Probabilistic Near-Real-Time Retrievals of Rain Over Africa Using Deep Learning

(article starts on next page)

Probabilistic Near-Real-Time Retrievals of Rain Over Africa Using Deep Learning

Adrià Amell[1] [ID], Lilian Hee[1] [ID], Simon Pfreundschuh[1,2] [ID], and Patrick Eriksson[1] [ID]

[1]Department of Space, Earth and Environment, Chalmers University of Technology, Gothenburg, Sweden, [2]Department of Atmospheric Science, Colorado State University, Fort Collins, CO, USA

Abstract We introduce Rain over Africa (RoA), a public retrieval algorithm providing near-real-time precipitation estimates over the African continent. The retrievals are based on Meteosat thermal infrared observations. Therefore, rain can be monitored constantly, minutes after input data dissemination. Despite this low latency, RoA accuracy is comparable to estimates requiring hours or more to obtain. Consequently, RoA is of particular interest where a rapid response is critical, such as for disaster preparedness. RoA retrievals employ a convolutional and quantile regression neural network: the latter enables detailed case-specific descriptions of the retrieval uncertainty. Four years of data from the calibration satellite in the GPM mission were used as training and evaluation labels. With this setup, limitations in earlier near-real-time retrievals for Africa were overcome. Moreover, the RoA network runs on regular workstations. With a 30-km effective resolution, RoA retrievals over land are more timely and detailed than the established IMERG precipitation estimates. RoA is also applicable over the surrounding ocean regions, maintaining a similar performance. However, there IMERG exhibits a better effective resolution, at least for its more favorable conditions. Additionally, RoA's probabilistic nature enables addressing the inherent uncertainties of satellite precipitation retrievals by using probabilities of exceeding precipitation thresholds. Further assessment reveals similar diurnal cycles between RoA and IMERG, although IMERG shows some instability. Visual inspection of rain evolution patterns also indicates that RoA is more consistent. Finally, an annual mean analysis including CHIRPS estimates shows regional differences among the three, with no clear outlier behavior for RoA.

Plain Language Summary Satellites are the only way to continuously monitor rainfall across all of Africa. However, current methods for estimating rain from space can take a long time because they are combining data from different sources. We introduce and test a new public method called RoA that provides near-real-time precipitation estimates for Africa. This is done by inputting one image from the Meteosat satellite for each time step into an artificial neural network, which we trained against precipitation estimates from the calibration satellite in the Global Precipitation Measurement mission, which has less continuous coverage of Africa than Meteosat. We found that the RoA estimates are as accurate as those from dedicated precipitation sensors. Moreover, while the latter are available every few hours at best, the RoA estimates can be provided every 15 min. This makes RoA useful for disaster preparedness and water management. Additionally, RoA can give practical probabilities of rain to help predict different scenarios.

1. Introduction

Accurate precipitation estimates are essential for a wide range of societal and scientific applications, including agricultural planning, water resource management, disaster mitigation, and climate research. Spaceborne precipitation retrievals overcome many of the challenges imposed by rain gauges and ground-based radars, which are characterized by their sparsity over large regions such as Africa (Kidd et al., 2017; Saltikoff et al., 2019). Microwave instruments are arguably the best suited for deriving near-surface precipitation estimates from space, as the frequencies at which they operate are only mildly affected by clouds, allowing them to see through. However, measuring in the microwave range imposes a low Earth orbit (LEO), resulting in large revisit periods, the time period elapsed between observations. Low Earth orbit satellite constellations, such as the one used in the Global Precipitation Measurement mission (GPM, Hou et al., 2014), can be exploited to reduce the revisit time, but they can remain worse than 1 hr for about half of the globe (Hou et al., 2014). Moreover, while the combination of operational microwave sensors reduces revisit times, the resulting precipitation estimates are generally less accurate than those of dedicated precipitation sensors such as GMI or AMSR2 (Pfreundschuh et al., 2024).

Different qualities are offered by satellites in geostationary orbit (GEO), which appear static when viewed from Earth. These platforms carry instruments that are used for measurement in the visible and infrared (IR) regions. Information at these wavelengths are strongly governed by cloud top properties, which complicates inferring near-surface precipitation, since the retrieval problem becomes notably ill posed or, in other words, suffers from a high degree of uncertainty. Nevertheless, geostationary imagery offers a high temporal resolution, acquiring data at least every 15 min for all of Africa. This represents a relatively continuous temporal coverage and can be used for near-real-time precipitation retrievals. The dissemination of information to the end user takes time: for high-impact precipitation events, reducing the latency of spaceborne rainfall estimates is crucial to assess risks.

There are several low-latency satellite precipitation retrieval approaches for large regions including Africa, which necessarily consider geostationary imagery. For instance, one can find that the closed-source operational products ERR/SCaMPR (Kuligowski, 2002; Kuligowski et al., 2016) from NOAA, P-IN-SEVIRI (Mugnai et al., 2013) from EUMETSAT hydrology SAF, GSMaP_NOW (Kubota et al., 2017) from JAXA, or PDIR-Now (Nguyen et al., 2020), the latest addition to the PERSIANN family of precipitation estimates (Nguyen et al., 2018), all claim to have latencies of at most 1 hr. On the open-source side of low-latency products covering Africa, we are only aware of the Convective Rainfall Rate (CRR) product from the EUMETSAT nowcasting SAF: this product provides categorical estimates, based on functions calibrated over three different radar networks in Europe.

Similarly to other domains, machine learning is also revolutionizing precipitation retrievals, with deep learning approaches increasingly adopted. In this context, U-net-like network architectures (Ronneberger et al., 2015) have proven particularly useful. These architectures employ convolutional layers arranged hierarchically, enabling the extraction of features at multiple spatial resolutions. This design allows the network to capture both fine-grained spatial patterns and broader large-scale structures. Consequently, U-net variants have been employed in recent studies of spaceborne precipitation estimation, including approaches that rely solely on geostationary infrared data (D'Adderio et al., 2023; Gao et al., 2022; Ji et al., 2024; Jiang et al., 2023; Pfreundschuh, Ingemarsson, et al., 2022; Wang et al., 2021), methods that fuse geostationary and passive microwave observations (Afzali Gorooh et al., 2022; Rahimi et al., 2024; Sadeghi et al., 2020), and techniques based exclusively on passive microwave inputs (Pfreundschuh, Brown, et al., 2022; Viltard et al., 2025).

Two of these studies are of particular relevance for the present work. Pfreundschuh, Ingemarsson, et al. (2022) developed Hydronn, a deep learning approach to improve HYDRO, an operational precipitation retrieval algorithm used at the National Institute for Space Research in Brazil. D'Adderio et al. (2023), in the context of an EUMETSAT hydrology SAF study, observed that through deep learning they obtained retrievals that compare favorably to their current algorithms in terms of rain rate errors. Notably, both studies employed U-net-like architectures to leverage spatial patterns. While these share several methodological similarities in that they depend only on data from geostationary platforms and employ deep learning with the label data from the same precipitation product, they differ in key aspects. Where D'Adderio et al. (2023) rely on the Meteosat SEVIRI instrument (Aminou et al., 1997; Schmid, 2000), which covers the 0° longitude, Pfreundschuh, Ingemarsson et al. (2022) use the imager onboard GOES-16 (Schmit et al., 2005), placed at approximately 75°. In addition, Hydronn employs probabilistic deep learning and provides a case-specific description of the irreducible uncertainty inherent to the precipitation retrieval problem, exacerbated in geostationary imagery.

Although these two works highlight complementary advances in deep learning for precipitation retrieval using geostationary imagery, important gaps remain in the broader landscape. Firstly, with the exception of Hydronn, existing models either rely on closed-source implementations or lack probabilistic uncertainty estimates, which limits transparency and restricts their adaptability for research and operational applications from third parties. Second, Africa is particularly vulnerable to extreme precipitation events, as these amplify existing socioeconomic challenges. Despite this, there is a notable absence of publicly available, accurate, and low-latency precipitation retrievals specifically targeting the African continent.

In this work, we address these gaps. We present a retrieval model, based on the Hydronn approach, but using exclusively Meteosat SEVIRI imagery to obtain near-real-time precipitation estimates covering all of Africa. We refer to these retrievals as RoA, an acronym for Rain over Africa. RoA retrievals are trained only with thermal IR information, which makes them independent from the presence of solar radiation. RoA retrievals use the same imager data as the closed-source work of D'Adderio et al. (2023), but have the added value of offering a probabilistic description of the retrieval, combined with a publicly available version of RoA (Amell, 2025); the

Table 1
Retrievals and Data Sets Used in This Work

Data	Source	Resolution[a]	Latency
SEVIRI IR images (brightness temp.)	Meteosat (2nd gen.)	3 km (0.027°), 15 min	A few minutes
GPM 2BCMB	GPM-CO (GMI and DPR-Ku)	5 km (0.045°), not applicable[b]	Not relevant[c]
RoA	SEVIRI (IR-only)	3 km (0.027°), 15 min	A few minutes[d]
IMERG Final	GEO and LEO platforms in GPM constellation and rain gauges	10 km (0.1°), 30 min	3.5 months
IMERG IR-only	GEO platforms	10 km (0.1°), 30 min	4 hr
CHIRPS	GEO platforms and rain gauges	5.5 km (0.05°), 1 day	3 weeks[e]

Data	Resampled	Purpose
SEVIRI IR images (brightness temp.)	No (own grid)	Input data for RoA
GPM 2BCMB	To SEVIRI to train/test and to IMERG for comparison in Section 4.4	Reference, training labels
RoA	No (preserves SEVIRI grid), except for Section 4.4 where it is mapped to IMERG	Retrieval presented here
IMERG Final	To SEVIRI for Sections 4.2 and 4.3, preserves its grid elsewhere	Comparison and baseline
IMERG IR-only[f]	Same as for IMERG Final	Comparison and baseline
CHIRPS	To IMERG grid	Comparison

[a]As specified by the data provider or approximated at 0° latitude and longitude. [b]Temporal resolution, here, should be understood as the revisit period: GPM 2BCMB depends on the nonsun-synchronous orbit of GPM-CO and therefore samples at different intervals as a function of the coordinates. [c]Its use in this study is not intended for a near-real-time application. [d]Produced by the end user. [e]According to Funk et al. (2015). [f]The IR precipitation field of IMERG files.

RoA code enables obtaining estimates within minutes of the dissemination of Meteosat data and, therefore, makes it suitable for nowcasting applications.

The input and reference data used to train the network are described in Section 2, together with two gauge-adjusted spaceborne precipitation data sets used to compare the RoA retrievals. Section 3 summarizes the deep learning retrieval method, including the neural network architecture as well as the training parameters. Afterward, in Section 4, the RoA retrievals are evaluated against a test data set and compared with the two gauge-adjusted data sets over one full year. Finally, a discussion is presented in Section 5, followed by conclusions.

2. Data

Table 1 summarizes the different retrievals and data sets used in this study. This section provides further details on each of them, with the exception of RoA, which is described in detail in Section 3. We describe how the data sets are derived, the strategy used for assembling the training database, and the versions employed.

2.1. Training and Test Database

2.1.1. Input: Meteosat SEVIRI

The Meteosat Second Generation (MSG, Schmetz et al., 2002) satellites are a series of four geostationary satellites that have been operational since the beginning of 2004 (Rota, 2004; Stark & Schumann, 2004) at longitudes close to 0°. The MSG satellites carry the Spinning Enhanced Visible and InfraRed Imager (SEVIRI) instrument (Aminou et al., 1997; Schmid, 2000), observing the Earth in 12 spectral channels with a repeat cycle of 15 min for the full Earth disc scan. All thermal IR channels have a sampling distance of 3 km at the subsatellite point.

In this work, we consider seven of the eight thermal IR channels from SEVIRI level 1.5 data (EUMET-SAT, 2009), ranging from 6.25 to 13.40 μm, as the input data; the 3.9 μm channel was not considered to avoid any solar contamination. In addition, we also use the satellite zenith angle as an input, to account for the degradation of spatial resolution when deviating from the subsatellite point. The RoA retrievals are obtained on the native grid of the SEVIRI observations, that is, on grid cells of 3 km at the subsatellite point.

2.1.2. Reference: GPM Combined Radar and Microwave Radiometer Retrievals

The GPM mission is a satellite mission consisting of an international constellation of spaceborne measurements from different platforms, which has the frequent sampling of precipitation as a primary objective (Hou et al., 2014). Central to the GPM constellation is the GPM Core Observatory (GPM-CO), a satellite which is used as a calibration reference. Launched in early 2014, this satellite follows a 65° low Earth orbit, sampling from 68°S to 68°N at different times with consecutive overpasses, intersecting the paths of the other GPM satellites.

The GPM-CO carries the multifrequency microwave radiometer GPM Microwave Imager (GMI, Draper et al., 2015), which can measure light to heavy precipitation, and the Dual-frequency Precipitation Radar (DPR, Kojima et al., 2012), which can provide precipitation characteristics in three dimensions and discriminate between rain and snow; the DPR is equipped with a Ku-band (13.6 GHz) radar and a Ka-band (35.55 GHz) radar. While the GMI provides a swath of 885 km, the DPR results in narrower overlapping swaths lagging that of GMI, with the Ka band having the smallest swath.

