefined accumulation thresholds for rain gauge data. At the hourly scale, the selected thresholds were 2, 10, 15, and 30 mm or more in one hour. Similarly, at the daily scale, the thresholds were 10, 30, 50, and 100 mm or more in one day. The maximum

values of these thresholds (30 mm/h and 100 mm/24 h) are taken from heavy rainfall warning thresholds set by the Spanish Meteorological Service at hourly scale (AEMET, 2023) and the Meteorological Service of Catalonia at daily scale (SMC, 2025), respectively. The rest of the values considered are divisors of those thresholds.

3. Results and discussion

3.1. Performance of satellite products at hourly and daily scales

This section presents the evaluation of precipitation retrievals from all satellite products using rain gauge data as a reference. Fig. 2 shows the probability distributions of precipitation occurrence from rain gauge observations and estimates from each satellite product at both hourly and daily resolutions. The number of data points considered in each analysis is indicated in the title of each panel. As shown, both the hourly-scale H68 and the daily-scale H64 represent significantly smaller samples compared to the other products. For H68, this is attributed to its reliance on recoveries determined by orbital geometry, while for H64, it stems from the application of the concordance method to a product with substantially lower spatial resolution.

While the distributions are generally similar in shape across all cases, precipitation accumulations are often overestimated, except for H61B at the hourly scale and H64 at the daily scale, where the estimated median value closely aligns with the observed value. Table 6 summarizes the continuous metrics (KGE, Rbias, and RMSE) comparing rain gauge data and satellite products at hourly and daily scales.

Table 5
List of categorical verification scores used to evaluate satellite products.

Name of metric	Formula	Perfect score
Probability of Detection (POD)	$POD = \frac{H}{H+M_F}$	1
Probability of False Detection (POFD)	$POFD = \frac{F}{F+CN}$	0
Hanssen and Kuipers (HK)	$HK = POD - POFD$	1
Critical Success Index (CSI)	$CSI = \frac{H}{H+M+F}$	1

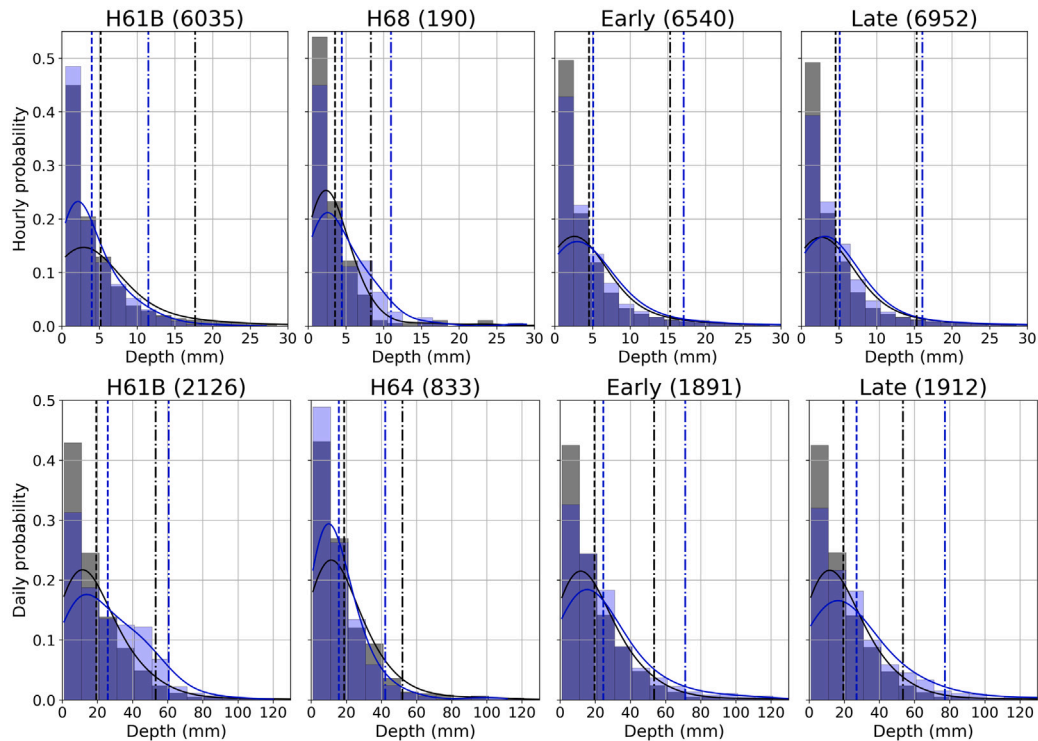


Fig. 2. Probability distribution of precipitation for each satellite product and the corresponding rain gauge records at the hourly scale (top panel) and at the daily scale (bottom panel). Solid lines correspond to kernel density estimates (KDE) computed using Gaussian kernels with a bandwidth of 0.6. Dashed lines represent the mean of the rain gauge (black) and satellite (blue) distributions. Dash-dotted lines indicate the 95th percentile of each distribution.

Table 6
Continuous and categorical verification scores applied at hourly and daily resolutions. Best and worst values for each score are indicated in green and red, respectively.

Hourly					
Product	KGE	Rbias	RMSE	HK	CSI
H61B	0.26	−23.47	6.67	0.48	0.37
H68	0.22	25.77	4.96	0.79	0.52
Early	0.18	12.36	8.01	0.57	0.41
Late	0.22	11.68	7.50	0.60	0.41
Daily					
Product	KGE	Rbias	RMSE	HK	CSI
H61B	0.21	33.91	24.77	0.27	0.72
H64	0.38	−15.81	19.54	0.27	0.72
Early	0.32	26.09	26.31	0.30	0.72
Late	0.33	38.60	26.74	0.36	0.75

Considering the thresholds for detecting a precipitation event (0.5 mm at the hourly scale and 1 mm at the daily scale), the CSI and HK indices obtained for each product are also included.

The variability in KGE, Rbias, and RMSE values among the products suggests that the accuracy and efficiency of satellite products can vary significantly depending on the product and the temporal scale examined. KGE values were relatively low for all products regardless of the temporal aggregation, indicating challenges in replicating the temporal variability and magnitude of precipitation measured by rain gauges. The highest value was observed for H64 at the daily scale, although it does not exceed 0.4. IMERG showed improved efficiency in retrievals with increased temporal aggregation. Several studies have noted this behavior (Peinó et al., 2022; Chen et al., 2018; Xu et al., 2019),

as longer temporal aggregations smooth out short-term precipitation variations, leading to better alignment with rain gauge observations.

H61B exhibits a significant negative bias (−23.47%) at the hourly scale, nearly equivalent in magnitude but opposite in sign to H68 (25.77%). IMERG products at this scale showed similar values and the best performance, although they overestimated the observations. RMSE values were fairly consistent across products. While H68 exhibited the best performance, this metric is more sensitive to larger errors. Consequently, the limited number of records with high precipitation accumulations for this product may have positively skewed the result.

At the daily scale, relative bias indicated substantial overestimation, especially for H61B and Late, which exceeded 30%, while H64 showed underestimation to a lesser extent and exhibited the best performance.

The products reflected substantial systematic errors in daily precipitation estimation algorithms (between 19 and 26 mm/day). These discrepancies are likely influenced by precipitation accumulations and limitations in spatial and temporal resolution used in comparisons between satellite products and rain gauges.

Overall, notable differences were observed between hourly and daily scales concerning the products' ability to detect precipitation events. CSI values increased at the daily scale compared to the hourly scale, while the products' ability to discriminate between events and non-events decreased. At the hourly scale, H68 demonstrated superior capability to distinguish between events and non-events ($HK = 0.79$) with moderately high accuracy in predicting correct events ($CSI = 0.52$). H61B showed slightly lower performance in these aspects. Conversely, at the daily scale, Late stood out with the best discrimination capability ($HK = 0.36$), though it did not exceed 0.5, while all products maintained high accuracy in event prediction ($CSI = 0.72$). This indicates that while temporal aggregation improves event detection accuracy, the ability to discriminate these events decreases due to a high false alarm rate affecting the index.

3.2. Spatial distribution of errors

Fig. 3 presents the validation results at the hourly scale for each pixel according to the spatial resolution of each satellite product in terms of KGE, based on all the case studies considered in the analysis. To better interpret the results, a scale was added where reddish colors represent poor product performance in these pixels, and green colors indicate better skill according to previous studies (EUMETSAT, 2022). H68 shows poor performance with a low KGE index in several regions (more than 70% of pixels rated unsatisfactory), particularly in the Pyrenees area. Compared to the other products, H61B shows the highest percentage of pixels classified as having acceptable or good performance in inland areas, including high mountain regions. Areas with unsatisfactory performance (68% of pixels) were primarily concentrated along the coast. IMERG products exhibit globally poor values, with only a few acceptable (4% and 8% of good pixels for Early and Late, respectively), indicating a limited ability to provide accurate and consistent estimates at this scale based on this index.

Fig. 4 illustrates the behavior of KGE at the daily scale, also considering all case studies. Although it is challenging to establish clear spatial patterns of better performance among these products, it is noteworthy that all of them show the best results in the southwest of the region. The H64 product stands out as the most suitable, with 31% of pixels showing KGE values considered "acceptable" or "very good". The main limitations of this product are evident in the Pyrenees and the northeastern part of the region. H61B exhibits substantially lower KGE values across almost the entire domain, except near the Ebro Valley, where IMERG products also show the best results. Coastal areas north of this basin are notably affected by the largest errors. It is important to mention that many case studies (see Appendix B) coincidentally recorded the highest precipitation accumulations in this region, suggesting that this behavior might be linked to the accumulation of precipitation during events.

The comparison between hourly and daily results shows that products with lower spatial resolution, such as H64, demonstrate the highest accuracy in terms of KGE. This may be primarily due to the limitations of the point-to-pixel method, which can struggle with high-resolution data due to the mismatch between the spatial resolution of satellite products and the spatial scale of ground-based measurements. IMERG products consistently perform poorly at both scales, indicating substantial limitations in their retrievals; however, the Late Run generally outperforms the Early Run, especially at the daily scale. In contrast, H61B demonstrates better performance at the hourly scale, likely due to its high spatial resolution, which makes it more sensitive to the high variability of precipitation. Additionally, its significantly lower

latency compared to the other evaluated products may further enhance its performance.

Notably, the largest errors in satellite precipitation estimates occur over coastal and mountainous regions, which aligns with previous studies on SPPs. In complex terrain, precipitation retrieval is particularly challenging due to the influence of orographic enhancement, rapid spatial variability of precipitation, and the limitations of PMW and IR retrieval techniques. Studies have shown that satellite-based PMW retrievals tend to underestimate precipitation associated with shallow orographic systems, as these systems often lack strong ice scattering signatures, which are the primary basis for PMW-based rainfall estimation (Adhikari and Behrangi, 2022). Similarly, IR-based algorithms struggle to detect warm orographic rainfall, leading to significant underestimations in mountainous regions (Bitew and Gebremichael, 2010). Over coastal areas, the transition between ocean and land surfaces introduces additional uncertainties, as satellite sensors must differentiate radiometric signals from contrasting surface emissivities, a well-known challenge for PMW retrievals (Derin et al., 2022; Turk et al., 2021). This issue is particularly relevant in Catalonia, where the proximity of mountainous regions to the Mediterranean coastline further complicates satellite precipitation retrievals. Similar biases have been reported in coastal mountain regions of the western United States, where SPPs struggle to capture rapid changes in precipitation due to land-ocean-atmosphere interactions (Chen et al., 2022).

However, recent advancements in satellite retrieval algorithms aim to mitigate these biases. For instance, IMERG Version 07 incorporates an upgraded GPROF retrieval algorithm, which introduces new mountain-specific surface classes and accounts for different orographic enhancement conditions. These modifications are expected to improve precipitation estimates over mountainous terrain; however, their effectiveness remains to be validated (Portier et al., 2025). Previous studies have demonstrated that incorporating terrain features into machine learning-based bias correction models significantly improves satellite precipitation estimates over complex terrain (Chen et al., 2022). While some efforts have been made to evaluate IMERG in coastal regions (Sui et al., 2020), similar validation efforts for H SAF products are still lacking. Given that both coastal and orographic precipitation processes play a crucial role in the precipitation climatology of Catalonia, future studies should further investigate how recent algorithmic improvements impact the performance of IMERG and H SAF products in these challenging environments.

3.3. Evaluation according to different hourly and daily rainfall amounts

Fig. 5 shows the distribution of hourly (top panel) and daily (bottom panel) precipitation accumulations for each satellite product and the corresponding rain gauge records used for validation. The four sections of the figure are identified based on the different precipitation accumulation thresholds (th) reached or exceeded in the reference rain gauges: 2, 10, 15, and 30 mm (hourly scale) and 10, 30, 50, and 100 mm (daily scale).

According to the results found here, Fig. 5 illustrates clearly that satellite products systematically underestimate precipitation accumulations, both hourly and daily, compared to rain gauge data, with this discrepancy becoming more pronounced as precipitation accumulations increase (see Appendix C). Notably, for the highest accumulations, both H61B and H64 exhibit minimal detection within these ranges, while H68 lacks sufficient data beyond these accumulations to establish a reliable distribution. IMERG products perform similarly to each other and are the only ones capable of recording precipitation accumulations comparable to those recorded by rain gauges, although this does not imply perfect spatial and temporal alignment. The greater variability in rain gauge measurements at all accumulation thresholds indicates that satellite products do not adequately capture this variability.

