Meteorological Applications

RMetS Royal Meteorological Society

RESEARCH ARTICLE OPEN ACCESS

Evaluation of Satellite Rainfall Products for Nowcasting Over Southern Africa Using Rain-Gauges and Lightning Data

Simon Ageet[1] | John H. Marsham[2] | Douglas J. Parker[2,3] | Abhilash Singh[4] | Christopher M. Taylor[5,6] | Alexander J. Roberts[7] | James Groves[7] | Edson Nkonde[8] | Adrià Amell[9]

[1]Institute of Meteorology and Climate Research, Karlsruhe Institute of Technology, Karlsruhe, Germany | [2]School of Earth, Environment and Sustainability, University of Leeds, Leeds, UK | [3]Climate Risk, NORCE Norwegian Research Centre AS, Bergen, Norway | [4]School of Mathematics, University of Leeds, Leeds, UK | [5]UK Centre for Ecology & Hydrology, Wallingford, UK | [6]National Centre for Earth Observation, Wallingford, UK | [7]National Centre for Atmospheric Science, University of Leeds, Leeds, UK | [8]Zambia Meteorological Department, Lusaka, Zambia | [9]Department of Environmental and Energy Sciences, Chalmers University of Technology, Gothenburg, Sweden

Correspondence: Simon Ageet (simon.ageet@kit.edu)

Received: 8 February 2026 | **Revised:** 26 July 2026 | **Accepted:** 22 September 2026

Keywords: CRR | GPM IMERG | RoA | satellite nowcasting | verification

ABSTRACT

Satellite-based 'nowcasting', which extrapolates the latest observations to provide forecasts from minutes to 6 h ahead, offers an accurate and timely approach for monitoring extreme weather. We evaluate the widely used NWCSAF Convective Rainfall Rate (CRR) nowcasting product using instantaneous (non-propagated) estimates against a network of high-temporal-resolution rain gauges in Zambia during the wet seasons from 2019 to 2024. We also assess two alternative satellite-based rainfall products: IMERG microwave-only (IMERG-HQ) and Rain over Africa (RoA), a machine learning–generated nowcasting product. Additionally, lightning data from the Meteosat Third Generation mission and cloud-top temperature data were used to evaluate the products. These datasets formed part of the products used by forecasters during nowcasting testbeds conducted under the WISER EWSA project and positive feedback from users was widely reported. Our analysis shows that the products have distributions similar to the in situ data, but all exhibit biases, with CRR and IMERG-HQ generally overestimating rainfall rates while RoA tends to underestimate. Across different thresholds and aggregation windows, RoA outperforms IMERG-HQ and CRR for rain occurrence on most metrics, including hit rate, probability of false alarm and Heidke skill score. However, all products show reduced skill for extreme rainfall. CRR also shows weak correspondence with lightning activity and does not detect rainfall from cloud tops warmer than −30°C. In contrast, RoA and IMERG-HQ exhibit greater sensitivity to cloud top temperature, including warmer cloud regimes. The generally superior performance of RoA compared to IMERG-HQ is remarkable, given that RoA only takes infrared satellite channels as input, whereas IMERG-HQ uses microwave channels. We attribute RoA's greater skill to the ability of ML to optimally exploit the data and bring spatial information into the algorithm. Our findings emphasise the value of satellite-based monitoring and nowcasting, while underscoring the need for training and calibration to improve accuracy.

1 | Introduction

Rainfall in the tropics is predominantly convective, resulting in precipitation that is often intense and spatially localised. Such events frequently lead to hydrometeorological disasters, primarily due to flooding. Klein et al. (2018) found that, in West Africa, convective cloud tops colder than −80°C are associated with an 80% likelihood of extreme rainfall. Under global warming, both the intensity and frequency of these localised extreme events increase (Taylor et al. 2017; Kendon et al. 2019).

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Meteorological Applications* published by John Wiley & Sons Ltd on behalf of Royal Meteorological Society.

The impacts of such events are exacerbated by rapid population growth, particularly in African countries, where urban populations are projected to triple between 2020 and 2050 (UN DESA 2018), increasing exposure and vulnerability to hydrometeorological hazards.