The combined GMI and DPR retrievals (Grecu et al., 2016) reduce uncertainties in the precipitation estimates compared to DPR-only or GMI-only retrievals. This is achieved by using complementary information in a physical model solved by an optimal estimation framework: the attenuation on the radar signal is constrained by the radiometer signal, which has more degrees of freedom. The different DPR swaths offer two practical possibilities for the combined retrievals: Ku-only or Ku and Ka band. This work uses the combined retrieval of GMI and Ku, employing the GPM 2BCMB product (Olson, 2018) as a reference, which offers estimates at a 5-km resolution on a 245-km swath. The choice of GPM 2BCMB as a ground truth is further justified by its role in calibrating GPM products such as IMERG (Section 2.2.1). GPM 2BCMB data are also used in this study for comparison in Sections 4.1 and 4.4. We employed version V06 for the training phase, while the successor version V07 is used for the evaluation and comparisons. The use of different versions is due to the data availability at the time of each usage.

It should be noted that ground-based measurements, mostly confined to land areas, are unevenly distributed and excessively sparse, particularly in Africa (Dinku, 2019; Saltikoff et al., 2019). In addition, the coarse temporal resolution of public rain gauge data sets hinders subhourly comparisons of rainfall estimates. Furthermore, in situ measurements introduce additional uncertainties when compared with spaceborne retrievals. Therefore, space-borne products such as GPM 2BCMB are the only viable reference for the full ROI. For these reasons, ground-based data are not considered in this work as a reference neither for training nor for evaluation.

2.1.3. Collocations

A database of collocations was assembled to train and evaluate the retrieval algorithm. Near-surface precipitation estimates from the GPM 2BCMB data product were collocated to the SEVIRI brightness temperatures using a nearest neighbor approach, both in space and time, on the geostationary projection provided by the SEVIRI level 1.5 data (EUMETSAT, 2009; Wolf, 1999). The specific implementation is described in Text S1 in Supporting Information S1. The collocated data were then folded in nonoverlapping tiles of $256 \times 256 \text{ px}^2$, with the precipitation estimates centered in the tile (Figure S1a in Supporting Information S1). This tile size was chosen as a balance between the 245-km swath of the reference data and the area covered: around 65% of the tile generally has no matched precipitation estimate. We defined the region of interest (ROI) of this work as the extent 40°S–40°N by 20°W–55°E, covering all of Africa. The database includes only tiles within this ROI, except for latitudes south of 25°S due to a silent software bug.

This database spans the period between 2018-03-01 and 2021-11-30. During this period, Meteosat-11 (MSG-4) was the primary operational satellite, and only data from this platform are included in the database. The data were systematically split into training (65%) and validation (16%) to mitigate any data leakage and account for any possible year-specific anomalies. The remaining 18% of the data were originally held out as a test set, but we discarded it in favor of the test set described below. Days 3, 7, 13, 19, and 29 of each month were assigned to the validation set, while the remaining nonprime numbered days to the training set. The training and validation data contain 40,609 tiles and almost 1 billion pixels with reference data.

At the time of writing, the MSG satellite at 0° covering Africa is once again MSG-3, the predecessor of MSG-4. This change occurred in 2023. In that year, more than 78% of the observations originated from MSG-3. We assembled the test set by collocating the GPM 2BCMB data on all matching MSG observations for 2023, with the

nearest neighbor approach (Text S1 in Supporting Information S1) within the ROI. The test set contains almost 500 million pixels with reference data.

Note that evaluating on this test set not only characterizes the performance under a different MSG platform but also introduces a temporal gap of over 1 year between training and testing of the retrieval, effectively removing data leakage. This evaluation strategy provides a more realistic assessment of the generalization ability of the model. The time span was chosen to align with the annual comparisons with the precipitation gauge-adjusted data sets in Sections 4.1 and 4.4: 2023 was the first complete calendar year available when we began the analyses.

Furthermore, since the collocations for training and evaluating the retrieval use GPM 2BCMB retrievals as the reference, it is also noteworthy that any retrieval based on this database will inherit GPM 2BCMB characteristics and assumptions. Consequently, RoA is expected to reproduce not only the GPM 2BCMB precipitation patterns but also any systematic biases present in GPM 2BCMB. These inherited errors can be amplified in the RoA retrieval process together with other IR-specific uncertainties, such as attenuation effects.

2.2. Satellite Precipitation Gauge-Adjusted Data Sets

2.2.1. IMERG V07

The Integrated MultisatellitE Retrievals for GPM (IMERG, Huffman, Bolvin et al., 2023) is a readily available precipitation GPM data set by NASA. The algorithm to produce IMERG integrates many precipitation estimates, derived from microwave spaceborne instruments complemented with single-channel IR data from geostationary platforms, aided by ancillary products, to offer a readily available global high-resolution precipitation product (Turk et al., 2008) of resolutions 30 min and $0.1° \times 0.1°$ (roughly 10×10 km^2). The IMERG estimates are obtained through intercalibration, propagation, and merging of the precipitation estimates, mostly obtained with the GPROF algorithm (Kummerow et al., 2015), which uses a Bayesian inversion method with GPM-CO data as the prior; during the GPM era, precipitation estimates from the combined DPR and GMI product are used to calibrate the IMERG estimates but not directly included in the product. Note that IMERG uses the closest satellite of opportunity from the GPM constellation for the precipitation estimates.

IMERG is produced in an iterative fashion, as the different data become available. The precipitation estimates are offered on three versions: the Early, Late and Final runs, each with a latency of about 4 hr, 14 hr, and 3.5 months, respectively, since the observation time. In addition to the data availability, the main differences between the runs is that the Early and Late include climatological adjustments based on historical Final runs, while the Final run includes calendar month adjustments in combination with calibration over land against Global Precipitation Climatology Centre (GPCC) gauge analyses products. Therefore, the Final run should be the most accurate, as suggested by the name, and is the version used here; specifically, the IMERG V07 Final run. The document describing the IMERG algorithm (Huffman, Bolvin, et al., 2023) states that its error estimate, based on the work of Huffman (1997) for monthly data, needs to be replaced, and we understand that as a shortcoming of IMERG.

In Section 4, statistics between IMERG and the reference in the test set are presented. In this case, IMERG is resampled to the finer grid cells of RoA with the nearest neighbor approach, enabling an evaluation against the same reference data. Since we use GPM 2BCMB as a reference, the collocated IMERG estimates can consistently rely on GMI retrievals (Huffman, Bolvin, et al., 2023); our notation 'IMERG (GMI)' refers to this situation. Given that the GMI sensor is the most capable sensor of the GPM constellation (Pfreundschuh et al., 2024), these IMERG estimates constitute an upper bound for its overall accuracy: the IMERG accuracy varies with the availability of the observations from different sensors, since it merges precipitation estimates from geostationary IR observations and the GPM constellation of passive microwave sensors. To establish the resulting variability in the accuracy of IMERG, we also assess the IR-only precipitation fields provided by IMERG. Due to the limited information content of these single-channel IR observations, they represent a lower bound for the accuracy of IMERG.

2.2.2. CHIRPS

The Climate Hazards group Infrared Precipitation with Stations (CHIRPS, Funk et al., 2015) data set is a quasi-global data set at a resolution of 0.05° with daily, pentadal, and monthly precipitation accumulations for land areas. Its process of obtaining precipitation estimates can be considered different than the IMERG algorithm or RoA therefore providing another retrieval framework with potentially different error characteristics. The

Table 2
Two Pixels, A and B, From the Collocation Database

Channel [μm]	Pixel value A	Pixel value B	
6.2		209.7 K	
7.3		210.0 K	
8.7		211.0 K	
9.7		218.6 K	RoA input
10.8		209.9 K	
12.0		209.1 K	
13.4		209.4 K	
Sat. angle	8.0°		
Rain rate [mm h^{-1}]	2.9	27.5	Label

algorithm behind CHIRPS is built around a climatology derived from station data, rain gauges, and intercalibrated geostationary IR observations. In a simplified explanation, the algorithm firstly detects the brightness temperatures below a fixed threshold to compute a precipitation estimate, based on historical statistics for each grid cell. Afterward, station data are used to blend rain gauge data into the precipitation estimate. As in IMERG, CHIRPS lacks a practical uncertainty estimate. In addition, CHIRPS is only provided over land.

3. Retrieval Method

Table 2 serves to illustrate one main challenge of precipitation retrievals: two pixels with the same brightness temperatures (to one decimal precision) can correspond to different reference rain rates, and the converse also holds. That is, the observations contain an uncertainty that, even when contextualized with spatial information, is irreducible. There is no unique solution for the retrieval problem; it is ill posed.

In this section, we outline the key ideas of the retrieval method employed for RoA, as well as the rationale behind the neural network choice. A detailed explanation has previously been provided in Pfreundschuh et al. (2018) and Pfreundschuh, Ingemarsson, et al. (2022).

3.1. Quantile Regression

Retrieval methods to estimate precipitation from satellites should consider the inherent irreducible uncertainty. The optimal estimation method (OEM, Rodgers, 2000), well-established for atmospheric retrievals, iteratively updates an initial uncertainty estimate following Bayesian statistics. However, OEM imposes a normal distribution, which can represent an inappropriate description of the uncertainties in spaceborne precipitation retrievals. A Markov chain Monte Carlo approach is arguably more accurate but too computationally heavy.

Tangent to these methods is quantile regression. This statistical method can characterize the irreducible uncertainty of retrieval problems, accommodate non-normal uncertainties, and requires no explicit assumptions regarding the relationship between brightness temperatures and precipitation. The root idea is that given a set of collocations, one can find a function f such that it yields the precipitation quantile $\hat{y}_\tau = f(\mathbf{x}, \tau)$, for a quantile level $\tau \in (0, 1)$, using a set of brightness temperatures $\mathbf{x}$. For example, if $\tau = 50\%$, one can identify all $\mathbf{x}$ that are the same in the collocations and compute the median rain rate they map to, obtaining $\hat{y}_{50\%}$.

In practice, quantile regression is achieved by minimizing the loss function

$$L_\tau\left(y, \hat{y}_\tau\right) = \max\left[\tau\left(y - \hat{y}_\tau\right), (1 - \tau)\left(y - \hat{y}_\tau\right)\right] \tag{1}$$

over all collocated precipitation estimates y to the input values $\mathbf{x}$. This formulation provides a nonparametric marginal cumulative distribution function that can be used to describe the probability of, among others, extreme events. A useful property of quantile regression is that it is invariant to monotonic transformations φ, since if $\Pr(y \leq y')$, then $\Pr(\varphi(y) \leq \varphi(y'))$. In addition, it is resistant to outliers in the data.

Figure 1 shows an RoA retrieval example for a pixel identified as heavy precipitation. The quantile regression results in a retrieval uncertainty that captures the reference value y, since τ reaches 1 after the y value. We can make three main observations for this particular case. First, the retrieval distribution is strongly non-Gaussian. Second, the retrieval shows some uncertainty (about 20%) in that there is precipitation in the pixel. Third, there is a long distribution tail, which suggests that similar $\mathbf{x}$ values can contain heavier precipitation. It is, however, not straightforward to compare the RoA retrieval, in blue, with the reference value point, in green. Therefore, it is practical to use a point estimate, where we employ the expected value of the RoA retrieval distribution, $\hat{y}_\mu$; this value is equivalent to the mean of an infinite number of samples from the retrieval distribution. In the example of Figure 1, $\hat{y}_\mu$ is approximately equal to the reference.
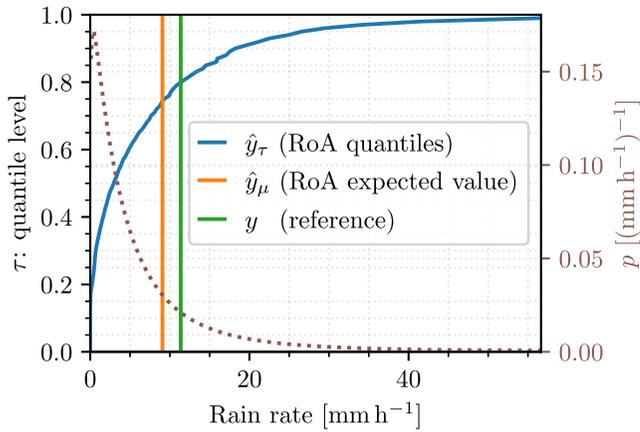
Figure 1. An RoA retrieval example. The brown dotted line shows an approximated probability density function p to aid interpretation; it was obtained by fitting a mixture distribution to produce a smoother curve.

3.2. Neural Network

There are multiple approaches to find f, but artificial neural networks tend to be advantageous due to their flexibility. In the field of remote sensing, Pfreundschuh et al. (2018) can be considered to have introduced quantile regression with neural networks, a term coined as QRNN. Advances in the field of neural networks are also applicable to QRNNs: for example, spatial information can be incorporated through convolutional layers in the network architecture which, in the context of precipitation retrievals, can be argued to improve the results (Pfreundschuh, Brown, et al., 2022).

Due to the similarity of this work with Pfreundschuh, Ingemarsson, et al. (2022), the retrieval method is based on a QRNN using the generic Hydronn architecture provided as part of its accompanying code (Pfreundschuh, 2022). Our QRNN is a fully convolutional neural network, which can extract contextual information at different scales, since it is based on a U-net-like architecture. For a deeper explanation on the architecture, we refer to Section 3.3 of Pfreundschuh, Ingemarsson, et al. (2022), noting that our implementation substitutes batch normalization with layer normalization. Our architecture configuration consists of two blocks per encoding stage, two head layers, 128 internal features where needed, and 99 output neurons; hyperparameter tuning was deemed too expensive for any potential improvement.