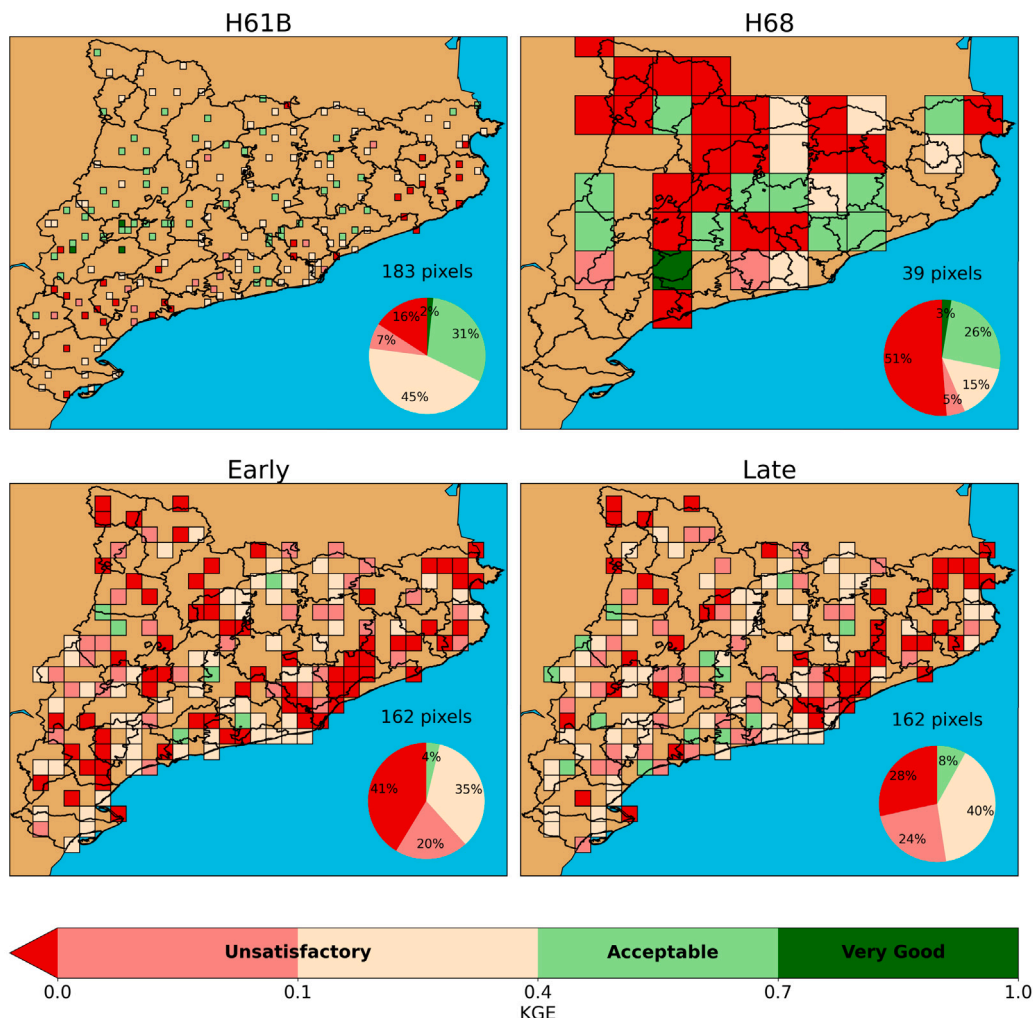


Fig. 3. Distribution of KGE values at hourly scale in each pixel of H61B, H68, Early and Late products. The pie charts illustrate the distribution of pixels classified based on the KGE, using the total number of evaluated pixels for each product as a reference. The total number of pixels assessed for each product is displayed in the corresponding panel of the figure.

Table 7 summarizes the key continuous statistics calculated for each satellite product at both hourly and daily temporal resolutions for the four predefined accumulation thresholds. Generally, all products show a decrease in KGE, and an increase in negative bias and RMSE as the precipitation threshold increases, though values are slightly better at the daily scale. At the hourly scale, H61 consistently achieves the highest KGE values, regardless of temporal aggregation, while Early does not excel in any performance metric. H68 outperforms the other products in terms of Rbias and RMSE, whereas Late shows less underestimation in daily estimates for events with medium to high accumulations. However, Late is distributed with a 14-hour latency due to the need for waiting for the next PMW overpass to allow back-interpolation in time. While this study primarily evaluates the accuracy of satellite precipitation products rather than their real-time applicability, it is important to acknowledge the implications of this delay. Despite IMERG Late's superior performance in detecting extreme rainfall, its latency limits its usefulness for time-sensitive applications such as early warning systems. Moreover, our results indicate that the improvements in IMERG Late over other products are only marginal, and in some cases, non-existent. This suggests that the underestimation of heavy rainfall events is not solely due to the availability of additional satellite data but rather an intrinsic limitation of the retrieval algorithm itself. Consequently, for near-real-time applications, H61B remains the most practical choice despite its known biases.

According to the categorical statistics shown in Fig. 6, all satellite products significantly decrease their ability to detect correct precipitation events at thresholds above 2 mm in one hour and 10 mm in one day. Although there are no dramatically marked differences, indicators improve with increased temporal aggregation, which is particularly notable in the IMERG products. Similar behavior has been reported in other studies (Peinó et al., 2022), which have also shown significant improvements when dealing with precipitation values over longer time scales, such as monthly, seasonal, and annual periods.

SPPs typically exhibit decreasing accuracy as rainfall accumulation increases, particularly for extreme precipitation events. This decline in performance is associated with multiple factors, including the physical characteristics of heavy rainfall and the limitations of satellite retrieval algorithms. Extreme precipitation often occurs in deep convective systems with large ice particles and strong vertical development, which introduce retrieval challenges for both PMW and IR sensors. PMW-based retrievals may suffer from heavy attenuation, particularly in high rainfall rates, leading to underestimations (Chen et al., 2022). Additionally, the IR-based algorithms, in the absence of PMW data, can misclassify cold cloud tops as precipitation, even in regions where no rainfall is occurring (Cui et al., 2020). Further uncertainties arise from the priori databases used in precipitation retrievals, which typically contain fewer extreme precipitation cases, leading to biases in satellite estimates. Spatial mismatches, temporal sampling issues, and errors in

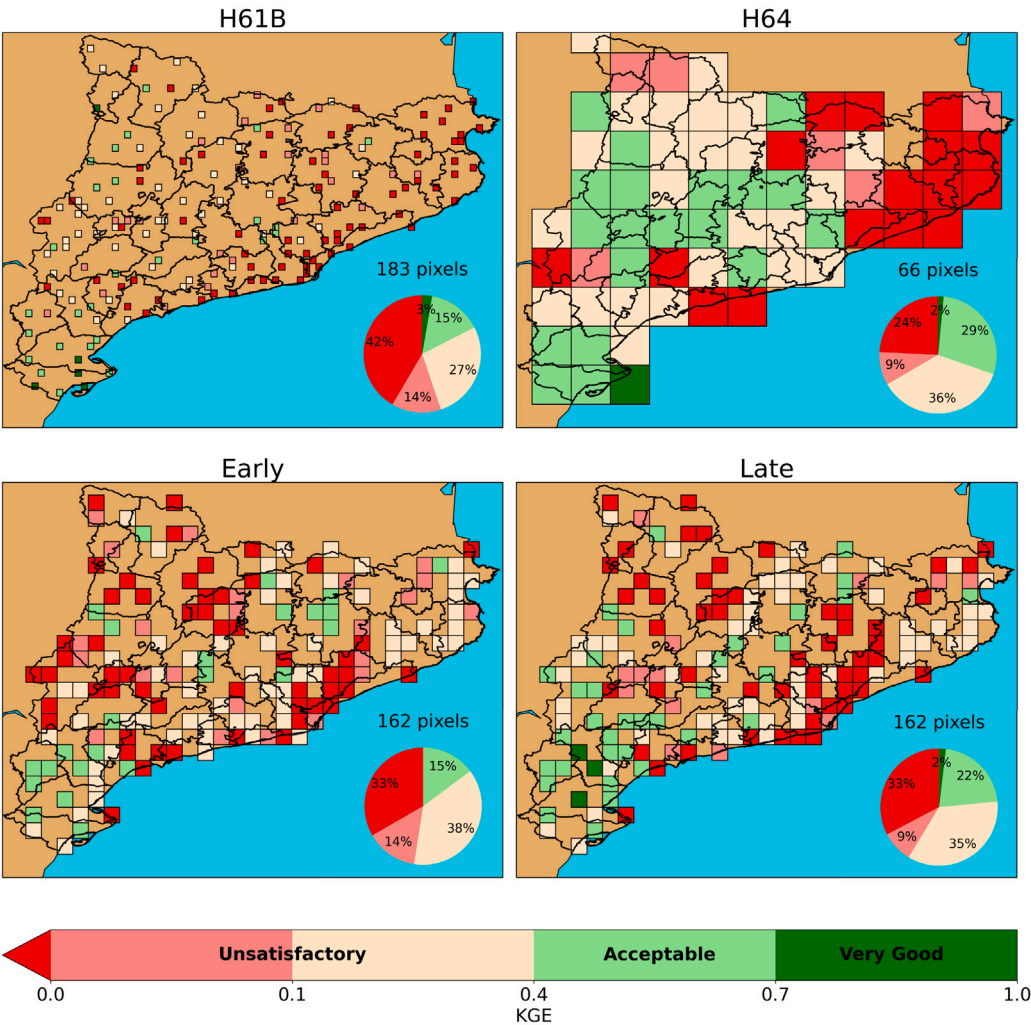


Fig. 4. Same as Fig. 3 but at daily scale.

Table 7
Continuous and categorical scores for different accumulation thresholds (th). Best and worst values for each score are indicated in green and red, respectively.

Hourly				Daily			
Product	KGE	Rbias	RMSE	Product	KGE	Rbias	RMSE
th ≥ 2 mm				th ≥ 10 mm			
H61B	0.23	-38.17	8.01	H61B	0.32	1.41	24.82
H68	0.22	2.41	5.67	H64	0.30	-33.17	22.23
Early	0.13	-11.19	9.36	Early	0.29	2.34	29.28
Late	0.18	-10.11	8.81	Late	0.37	15.24	29.11
th ≥ 10 mm				th ≥ 30 mm			
H61B	-0.07	-61.90	16.22	H61B	0.31	-31.70	31.14
H68	-0.75	-45.99	13.01	H64	0.09	-50.64	33.43
Early	-0.58	-54.55	17.74	Early	0.03	-19.53	39.74
Late	-0.40	-51.93	17.38	Late	0.22	-6.29	36.98
th ≥ 15 mm				th ≥ 50 mm			
H61B	-0.25	-67.16	21.08	H61B	0.19	-43.40	47.46
H68	-1.01	-59.19	16.11	H64	-0.39	-55.65	51.66
Early	-0.90	-64.70	22.19	Early	-0.27	-29.33	59.96
Late	-0.74	-63.73	22.12	Late	0.03	-16.28	52.73
th ≥ 30 mm				th ≥ 100 mm			
H61B	-0.89	-79.93	37.22	H61B	-0.14	-51.72	81.33
H68		-87.54	27.99	H64	-0.91	-54.56	73.54
Early	-1.88	-81.83	36.76	Early	-1.01	-59.99	96.88
Late	-1.41	-80.23	37.17	Late	-0.72	-45.96	84.20

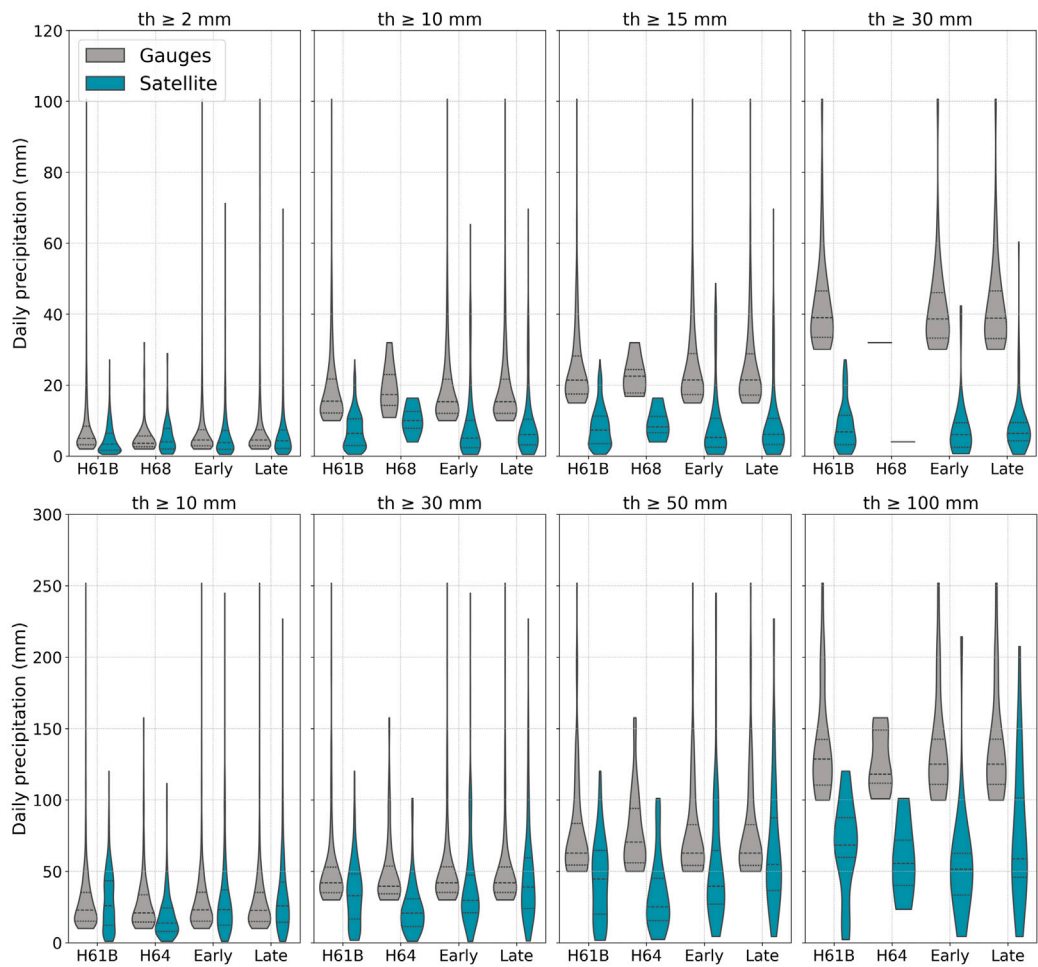


Fig. 5. Distribution of rainfall observations (gray) and satellite products (green) at hourly (top row) and daily (bottom row) scales, corresponding to cases where rain gauge measurements met or exceeded the precipitation accumulation (mm) thresholds (th) indicated. The violin plots represent the distribution of the data, showing the probability density at different values, with the width of each plot indicating the concentration of values. The lines within the plots represent the median and the first and third quartiles (25th and 75th percentiles) of the sample.