Weather forecasting in tropical Africa remains challenging due to persistent limitations in the performance of numerical weather prediction (NWP) systems. Accurate and timely forecasts are nevertheless essential for reducing the impacts of weather-related disasters. Across Africa, operational forecasts are primarily produced at daily and seasonal timescales using NWP models (Milton et al. 2017). However, several studies have shown that forecast skill in tropical regions often does not substantially exceed that of simple climatological forecasts (e.g., Vogel et al. 2018; Woodhams et al. 2018; Ageet et al. 2023). This reduced skill is largely attributed to the poor representation of key atmospheric processes, particularly convection (Vogel et al. 2018), together with the scarcity of observational data available for data assimilation (Žagar 2017).

Nowcasting, defined by the World Meteorological Organisation (WMO; WMO 2017) as 'forecasting with local detail, by any method, over a period from the present to 6 h ahead, including a detailed description of the present weather', offers a complementary solution to daily NWP forecasts. By leveraging the most recent observational data and propagating it forward over short lead times (up to 6 h), nowcasts are typically more accurate than daily and longer-range NWP forecasts (e.g., Bližňák et al. 2017; White et al. 2017) and are therefore well suited for predicting high-impact weather events at very short lead times. This is particularly valuable in (sub)tropical regions of Africa, where weather systems are highly dynamic and stochastic in nature (Lamptey et al. 2024). Motivated by these challenges and opportunities, the Weather and Climate Information Services Early Warning for Southern Africa (WISER-EWSA) project, under which the verification work presented in this article was conducted, aimed to introduce and enhance nowcasting across Southern Africa.

To achieve the project's aim, WISER-EWSA aimed to implement two nowcasting testbeds (Ageet et al. 2024; San Roman Pineda et al. 2025), intensive two-week events designed to bring together key stakeholders, including meteorological services, researchers, disaster preparedness agencies, media representatives, intermediary non-governmental organisations and urban communities. These testbeds facilitated the co-design, co-production and trial of nowcasting products for three urban settings: Kanyama (Zambia), Boane (Mozambique) and Katlehong (South Africa). Testbed 1 (T1; Ageet et al. 2024) was conducted from 29 January to 9 February 2024 in Lusaka, Zambia. Testbed 2 (T2; San Roman Pineda et al. 2025), originally planned for January–February 2025 in Maputo, Mozambique, was subsequently divided into two components: T2-Zambia (T2-Z), held from 27 January to 7 February 2025 and T2-Maputo (T2-M), conducted from 20 to 27 March 2025. Additionally, in response to insights gained from T1, an extended 'King-size' testbed was introduced. Conducted throughout the rain season, October 2024 to April 2025, this phase involved a series of intensive, event-triggered nowcasting activated when severe weather was forecasted by medium-range forecasts. Across all testbeds, a suite of nowcasting tools and resources was used to create the nowcasts.

In Africa, where weather radar coverage remains sparse and uneven, satellite-based nowcasting systems have become the primary operational approach for monitoring and forecasting short-term convective rainfall. Although weather radars provide the most accurate remote sensing of rainfall area, intensity and temporal evolution, their deployment across much of the continent remains limited due to the high costs associated with acquisition, operation and maintenance (World Bank 2021; WMO 2025). Consequently, operational nowcasting in many African countries relies heavily on satellite observations and satellite-derived products. Roberts et al. (2022) reviewed the principal nowcasting methodologies applied across Africa and highlighted the widespread adoption of products developed by the EUMETSAT Satellite Application Facility on Support to Nowcasting and Very Short-Range Forecasts (NWC SAF). These products have been operationally implemented by institutions such as the South African Weather Service (SAWS; de Coning et al. 2015, Gijben and De Coning 2017) and within regional forecasting initiatives including the Global Challenges Research Fund African Science for Weather Information and Forecasting Techniques project (GCRF SWIFT; Parker et al. 2022). Similarly, the WISER-EWSA programme employed NWC SAF products as core operational tools during forecasting testbeds. Among the available products, the Convective Rainfall Rate (CRR) and Rapidly Developing Thunderstorms (RDT) products were the most widely used during the testbeds. The forecasting activities also provided an opportunity to evaluate emerging machine-learning-based products, including the Rain over Africa (RoA; Amell et al. 2025) satellite rainfall estimates and machine-learning nowcasting of satellite imagery (Hermes et al. 2025).