3.3. Training

In our implementation of the QRNN, we fixed the quantile levels to be all levels mapping to percentiles, that is, $\tau = 1\%, 2\%, \ldots, 99\%$. The eight input features, consisting of the brightness temperatures from the seven spectral channels and the satellite zenith angle (Table 2), are standardized with the training set statistics. On the other end, the reference precipitation estimates y [mm h^{-1}] are treated as absolute ground truth. Defining $z = \varphi(y)$, we used

$$z = \begin{cases} \log y, & y \neq 0 \\ \text{Uniform}(\log 10^{-3}, \log 10^{-2}), & \text{otherwise} \end{cases} \quad (2)$$

as the actual training labels for the loss function, where the last expression indicates sampling a random value, where 10^{-3} and 10^{-2} shall be interpreted as rain rates in mm h^{-1}. The threshold of 10^{-2} mm h^{-1} was determined based on the database. The averaged quantile loss function for all quantile levels was used as the loss function, and output values with no reference data were masked out. Note that to make use of the outputs of the network, we apply $\varphi^{-1}(\hat{z}_\tau) = \mathbf{1}_{\hat{z}_\tau > 10^{-2}} \exp \hat{z}_\tau = \hat{y}_\tau$, which also helps to constrain retrievals with no precipitation.

The use of a logarithmic transform on the reference data to use as training labels is intended to diminish the skewness of the training distribution. This is expected to help the network learn any details that can characterize the larger precipitation values. It is noteworthy that both at training and inference time, the quantile regression property of invariance to monotonic transformations is employed, since both the transform in Equation 2 and its inverse are monotonic.

At training time, we applied data augmentation on the tiles by random flips, quarter-turn rotations, and random crops of 128×128 px^2 (approximately 384×384 km^2 on ground); this tile size was chosen as a compromise between the network architecture and the fraction of reference data in a tile. In addition to the loss function, the Spearman correlation coefficient between $\hat{y}_\tau$ and τ was used to monitor the level of quantile crossing. However, we observed that this measure quickly tended to the best value, although struggling to reach it. As the optimizer method we used Adam, with an initial learning rate of 5×10^{-4}, combined with a cosine annealing scheduler with warm restarts, with an initial period of 20 epochs and a multiplicative factor of 2 for subsequent periods. We set the batch size to 100 samples and determined the state of the network at epoch 1,260, reached after training for more than two training weeks on an NVIDIA A100 40 GiB, as the final state based on the loss value of the validation set, since overfitting of the training set was observed afterward.

4. Evaluation

The purpose of RoA is to obtain probabilistic precipitation estimates shortly after the satellite scan is completed. This section not only contains evaluations against the test set but also a comparison with IMERG and CHIRPS estimates. Neither of these two data sets offer the low latency that RoA can provide, but they take into account information from rain gauges in addition to satellite measurements; thus, they may be considered better suited for certain hydrological applications.

We begin by comparing the diurnal cycles of RoA and IMERG over three different regions, motivated by the single-source input to RoA as well as its better temporal resolution and latency. This comparison assesses how well these two retrievals, which offer diurnal sampling, align in capturing precipitation events, an important aspect for interpreting subdaily variability, especially in regions with pronounced convective activity. Note that GPM 2BCMB data are too sparse to be included in the diurnal cycle analysis, as will be shown. However, we use the mean of the GPM 2BCMB record, excluding timestamps used during training, to frame one case.

Afterward, we present pixelwise statistics from the test set to evaluate how accurately RoA reproduces observed precipitation at the pixel level. This provides a general quantitative assessment of retrieval performance. Statistics from IMERG are included as a baseline for comparison. Additionally, we compare these bulk statistics separately over land and ocean.

To go beyond pointwise accuracy, we compute the effective spatial resolution of the retrievals, a metric which quantifies the agreement in spatial patterns between the retrieval and the reference precipitation fields. This aspect is relevant for applications where the realism of spatial structure matters in addition to bulk statistics. As in the pixelwise analysis, we include statistics from IMERG to frame the RoA performance and separate land from ocean.

Finally, we compare the annual mean precipitation from RoA with estimates from IMERG and CHIRPS, using the GPM 2BCMB record (again excluding training timestamps) as a climatological reference. As in the diurnal cycle comparison, the GPM 2BCMB data remain excessively sparse, and we apply a smoothing technique, as it will be explained. In any case, this allows us to examine potential long-term biases or discrepancies and offers a broader context for understanding RoA's behavior across different timescales.

Where necessary, we use $0.2 \, \text{mm hr}^{-1}$ as a threshold for no rain, based on the required detection limit of the GPM-CO instruments (Kojima et al., 2012; Skofronick-Jackson et al., 2017). A few unreasonable values were detected in GPM 2BCMB and RoA data: we discarded values exceeding $200 \, \text{mm hr}^{-1}$. The fraction of data points over land is about 50%, which should be kept in mind when interpreting the results.

4.1. Diurnal Cycles

We look at diurnal cycles using RoA and IMERG estimates: the highest temporal resolution of the CHIRPS data set is a daily estimate, and consequently, it is unfeasible to obtain such information from this data set. The cycles are computed using retrievals for all of 2023, where RoA estimates are collocated on the coarser IMERG grid, using the nearest neighbor in space and averaging in time. In addition, we use the GPM 2BCMB data record, excluding observations used in the training phase, to elaborate on one case.

Figure 2 shows diurnal cycles for three areas: two land areas of dry and wet characteristics and an ocean area expected to be wet (based on Figure S2 in Supporting Information S1). Starting with the most dry region (Figure 2a), it stands out the magnitude and diurnal cycle of the RoA estimate with respect to IMERG from June through November. However, IMERG also exhibits a diurnal cycle for these months, albeit lagging a few hours that of RoA. The period of June through August (JJA) exhibits the largest rain rate. Comparing with the expected diurnal cycle, based on approximately 7 years of GPM 2BCMB data, we observe that RoA shows a better agreement in both magnitude and time: GPM 2BCMB and RoA alike place a precipitation peak around 16:00 local solar time (LST), despite temporal information not being provided to the retrieval. In other words, the larger RoA estimate is reasonable based on the GPM 2BCMB average. As hinted by the JJA period in Figure 2a, the GPM 2BCMB record does not provide enough data for stable statistics, even when averaged over 3 months of multiple years. Consequently, more diurnal cycles of GPM 2BCMB data are not shown in Figure 2.

Following with the wet ocean area from Figure 2b, the patterns match relatively well with similar magnitudes, with a clear dry period. Finally, the rainforest area presented in Figure 2c shows a pronounced diurnal cycle.
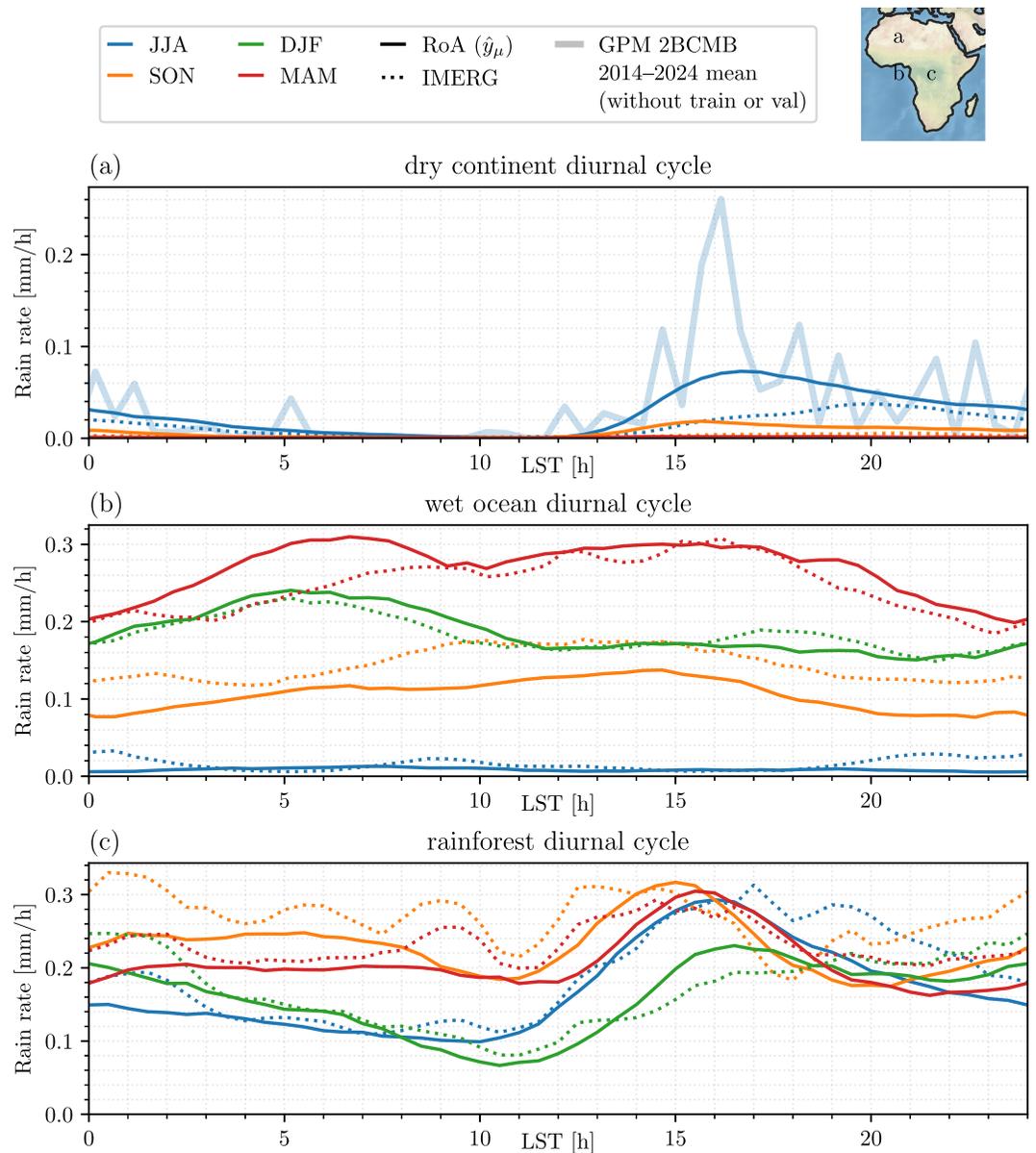
Figure 2. Diurnal cycles for 2023 for IMERG and RoA for three different areas, indicated in the legend (map based on Natural Earth data) and aggregated by 3-month periods (for instance, June, July, and August: JJA). The areas are $5° \times 5°$ large, with centers at (a) 22.5°N, 2.5°E, (b) 0°N, 2.5°E, and (c) 0°N, 22.5°E. All points in the same 30-min time stratum and area are averaged to obtain a single value. The colors represent groups of three consecutive months, while the line style indicates the retrieval method. Shown in a translucent line in (a) is the GPM 2BCMB estimate from 2014 through 2024, employing the same bins, and excluding data from the training and validation data sets.

Although in this case the magnitudes are also similar, IMERG shows small oscillations within the diurnal cycle, clearly visible for two noncontiguous periods, which are not observed with the RoA estimates.
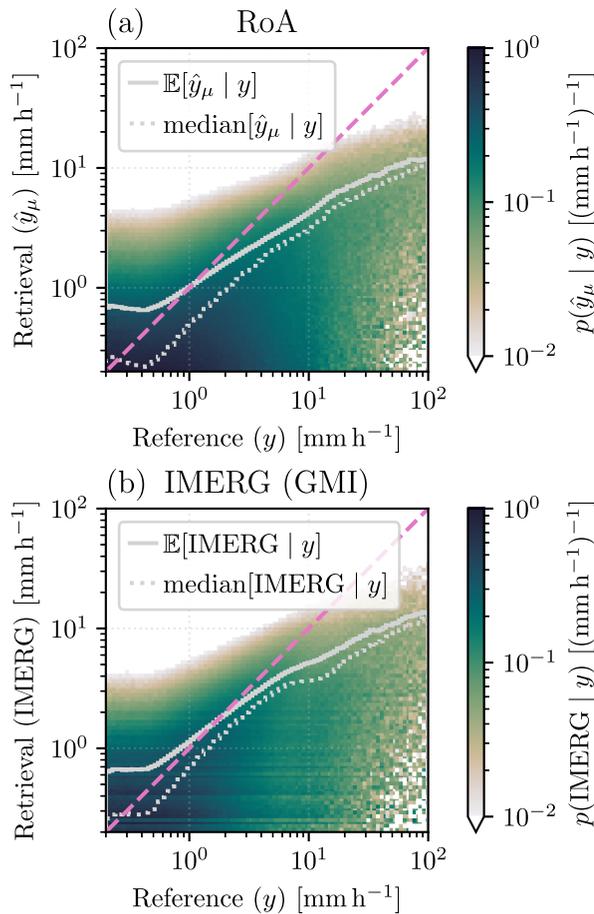
4.2. Pixelwise Test Set Statistics

An evaluation against the reference is essential to characterize the retrieval skills. Notwithstanding that the reference data have some uncertainty, we consider it a ground truth and perform an evaluation with the test set.