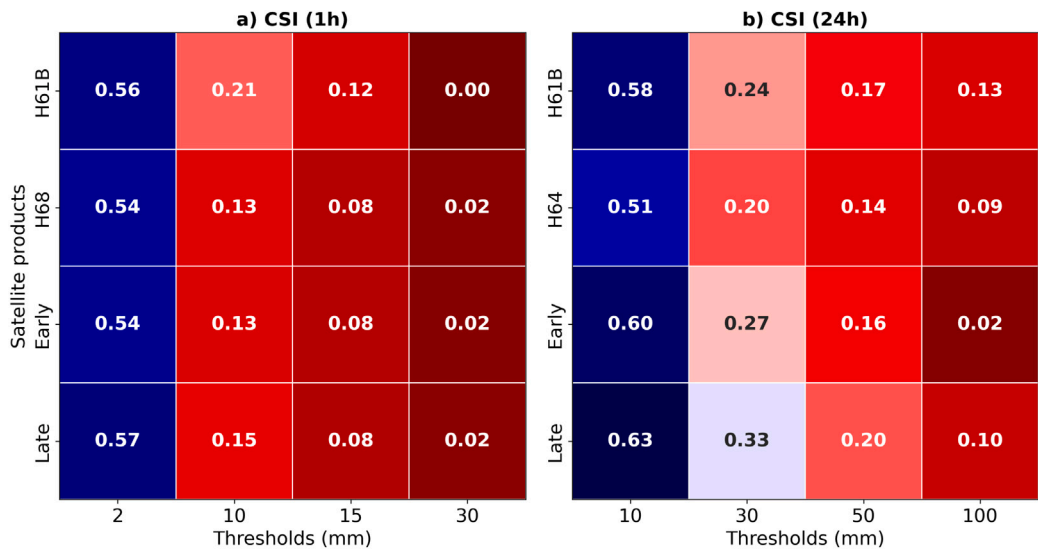


Fig. 6. Critical Success Index score for each satellite product on an hourly scale (left panel) and on a daily scale (right panel) considering different rainfall accumulation thresholds. Darker bluish colors (darker reddish) indicate good (bad) performance of the satellite estimates.

gauge-based reference datasets further complicate the validation of extreme precipitation retrievals (Tan et al., 2016; Kirstetter et al., 2020). Studies have shown that IMERG products tend to have reduced performance at finer temporal resolutions, partly due to the interpolation processes and the limited revisit time of the Global Precipitation Measurement (GPM) constellation (Chen et al., 2018; Xu et al., 2019; Gaona et al., 2016). Orographic influences and mesoscale factors further contribute to the complex spatial variability of precipitation, increasing the challenge of accurately capturing extreme events (Tapiador et al., 2020).

Recent advancements in the IMERG V07 algorithm have sought to improve the detection and quantification of extreme precipitation events. One major modification is the removal of internal precipitation rate caps, which previously limited estimates to 50 mm/h in the IR module and 120 mm/h in the PMW module (Portier et al., 2025). In V07, the cap has been increased to 200 mm/h, allowing for a more realistic representation of extreme rainfall when retrievals support such values. Additionally, the implementation of the SHARPEN scheme (Tan et al., 2021) addresses biases introduced by the Kalman filter, which previously smoothed extreme values and artificially expanded precipitation areas. This refinement enables V07 to better preserve the statistical distribution of extreme precipitation rates. Furthermore, improvements in the GPROF V07 retrieval algorithm have enhanced precipitation estimation over challenging surface types, including frozen terrains, mountainous regions, and coastal zones. While these updates represent significant progress, their effectiveness in real-world scenarios requires further validation, particularly in complex environments such as those analyzed in this study. Future research should focus on assessing the impact of these algorithmic advancements on the detection of extreme precipitation, as well as their implications for hydrological applications.

3.4. Relationship with cloud-top microphysical properties

This section analyzes the characteristics of cloud-top microphysical variables associated with precipitation observed or estimated at the surface, using both rain gauge data and satellite products on an hourly time scale. The CMIC product provides detailed information on the microphysics of cloud tops for all pixels identified as cloudy in the satellite scene. It is important to note that while the cloud phase is available for any time of the day, the other variables are only calculated under daytime conditions, reducing the size of the initial sample. Since CMIC only contains information about the upper part of the clouds, this analysis can provide additional and independent insights into the behavior of these variables when precipitation is observed or estimated.

Fig. 7 shows the distribution of cloud-top microphysical variables, including R_{eff} , COT, IWP, and LWP, under two thresholds of observed or estimated surface precipitation at hourly scale. The left column displays cases with precipitation equal to or greater than 2 mm, while the right column corresponds to values equal to or greater than 10 mm. These distributions consider data from rain gauges as well as the four satellite products analyzed on an hourly scale: H61B, H68, Early, and Late.

The R_{eff} associated with precipitation cases exhibits very similar distributions for both satellite estimates and rain gauge observations. The medians range around 0.02–0.03 mm with narrow inter-quartile ranges, indicating low variability in this variable overall. However, in those cases where 10 mm or more accumulated, a slight trend towards higher R_{eff} values is observed. Previous studies have identified a threshold of 0.013–0.014 mm as an indicator of precipitation onset (Freud and Rosenfeld, 2012; Rosenfeld and Gutman, 1994; Braga et al., 2017), associated with an increase in both droplet sweep volume and collision efficiency. This behavior is reflected in a significant increase in precipitation probability when R_{eff} reaches 0.014 mm, although light precipitation has been documented for values between 0.01 and 0.012 mm (Braga et al., 2021).

The COT values vary widely, ranging from less than 50 to more than 200. However, the medians of rain gauge observations and satellite estimates are similar for precipitation thresholds exceeding 2 mm. At higher hourly accumulations, above 10 mm, more pronounced differences emerge between the distributions associated with precipitation. Higher COT values observed under conditions of greater precipitation accumulations may be linked to denser and optically thicker clouds, which could contribute to increased precipitation accumulations. This aligns with previous studies that indicate that decreasing cloud-top pressure and increasing COT are often associated with higher rainfall rates due to deeper and more vigorous cloud development (Tapiador et al., 2019). This behavior is evident in rain gauge observations, where an increase in the median COT values is observed, surpassing those associated with satellite-estimated rainfall cases.

The H68 and IMERG products, while showing a slight increase relative to the 2 mm threshold, present lower medians and narrower interquartile ranges compared to rain gauges, suggesting a limited relationship between their retrieval algorithms and variations in COT. On the other hand, the H61B product stands out as the only one with a median higher than the rain gauge observations, showing higher values and, therefore, a stronger relationship between denser clouds and higher precipitation accumulation estimates.

The IWP shows high values across all distributions, with a notable increase as precipitation accumulations rises, except in the case of the H68 product. This may be attributed to the smaller dataset available when applying a threshold above 10 mm. Most of the precipitation cases analyzed in this study are associated with clouds containing ice at their tops, while cases with a liquid phase are scarce. This is reflected in the low values and reduced variability observed in the IWP distribution graphs, with limited or no representation in some of the satellite products evaluated.

Additionally, previous studies have indicated that the correlation between total cloud water content and surface precipitation rates is weaker in arid regions (You and Liu, 2012). This factor may be relevant in interpreting the low sensitivity of certain satellite products to IWP in scenarios with high precipitation accumulations.

While PMW retrieval algorithms indirectly account for cloud properties through brightness temperature measurements at different frequencies, the direct integration of COT and IWP into satellite rainfall estimation remains an open challenge. Traditional retrieval methods often struggle to establish non-linear relationships between cloud properties and precipitation rates, potentially overlooking key meteorological parameters. Recent advancements in machine learning and deep learning techniques, such as Random Forest (RF), XGBoost, Support Vector Machine (SVM), and Convolutional Neural Networks (CNN), offer promising avenues for addressing these limitations (Wang et al., 2020; Li et al., 2021; Yang et al., 2022). These methods can efficiently handle large datasets, capture complex relationships, and improve the accuracy of precipitation retrievals without requiring priori assumptions. This study identifies specific value ranges of cloud top properties and their relationship to satellite-based precipitation estimates. Future efforts should focus on integrating these variables effectively into retrieval algorithms, leverage machine learning approaches to improve precipitation estimation, and explore transfer learning techniques to optimize variable selection in next versions of H SAF and IMERG products.

3.4.1. Thresholds of microphysical variables for Hits, False Alarms and Misses

Fig. 8 illustrates the behavior of cloud-top microphysical variables across three categories: hits, false alarms, and misses, comparing satellite precipitation estimates to rain gauge observations, using the same accumulation thresholds as previously (2 and 10 mm). Examining these categories enables identifying the optimal ranges of microphysical variables where satellite products accurately capture observed precipitation, as well as those where false alarms occur or precipitation events exceeding a specific threshold are missed.

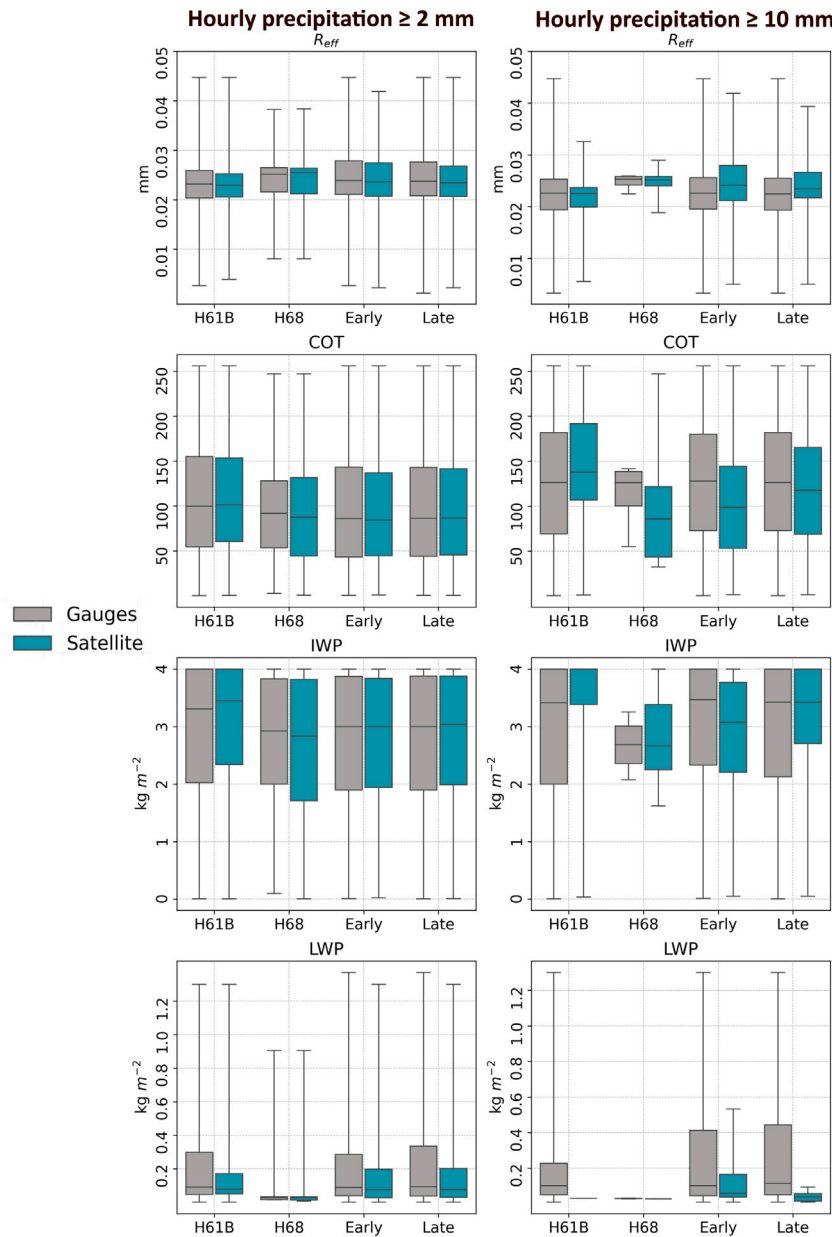


Fig. 7. Boxplots of microphysical variables at the cloud top for satellite rainfall product when rain gauges observed and satellite products estimated precipitation accumulated reaching or exceeding (left column) 2 mm and (right column) 10 mm in one hour. Note that the distributions for each product are associated with a different rain gauge data distribution, as the samples are different in each case.