While all nowcasting products demonstrated both strengths and limitations, primarily assessed through day-to-day, subjective, case-based evaluations, a systematic and objective verification is essential to determine their suitability for Southern Africa. Many of these products are not optimised for tropical or subtropical regions; for instance, the CRR product is calibrated using European radar networks (Hernanz and Lahuerta 2019), which may not fully capture the convective characteristics of Africa. Unfortunately, verification efforts over the African continent remain limited, with the few studies done for West Africa, East and Southern Africa. Anderson et al. (2024) evaluated a novel convective activity nowcasting method based on satellite-derived cloud-top temperatures, against daily rain gauge observations over the Sahel. Similarly, Hill et al. (2020) assessed the performance of CRR and Rapidly Developing Thunderstorm (RDT) products over West and East Africa using the Global Precipitation Measurement (GPM) mission's IMERG dataset (Hou et al. 2014). In South Africa, de Coning et al. (2015) and Gijben and De Coning (2017) verified RDT using ground-based radar and lightning data. Recently, machine learning has provided significant opportunities to improve nowcasting, including in Africa, both by enhancing rainfall estimation and improving forward predictions (e.g., Ayzel and Heistermann 2025; Hermes et al. 2025). There is therefore a need to assess and quantify the skill of satellite-based nowcasting products over Africa, particularly at

14698080, 2026, 5, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70253 by UK Centre For Ecology & Hydrology, Wiley Online Library on [16/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

sub-daily rainfall accumulations and intensities, using in situ rain gauge observations and to evaluate whether alternative products, including ML-based approaches, offer improved performance compared to the commonly used CRR.

This study therefore aims to systematically assess the skill of three near-real time satellite products over Southern Africa at sub-hourly to hourly timescales and to identify the most suitable for nowcasting for the region. We evaluate instantaneous estimates from the benchmark NWCSAF CRR nowcasting product, the microwave-only component of IMERG, which is not available in real time for nowcasting but is often used as a reference (e.g., for training machine-learning algorithms) and RoA, a newly developed machine-learning-based satellite rainfall product, during the 2019–2024 wet seasons. We use a network of 27 Automatic Weather Stations (AWS) provided by the Zambia Meteorological Department as part of the WISER EWSA initiative, alongside lightning data from Meteosat Third Generation (MTG). Importantly, unlike most previous studies, our evaluation is conducted on a sub-daily timescale, allowing us to assess the suitability of these products for the nowcasting timescales of minutes to 6 h of great importance to many users.

The rest of the manuscript is structured as follows: Section 2 introduces the datasets used and the methods of analysis applied. Section 3 presents the results, including the evaluation of rainfall distributions and intensities in the three products against gauge observations, an assessment of the potential to infer rainfall rates from the distribution of satellite rainfall estimates (SREs) and a comparison of these products with rainfall proxies such as lightning and cloud-top temperatures. The results are discussed in Section 4 and the conclusions presented in Section 5.

2 | Data and Methods

2.1 | Data

2.1.1 | Satellite-Based Products Evaluated

2.1.1.1 | Convective Rainfall Rate. The Convective Rainfall Rate (CRR; Hernanz and Lahuerta 2019), hereafter referred to as CRR, is a near-real-time product developed for nowcasting applications. It provides instantaneous rainfall rates and hourly accumulations associated with convective and convectively induced stratiform cloud systems. The product is derived from a combination of SEVIRI (Spinning Enhanced Visible and Infrared Imager) radiances, specifically using either a two or three-variable (irradiances) approach. The two-variable approach configuration employs the 10.8 μm infrared (IR) band and the difference between the 10.8 μm IR and 6.2 μm water vapour (WV) channels, while the three-variable approach additionally uses the visible channel (0.6 μm VIS). While the three-channel configuration generally yields more accurate results and is therefore preferred, it is only available during daytime. More precisely, for each pixel, if the solar zenith angle is below a user-defined threshold (default: 80°) and the visible channel is usable, the three-channel algorithm is applied. Otherwise, typically at night or under low solar illumination, the algorithm uses the two-channel configuration.