Examining the probability distribution function p of $\hat{y}_\mu$ and reference values in Figure 3, we observe that the RoA retrieval $\hat{y}_\mu$ closely follows the reference until about 10 mm hr^{-1} and progressively deviates from the reference
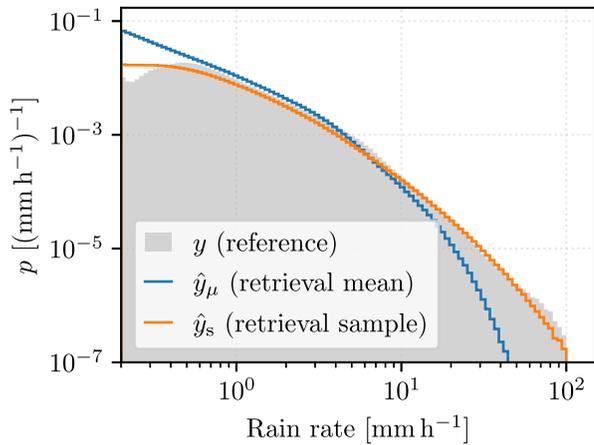
Figure 3. Probability distribution functions on log-spaced bins for the test set. The probability not displayed is higher than 96% for all three cases.

distribution for larger values, which are less frequent. Nevertheless, we argue that these extreme values of the reference are well captured by the RoA retrieval distribution: if $\hat{y}_\mu$ is replaced by a sample from the retrieval distribution as the point estimate, denoted as $\hat{y}_s$, then the distribution closely matches the reference even for the less frequent values. On the other end, the distributions deviate when approaching the threshold of 0.2 mm hr^{-1}.

To reveal more information about the retrieval accuracy, we inspect the distributions of retrieved rain rates conditioned on the reference, presented in Figure 4a. It is naive to expect no spread in the plot: in addition to uncertainties from the data collocation and the reference data, the thermal IR SEVIRI radiances are not ideal for sounding through thick clouds due to attenuation. This spread is clearly visible. Lack of familiarity with the retrieval problem can lead to perceive that the RoA performance is poor. However, when comparing with the analogous plot for IMERG (GMI) data in Figure 4b, which incorporates retrievals from microwave instruments and represents the best-case IMERG estimates (Section 2.2.1), we observe a similar performance. In any case, the retrievals show a tendency to be close to the identity line: for slight precipitation, the retrieved $\hat{y}_\mu$ tends to align with the identity line despite over-estimation on average, while underestimation is observed for larger reference values, a situation also observed for IMERG (GMI).

Figure 4. Scatterplot of the RoA retrieval values $\hat{y}_\mu$ versus the reference y, summarized as conditional probabilities in (a) for the test set. Panel (b) shows the analogous plot when using IMERG (GMI) as the retrieval value.

It is often desirable to have summary statistics such as the values presented in Table 3, where we also provide the analogous statistics computed for IMERG (GMI) and the IMERG IR-only estimates. Nonetheless, it is arguable that summary statistics provide overly general information on the skill of the retrieval. Therefore, we also provide the summary statistics based on the rain rate coding from the World Meteorological Organization (WMO, 2023), although we define no rain as precipitation below 0.2 mm hr^{-1} for our data, the same value used for the detection threshold.

Furthermore, to quantify the probabilistic aspects of the RoA retrievals, we also show in Table 3 the frequency in which the reference y is captured between the retrieval quantiles $\hat{y}_\tau$ at levels $\tau = 16\%$ and 84%; the average continuous ranked probability score, a generalization of the mean absolute error (MAE) to probability distributions; and the average difference between quantile levels 16% and 84%. Not shown in the table is that the RoA retrieval distributions show $\hat{y}_\mu \geq \hat{y}_{50\%}$ when there is precipitation, that is, they are right-tailed, which is expected in this context: the retrieval distribution is different from a normal distribution. Figure 5 shows the reliability diagrams, which quantify how often an event occurs for a given predicted probability. For this reason, we present these diagrams for different precipitation exceedance thresholds and separate them by the surface type. The curves are close to the identity line in most cases, indicating that the retrieval distributions are, on average, well calibrated and provide reliable inputs for threshold-based decision-making.

The summary statistics in Table 3 reveal several key findings. While IMERG (GMI) demonstrates superior bulk statistics overall, RoA exhibits comparable performance. Notably, when contextualized with IMERG IR-only estimates, the similarities between IMERG (GMI) and RoA become even more apparent: RoA performs significantly better than the IMERG IR-only estimates. These results can be intuited from a single retrieval example, presented in the right column of Figure 8. The influence of the surface type on these statistics is discussed later.

Table 3
*Summary Statistics Computed With the Test Set*

Intensity [mm h^{-1}]	No rain <0.2	Slight [0.2, 2.5)	Moderate [2.5, 10)	Heavy [10, 50)	Violent ≥50	All
Samples	97.4%	1.9%	0.6%	0.1%	0.01%	>10^8
RMSE RoA	0.18	**1.58**	**3.75**	14.29	65.33	0.79
IMERG (GMI)	**0.18**	1.60	3.86	**14.18**	**64.23**	0.79
IMERG IR-only	0.40	2.12	4.61	16.17	68.75	0.96
Bias	0.02	**0.02**	−2.07	−11.45	−61.36	−0.012
	0.02	0.14	**−1.30**	**−10.83**	**−60.12**	**−0.004**
	0.04	−0.06	−2.58	−13.48	−64.99	0.007
MAE	0.02	0.95	3.04	12.00	61.36	0.07
	0.02	**0.95**	**2.92**	**11.79**	**60.12**	**0.07**
	0.04	1.24	3.82	14.01	64.99	0.11
Correlation	**0.10**	0.22	0.20	**0.23**	0.08	0.54
	0.09	**0.28**	**0.22**	0.21	**0.11**	**0.56**
	0.07	0.12	0.10	0.15	0.07	0.32
$\mathbb{E}\left[y \in \left[\hat{y}_{16\%}, \hat{y}_{84\%}\right]\right]$ [%]	99.0	44.4	35.6	20.2	1.1	97.4
$\mathbb{E}\left[\hat{y}_{84\%} - \hat{y}_{16\%}\right]$	0.03	1.62	4.05	9.39	17.56	0.09
Average CRPS	0.01	0.64	2.52	10.08	45.92	0.05

Note. The statistics are reported for RoA $\hat{y}_\mu$, IMERG (GMI), and IMERG IR-only precipitation estimates, respectively. The best statistics are highlighted in bold, where decimals were used to resolve ties. RMSE: root mean squared error; MAE: mean absolute error; and CRPS: continuous ranked probability score. The correlation is computed using the Pearson correlation coefficient. The last three rows only apply to RoA. Units for RMSE, Bias, MAE, quantile spread, and CRPS: mm h^{-1}. Strata defined by the reference values y in the test set. About 50% of the data points are over ocean.

An analysis of Table 3 draws attention to the average size of the errors for no and slight precipitation, with root mean squared error (RMSE) values of 0.18 and 1.58 mm hr^{-1}, respectively, with similar values for IMERG (GMI). This can indicate difficulties in clearly delimiting a precipitation field, which will be further discussed in Section 5.1. Based on the last two rows of Table 3, the RoA retrieval distributions capture, to a certain extent, the reference value for cases below heavy precipitation within 68% of the central part of the distribution, which is moderately narrow, in relative terms. This last observation is especially useful for classifying precipitation
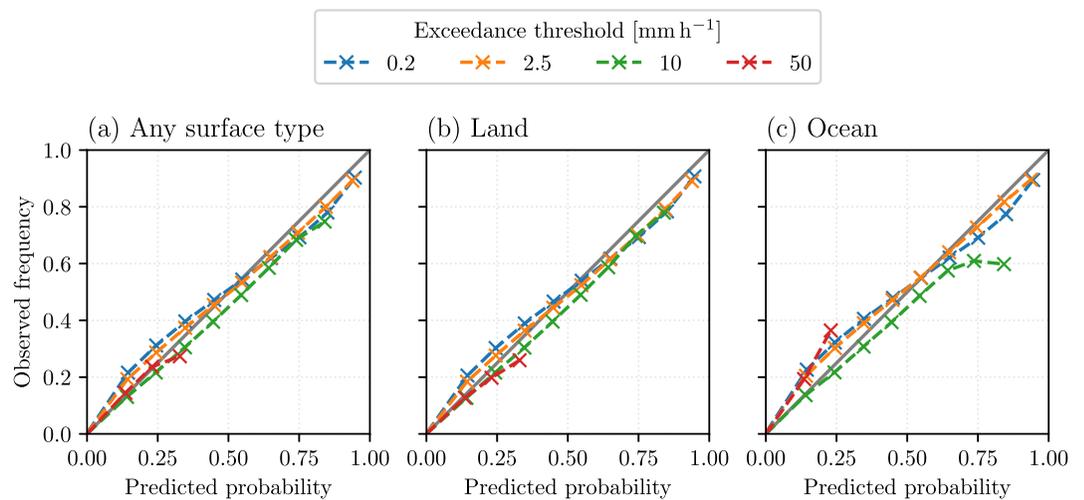
Figure 5. Reliability diagrams for all RoA retrieval distributions in the test set, using the precipitation exceedance thresholds listed in the legend and separated by the surface type. Each marker in the plot represents averaged data in 10 equally spaced bins; bins containing less than 200 data points were excluded to ensure reliable statistics.
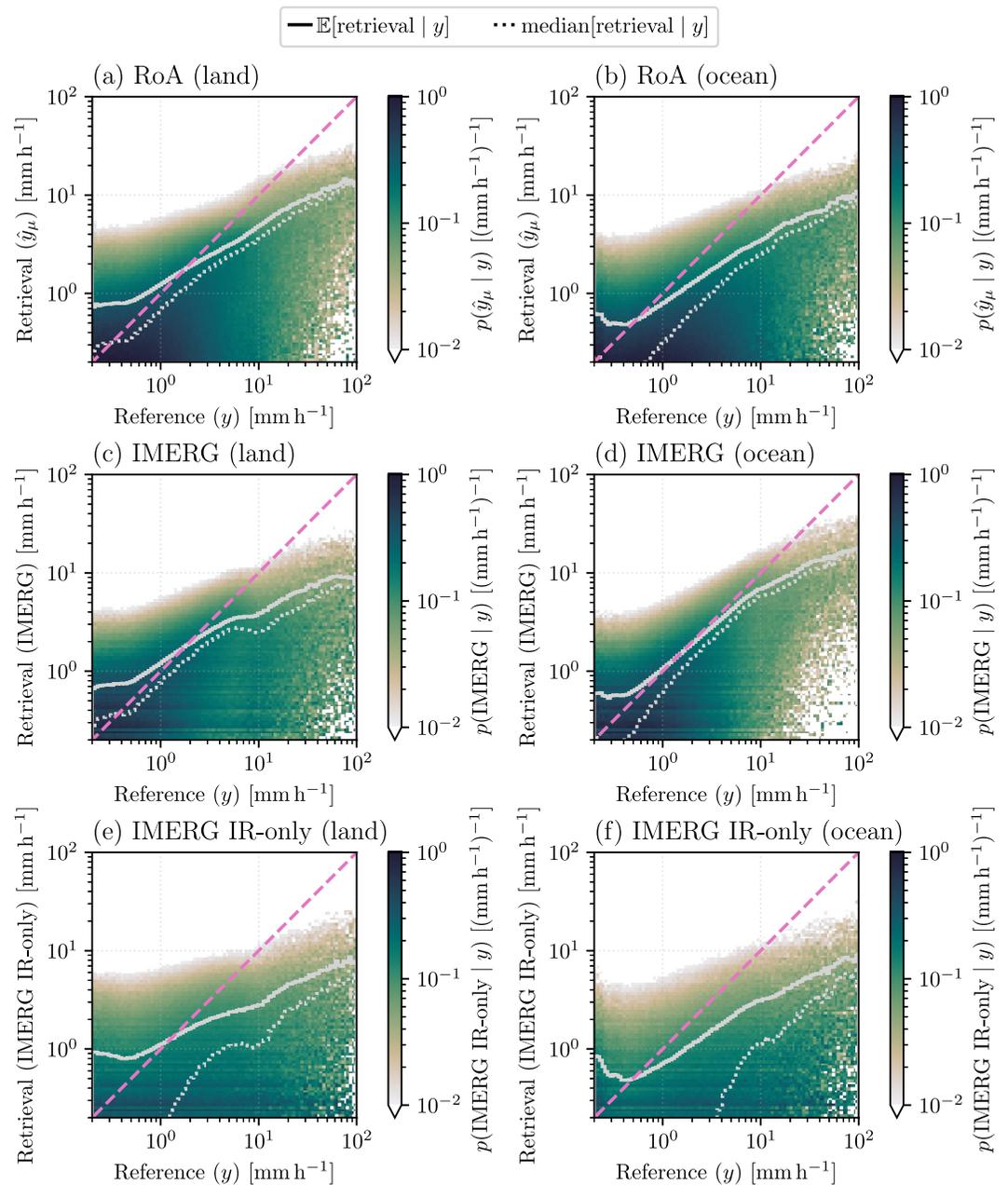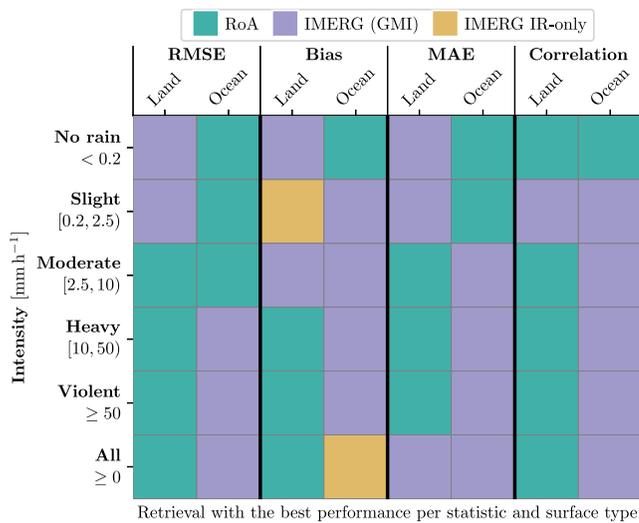
Figure 6. As in Figure 4, but separated by the surface type and including IMERG IR-only.

regimes, that is, information from the retrieval distribution can be used to provide uncertainty for different precipitation ranges.