Consistently with previous results, the microphysical variables that exhibit distinctive patterns under different precipitation accumulations thresholds are COT and IWP. For thresholds exceeding 2 mm, hits are associated with COT values greater than 100 and IWP values exceeding 3 kg/m². As precipitation accumulations increases, these thresholds reach COT values near 150 and IWP values close to 4 kg/m², particularly in the H61B and IMERG products. However, these variables values are notably lower in cases of false alarms and even lower when the products fail to detect precipitation, especially for accumulated above 2 mm.

The H61B and Late products show similar distributions, whereas Early exhibits limited variability in the medians associated with false alarms and misses for accumulated above 2 mm. For H68, false alarms are linked to COT values below 50, the lowest threshold recorded for this category. For accumulated exceeding 10 mm, establishing optimal thresholds to characterize each category becomes more challenging. Nevertheless, IMERG products display false alarms associated with low

COT values, sometimes below 100, and relatively low IWP values compared to misses and hits, consistent with findings by Maranan et al. (2020).

According to Dai et al. (2013), deep convective clouds are often capped by ice-phase anvil clouds. Under such conditions, raindrops may undergo displacement during descent due to strong horizontal wind shear, leading to a mismatch between surface observations and the characteristics of upper-level clouds.

3.4.2. Precipitation estimation errors associated with cloud top phase

To quantify differences based on the cloud-top phase, Fig. 9 presents the relative bias (Rbias) and root-mean-square error (RMSE) for each product as a function of the cloud phase associated with accumulated precipitation exceeding 2 mm and 10 mm. The results highlight significant limitations of the products for higher hourly accumulations, with underestimations exceeding 40% and errors surpassing 10 mm.

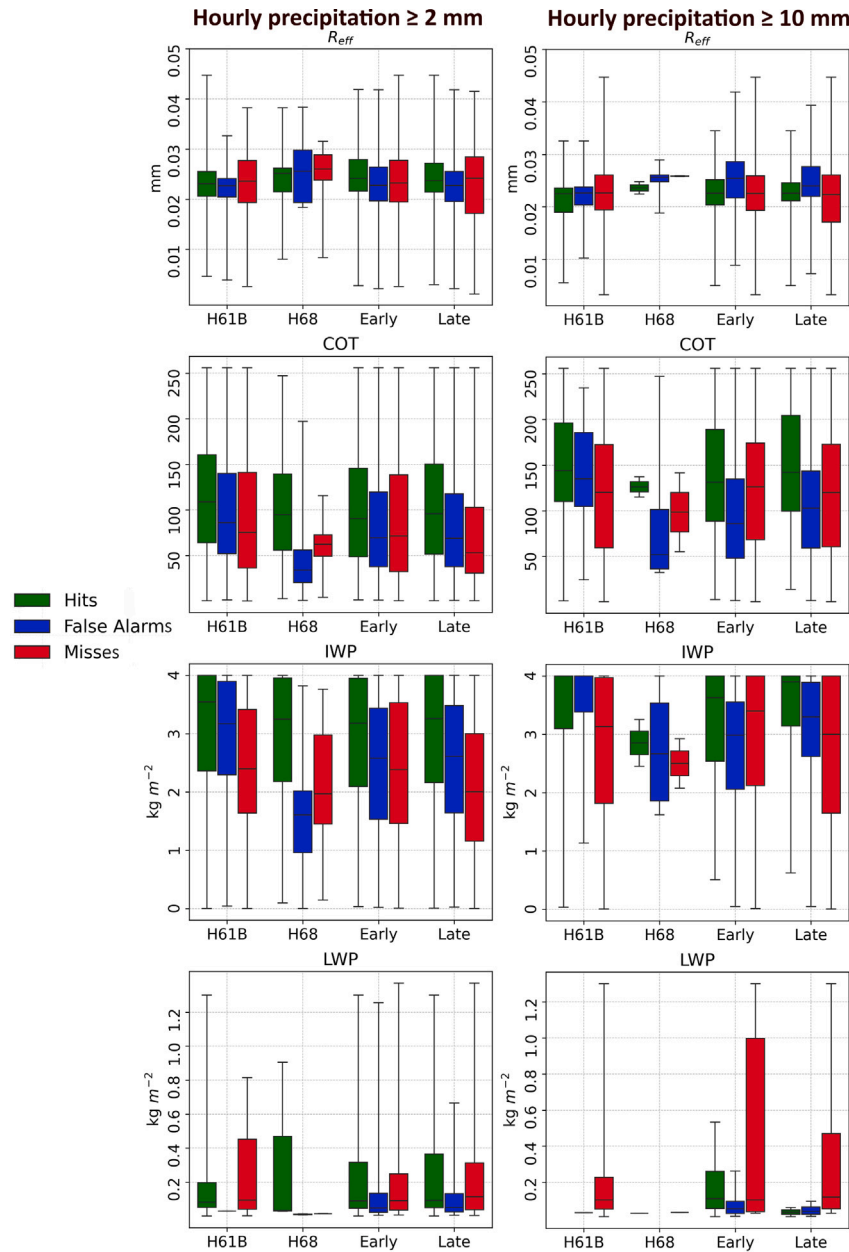


Fig. 8. Boxplots of microphysical variables during precipitation events classified as Hits, False Alarms, and Misses using thresholds of 2 mm (left column) and 10 mm in one hour (right column).

For thresholds above 2 mm, the best results are observed when the cloud-top phase is ice, particularly in the H68 and IMERG products. Conversely, under liquid or mixed phases, the estimates exhibit underestimations close to or above 40% and errors exceeding 7.5 mm. As noted by Maranan et al. (2020), the IMERG products and their sources are very sensitive to low LWP values, explaining the larger errors in warm clouds. In contrast, these products perform significantly better for glaciated clouds.

These findings align with those reported by Peinó et al. (2024), who emphasized that IMERG sources, particularly passive microwave (PMW) data, are sensitive to the cloud phase and microphysical characteristics. The poorest performance is typically associated with warm and mixed-phase clouds, underscoring IMERG's dependence on the formation of ice particles in convective clouds. PMW sensors are generally more effective at detecting precipitation from deep convective systems associated with cold clouds over land, as they leverage ice scattering signatures at high frequencies (e.g., 89 GHz and above) to

infer precipitation rates. This results in better detection of moderate to heavy rainfall but can lead to underestimations in warm rain scenarios, where ice-phase hydrometeors are absent. In contrast, IR-based retrievals infer rainfall indirectly from cloud-top temperatures, which can be problematic for warm rain processes (e.g., shallow convection or stratiform precipitation), where cloud tops remain relatively warm despite precipitation is present below cloud base. The poorest performance of satellite products is often associated with warm and mixed-phase clouds, reflecting their reliance on ice particle formation within convective clouds. Although our study focuses on a semiarid climate in the Western Mediterranean, our findings align with those of Maranan et al. (2020) in a forested region of West Africa, suggesting that observed biases, particularly for intense rainfall, stem primarily from limitations inherent to satellite precipitation algorithms rather than regional precipitation dynamics.

The results underscore the close relationship between clouds and precipitation but also reveal notable variability in precipitation within

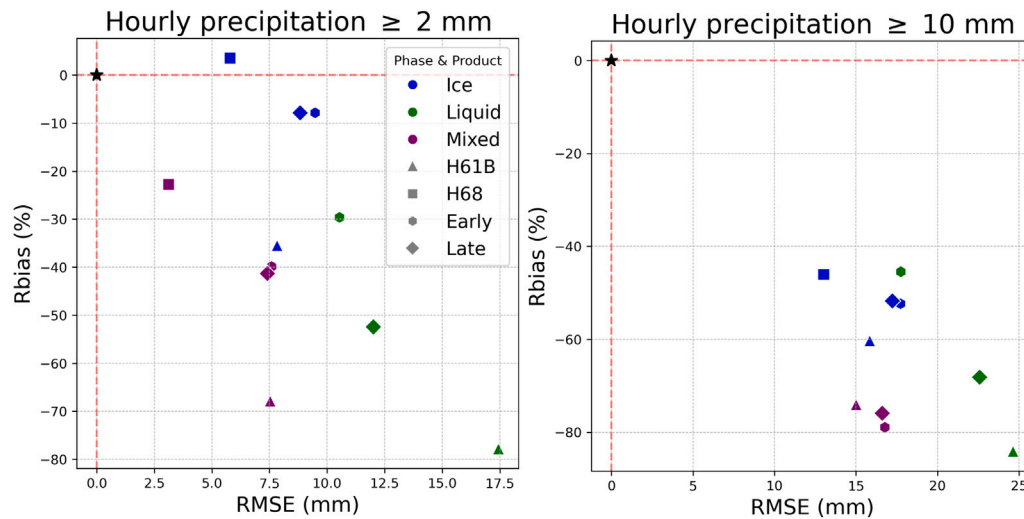


Fig. 9. Scatter plot of Rbias (%) versus RMSE (mm) for each satellite product stratified according to the upper cloud phase giving rise to accumulated precipitation (left panel) ≥ 2 mm (right panel) ≥ 10 mm. The marked point (black star) is the reference for no errors, whereas the bottom right is the location with the highest errors. The colored circles represent the three possible cloud top phases and are used to stratify the errors associated with the product estimates (depicted by the remaining symbols). For example, the blue square represents the error associated with H68 estimates when the cloud top phase was ice.

clouds with similar microphysical characteristics—and vice versa. Although ice-phase processes appear to dominate in the analyzed cases of extreme precipitation, this may reflect the selection bias introduced by focusing on intense events, which are more likely to involve deep convection. Warm rain processes, which can be relevant in orographic precipitation and may occur even under cloud tops with ice, are likely underrepresented in our dataset. This highlights the potential decoupling between surface precipitation and the ice-phase signatures detected by PMW sensors. Therefore, while our findings emphasize the role of ice-phase processes in satellite precipitation estimates, they also point to the need for further studies that incorporate a broader range of precipitation regimes, including warm rain events.

4. Conclusions

The performance of three H SAF products (H61B, H64, H68) and the IMERG Early and Late Runs V06B were validated using hourly and daily precipitation records from 186 rain gauges distributed over a Mediterranean coastal region using eighteen rainfall days. A point-to-pixel method was used to establish spatial matches between the two datasets, and several metrics were analyzed to quantify the discrepancies in precipitation accumulations. The validation and intercomparison of the performance of these satellite products were conducted considering four precipitation accumulations thresholds at both hourly and daily scales. This analysis also examined the relationship between surface precipitation observations/estimates and microphysical properties at the cloud top. The main findings can be summarized as follows:

- Satellite estimates tend to overestimate low precipitation accumulations, with this overestimation being more evident in the H68 product at the hourly scale and in IMERG Late at the daily scale. Nevertheless, both products demonstrate remarkable skill in detecting precipitation occurrences at accumulations of 0.5 mm in one hour and 1 mm in one day, respectively. The H64 product showed the lowest errors at the daily scale, albeit with limitations in identifying events.
- At the hourly scale, the H61B and H68 products exhibited a higher proportion of pixels with high KGE values compared to the IMERG products, with IMERG Early showing the poorest performance. While IMERG estimates slightly improve at the daily scale, the H61B product loses skill as the temporal aggregation increases, and H64 stands out as the best-performing product.

At this scale, the worst results were observed in coastal areas, particularly in northeastern Catalonia, while the highest KGE values were found in the southwestern region.

- Intense precipitation is rarely captured by satellite products. In fact, all satellite products show a significant decline in their ability to accurately detect precipitation once thresholds exceed 2 mm in one hour and 10 mm in one day. IMERG products are more sensitive to detecting extreme accumulation values, although with low temporal and spatial agreement with rain gauge observations. The H68 product exhibited the largest errors in such events, while IMERG Late proved to be the most effective despite maintaining notable biases.
- The accurate precipitation estimates from the satellite products were consistently associated with high values of COT and IWP, whereas false alarms tended to correspond to low values of these variables. Although the analyzed cases mostly involved clouds with significant ice-phase signatures, the results indicated poorer performance in estimating precipitation associated with the presence of liquid or mixed-phase water at the cloud top. This may reflect the challenges in detecting warm rain processes with PMW sensors and the under-representation of such events due to the focus on extreme precipitation episodes.

The results of this study not only validate a range of satellite precipitation estimation products but also facilitate comparisons among them, providing valuable information to users for various applications. The overall agreement between the rain gauge and satellite estimates was weak, indicating that although the products can capture extreme precipitation values, they do not always align with ground observations. This study highlights the need for continued detailed validation efforts to enhance the accuracy and effectiveness of precipitation retrieval algorithms. One of the main challenges remains the difficulty of detecting and quantifying extreme precipitation events, particularly in regions with complex topography, where the performance of satellite products declines significantly. Future research should explore ways to improve retrieval algorithms for these cases, including refining bias correction methodologies and incorporating additional atmospheric and cloud microphysical parameters into the estimation process.

Despite this limitation, the use of these products adds value in various operational contexts depending on specific needs. For instance, in risk and emergency management, it is crucial to have data that is spatially homogeneous, with a high update frequency and the lowest

possible latency, such as H61B. In hydrometeorological applications, like modeling for river basins, a less stringent latency and a spatial resolution adaptable to the basin's size might be required (products like H68, Early, or H64). Finally, in post-event analyses, the focus shifts to precipitation estimates that are as close as possible to observed measurements, with higher spatial resolution being more critical than lower latency, as in the Late product.