The CRR algorithm is calibrated through pixel-level comparison of SEVIRI imagery with collocated ground-based radar observations. Calibration has been primarily performed using radar data from European networks, notably those in the Baltic region, Hungary and Spain. As a result, the algorithm is optimised for the mid-latitudes conditions, potentially limiting its accuracy in tropical environments, which are the focus of this study. As outlined in Hill et al. (2020), the basic CRR rain rates obtained from the above steps undergo several post-processing corrections aimed at improving the fidelity of the estimates. These include: (a) Filtering is applied to exclude non-convective rainfall by setting to zero the CRR value for any pixel where the surrounding 7 × 7-pixel neighbourhood contains no values exceeding 3 mm/h. Although the initial filtering removes low-intensity non-convective rainfall (i.e., < 3 mm/h), subsequent corrections discussed below, that happen after the filtering, may reduce rain rates below this threshold. Consequently, the final CRR product may include values under 3 mm/h. (b) Adjustments applied to account for the spatial and temporal variability of cloud properties. These include moisture corrections (increasing rainfall rates in moist areas and reducing them in dry areas), cloud-top evolution corrections (enhancing or suppressing rainfall based on cloud growth or decay) and orographic corrections (modifying rain rates based on terrain-induced vertical motion using surface pressure gradients aligned with near-surface NWP wind fields). (c) For some regions, where available, lightning and/or radar data are incorporated to enhance the detection and quantification of convective rainfall, given their strong association with convective processes. Finally, the rain rates are extrapolated at intervals of 15 min for short-term nowcasting purposes. This is achieved through the application of the NWCSAF product generating element (PGE), EXIM (EXtrapolted IMagery) that makes use of atmospheric motion vectors to provide information on the direction and speed of features in both geostationary satellite imagery and NWCSAF products like CRR (Hernanz and Lahuerta 2019). In this study, only instantaneous CRR estimates prior to propagation are used.

2.1.1.2 | Rain Over Africa. Rain of Africa (RoA; Amell et al. 2025) is a near-real-time (NRT), satellite-based rainfall product that provides continent-wide coverage across Africa every 15 min on a 3-km grid, the native MSG resolution. The RoA estimates are derived only from SEVIRI thermal infrared radiances, enabling continuous rainfall monitoring with only a latency of a few minutes.

The retrieval algorithm underlying RoA combines a quantile regression neural network (QRNN; Pfreundschuh et al. 2018) with a convolutional neural network. The QRNN component enables RoA to address the irreducible uncertainty inherent in SREs, where identical infrared temperatures can correspond to different rainfall intensities or vice versa. Therefore, RoA is also able to provide case-specific retrieval errors, but which we do not analyse here. The model is trained and validated using 4 years of collocations of Meteosat SEVIRI and the GPM Core Observatory combined radar-microwave radiometer product.

14698080, 2026, 5, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70253 by UK Centre For Ecology & Hydrology, Wiley Online Library on [16/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

No gauge observations are incorporated at any stage of the model development, training or validation. Details on the algorithmic methodology are presented in (Amell et al. 2025). In this work, we use a version of RoA that predates its publication, where the main differences reside in the projection used (a $0.027^{\circ} \times 0.027^{\circ}$ grid) and the extent of the training and evaluation database. An internal evaluation showed minimal differences between the two estimates and we consider the version used functionally equivalent to the published RoA.

Compared to widely used SREs such as IMERG and Climate Hazard Infrared Precipitation with Stations (CHIRPS; Funk et al. 2015), RoA offers substantially higher temporal resolution and NRT capability. When compared with IMERG, RoA reproduced similar diurnal cycles but showed better consistency, including a better resolution of small-scale structures over land, underscoring the RoA potential for real-time rainfall monitoring and nowcasting over Africa.

2.1.1.3 | Integrated Multi-Satellite Retrievals for Global Precipitation Measurement (IMERG). Integrated Multi-satellitE Retrievals for the Global Precipitation Measurement (GPM; Hou et al. 2014) mission (IMERG; Huffman et al. 2020), provides global satellite-based precipitation estimates at a spatial resolution of 0.1° and a temporal resolution of 30 min. The IMERG product suite includes multiple variants based on latency (i.e., Early, Late and Final runs) and input sources (e.g., microwave-only, infrared-only and infrared-calibrated with microwave). In this study, we use the Final run of the microwave-only product, referred to as IMERG High-Quality (IMERG-HQ), from version 7 (IMERG v07; Huffman et al. 2023).