Figure 6 shows how the distributions in Figure 4 vary with the surface type. RoA remains relatively stable regardless of the surface type, albeit it performs better over land than ocean for this test data. IMERG (GMI), however, varies considerably with the surface type. This stability in RoA is unsurprising, as its information primarily derives from cloud top properties. A similar pattern is seen in the IMERG IR-only retrievals: although their overall performance is significantly worse, they are less sensitive to the surface type.

The accuracy of microwave-based precipitation retrievals depends significantly on the surface type. This stems from the fundamental principles of microwave remote sensing: microwaves can penetrate clouds and interact with the surface below. Over oceans, the surface is radiometrically cold in the microwave spectrum and has an

Figure 7. Scorecard summarizing Tables S1 and S2 in Supporting Information S1.

approximately uniform emissivity, making it easier to retrieve precipitation. Over land, however, the surface is more heterogeneous and radiometrically complex, which interferes with the discerning precipitation signal from the radiative background. In other words, a rain cell appears as a deviation in an otherwise relatively smooth radiance field over ocean, while over land complex horizontal structures are present already during rain-free conditions, making the retrieval significantly more difficult over land. As a result, the accuracy of ideal microwave retrievals should surpass IR-only methods such as RoA over ocean surfaces, while it is reasonable that they worsen over land. This agrees with the IMERG (GMI) variability by the surface type observed in Figures 6c and 6d and aligns with previous IMERG analyses (Derin et al., 2022).

These remarks are also consistent with the scorecard in Figure 7, which summarizes the bulk statistics from Table 3 calculated by the surface type. The data suggest a consistent trend: IMERG (GMI) tends to outperform RoA over ocean, whereas RoA performs better over land surfaces. When carefully inspecting the actual values, provided in Tables S1 and S2 in Supporting Information S1, the advantage of IMERG (GMI) over ocean is more pronounced than any advantage of RoA over land. Given that ocean and land surfaces each comprise approximately 50% of the data points, this positively impacts the statistics of IMERG (GMI) when surface types are not separated, as in Table 3.

Figure 7 also helps to visualize an answer to this important question: What is the performance of RoA when evaluated on data from outside the training period and acquired using a different Meteosat platform? As explained in Section 2.1.3, the test set consists of data collected over a year after the training period, primarily from the MSG-3 platform, whereas training used data from MSG-4. The frequency with which RoA appears in the scorecard highlights its robustness in generalizing to unseen scenarios.

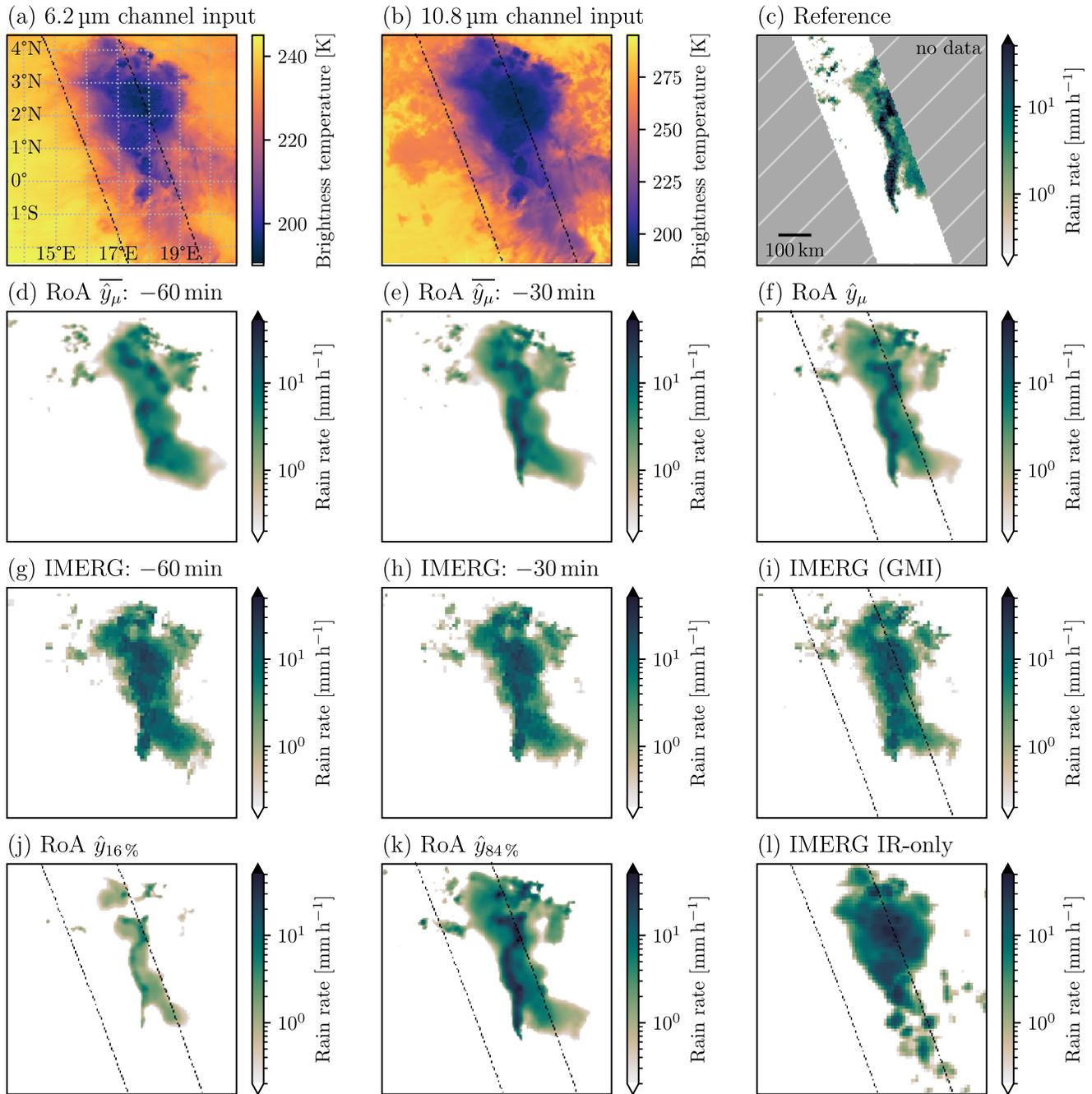## 4.3. Spectral Analysis for Spatial Resolution

The retrievals can suffer a blur effect when compared to the reference, both inside the precipitation field and in its delimitation. This can be appreciated in Figure 8 and is discussed later in Section 5.1. In any case, this smearing of the retrievals challenges point-to-point comparisons to the reference data, and techniques that take into account geometrical properties at different scales can be relevant. That is, conventional pixelwise metrics such as the previous sections do not necessarily reflect the performance of precipitation retrieval methods in capturing spatial structures: for example, how the convective squall line present in Figure 8c is resolved in and Figures 8f and 8i or Figure 8l. Moreover, these pixelwise metrics overlook the spatial correlation of precipitation, failing to provide an indication of the accuracy in capturing precipitation patterns. We consider appropriate to determine an "effective (spatial) resolution," as a means to quantify the finest spatial details that the precipitation retrieval can reliably reproduce when compared to a reference.

Guilloteau et al. (2021) present and apply a methodology based on the Fourier spectral analysis of precipitation fields to compare different products. Their method consists of computing the Fourier spectral coherence between a reference and an estimate. This metric can be interpreted as the correlation in the spatial frequency domain between the Fourier power spectral densities (PSDs) of the reference and the estimate. In the general case, the spectral coherence is a complex number, whose modulus takes values between 0 and 1: the greater the spectral coherence, the more the patterns of the fields match at the given spatial sampling frequency. That is, the spectral coherence can be used to compare the patterns at different scales. In addition, Guilloteau et al. (2021) argue for using the value of $1/\sqrt{2}$ as a threshold to determine the effective resolution relative to the reference: the effective resolution then corresponds to half the shortest wavelength (inverse of the spatial frequency k) for which the spectral coherence reaches the threshold (Pfreundschuh et al., 2024).

Consequently, we followed the formulas in Guilloteau et al. (2021) to compute the spectral coherence. In particular, we computed the Fourier PSDs by averaging the PSDs obtained from two-dimensional fast Fourier transforms on more than 5×10^4 randomly sampled, nonoverlapping windows of 50×50 px^2; the window size is a trade-off

Figure 8. (a) and (b) are two of the eight input features to the network, for which RoA retrieves quantiles, such as (j) and (k), which are used to compute the point estimate $\hat{y}_\mu$ presented in (f); all 99 retrieval quantiles are shown in Figure S5 in Supporting Information S1. Panel (c) shows the reference data and (i) the corresponding IMERG (GMI) estimate. Panel (l) shows the IMERG IR-only precipitation estimate. The dashed lines indicate the limits of the reference data. Panels (d) and (e) and (g) and (h) show two previous retrievals for RoA and IMERG, respectively; the notation $\overline{\hat{y}_\mu}$ highlights that the RoA retrievals were averaged in time to match the IMERG temporal resolution. The finer RoA time series is given in Figure S4 in Supporting Information S1. This example is over a land region, with coordinates indicated in (a).

imposed by the relatively narrow swath of the reference. We also ensured that all RoA estimates in a window originated from the same observation. The sampling frequencies in both directions are $k_x = k_y = 3^{-1}$ km^{-1}, the inverse of the SEVIRI data resolution. We did not observe any strong anisotropy of the PSDs, besides along $k_x = k_y = 0$, which we could associate with unavoidable discontinuities when computing the Fourier transform. Finally, we summarized the information by computing the modulus of the azimuthally averaged spectral coherence. The actual implementation is given by Amell and Hee (2025).
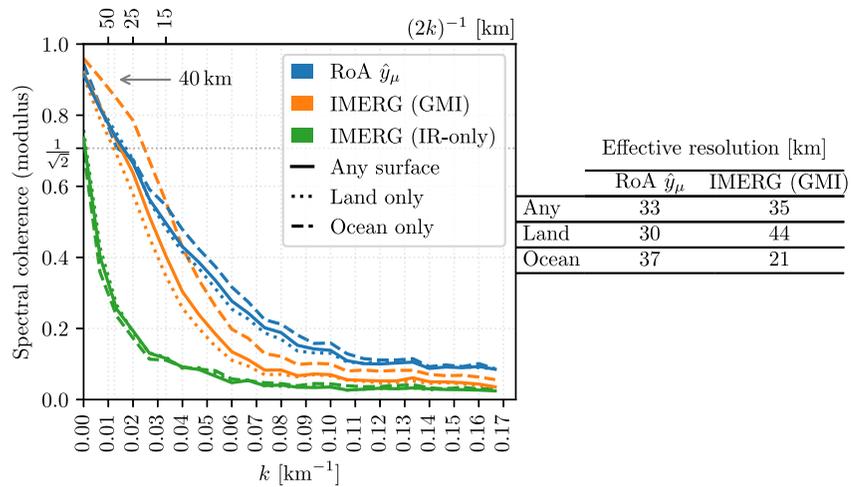
Figure 9. Azimuthal average of the 2D spectral coherence of RoA $\hat{y}_\mu$, IMERG (GMI), or the IMERG IR-only variable to the reference data in the test set. The curves stop at the Nyquist frequency. The scale $(2k)^{-1}$ for which the spectral coherence matches $1/\sqrt{2}$ is the effective resolution (Pfreundschuh et al., 2024). The exact value is computed by linearly interpolating the spectral coherence and the $(2k)^{-1}$ values.

We also computed the spectral coherence of the IMERG (GMI) and IMERG IR-only estimates to the reference in the test set, employing the same method, and with all IMERG estimates in each window originating from the same timestamp. The resulting spectral coherence curves provide bounds for the IMERG spectral coherence. In particular, the IMERG (GMI) curve should be regarded as the best-case spectral coherence, since GPROF GMI retrievals are systematically used in these IMERG collocations (Section 2.2.1).

Figure 9 indicates an effective resolution around 35 km for both RoA and IMERG (GMI) regardless of the surface type. More precisely, a linear interpolation reveals that the overall effective resolution of RoA is 33 km, while IMERG (GMI) stays at 35 km. As we exposed in the pixelwise statistics, IMERG (GMI) retrievals are severely dependent on the surface type: while RoA maintains a stable effective resolution when separating surface types (land: 30 km and ocean: 37 km), IMERG (GMI) achieves a much better 21 km over ocean but deteriorates to 44 km over land. These findings align with Pfreundschuh et al. (2024), who analyzed GPROF GMI retrievals and obtained effective resolutions of 46 km over the conterminous United States (CONUS) and 18 km over the tropical Pacific, while also demonstrating that neural network versions of GPROF yield improved resolutions.

4.4. Annual Mean Precipitation

Using the full year 2023, we compare RoA, IMERG, and CHIRPS retrievals for precipitation estimates over African land. Given that the retrievals are provided on different grids, we resample the data to the coarser grid, which corresponds to IMERG, using the nearest neighbor algorithm. The annual means are then obtained by averaging all values binned in one spatial grid cell. For consistency with RoA and IMERG, CHIRPS daily accumulations are converted to hourly rates by dividing by 24; since this conversion is applied only to the overall annual mean and no subdaily variability is analyzed, this does not introduce any methodological bias.

We also integrate the GPM 2BCMB mean in the comparison. However, due to its sparsity, we computed this statistic with the record from 2014 through 2024, but excluded data used in the training of RoA. The GPM 2BCMB data were first averaged in the IMERG grid cells and subsequently smoothed with a moving average of 5×5 px^2. We indicate the resulting value as $\overline{\text{GPM 2BCMB}}$ and use it as a climatological reference to frame the analysis.