During the course of this study, the available IMERG data corresponded to version 06B; however, IMERG was later updated to V07B, which introduced significant improvements, including three corrections to sources of bias and the implementation of two new advanced schemes, SHARPEN and SPEEDe. These advancements should be considered in future research to assess their impact on precipitation estimates. Additionally, our findings highlight key areas for improving satellite precipitation estimation algorithms, particularly in capturing warm cloud rainfall and extreme precipitation events. Incorporating cloud microphysical properties, such as COT and IWP, into retrieval algorithms could enhance the detection of warm rain processes, which remain a challenge for current satellite products. Furthermore, given that precipitation typology significantly influences satellite product performance, integrating GPROF precipitation type information into IMERG could improve its accuracy and consistency across different climatic and geographic regions. Another fundamental challenge is the detection of ice-phase hydrometeors, as current sensors have limited capability to distinguish between different ice crystal types and their impact on upwelling radiation. Addressing these uncertainties is essential for improving satellite-based precipitation retrievals, initializing numerical weather models, and refining microphysical parameterizations in future algorithm developments.

This study also has limitations that should be acknowledged. One key challenge is the spatial scale mismatch between point-based rain gauge measurements and satellite grid data, which can introduce biases, particularly in regions with complex topography such as Catalonia. Additionally, rain gauge measurements were not corrected for wind effects, which may lead to systematic biases due to evaporation loss and underestimation of trace precipitation values. The pixel-to-point validation method, while widely used, assumes that a single rain gauge represents the mean precipitation of an entire satellite grid cell, a limitation that becomes more problematic at sub-daily timescales. Future research should explore improving individual PMW retrievals before merging them into composite products such as IMERG. A precipitation classification scheme incorporating mixed-phase precipitation retrievals could enhance detection accuracy, particularly for orographic precipitation. Moreover, the integration of high-resolution

atmospheric models and machine learning techniques could refine retrieval algorithms, especially for extreme rainfall events. As part of future efforts, the validation of the recently released IMERG V07 will be essential to assess its improvements in capturing both light and extreme precipitation. These advancements will contribute to more accurate and timely satellite-based precipitation products, ultimately benefiting hydrometeorological forecasting and disaster risk management.

CRediT authorship contribution statement

Eric Peinó: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Marco Petracca:** Writing – review & editing, Methodology, Investigation, Data curation, Conceptualization. **Francesc Polls:** Writing – review & editing. **Mireia Udina:** Writing – review & editing, Supervision. **Joan Bech:** Writing – review & editing, Supervision, Methodology, Investigation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

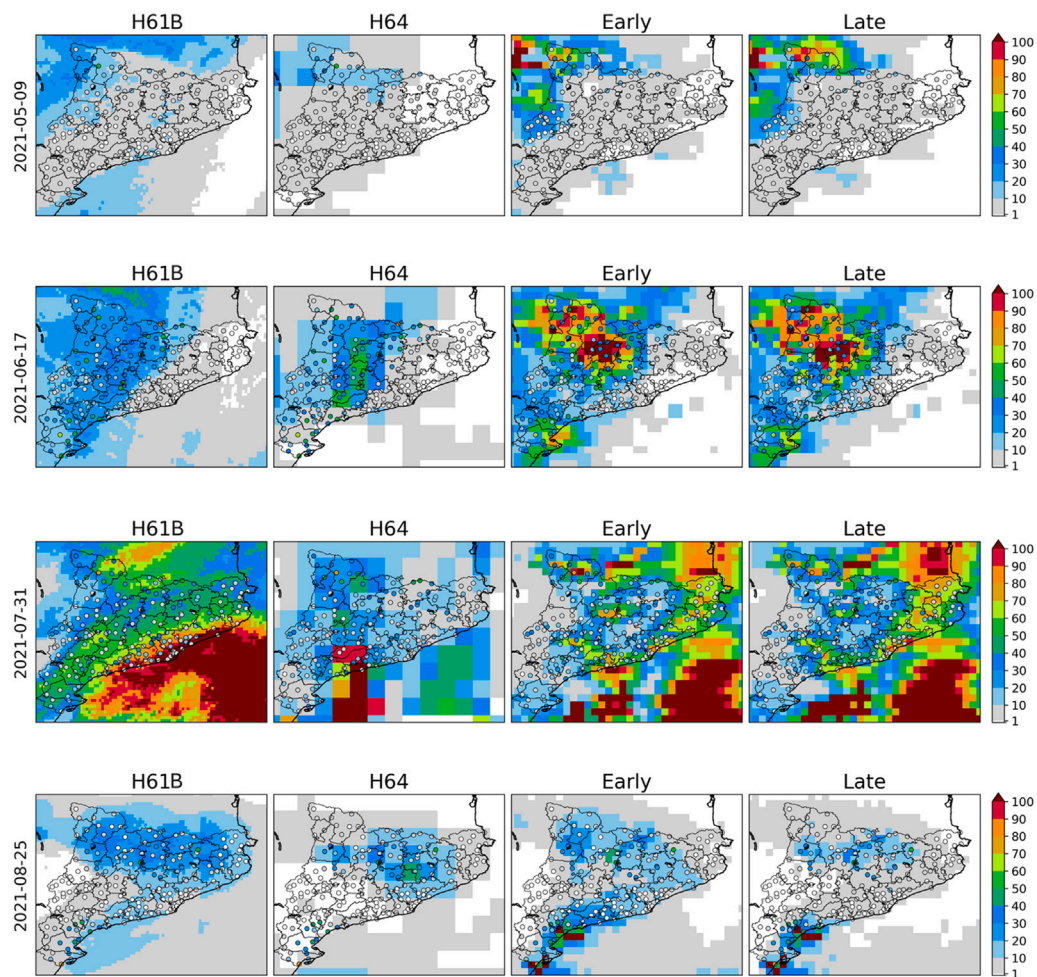
Satellite data for this study were provided by NASA/JAXA and the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT). Gauge data were obtained from the Meteorological Service of Catalonia and the Spanish Meteorological Agency (AEMET). The authors also acknowledge support from the Secretariat for Universities and Research of the Ministry of Enterprise and Knowledge of the Government of Catalonia, co-funded by the European Social Fund (ESF), and from the Water Research Institute (IdRA) of the University of Barcelona. This research was additionally funded by Spanish Ministry of Science and Innovation through the projects ARTEMIS (PID2021-124253OB-I00) and PANDORA (PDC2022-133207-I00).

Appendix A

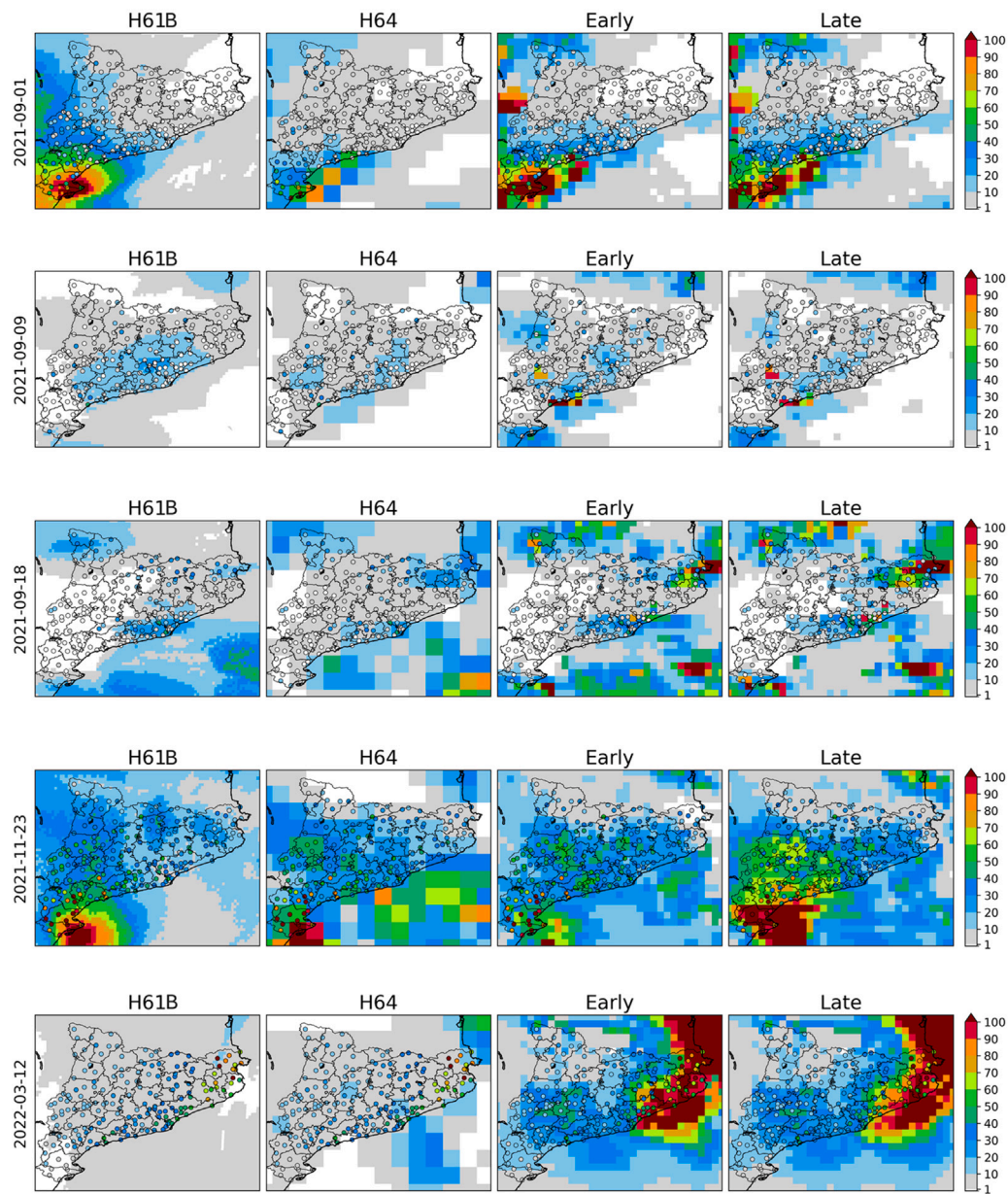
See [Table A.1](#).

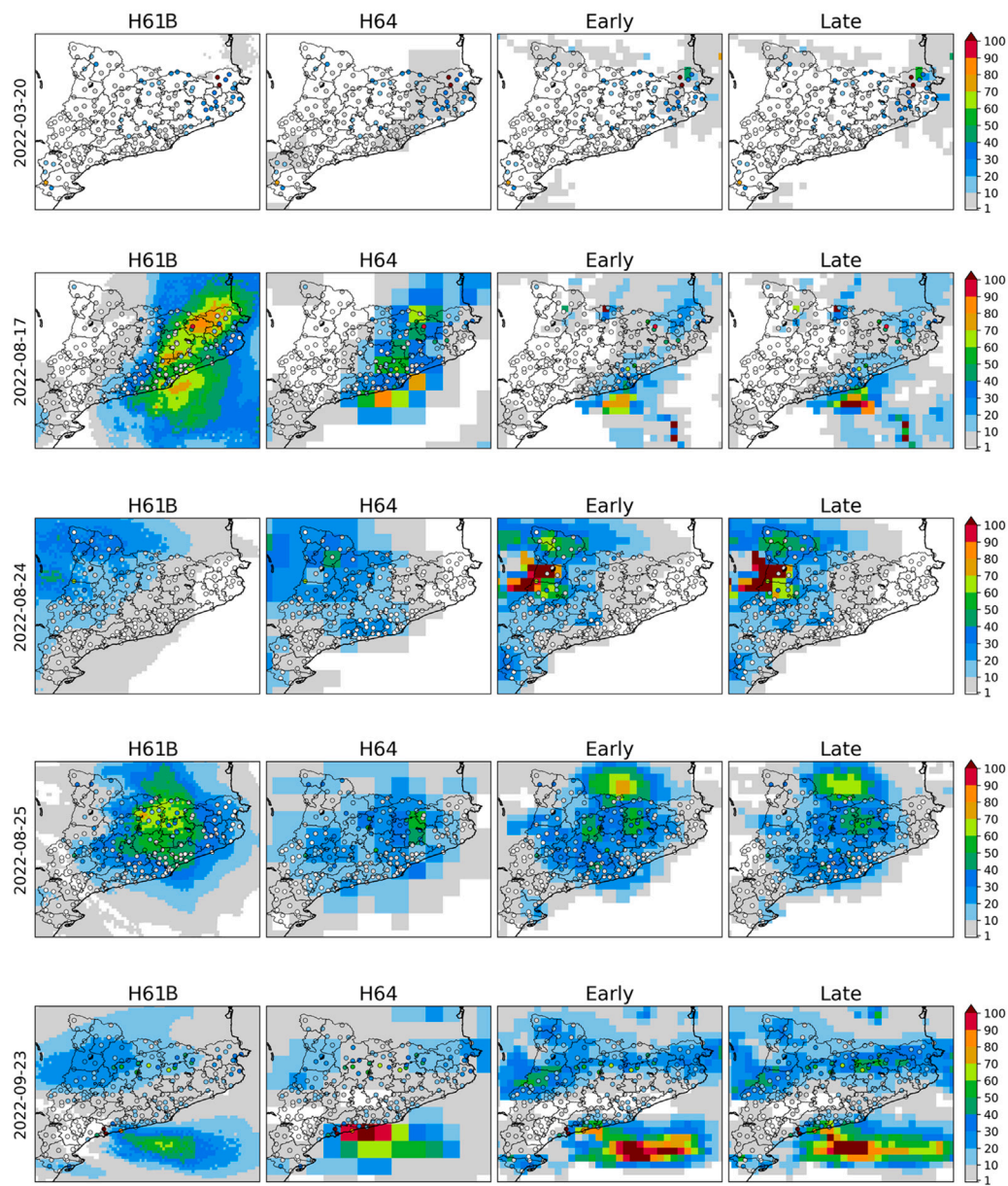
Table A.1
List of maximum rainfall observed by raingauges (RG) and estimated by satellite products at hourly and daily scale.