Figure 10 shows the statistics for 2023. It can be appreciated that the three data sets compare similarly on a general level, particularly in the wet tropical areas. However, neither of the three estimates completely agree on smaller scales: there are noticeable differences between RoA and the gauge-adjusted data sets, although these also show differences between themselves (Figure 10c). Particularly for IMERG, it stands out its high precipitation estimates for Lake Victoria, which is neither observed in RoA or CHIRPS. When analyzing the panels with
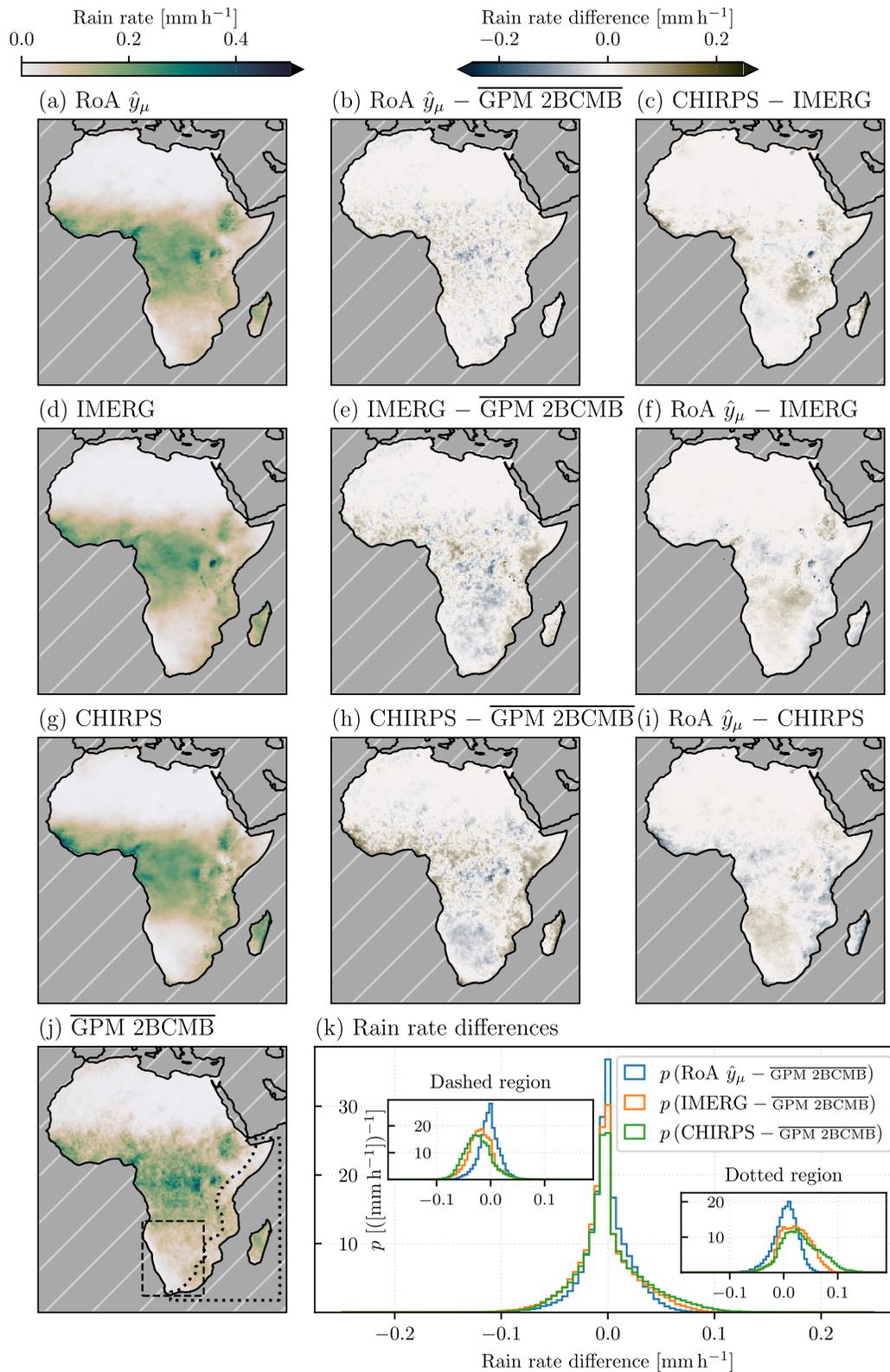
Figure 10. Mean precipitation values for year 2023 in (a), (d), and (g) and their differences in (c), (f), and (i), compared on the IMERG grid with a nearest neighbor resampling approach and restricted to African land. Panel (j) shows the GPM 2BCMB mean, as described in the text, and panels (b), (e), and (h) show the respective differences. Panel (k) presents the area-weighted distributions of the differences between the central columns, and the two subplots within panel (k) show the analogous distributions computed only over the regions indicated in panel (j). All values in each spatial grid cell are averaged to obtain the annual mean.
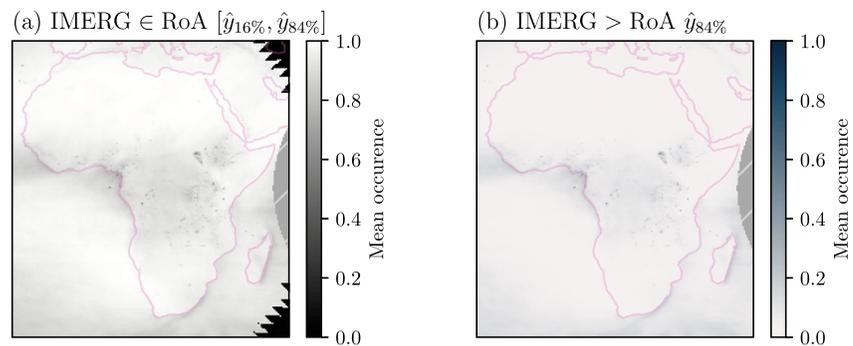
Figure 11. Frequency in which the IMERG half-hourly estimate is in the interval given by the RoA quantiles $\hat{y}_{16\%}$ and $\hat{y}_{84\%}$ during year 2023 in (a) and when the IMERG estimates are greater than $\hat{y}_{84\%}$ in (b). The same data collocation as in Figure 10 was used. Text S2 in Supporting Information S1 provides the explanation for the visible artifacts at the corners.

differences to the GPM 2BCMB mean (b, e, and h), we observe clear structures in IMERG and CHIRPS, but these structures are less evident in RoA. These panels also highlight that RoA estimates are more consistent with the GPM 2BCMB mean toward the southern Africa and the Indian Ocean coast, where there are clear disagreements with the other two data sets (panels f and i). Consequently, this is reflected in Figure 10k, where the differences between the RoA mean and $\overline{\text{GPM 2BCMB}}$ are closer to zero than the respective differences of the two other data sets.

Not shown in Figure 10, but can be appreciated in an analogous map including ocean in Supporting Information S1 (Figure S2), is that there are significant differences between IMERG and RoA over the oceans, where IMERG can be more accurate. The largest visible disagreement also coincides with the southernmost latitudes, which the training data did not include (Section 2.1.3). Overall, IMERG estimates larger precipitation values over water, toward the western Africa coasts, around the Mozambique channel, and toward the Southern Ocean. A similar pattern is also observed with $\overline{\text{GPM 2BCMB}}$.

A feature of RoA not offered by IMERG or CHIRPS is the possibility to provide practical uncertainty estimates. That is, RoA can provide precipitation quantiles to describe the rain rate probability. Figure 11a reveals that the RoA retrieval distribution tends to contain the IMERG estimate between the quantiles at levels 16% and 84% and that these half-hourly IMERG estimates are seldom larger than the RoA 84% in regions of high precipitation. That the latter event occurs reveals that the RoA retrieval distributions are not excessively wide. The opposite situation, where IMERG is less than the RoA 16%, is virtually negligible.

5. Discussion

5.1. Retrieval Skills

Figures 6 and 7 and the associated tables demonstrated the accuracy of RoA retrievals when compared to the best-case IMERG estimates across different metrics and rain rates, with RoA estimates particularly skilled over land and showing clear advantages over the IMERG IR-only estimates. This has two main implications. On the one hand, retrieval algorithms that merge data from different instruments, such as IMERG, can be greatly improved by making better use of the geostationary measurements. On the other hand, accurate retrievals can be obtained with a significantly better latency.

These improvements are observed despite a challenge in the RoA retrievals: IR attenuation caused by clouds. The RoA neural network needs to rely on contextual information to accurately determine precipitation estimates, given that they mostly provide cloud top properties. This was observed in the scatterplot from Figure 4, where large reference values are underestimated. Retrieval accuracy can likely increase by incorporating the time dimension into the network, that is, employing information from more than one geostationary image. This would enable exploiting precipitation dynamics and is a possible future line of research.

Pixelwise metrics for evaluating the retrieval accuracy are well established, but they ignore spatial and temporal correlations of the precipitation. For instance, qualitatively comparing the retrieval example in Figure 8f (RoA) and Figure 8i (IMERG (GMI)) with the reference in Figure 8c, one can observe that RoA captures features that

IMERG (GMI) does not, rendering RoA closer to the reference. In particular, RoA performs well in resolving the convective squall line, especially compared to IMERG (GMI) and even more so to the IMERG IR-only estimate. Moreover, when looking at two previous retrievals, the formation of the convective squall line is more noticeable in RoA (Figures 8d and 8e) than in IMERG (Figures 8g and 8h); note that there is no reference for the time series.

Additional details for Figure 8 are provided in Movie S1, which presents a longer time series and displays the temporal evolution of IMERG IR-only estimates. Movie S1 also shows the time series of the RoA quantiles for an arbitrary pixel, emphasizing that RoA provides a full probabilistic description of the retrieval. Overall, a qualitative assessment shows better agreement between RoA and IMERG than between IMERG and its IR-only estimates. Beyond the higher agreement in precipitation details and structures, the movie reveals sparse, small precipitation cells in IMERG IR-only that appear inconsistently in IMERG and are absent in RoA, thus resembling a contamination effect in IMERG.

Continuing with the example in Figure 8, if the RMSE, bias, MAE, and correlation metrics for RoA and IMERG (GMI) are computed, then one could argue that IMERG (GMI) approximates better this case than RoA (RMSE: 7.89 mm hr^{-1}, 8.14 mm hr^{-1}; bias: -0.48 mm hr^{-1}, -0.98 mm hr^{-1}; MAE: 2.27 mm hr^{-1}, 2.21 mm hr^{-1}; and correlation: 0.62, 0.61; values for IMERG (GMI) and RoA, respectively). To summarize, we consider there is the need to establish metrics that also take into account the correlations in the precipitation fields at different scales. In this work, we have used the methodology of Guilloteau et al. (2021) to obtain an effective resolution, since it offers a compelling rationale, and, to our knowledge, is the most recent work addressing this challenge.

5.1.1. Effective Resolution

RoA was found to have an overall effective resolution of 33 km, performing better over land (30 km) than over ocean (37 km). This compares favorably to the best-case IMERG resolution over land, which is 44 km. In practice, IMERG Final run estimates likely experience an even worse effective resolution: when comparing IMERG with retrievals that do not constantly incorporate GMI measurements, its effective resolution was found closer to 80 km over CONUS (Guilloteau & Foufoula-Georgiou, 2020). Thus, we consider that the IR-only RoA retrievals resolve precipitation structures significantly better than IMERG over land. Still, there can exist room for improvements with a machine learning approach.

The backbone of RoA is an artificial neural network. In the field of machine learning, it is known that pixelwise p-norm training losses, such as the RMSE loss ($p = 2$), suffer from a blur effect, which influences the effective resolution: the loss enforces per-pixel accuracy and does not prioritize capturing structural details. As a consequence, in complex scenes or boundaries, the model will tend to predict a generalized mean of possible outputs, smoothing the limits. Apart from the quantile level $\tau = 0.5$ ($p = 1$), the quantile loss is not exactly a p-norm loss, but it shares similarities, therefore leading to a smearing effect, which can be appreciated in Figure 8f.

Consequently, future work to improve RoA retrievals can consider adding regularization terms to the training loss to mitigate this smearing of the retrievals. Two examples of regularization terms are an adversarial loss and a structural loss. The former example is obtained through an independent network, the discriminator, which learns to penalize predicted precipitation fields, which do not resemble the reference during the training phase. There are multiple examples for structural losses, but one can think of transforming the precipitation field into a binary value of no rain and rain, and then computing a similarity coefficient of the binary precipitation fields. Nonetheless, such regularization terms involve increasing the number of hyperparameters, beginning with their importance in the total training loss function. Moreover, it is not trivial how to preserve the probabilistic description of quantile regression with such regularization schemes, which may also compromise pixel accuracy. Accordingly, further experimentation is warranted to determine an optimal balance and to understand its implications.

The ideas above aim to improve the sharpness of the precipitation fields, that is, to enhance the effective resolution of the RoA retrievals. Using Figures 8c, 8f, and 8i, and l to help the discussion, we can observe that there can be overestimation beyond the reference precipitation field, which are likely stratiform parts of the cloud field. The limited information from the geostationary radiances challenges RoA to obtain a sharper delineation of the precipitation field. Nonetheless, this difficulty is reflected in the uncertainties: Figure 8j shows the retrieval quantile at the 16% level, where this overestimation is not visible, yet the convective squall line is already detected.

5.2. Probabilistic Aspects of the Retrievals

The RoA retrievals provide uncertainty at different levels for each individual estimate. This is illustrated in Figures 8j and 8k, where the RoA quantiles at levels $\tau =$16% and 84% are depicted; all 99 percentiles retrieved are given in Figure S5 in Supporting Information S1. These quantiles can be used to compute probabilities of precipitation at different levels.

One limitation of the quantile regression employed is that these uncertainty estimates are marginal distributions. That is, there is no information available on the nature of the retrieval uncertainties from the neighboring locations for each estimate. If a sample $\hat{y}_s$ was used instead of the expected value of the distribution $\hat{y}_\mu$, as in Figure 3, then a precipitation field resembles a signal with noticeable noise (Figure S3 in Supporting Information S1). This is a consequence that there is no joint distribution between the samples: the RoA retrieval distributions are marginals; they do not include correlations. Obtaining random precipitation fields with realistic correlations can be advantageous to evaluate different possible scenarios, as well as to obtain an uncertainty for area-averaged values.

Techniques such as vector quantile regression, which can be considered an extension of quantile regression, have emerged and can provide joint estimates. Nonetheless, reference data such as the GPM 2BCMB product, which incorporates an uncertainty estimate, does not provide correlations with the nearby pixels either. This also has the consequence that sampling a random precipitation field from the reference data would result in a pattern that resembles a signal with noise. For this reason, the reference data have been considered as an absolute ground truth, although GPM 2BCMB is a retrieval itself, which means it also suffers an irreducible uncertainty, in addition to incorporating microphysical assumptions. Furthermore, the uncertainty in the GPM 2BCMB estimates is given by the standard deviation of the optimal estimation ensemble used (Grecu et al., 2016). Therefore, if a machine learning were trained to minimize the error at different uncertainty levels of the reference, then, in the optimal case, the retrievals should resemble bell-shaped curves: this is unlikely to be satisfied for precipitation retrievals using IR information due to attenuation.

5.3. Next-Generation Data Sources

As mentioned in the introduction, RoA retrievals aim to obtain precipitation estimates shortly after the SEVIRI scan is completed and disseminated. That is, these retrievals are best suited for nowcasting applications, with long-term hydrological applications playing a secondary but relevant role. RoA is currently dependent on the MSG satellites, but the first satellite of the Meteosat Third Generation imaging series (MTG-I, Holmlund et al., 2021), which features a similar imager to SEVIRI but offers finer spatiotemporal resolutions, was launched at the end of 2022 (CEOS, 2024; WMO, 2024).

Currently, the end of life of MSG at 0° longitude is in 5 years (CEOS, 2024; WMO, 2024), which renders these RoA retrievals applicable for almost half a decade, regardless of the state of the MTG mission. However, it is sensible to consider updating RoA to support MTG-I, given that it succeeds MSG and offers a finer time resolution.

A pragmatic approach to use MTG-I imagery with RoA is to resample MTG-I imagery to match the resolution of MSG, given that the thermal IR channels are virtually equivalent, and execute the retrieval with this resampled data. A more involved approach is to train a network with MTG-I imagery. In this case, there are two main remarks. First, even if the new generation has a finer spatial resolution, the reference data are coarser, which must be borne in mind when constructing collocations and evaluating the retrieval skill. Second, at the time of writing, operational MTG-I imagery spans less than 1 year. This requires a waiting period to obtain many collocations with MTG, which should be feasible given that the GPM-CO satellite, used to obtain the reference data, is expected to last at least until 2030 (WMO, 2024). Consequently, for an immediate update, either the pragmatic approach or a fine-tuning approach with resampled MSG data to the MTG-I resolution are reasonable. The latter approach could also benefit from augmenting the training data set with collocations since 2014 to enhance the retrieval robustness.

The new MTG-I also creates new opportunities with its Lightning Imager instrument (Holmlund et al., 2021), whose data products are already being disseminated. As the name suggests, this instrument detects any of the cloud-to-cloud, intracloud, and cloud-to-ground lightnings. Precipitation and lightning are closely related (Soula, 2009); therefore, the use of such information in a retrieval is expected to be beneficial.

Finally, this new Meteosat generation does not only encompass a successor to the MSG imager and a lightning imager but also the first hyperspectral IR instrument on a geostationary platform (Holmlund et al., 2021). This instrument is on the sounder series (MTG-S), whose first satellite has recently been launched. Although it has a larger temporal resolution, horizontally and vertically resolved atmospheric information from this instrument can further improve the accuracy of satellite-based precipitation estimates. In any case, it will be some time before any public operational MTG-S data become available.

6. Using RoA

The RoA retrieval algorithm is publicly available as a small Python package (Amell, 2025), which contains an illustrative example for executing retrievals for all of continental Africa. MSG data can be acquired using the EUMETSAT near-real-time primary dissemination mechanism, EUMETCast. However, most users will need to download MSG data from the EUMETSAT Data Store. In this case, the EUMETSAT Data Access Client (EUMDAC, EUMETSAT, 2024) can be used to search and download the data. A RoA retrieval pipeline, which acquires data from the EUMETSAT Data Store can be designed as follows:

(a) 6.7 s: download an MSG observation.
(b) 15.1 s: prepare MSG observation; read binary file, crop to ROI, standardize values, and create input tiles.
(c) 23.7 s: run retrieval, obtaining a vector of $\hat{y}_\tau$'s for each pixel for each tile.
(d) 4.0 s: assemble tiles.
(e) 16.7 s: compute $\hat{y}_\mu$ or the relevant statistic for each pixel.
(f) 0.3 s: assign coordinates to each pixel and save to disk.

The numbers indicate the elapsed time for each step in a Jupyter Notebook, available as part of the companion code (Amell, 2025), executed with a free cloud-based GPU service (Google Colab; T4 GPU runtime with ~16 GiB RAM and VRAM, and 2 CPUs at 2 GHz). Using a CPU instead of a GPU consumes significantly more time. In our experience, given that MSG data are available on the EUMETSAT Data Store shortly after SEVIRI finishes its acquisition, we argue that the retrievals can be obtained in a near-real-time manner, particularly when there is access to a GPU. Therefore, the main limitation for near-real-time retrievals is the process of downloading MSG data, which will depend on the Internet setup of the user and the EUMETSAT Data Store.

The pipeline presented enables obtaining a full probabilistic RoA retrieval and manipulating data programmatically for any researcher. Nevertheless, at the time of writing, there are already public early RoA implementations at https://africa-hydrology.ceh.ac.uk/nowcasting (last accessed: 2025-09-19) and https://science.ncas.ac.uk/wiserewsa (last accessed: 2025-09-19), which provide RoA estimates that can meet the needs of many users.

6.1. Technical Remarks on Inference

6.1.1. Tile Size

The network was trained with tiles of 128×128 px^2. Consequently, it is sensible to only use this tile size as the input size. Nevertheless, the tiles have to be strategically assembled to mitigate edge effects (Figure S6a in Supporting Information S1). Here, we opt for a weighted average with 50% of overlap with the adjacent tiles, for which we have manually not detected edge effects (Figure 8f or Figure S6b in Supporting Information S1), as well as that it can align better with reference data (Figure 8c or Figure S6c in Supporting Information S1); this tiling has been employed for any result presented in this paper. Mathematically, the fully convolutional nature of the RoA network can be used to exploit larger input tile sizes: for example, 256×256 px^2 (Figure S6d in Supporting Information S1), where this larger tile size can even yield better pixelwise statistics (not shown). Given this situation, one can refrain from using a tiled approach for the retrieval over large areas. On the other hand, we conducted different experiments, which suggested that larger tile sizes can result in underestimated precipitation, in particular when the tile includes arid regions.

6.1.2. Quantile Crossing

The training of the RoA network does not enforce monotonicity in the quantiles. This can cause a situation where $\hat{y}_{\tau_1} > \hat{y}_{\tau_2}$ with $\tau_2 < \tau_1$, that is, quantile crossing. It is out of the scope of this work to address this problem, and solutions have been proposed and used in many works elsewhere, including corrections based on postprocessing

the retrieved quantiles. We decided to no resolve this situation for two reasons. On the one hand, in the context of this work, we are not required that the retrieved quantiles follow their expected order: this could be required in contexts where retrieval distributions are compared instead of point estimates. On the other hand, based on a number of qualitative analyses, we observed that the quantile crossing largely resembles jitter on the output percentiles, which also suggest that the training converged well. Moreover, we compute the expected value $\hat{y}_\mu$ by integration of the cumulative distribution function, approximated by the retrieved quantiles, with levels 0 and 1 extrapolated, and the numerical method does not require monotonicity of the quantiles. The curious reader can refer to Text S3 in Supporting Information S1 for further details.

7. Conclusions

This work describes and evaluates a retrieval model to obtain precipitation estimates only from the geostationary MSG thermal IR imagery. The region of interest was confined to contain all of Africa; therefore, we refer to the retrieval as Rain over Africa, or RoA for short. The RoA retrievals rely on a convolutional neural network, trained against GPM 2BCMB estimates, capable of providing pixelwise, case-specific uncertainties aiming to describe the irreducible uncertainty when estimating precipitation from IR measurements.

The RoA retrievals have varying degrees of accuracy: performance significantly degrades for extreme precipitation events, which can be attributed to IR attenuation, but this is also observed in the established IMERG estimates. Additionally, a spatial blur effect is noticeable, partly due to inherent uncertainties and partly because of the loss function used. Nonetheless, it is observed that it can be expected that the reference precipitation estimates are contained in the RoA retrieval uncertainties. In addition, when contrasted with the gauge-adjusted IMERG precipitation data set, it is revealed that RoA generally exhibits a favorable pixel-level performance over land, whereas IMERG benefits from the inclusion of microwave precipitation retrievals over oceanic regions. More generally, RoA demonstrates that it can capture finer precipitation structures over land, with an effective resolution that is better to the best-case IMERG retrievals. Moreover, 1-year precipitation averages show similar patterns when comparing RoA retrievals to IMERG and a second gauge-adjusted data set, CHIRPS, with RoA being the closest to a long-term average of GPM 2BCMB estimates over African land. Furthermore, the performance of RoA suggests making better use of geostationary data in data sets such as IMERG and CHIRPS.

The main virtue of RoA is the low latency with which detailed and accurate probabilistic precipitation estimates can be obtained for all of Africa. Even when produced using standard computing resources, RoA estimates become available approximately 100 times faster than those from IMERG Early, the most rapid version of IMERG. In practice, this means that RoA estimates can be derived from the most recent Meteosat scan even before the subsequent one begins, without requiring an advanced computational infrastructure, and using open-source code.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

The RoA retrieval is publicly available at https://github.com/SEE-GEO/roa (Amell, 2025). The training, validation, and test data sets are available at https://doi.org/10.5281/zenodo.14711540 (Hee & Amell, 2025), with scripts and notebooks supporting the results and conclusions found at https://doi.org/10.5281/zenodo.14738647 (Amell & Hee, 2025). Meteosat data were downloaded from EUMETSAT (2009), GPM 2BCMB from Olson (2018, 2022), IMERG from Huffman, Stocker, et al. (2023), and CHIRPS from https://data.chc.ucsb.edu/products/CHIRPS-2.0/africa_daily/tifs/p05/ (last accessed 2025-06-27).

References

Afzali Gorooh, V., Akbari Asanjan, A., Nguyen, P., Hsu, K., & Sorooshian, S. (2022). Deep neural network high spatiotemporal resolution precipitation estimation (deep-step) using passive microwave and infrared data. *Journal of Hydrometeorology*, 23(4), 597–617. https://doi.org/10.1175/JHM-D-21-0194.1

Amell, A. (2025). RoA [Software]. *Zenodo*. https://doi.org/10.5281/zenodo.14710246

Amell, A., & Hee, L. (2025). Scripts and notebooks used for the RoA paper [Software]. *Zenodo*. https://doi.org/10.5281/zenodo.14738647

Aminou, D. M. A., Jacquet, B., & Pasternak, F. (1997). Characteristics of the Meteosat second generation (MSG) radiometer/imager: SEVIRI. In H. Fujisada (Ed.), *Sensors, systems, and next-generation satellites* (Vol. 3221, pp. 19–31). SPIE / International Society for Optics and Photonics. https://doi.org/10.1117/12.298084

CEOS. (2024). The CEOS database ([Database]). Retrieved from https://database.eohandbook.com/index.aspx

D'Adderio, L. P., Casella, D., Dietrich, S., Panegrossi, G., & Sanò, P. (2023). A first step towards Meteosat third generation day-2 precipitation rate product: Deep learning for precipitation rate retrieval from geostationary infrared measurements. *Remote Sensing*, *15*(24), 5662. https://doi.org/10.3390/rs15245662

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. *Journal of Hydrometeorology*, *23*(8), 1297–1314. https://doi.org/10.1175/JHM-D-21-0234.1

Dinku, T. (2019). Chapter 7—Challenges with availability and quality of climate data in Africa. In A. M. Melesse, W. Abtew, & G. Senay (Eds.), *Extreme hydrology and climate variability* (pp. 71–80). Elsevier. https://doi.org/10.1016/B978-0-12-815998-9.00007-5

Draper, D. W., Newell, D. A., Wentz, F. J., Krimchansky, S., & Skofronick-Jackson, G. M. (2015). The global precipitation measurement (GPM) microwave imager (GMI): Instrument overview and early on-orbit performance. *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *8*(7), 3452–3462. https://doi.org/10.1109/JSTARS.2015.2403303

EUMETSAT. (2009). High rate Seviri level 1.5 image data—msg—0 degree. Retrieved from https://data.eumetsat.int/product/EO:EUM:DAT:MSG:HRSEVIRI

EUMETSAT. (2024). eumdac. Retrieved from https://user.eumetsat.int/resources/user-guides/eumetsat-data-access-client-eumdac-guide