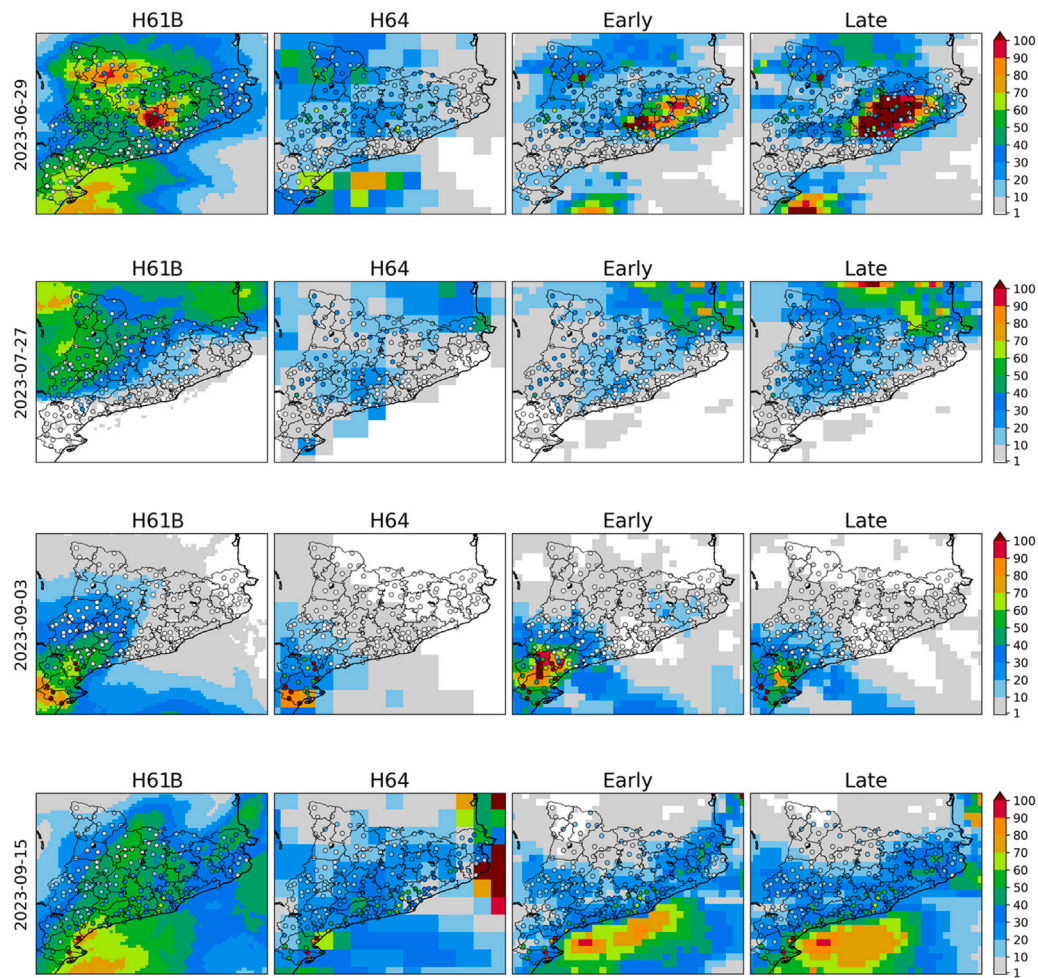
Dates	Hourly maximum (mm)					Daily maximum (mm)				
	RG	H61B	H68	Early	Late	RG	H61B	H64	Early	Late
09/05/2021	14.4	5.9	2.5	38.1	40.1	50.8	28.4	21.1	80.3	80.5
17/06/2021	49.1	10.2	17.4	48.7	42.6	86.8	37.3	55.2	164.7	136.2
31/07/2021	30.7	15.0		47.0	41.0	70.3	115.0	106.2	98.0	89.8
25/08/2021	51.4	8.5	0.7	48.2	49.9	78.8	33.5	41.7	117.8	132.6
01/09/2021	100.6	17.3		50.1	49.2	251.9	118.5	85.1	244.9	226.9
09/09/2021	39.0	9.4	0.9	71.3	60.6	44.7	25.3	16.5	110.5	99.9
18/09/2021	43.5	8.6	6.1	65.8	69.6	48.8	27.6	32.8	121.8	141.9
23/11/2021	77.2	11.8		18.9	35.3	191.9	110.2	101.2	64.5	150.4
12/03/2022	20.4	2.5	2.2	21.3	19.8	110.9	11.8	12.9	147.9	152.5
20/03/2022	20.7	0.1	0.3	9.6	9.6	113	1.1	3.6	40.6	54.7
17/08/2022	49.4	13.3	6.6	59.6	59.6	97.2	80.8	61.5	117.5	162.1
24/08/2022	29.7	11.3	0.7	44.1	53.3	62.3	37.9	40.2	123.9	143.2
25/08/2022	46.2	13.5	4.2	26.2	22.3	53	60.6	54.4	60.7	54.1
23/09/2022	54.9	9.0	0.0	31.8	22.1	130.4	40.6	111.6	71.1	62.9
29/06/2023	53.6	27.1	10.9	36.5	42.8	139.3	120.2	64.7	118.5	199.4
27/07/2023	29.5	23.3	15.2	30.1	30.1	46.8	59.1	43.0	44.9	60.5
03/09/2023	99.2	11.5	0.4	21.1	16.1	206.0	87.7	88.8	90.4	67.1
15/09/2023	50.5	24.3	28.9	19.1	19.5	120.0	80.3	105.9	46.7	64.2

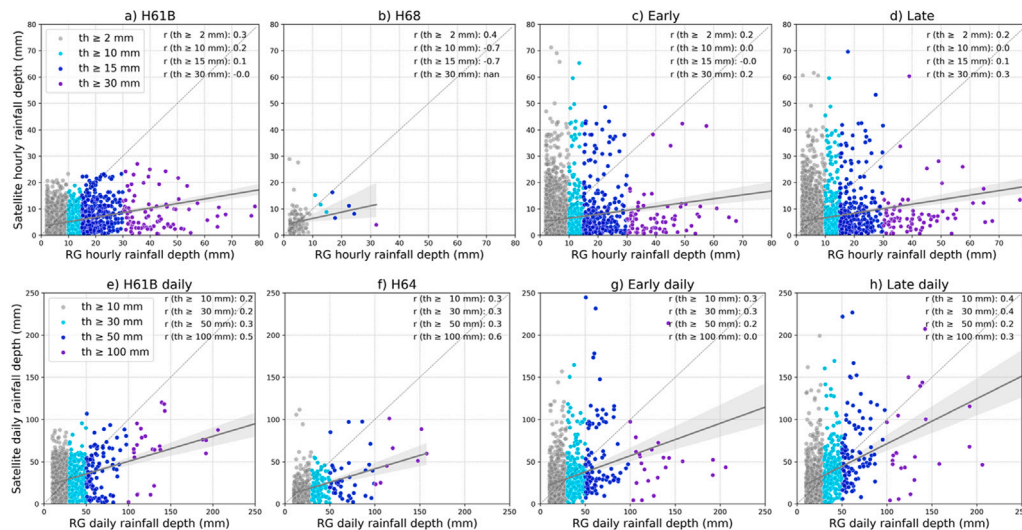


Appendix B. Performance of daily precipitation (mm) in all selected cases of each satellite product versus the records of all rain gauges in the domain









Appendix C. Scatter plots comparing each satellite product against rain gauge observations at hourly (top row) and daily (bottom row) scales. Rain gauge precipitation intensities are stratified into different accumulation thresholds and represented by different colors. Only gauge values meeting or exceeding each threshold are included, while satellite values are unrestricted. The regression fit line, corresponding to the lowest accumulation threshold for each timescale, is also shown with a 95% confidence interval (gray line with shaded area)

Data availability

All datasets used in this study are publicly available and properly cited in the Data Availability section of the manuscript, including links to the original sources.

References

- Adhikari, A., Behrangi, A., 2022. Assessment of satellite precipitation products in relation with orographic enhancement over the Western United States. *Earth Space Sci.* 9, e2021EA001906. <https://doi.org/10.1029/2021EA001906>.
- AEMET, 2023. Manual de uso de términos meteorológicos. URL: https://www.aemet.es/documentos/es/conocermas/recursos_en_linea/publicaciones_y_estudios/publicaciones/Manual_de_uso_de_terminos_met_2023/Manual_de_uso_de_terminos_met_2023.pdf.
- Baez-Villanueva, O.M., Zambrano-Bigiarini, M., Ribbe, L., Nauditt, A., Giraldo-Osorio, J.D., Thinh, N.X., 2018. Temporal and spatial evaluation of satellite rainfall estimates over different regions in Latin-America. *Atmos. Res.* 213, 34–50. <https://doi.org/10.1016/j.atmosres.2018.05.011>.
- Bitew, M.M., Gebremichael, M., 2010. Evaluation through independent measurements: Complex terrain and humid tropical region in Ethiopia. In: *Satellite Rainfall Applications for Surface Hydrology*. Springer, Dordrecht, pp. 205–214. https://doi.org/10.1007/978-90-481-2915-7_12.
- Braga, R.C., Rosenfeld, D., Krüger, O.O., Ervens, B., Holanda, B.A., Wendisch, M., Krisna, T., Pöschl, U., Andreae, M.O., Voigt, C., Pöhlker, M.L., 2021. Linear relationship between effective radius and precipitation water content near the top of convective clouds: measurement results from ACRIDICON-CHUVA campaign. *Atmospheric Chem. Phys.* 21, 14079–14088. <https://doi.org/10.5194/acp-21-14079-2021>.
- Braga, R.C., Rosenfeld, D., Weigel, R., Jurkat, T., Andreae, M.O., Wendisch, M., Pöschl, U., Voigt, C., Mahnke, C., Borrmann, S., Albrecht, R.I., Molleker, S., Vila, D.A., Machado, L.A.T., Grulich, L., 2017. Further evidence for CCN aerosol concentrations determining the height of warm rain and ice initiation in convective clouds over the Amazon basin. *Atmospheric Chem. Phys.* 17, 14433–14456. <https://doi.org/10.5194/acp-17-14433-2017>.
- Brocca, L., Ciabatta, L., Massari, C., Moramarco, T., Hahn, S., Hasenauer, S., Kidd, R., Dorigo, W., Wagner, W., Levizzani, V., 2014. Soil as a natural rain gauge: Estimating global rainfall from satellite soil moisture data. *J. Geophys. Res.: Atmospheres* 119, 5128–5141.
- Caracciolo, D., Francipane, A., Viola, F., Noto, L.V., Deidda, R., 2018. Performances of GPM satellite precipitation over the two major Mediterranean islands. *Atmos. Res.* 213, 309–322. <https://doi.org/10.1016/j.atmosres.2018.06.010>.
- Cenci, L., Laiolo, P., Gabellani, S., Campo, L., Silvestro, F., Delogu, F., Boni, G., Rudari, R., 2016. Assimilation of H-SAF soil moisture products for flash flood early warning systems. Case study: Mediterranean catchments. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* 9, 5634–5646. <https://doi.org/10.1109/JSTARS.2016.2598475>, conference Name: IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.
- Chen, C., Chen, Q., Duan, Z., Zhang, J., Mo, K., Li, Z., Tang, G., 2018. Multiscale comparative evaluation of the GPM IMERG v5 and TRMM 3B42 v7 precipitation products from 2015 to 2017 over a climate transition area of China. *Remote. Sens.* 10, 944. <https://doi.org/10.3390/rs10060944>.
- Chen, H., Sun, L., Cifelli, R., Xie, P., 2022. Deep learning for bias correction of satellite retrievals of orographic precipitation. *IEEE Trans. Geosci. Remote Sens.* 60, 1–11. <https://doi.org/10.1109/TGRS.2021.3105438>, conference Name: IEEE Transactions on Geoscience and Remote Sensing.
- Chiaravallotti, F., Brocca, L., Procopio, A., Massari, C., Gabriele, S., 2018. Assessment of GPM and SM2RAIN-ASCAT rainfall products over complex terrain in southern Italy. *Atmos. Res.* 206, 64–74. <https://doi.org/10.1016/j.atmosres.2018.02.019>.
- Ciabatta, L., Massari, C., Brocca, L., Gruber, A., Reimer, C., Hahn, S., Paulik, C., Dorigo, W., Kidd, R., Wagner, W., 2018. SM2RAIN-CCI: A new global long-term rainfall data set derived from ESA CCI soil moisture. *Earth Syst. Sci. Data* 10, 267–280. <https://doi.org/10.5194/essd-10-267-2018>.
- Cui, W., Dong, X., Xi, B., Feng, Z., Fan, J., 2020. Can the GPM IMERG final product accurately represent MCSs' precipitation characteristics over the Central and Eastern United States? *J. Hydrometeorol.* 21, 39–57. <https://doi.org/10.1175/JHM-D-19-0123.1>, publisher: American Meteorological Society Section: Journal of Hydrometeorology.
- Dai, Q., Han, D., Rico-Ramirez, M.A., Islam, T., 2013. The impact of raindrop drift in a three-dimensional wind field on a radar-gauge rainfall comparison. *Int. J. Remote Sens.* 34, 7739–7760. <https://doi.org/10.1080/01431161.2013.826838>.
- Derin, Y., Kirstetter, P.E., Brauer, N., Gourley, J.J., Wang, J., 2022. Evaluation of IMERG satellite precipitation over the land-coast-ocean continuum. Part II: Quantification. *J. Hydrometeorol.* 23, 1297–1314. <https://doi.org/10.1175/JHM-D-21-0234.1>.
- Ebert, E.E., 2008. Fuzzy verification of high-resolution gridded forecasts: a review and proposed framework. *Meteorol. Appl.* 15, 51–64. <https://doi.org/10.1002/met.25>.
- EUMETSAT, 2020. Algorithm Theoretical Baseline Document (ATBD) for product H61 P-AC-SEVIRI-PMW and H90 P-AC-SEVIRI e. URL: https://hsaf.meteoam.it/CaseStudy/GetDocumentUserDocument?fileName=saf_hsaf_atbd-61b-90b_1_0.pdf&tipo=ATBD.
- EUMETSAT, 2021. Algorithm Theoretical Baseline Document (ATBD) for product H68 (P-IN-PMW). URL: https://hsaf.meteoam.it/CaseStudy/GetDocumentUserDocument?fileName=SAF_HSAF_ATBD-H68_v3.0.pdf&tipo=ATBD.
- EUMETSAT, 2022. Algorithm Theoretical Baseline Document (ATBD) for product H64 P-AC-SM2RAIN. URL: https://hsaf.meteoam.it/CaseStudy/GetDocumentUserDocument?fileName=SAF_HSAF_ATBD-H64_2_1.pdf&tipo=ATBD.
- Freud, E., Rosenfeld, D., 2012. Linear relation between convective cloud drop number concentration and depth for rain initiation. *J. Geophys. Res.: Atmospheres* 117, <https://doi.org/10.1029/2011JD016457>.