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., et al. (2015). The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. *Scientific Data*, *2*(1), 150066. https://doi.org/10.1038/sdata.2015.66

Gao, Y., Guan, J., Zhang, F., Wang, X., & Long, Z. (2022). Attention-unet-based near-real-time precipitation estimation from Fengyun-4A satellite imageries. *Remote Sensing*, *14*(12), 2925. https://doi.org/10.3390/rs14122925

Grecu, M., Olson, W. S., Munchak, S. J., Ringerud, S., Liao, L., Haddad, Z., et al. (2016). The GPM combined algorithm. *Journal of Atmospheric and Oceanic Technology*, *33*(10), 2225–2245. https://doi.org/10.1175/JTECH-D-16-0019.1

Guilloteau, C., & Foufoula-Georgiou, E. (2020). Multiscale evaluation of satellite precipitation products: Effective resolution of IMERG. In V. Levizzani, C. Kidd, D. B. Kirschbaum, C. D. Kummerow, K. Nakamura, & F. J. Turk (Eds.), *Satellite precipitation measurement: Volume 2* (pp. 533–558). Springer International Publishing. https://doi.org/10.1007/978-3-030-35798-6_5

Guilloteau, C., Foufoula-Georgiou, E., Kirstetter, P., Tan, J., & Huffman, G. J. (2021). How well do multisatellite products capture the space–time dynamics of precipitation? Part I: Five products assessed via a wavenumber–frequency decomposition. *Journal of Hydrometeorology*, *22*(11), 2805–2823. https://doi.org/10.1175/JHM-D-21-0075.1

Hee, L., & Amell, A. (2025). RoA collocations database [Dataset]. *Zenodo*. https://doi.org/10.5281/zenodo.14711540

Holmlund, K., Grandell, J., Schmetz, J., Stuhlmann, R., Bojkov, B., Munro, R., et al. (2021). Meteosat third generation (MTG): Continuation and innovation of observations from geostationary orbit. *Bulletin of the American Meteorological Society*, *102*(5), E990–E1015. https://doi.org/10.1175/BAMS-D-19-0304.1

Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., et al. (2014). The global precipitation measurement mission. *Bulletin of the American Meteorological Society*, *95*(5), 701–722. https://doi.org/10.1175/BAMS-D-13-00164.1

Huffman, G. J., Bolvin, D. T., Joyce, R., Nelkin, E., Tan, J., Braithwaite, D., et al. (2023). *Nasa global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG) version 07 (Algorithm Theoretical Basis Document (ATBD) No. V07)*. NASA/Goddard Space Flight Center. Retrieved from https://arthurhou.pps.eosdis.nasa.gov/Documents/IMERG_V07_ATBD_final.pdf

Huffman, G. J. (1997). Estimates of root-mean-square random error for finite samples of estimated precipitation. *Journal of Applied Meteorology*, *36*(9), 1191–1201. https://doi.org/10.1175/1520-0450(1997)036⟨1191:EORMSR⟩2.0.CO;2

Huffman, G. J., Stocker, E., Bolvin, D., Nelkin, E., & Tan, J. (2023). GPM IMERG final precipitation l3 half hourly 0.1 degree x 0.1 degree v07 [Dataset]. https://doi.org/10.5067/GPM/IMERG/3B-HH/07

Ji, X., Song, X., Guo, A., Liu, K., Cao, H., & Feng, T. (2024). Oceanic precipitation nowcasting using a unet-based residual and attention network and real-time Himawari-8 images. *Remote Sensing*, *16*(16), 2871. https://doi.org/10.3390/rs16162871

Jiang, Y., Gao, F., Zhang, S., Cheng, W., Liu, C., & Wang, S. (2023). Mcspf-net: A precipitation forecasting method using multi-channel cloud observations of Fy-4A satellite by 3d convolution neural network. *Remote Sensing*, *15*(18), 4536. https://doi.org/10.3390/rs15184536

Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G., & Kirschbaum, D. B. (2017). So, how much of the Earth's surface is covered by rain gauges? *Bulletin of the American Meteorological Society*, *98*(1), 69–78. https://doi.org/10.1175/BAMS-D-14-00283.1

Kojima, M., Miura, T., Furukawa, K., Hyakusoku, Y., Ishikiri, T., Kai, H., et al. (2012). Dual-frequency precipitation radar (DPR) development on the global precipitation measurement (GPM) core observatory. In H. Shimoda, X. Xiong, C. Cao, X. Gu, C. Kim, & A. S. K. Kumar (Eds.), *Earth observing missions and sensors: Development, implementation, and characterization II* (Vol. 8528, p. 85281A). SPIE/International Society for Optics and Photonics. https://doi.org/10.1117/12.976823

Kubota, T., Aonashi, K., Ushio, T., Shige, S., Takayabu, Y. N., Arai, Y., et al. (2017). Recent progress in global satellite mapping of precipitation (GSMAP) product. In *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 2712–2715). https://doi.org/10.1109/IGARSS.2017.8127556

Kuligowski, R. J. (2002). A self-calibrating real-time GOES rainfall algorithm for short-term rainfall estimates. *Journal of Hydrometeorology*, *3*(2), 112–130. https://doi.org/10.1175/1525-7541(2002)003⟨0112:ASCRTG⟩2.0.CO;2

Kuligowski, R. J., Li, Y., Hao, Y., & Zhang, Y. (2016). Improvements to the GOES-r rainfall rate algorithm. *Journal of Hydrometeorology*, *17*(6), 1693–1704. https://doi.org/10.1175/JHM-D-15-0186.1

Kummerow, C. D., Randel, D. L., Kulie, M., Wang, N.-Y., Ferraro, R., Munchak, S. J., & Petkovic, V. (2015). The evolution of the Goddard profiling algorithm to a fully parametric scheme. *Journal of Atmospheric and Oceanic Technology*, *32*(12), 2265–2280. https://doi.org/10.1175/JTECH-D-15-0039.1

Mugnai, A., Casella, D., Cattani, E., Dietrich, S., Laviola, S., Levizzani, V., et al. (2013). Precipitation products from the hydrology SAF. *Natural Hazards and Earth System Sciences*, *13*(8), 1959–1981. https://doi.org/10.5194/nhess-13-1959-2013

Nguyen, P., Ombadi, M., Gorooh, V. A., Shearer, E. J., Sadeghi, M., Sorooshian, S., et al. (2020). PERSIANN dynamic infrared–rain rate (PDIR-now): A near-real-time, quasi-global satellite precipitation dataset. *Journal of Hydrometeorology*, *21*(12), 2893–2906. https://doi.org/10.1175/JHM-D-20-0177.1

Nguyen, P., Ombadi, M., Sorooshian, S., Hsu, K., AghaKouchak, A., Braithwaite, D., et al. (2018). The PERSIANN family of global satellite precipitation data: A review and evaluation of products. *Hydrology and Earth System Sciences*, *22*(11), 5801–5816. https://doi.org/10.5194/hess-22-5801-2018

Olson, W. (2018). Gpm dpr and gmi combined precipitation l2b 1.5 hours 5 km v06. Greenbelt, MD. https://doi.org/10.5067/GPM/DPRGMI/CMB/2B/06

Olson, W. (2022). Gpm dpr and gmi combined precipitation l2b 1.5 hours 5 km v07. Greenbelt, MD. https://doi.org/10.5067/GPM/DPRGMI/CMB/2B/07

Pfreundschuh, S. (2022). Hydronn [Software]. *Zenodo.* https://doi.org/10.5281/zenodo.6371712

Pfreundschuh, S., Brown, P. J., Kummerow, C. D., Eriksson, P., & Norrestad, T. (2022). Gprof-nn: A neural-network-based implementation of the Goddard profiling algorithm. *Atmospheric Measurement Techniques*, *15*(17), 5033–5060. https://doi.org/10.5194/amt-15-5033-2022

Pfreundschuh, S., Eriksson, P., Duncan, D., Rydberg, B., Håkansson, N., & Thoss, A. (2018). A neural network approach to estimating a posteriori distributions of Bayesian retrieval problems. *Atmospheric Measurement Techniques*, *11*(8), 4627–4643. https://doi.org/10.5194/amt-11-4627-2018

Pfreundschuh, S., Guilloteau, C., Brown, P. J., Kummerow, C. D., & Eriksson, P. (2024). GPROF V7 and beyond: Assessment of current and potential future versions of the GPROF passive microwave precipitation retrievals against ground radar measurements over the continental US and the Pacific Ocean. *Atmospheric Measurement Techniques*, *17*(2), 515–538. https://doi.org/10.5194/amt-17-515-2024

Pfreundschuh, S., Ingemarsson, I., Eriksson, P., Vila, D. A., & Calheiros, A. J. P. (2022). An improved near-real-time precipitation retrieval for Brazil. *Atmospheric Measurement Techniques*, *15*(23), 6907–6933. https://doi.org/10.5194/amt-15-6907-2022

Rahimi, R., Ravirathinam, P., Ebtehaj, A., Behrangi, A., Tan, J., & Kumar, V. (2024). Global precipitation nowcasting of integrated multi-satellite retrievals for GPM: A u-net convolutional LSTM architecture. *Journal of Hydrometeorology*, *25*(6), 947–963. https://doi.org/10.1175/JHM-D-23-0119.1

Rodgers, C. D. (2000). *Inverse methods for atmospheric sounding.* World Scientific. https://doi.org/10.1142/3171

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi (Eds.), *Medical image computing and computer-assisted intervention—MICCAI 2015* (Vol. 9351, pp. 234–241). Springer International Publishing. https://doi.org/10.1007/978-3-319-24574-4_28

Rota, S. (2004). Meteosat second generation mission status. In H. Lacoste (Ed.), *Second MSG RAO Workshop* (Vol. 582, p. 5).

Sadeghi, M., Nguyen, P., Hsu, K., & Sorooshian, S. (2020). Improving near real-time precipitation estimation using a u-net convolutional neural network and geographical information. *Environmental Modelling & Software*, *134*, 104856. https://doi.org/10.1016/j.envsoft.2020.104856

Saltikoff, E., Friedrich, K., Soderholm, J., Lengfeld, K., Nelson, B., Becker, A., et al. (2019). An overview of using weather radar for climatological studies: Successes, challenges, and potential. *Bulletin of the American Meteorological Society*, *100*(9), 1739–1752. https://doi.org/10.1175/BAMS-D-18-0166.1

Schmetz, J., Pili, P., Tjemkes, S., Just, D., Kerkmann, J., Rota, S., & Ratier, A. (2002). An introduction to Meteosat second generation (MSG). *Bulletin of the American Meteorological Society*, *83*(7), 977–992. https://doi.org/10.1175/1520-0477(2002)083⟨0977:AITMSG⟩2.3.CO;2

Schmid, J. (2000). The seviri instrument. In *Proceedings of the 2000 EUMETSAT Meteorological Satellite Data User's Conference*. Retrieved from https://www-cdn.eumetsat.int/files/2020-04/pdf_ten_msg_seviri_instrument.pdf

Schmit, T. J., Gunshor, M. M., Menzel, W. P., Gurka, J. J., Li, J., & Bachmeier, A. S. (2005). Introducing the next-generation advanced baseline imager on GOES-R. *Bulletin of the American Meteorological Society*, *86*(8), 1079–1096. https://doi.org/10.1175/BAMS-86-8-1079

Skofronick-Jackson, G., Petersen, W. A., Berg, W., Kidd, C., Stocker, E. F., Kirschbaum, D. B., et al. (2017). The global precipitation measurement (GPM) mission for science and society. *Bulletin of the American Meteorological Society*, *98*(8), 1679–1695. https://doi.org/10.1175/BAMS-D-15-00306.1

Soula, S. (2009). Lightning and precipitation. In H. D. Betz, U. Schumann, & P. Laroche (Eds.), *Lightning: Principles, instruments and applications: Review of modern lightning research* (pp. 447–463). Springer. Netherlands. https://doi.org/10.1007/978-1-4020-9079-0_20

Stark, H. R., & Schumann, W. (2004). Status of the ESA Meteosat Second Generation (MSG) Programme. In H. Lacoste (Ed.), *Second MSG RaO Workshop* (Vol. 582, p. 11).

Turk, F. J., Arkin, P., Ebert, E. E., & Sapiano, M. R. P. (2008). Evaluating high-resolution precipitation products. *Bulletin of the American Meteorological Society*, *89*(12), 1911–1916. https://doi.org/10.1175/2008bams2652.1

Viltard, N., Sambath, V., Lepetit, P., Martini, A., Barthès, L., & Mallet, C. (2025). Evaluation of drain, a deep-learning approach to rain retrieval from GPM passive microwave radiometer. *IEEE Transactions on Geoscience and Remote Sensing*, *63*, 1–7. https://doi.org/10.1109/TGRS.2023.3293932

Wang, C., Tang, G., Xiong, W., Ma, Z., & Zhu, S. (2021). Infrared precipitation estimation using convolutional neural network for Fengyun satellites. *Journal of Hydrology*, *603*, 127113. https://doi.org/10.1016/j.jhydrol.2021.127113

WMO. (2023). *Guide to instruments and methods of observation* (2023rd ed., Vol. I). WMO. (No. WMO-No. 8) (Chapter 14, Annex). Retrieved from https://library.wmo.int/idurl/4/68695

WMO. (2024). Oscar. ([Database]) Retrieved from https://space.oscar.wmo.int

Wolf, R. (1999). Lrit/hrit global specification (No. CGMS 03 2.6). Retrieved from https://www.cgms-info.org/wp-content/uploads/2021/10/pdf_cgms_03.pdf