- Furl, C., Ghebreyesus, D., Sharif, H.O., 2018. Assessment of the performance of satellite-based precipitation products for flood events across diverse spatial scales using GSSHA modeling system. *Geosciences* 8, 191. URL: <https://www.mdpi.com/2076-3263/8/6/191>, number: 6 Publisher: Multidisciplinary Digital Publishing Institute.
- Gaona, M.F.R., Overeem, A., Leijnse, H., Uijlenhoet, R., 2016. First-year evaluation of GPM rainfall over the Netherlands: IMERG day 1 final run (V03D). *J. Hydrometeorol.* 17, 2799–2814. <http://dx.doi.org/10.1175/JHM-D-16-0087.1>, publisher: American Meteorological Society Section: Journal of Hydrometeorology.
- Gentilucci, M., Barbieri, M., Pambianchi, G., 2022. Reliability of the IMERG product through reference rain gauges in Central Italy. *Atmos. Res.* 278, 106340. <http://dx.doi.org/10.1016/j.atmosres.2022.106340>.
- Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *J. Hydrol.* 377, 80–91. <http://dx.doi.org/10.1016/j.jhydrol.2009.08.003>.
- H SAF, 2022a. Product Validation Report (PVR) for Product H68 (P-in-PMW). Technical Report SAF/HSAF/PVR-68, EUMETSAT Satellite Application Facility on Support to Operational Hydrology and Water Management (H SAF), URL: https://hsaf.meteoam.it/CaseStudy/GetDocumentUserDocument?fileName=saf_hsaf_pvr-68.pdf&tipo=PVR.
- H SAF, 2022b. Product Validation Report (PVR) for Product P-AC-SM2RAIN (H64). Technical Report SAF/HSAF/PVR-64, EUMETSAT Satellite Application Facility on Support to Operational Hydrology and Water Management (H SAF), URL: https://hsaf.meteoam.it/CaseStudy/GetDocumentUserDocument?fileName=saf_hsaf_pvr-64.pdf&tipo=PVR.
- H SAF, 2022c. Product Validation Report (PVR) for Products P-AC-SEVIRI-PMW (H61B) and P-AC-SEVIRI-E (H90). Technical Report SAF/HSAF/PVR-61-90, EUMETSAT Satellite Application Facility on Support to Operational Hydrology and Water Management (H SAF), URL: https://hsaf.meteoam.it/CaseStudy/GetDocumentUserDocument?fileName=saf_hsaf_pvr-61-90.pdf&tipo=PVR.
- Hafizi, H., Sorman, A.A., 2023. Performance assessment of multi-source, satellite-based and reanalysis precipitation products over variable climate of Turkey. *Theor. Appl. Climatol.* 153, 1341–1354. URL: <https://link.springer.com/article/10.1007/s00704-023-04538-6>, company: Springer Distributor: Springer Institution: Springer Label: Springer Number: 3 Publisher: Springer Vienna.
- Huffman, G.J., Bolvin, D.T., Braithwaite, D., Hsu, K.L., Joyce, R.J., Kidd, C., Nelkin, E.J., Soroshian, S., Stocker, E.F., Tan, J., Wolff, D.B., Xie, P., 2020. Integrated multi-satellite retrievals for the global precipitation measurement (GPM) mission (IMERG). In: Levizzani, V., Kidd, C., Kirschbaum, D.B., Kummerow, C.D., Nakamura, K., Turk, F.J. (Eds.), *Satellite Precipitation Measurement*. In: *Advances in Global Change Research*, vol. 1, Springer International Publishing, Cham, pp. 343–353. http://dx.doi.org/10.1007/978-3-030-24568-9_19.
- Huffman, G.J., Bolvin, D.T., Braithwaite, D., Hsu, K., Joyce, R., Xie, P., Yoo, S.H., 2015. NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG). URL: https://gpm.nasa.gov/sites/default/files/2020-05/IMERG_ATBD_V06.3.pdf.
- Jansa, A., Alpert, P., Arbogast, P., Buzzi, A., Ivancan-Picek, B., Kotroni, V., Llasat, M.C., Ramis, C., Richard, E., Romero, R., 2014. MEDEX: a general overview. *Nat. Hazards Earth Syst. Sci.* 14, 1965–1984. <http://dx.doi.org/10.5194/nhess-14-1965-2014>.
- Jolliffe, I.T., Stephenson, D.B., 2012. *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. John Wiley & Sons, Google-Books-ID: DCxsKQeaBH8C.
- Katsanos, D., Retalis, A., Kalogiros, J., Psiloglou, B.E., Roukounakis, N., Anagnostou, M., 2024. Performance evaluation of satellite precipitation products during extreme events—The case of the Mediane Daniel in Thessaly, Greece. *Remote. Sens.* 16, 4216. URL: <https://www.mdpi.com/2072-4292/16/22/4216>, number: 22 Publisher: Multidisciplinary Digital Publishing Institute.
- Kazamias, A.P., Sapountzis, M., Lagouvardos, K., 2022. Evaluation of GPM-IMERG rainfall estimates at multiple temporal and spatial scales over Greece. *Atmos. Res.* 269, 106014. <http://dx.doi.org/10.1016/j.atmosres.2021.106014>.
- Kidd, C., Huffman, G., Maggioni, V., Chambon, P., Oki, R., 2021. The global satellite precipitation constellation: Current status and future requirements. *Bull. Am. Meteorol. Soc.* 102, E1844–E1861. <http://dx.doi.org/10.1175/BAMS-D-20-0299.1>.
- Kidd, C., Levizzani, V., 2019. Quantitative precipitation estimation from satellite observations. In: *Extreme Hydroclimatic Events and Multivariate Hazards in a Changing Environment*. Elsevier, pp. 3–39. <http://dx.doi.org/10.1016/B978-0-12-814899-0.00001-8>.
- Kidd, C., Tan, J., Kirstetter, P.E., Petersen, W.A., 2018. Validation of the Version 05 Level 2 precipitation products from the GPM Core Observatory and constellation satellite sensors. *Q. J. R. Meteorol. Soc.* 144, 313–328. <http://dx.doi.org/10.1002/qj.3175>, ISBN: 0035-9009 Publisher: Wiley Online Library.
- Kirschbaum, D.B., Huffman, G.J., Adler, R.F., Braun, S., Garrett, K., Jones, E., McNally, A., Skofronick-Jackson, G., Stocker, E., Wu, H., Zaitchik, B.F., 2017. NASA's remotely sensed precipitation: A reservoir for applications users. *Bull. Am. Meteorol. Soc.* 98, 1169–1184. <http://dx.doi.org/10.1175/BAMS-D-15-00296.1>, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society.
- Kirstetter, P.E., Petersen, W.A., Kummerow, C.D., Wolff, D.B., 2020. Integrated multi-satellite evaluation for the global precipitation measurement: Impact of precipitation types on spaceborne precipitation estimation. In: Levizzani, V., Kidd, C., Kirschbaum, D.B., Kummerow, C.D., Nakamura, K., Turk, F.J. (Eds.), *Satellite Precipitation Measurement*. In: *Advances in Global Change Research*, vol. 2, Springer International Publishing, Cham, pp. 583–608. http://dx.doi.org/10.1007/978-3-030-35798-6_7.
- Kling, H., Fuchs, M., Paulin, M., 2012. Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *J. Hydrol.* 424–425, 264–277. <http://dx.doi.org/10.1016/j.jhydrol.2012.01.011>.
- Koster, R.D., Brocca, L., Crow, W.T., Burgin, M.S., De Lannoy, G.J., 2016. Precipitation estimation using L-band and C-band soil moisture retrievals. *Water Resour. Res.* 52, 7213–7225. <http://dx.doi.org/10.1002/2016WR019024>.
- Kumah, K.K., 2023. H SAF precipitation products evaluation in Sub-Saharan Africa. URL: https://hsaf.meteoam.it/VisitingScientist/GetDocument?fileName=ReportV4_2023-Dec-14-4.pdf.
- Lahuerta García, J.A., 2021. Algorithm theoretical basis document for the precipitation product processors of the NWC/GEO.
- Lanza, L.G., Stagi, L., 2008. Certified accuracy of rainfall data as a standard requirement in scientific investigations. *Adv. Geosci.* 16, 43–48.
- Lei, H., Li, H., Zhao, H., Ao, T., Li, X., 2021. Comprehensive evaluation of satellite and reanalysis precipitation products over the eastern Tibetan plateau characterized by a high diversity of topographies. *Atmos. Res.* 259, 105661. <http://dx.doi.org/10.1016/j.atmosres.2021.105661>.
- Li, R., Qi, D., Zhang, Y., Wang, K., 2022. A new pixel-to-object method for evaluating the capability of the GPM IMERG product to quantify precipitation systems. *J. Hydrol.* 613, 128476. <http://dx.doi.org/10.1016/j.jhydrol.2022.128476>.
- Li, Y., Wang, W., Lu, H., 2018. Evaluation of the latest satellite-based precipitation products through pixel-point comparison and hydrological application over the Mekong River Basin. In: *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*. pp. 3043–3046. <http://dx.doi.org/10.1109/IGARSS.2018.8519448>.
- Li, Z., Yang, D., Hong, Y., 2013. Multi-scale evaluation of high-resolution multi-sensor blended global precipitation products over the Yangtze River. *J. Hydrol.* 500, 157–169. <http://dx.doi.org/10.1016/j.jhydrol.2013.07.023>.
- Li, X., Yang, Y., Mi, J., Bi, X., Zhao, Y., Huang, Z., Liu, C., Zong, L., Li, W., 2021. Leveraging machine learning for quantitative precipitation estimation from Fengyun-4 geostationary observations and ground meteorological measurements. *Atmospheric Meas. Tech.* 14, 7007–7023. URL: <https://amt.copernicus.org/articles/14/7007/2021/>, publisher: Copernicus GmbH.
- Llabrés-Brustenga, A., Rius, A., Rodríguez-Solà, R., Casas-Castillo, M.C., 2020. Influence of regional and seasonal rainfall patterns on the ratio between fixed and unrestricted measured intervals of rainfall amounts. *Theor. Appl. Climatol.* 140, 389–399. <http://dx.doi.org/10.1007/s00704-020-03091-w>.
- Llabrés-Brustenga, A., Rius, A., Rodríguez-Solà, R., Casas-Castillo, M.C., Redaño, A., 2019. Quality control process of the daily rainfall series available in Catalonia from 1855 to the present. *Theor. Appl. Climatol.* 137, 2715–2729. <http://dx.doi.org/10.1007/s00704-019-02772-5>.
- Llasat, M.C., Marcos, R., Llasat-Botija, M., Gilabert, J., Turco, M., Quintana-Seguí, P., 2014. Flash flood evolution in north-western Mediterranean. *Atmos. Res.* 149, 230–243. <http://dx.doi.org/10.1016/j.atmosres.2014.05.024>.
- Llasat, M.C., Marcos, R., Turco, M., Gilabert, J., Llasat-Botija, M., 2016. Trends in flash flood events versus convective precipitation in the Mediterranean region: The case of Catalonia. *J. Hydrol.* 541, 24–37. <http://dx.doi.org/10.1016/j.jhydrol.2016.05.040>.
- Maranan, M., Fink, A.H., Knippertz, P., Amekudzi, L.K., Atiah, W.A., Stengel, M., 2020. A process-based validation of GPM IMERG and its sources using a mesoscale rain gauge network in the West African Forest Zone. *J. Hydrometeorol.* 21, 729–749. <http://dx.doi.org/10.1175/JHM-D-19-0257.1>.
- Massari, C., Crow, W., Brocca, L., 2017. An assessment of the performance of global rainfall estimates without ground-based observations. *Hydrol. Earth Syst. Sci.* 21, 4347–4361. <http://dx.doi.org/10.5194/hess-21-4347-2017>.
- Moazami, S., Na, W., Najafi, M.R., de Souza, C., 2022. Spatiotemporal bias adjustment of IMERG satellite precipitation data across Canada. *Adv. Water Resour.* 168, 104300. <http://dx.doi.org/10.1016/j.advwatres.2022.104300>.
- Mugnai, A., Dietrich, S., Levizzani, V., Casella, D., Cattani, E., Di Paola, F., Formenon, M., Laviola, S., Sanò, P., 2013. Precipitation products from the hydrology saf. *Nat. Hazards Earth Syst. Sci.* 13, 1959–1981. <http://dx.doi.org/10.5194/nhess-13-1959-2013>.
- Navarro, A., García-Ortega, E., Merino, A., Sánchez, J.L., Tapiador, F.J., 2020. Orographic biases in IMERG precipitation estimates in the Ebro River basin (Spain): The effects of rain gauge density and altitude. *Atmos. Res.* 244, 105068. <http://dx.doi.org/10.1016/j.atmosres.2020.105068>.
- Peinó, E., Bech, J., Udina, M., 2022. Performance assessment of GPM IMERG products at different time resolutions, climatic areas and topographic conditions in Catalonia. *Remote. Sens.* 14, 5085. URL: <https://www.mdpi.com/2072-4292/14/20/5085>, number: 20 Publisher: Multidisciplinary Digital Publishing Institute.
- Peinó, E., Bech, J., Udina, M., Polls, F., 2024. Disentangling satellite precipitation estimate errors of heavy rainfall at the daily and sub-daily scales in the Western Mediterranean. *Remote. Sens.* 16, 457. URL: <https://www.mdpi.com/2072-4292/16/3/457>, number: 3 Publisher: Multidisciplinary Digital Publishing Institute.
- Petracca, M., D'Adderio, L.P., Porcù, F., Vulpiani, G., Sebastianelli, S., Puca, S., 2018. Validation of GPM dual-frequency precipitation radar (DPR) rainfall products over Italy. *J. Hydrometeorol.* 19, 907–925. <http://dx.doi.org/10.1175/JHM-D-17-0144.1>.

- Portier, A., Reed, J., Tan, J., Huffman, G., 2025. FAQ for IMERG V07 applications users. URL: https://gpm.nasa.gov/sites/default/files/2025-02/FAQ%20for%20IMERG%20V07%20Applications%20Users_202502.pdf.
- Pradhan, R.K., Markonis, Y., Vargas Godoy, M.R., Villalba-Pradas, A., Andreadis, K.M., Nikolopoulos, E.I., Papalexio, S.M., Rahim, A., Tapiador, F.J., Hanel, M., 2022. Review of GPM IMERG performance: A global perspective. *Remote Sens. Environ.* 268, 112754. <http://dx.doi.org/10.1016/j.rse.2021.112754>.
- Puca, S., Porcu, F., Rinollo, A., Vulpiani, G., Baguis, P., Balabanova, S., Campione, E., Ertürk, A., Gabellani, S., Iwanski, R., Jurašek, M., Kaňák, J., Kerényi, J., Koshinchanov, G., Kozinarova, G., Krahe, P., Lapeta, B., Lábó, E., Milani, L., Okon, L., Öztopal, A., Pagliara, P., Pignone, F., Rachimow, C., Rebora, N., Roulin, E., Sönmez, I., Toniazio, A., Biron, D., Casella, D., Cattani, E., Dietrich, S., Di Paola, F., Laviola, S., Levizzani, V., Melfi, D., Mugnai, A., Panegrossi, G., Petracca, M., Sanò, P., Zauli, F., Rosci, P., De Leonibus, L., Agosta, E., Gattari, F., 2014. The validation service of the hydrological SAF geostationary and polar satellite precipitation products. *Nat. Hazards Earth Syst. Sci.* 14, 871–889. <http://dx.doi.org/10.5194/nhess-14-871-2014>.
- Rojas, Y., Minder, J.R., Campbell, L.S., Massmann, A., Garreaud, R., 2021. Assessment of GPM IMERG satellite precipitation estimation and its dependence on microphysical rain regimes over the mountains of south-central Chile. *Atmos. Res.* 253, 105454. <http://dx.doi.org/10.1016/j.atmosres.2021.105454>.
- Rosenfeld, D., Gutman, G., 1994. Retrieving microphysical properties near the tops of potential rain clouds by multispectral analysis of AVHRR data. *Atmos. Res.* 34, 259–283. [http://dx.doi.org/10.1016/0169-8095\(94\)90096-5](http://dx.doi.org/10.1016/0169-8095(94)90096-5).
- Serrat-Capdevila, A., Valdes, J.B., Stakhiv, E.Z., 2014. Water management applications for satellite precipitation products: Synthesis and recommendations. *JAWRA J. Am. Water Resour. Assoc.* 50, 509–525. <http://dx.doi.org/10.1111/jawr.12140>.
- Skofronick-Jackson, G., Petersen, W.A., Berg, W., Kidd, C., Stocker, E.F., Kirschbaum, D.B., Kakar, R., Braun, S.A., Huffman, G.J., Iguchi, T., 2017. The Global Precipitation Measurement (GPM) mission for science and society. *Bull. Am. Meteorol. Soc.* 98, 1679–1695. <http://dx.doi.org/10.1175/BAMS-D-15-00306.1>.
- SMC, 2025. Situació meteorològica de perill. URL: <https://www.meteo.cat/wpweb/divulgacio/la-prediccio-meteorologica/situacio-meteorologica-de-perill/>. (Accessed: 29 March 2025).
- Sui, X., Li, Z., Ma, Z., Xu, J., Zhu, S., Liu, H., 2020. Ground validation and error sources identification for GPM IMERG product over the Southeast Coastal Regions of China. *Remote. Sens.* 12, 4154. URL: <https://www.mdpi.com/2072-4292/12/24/4154>, number: 24 Publisher: Multidisciplinary Digital Publishing Institute.
- Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., Hsu, K.L., 2018. A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Rev. Geophys.* 56, 79–107. <http://dx.doi.org/10.1002/2017RG000574>.
- Tan, J., Huffman, G.J., Bolvin, D.T., Nelkin, E.J., 2019. IMERG V06: Changes to the Morphing Algorithm. *J. Atmos. Ocean. Technol.* 36, 2471–2482. <http://dx.doi.org/10.1175/JTECH-D-19-0114.1>.
- Tan, J., Huffman, G.J., Bolvin, D.T., Nelkin, E.J., Rajagopal, M., 2021. SHARPEN: A scheme to restore the distribution of averaged precipitation fields. *J. Hydrometeorol.* 22, 2105–2116. <http://dx.doi.org/10.1175/JHM-D-20-0225.1>.
- Tan, J., Petersen, W.A., Tokay, A., 2016. A novel approach to identify sources of errors in IMERG for GPM ground validation. *J. Hydrometeorol.* 17, 2477–2491. <http://dx.doi.org/10.1175/JHM-D-16-0079.1>, publisher: American Meteorological Society Section: Journal of Hydrometeorology.
- Tapiador, F.J., Marcos, C., Sancho, J.M., 2019. The convective rainfall rate from cloud physical properties algorithm for meteosat second-generation satellites: Microphysical basis and intercomparisons using an object-based method. *Remote. Sens.* 11, 527. URL: <https://www.mdpi.com/2072-4292/11/5/527>, number: 5 Publisher: Multidisciplinary Digital Publishing Institute.
- Tapiador, F.J., Navarro, A., García-Ortega, E., Merino, A., Sánchez, J.L., Marcos, C., Kummerow, C., 2020. The contribution of rain gauges in the calibration of the IMERG product: Results from the first validation over Spain. *J. Hydrometeorol.* 21, 161–182. <http://dx.doi.org/10.1175/JHM-D-19-0116.1>.
- Tapiador, F.J., Turk, F.J., Petersen, W., Hou, A.Y., García-Ortega, E., Machado, L.A.T., Angelis, C.F., Salio, P., Kidd, C., Huffman, G.J., de Castro, M., 2012. Global precipitation measurement: Methods, datasets and applications. *Atmos. Res.* 104–105, 70–97. <http://dx.doi.org/10.1016/j.atmosres.2011.10.021>.
- Tapiador, F.J., Villalba-Pradas, A., Navarro, A., Martín, R., Merino, A., García-Ortega, E., Sánchez, J.L., Kim, K., Lee, G., 2021. A satellite view of an intense snowfall in Madrid (Spain): The storm ‘Filomena’ in January 2021. *Remote. Sens.* 13, 2702. URL: <https://www.mdpi.com/2072-4292/13/14/2702>, number: 14 Publisher: Multidisciplinary Digital Publishing Institute.
- Trapero, L., Bech, J., Duffourg, F., Esteban, P., Lorente, J., 2013. Mesoscale numerical analysis of the historical November 1982 heavy precipitation event over Andorra (Eastern Pyrenees). *Nat. Hazards Earth Syst. Sci.* 13, 2969–2990. URL: <https://nhess.copernicus.org/articles/13/2969/2013/>, publisher: Copernicus GmbH.
- Trapero, L., Bech, J., Rigo, T., Pineda, N., Forcadell, D., 2009. Uncertainty of precipitation estimates in convective events by the Meteorological Service of Catalonia radar network. *Atmos. Res.* 93, 408–418. <http://dx.doi.org/10.1016/j.atmosres.2009.01.021>.
- Trenberth, K.E., Fasullo, J.T., Kiehl, J., 2009. Earth’s global energy budget. *Bull. Am. Meteorol. Soc.* 90, 311–324. <http://dx.doi.org/10.1175/2008BAMS2634.1>.
- Turk, F.J., Ringerud, S.E., You, Y., Camplani, A., Casella, D., Panegrossi, G., Sanò, P., Ebtehaj, A., Guilleoteau, C., Utsumi, N., Prigent, C., Peters-Lidard, C., 2021. Adapting passive microwave-based precipitation algorithms to variable microwave land surface emissivity to improve precipitation estimation from the GPM constellation. *J. Hydrometeorol.* 22, 1755–1781. <http://dx.doi.org/10.1175/JHM-D-20-0296.1>.
- UK Met Office, 2025. Sea level pressure analysis charts. URL: https://digital.nmla.metoffice.gov.uk/SO_c9986ed3-e38f-4ed7-93c8-092de1a60cdb/. (Accessed 29 March 2025).
- Vicente-Serrano, S.M., Beguería, S., López-Moreno, J.I., 2010. A multiscale drought index sensitive to global warming: The standardized precipitation evapotranspiration index. *J. Clim.* 23, 1696–1718. <http://dx.doi.org/10.1175/2009JCLI2909.1>.
- Wang, S., Liu, J., Wang, J., Qiao, X., Zhang, J., 2019. Evaluation of GPM IMERG v05b and TRMM 3b42v7 precipitation products over high mountainous tributaries in lhasa with dense rain gauges. *Remote. Sens.* 11, 2080. URL: <https://www.mdpi.com/2072-4292/11/18/2080>, number: 18 Publisher: Multidisciplinary Digital Publishing Institute.
- Wang, C., Xu, J., Tang, G., Yang, Y., Hong, Y., 2020. Infrared precipitation estimation using convolutional neural network. *IEEE Trans. Geosci. Remote Sens.* 58, 8612–8625. <http://dx.doi.org/10.1109/TGRS.2020.2989183>, conference Name: IEEE Transactions on Geoscience and Remote Sensing.
- WMO, G., 1994. Guide To Hydrological Practices: Data Acquisition and Processing, Analysis, Forecasting and Other Applications. WMO, Geneva, Switzerland.
- Xie, W., Yi, S., Leng, C., Xia, D., Li, M., Zhong, Z., Ye, J., 2022. The evaluation of IMERG and ERA5-Land daily precipitation over China with considering the influence of gauge data bias. *Sci. Rep.* 12, 8085. <http://dx.doi.org/10.1038/s41598-022-12307-0>, publisher: Nature Publishing Group.
- Xu, J., Ma, Z., Tang, G., Ji, Q., Min, X., Wan, W., Shi, Z., 2019. Quantitative evaluations and error source analysis of Fengyun-2-based and GPM-based precipitation products over mainland China in summer, 2018. *Remote. Sens.* 11, 2992. <http://dx.doi.org/10.3390/rs11242992>, ISBN: 2072-4292 Publisher: MDPI.
- Yang, M., Liu, G., Chen, T., Chen, Y., Xia, C., 2020. Evaluation of GPM IMERG precipitation products with the point rain gauge records over Sichuan, China. *Atmos. Res.* 246, 105101. <http://dx.doi.org/10.1016/j.atmosres.2020.105101>.
- Yang, Y., Sun, W., Chi, Y., Yan, X., Fan, H., Yang, X., Ma, Z., Wang, Q., Zhao, C., 2022. Machine learning-based retrieval of day and night cloud macrophysical parameters over East Asia using Himawari-8 data. *Remote Sens. Environ.* 273, 112971. <http://dx.doi.org/10.1016/j.rse.2022.112971>.
- You, Y., Liu, G., 2012. The relationship between surface rainrate and water paths and its implications to satellite rainrate retrieval. *J. Geophys. Res.: Atmospheres* 117, <http://dx.doi.org/10.1029/2012JD017662>.
- Yu, C., Hu, D., Liu, M., Wang, S., Di, Y., 2020. Spatio-temporal accuracy evaluation of three high-resolution satellite precipitation products in China area. *Atmos. Res.* 241, 104952. <http://dx.doi.org/10.1016/j.atmosres.2020.104952>.
- Zhang, D., Yang, M., Ma, M., Tang, G., Wang, T., Zhao, X., Ma, S., Wu, J., Wang, W., 2022. Can GPM IMERG capture extreme precipitation in North China plain? *Remote. Sens.* 14, 928. URL: <https://www.mdpi.com/2072-4292/14/4/928>, number: 4 Publisher: Multidisciplinary Digital Publishing Institute.