IMERG-HQ is derived exclusively from passive microwave (PMW) sensors, including the GPM Microwave Imager (GMI) and additional partner sensors from the GPM constellation. Unlike other IMERG variants, IMERG-HQ therefore excludes both infrared data and gauge-based adjustments, making it well suited for the independent validation of precipitation estimates such as the SEVIRI-based CRR. The merging and intercalibration of the PMW instruments are performed using the GPROF-2017 retrieval algorithm (Kummerow et al. 2015), with reference to high-quality datasets such as the Combined Radar-Radiometer Algorithm (CORRA), TMI and GMI. A quantile-matching approach is employed to adjust precipitation histograms from individual PMW sensors to match the reference distributions, incorporating both climatological and seasonal corrections. The resulting fields are then gridded at 0.1° spatial resolution for each 30-min interval, resulting in the IMERG-HQ fields. IMERG-HQ has been shown to have superior performance over other IMERG variants (Tan et al. 2016) and this may be attributed to the strengths of PMW sensors, which are more sensitive to precipitation, including light rain and mixed-phase hydrometeors and are less affected by cloud cover. However, a key limitation of IMERG-HQ is the substantial reduction in temporal coverage; approximately 70%–90% of 30-min time steps are missing compared to the full IMERG product. This loss arises from the infrequent PMW satellite overpasses, while additional spatial gaps occur because PMW sensors operate on lower-orbit platforms with narrower swath widths, resulting in incomplete spatial coverage at each overpass.

2.1.2 | Observational Data

2.1.2.1 | Meteosat Third Generation Lightning Data. Lightning data is a reliable proxy for atmospheric convection and is widely used to monitor convective development (EUMETSAT 2025b; Erdmann and Poelman 2023). The launch of the Meteosat Third Generation (MTG) geostationary satellite, equipped with the Lightning Imager (LI), by European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT), enables real-time, continuous detection of total lightning activity (i.e., cloud-to-cloud and intra-cloud and cloud-to-ground lightning) across most of Earth disc. The MTG Level 2 LI data (MTG LI; EUMETSAT 2025a) offers new opportunities for improved monitoring of convective rainfall and hence, enhancing nowcasting efforts, particularly in tropical regions where radar systems are rarely available. Moreover, the performance of MTG LI has been demonstrated to be reliable and accurate across diverse geographical regions, various daylight conditions, viewing geometry and storm characteristics. Bližňák and Sokol (2026) reported a flash detection efficiency (FDE) exceeding 80%, a false alarm rate below 30% and millisecond-level timing agreement. Using 18 months of operational observations (Viticchié et al. 2026), reported an average FDE of 87.4%, a false flash fraction of just 0.16% and excellent timing and geolocation accuracy.

Lightning can be detected using a range of techniques across different parts of the electromagnetic spectrum. These include lightning mapping arrays, which utilise very-high-frequency radio emissions; ground-based lightning detection networks, which rely on lower-frequency signals; and space-borne LIs, which observe the optical pulses emitted by lightning. The MTG LI data used in this study are based on the latter approach, with the instrument launched in December 2022 onboard the MTG satellite. In a nutshell, the MTG LI detection process works as follows. First, the radiance in each pixel consisting of both a background component and a potential lightning signal is measured. The background is removed from the measurement and the residual 'lightning' signal is then compared to a detection threshold. Any residuals exceeding this threshold are flagged as candidate lightning events. Because several artefacts, such as radiometric noise, sunlight contamination, particle hits on the focal plane or platform micro-vibrations, can be mistakenly detected as lightning, additional processing is required to distinguish true lightning events from false detections. The LI instrument is also calibrated twice per year. This is done by analysing dark scenes during Earth-eclipse conditions at night, when the instrument view is free of sunlight and suitable for radiometric calibration. Further details of the processing chain are provided in EUMETSAT (2025a). To the best of our knowledge, MTG lightning data have not been incorporated in any way into the satellite-based products evaluated in this study. Although the CRR algorithm can, in principle, incorporate lightning information, the version used here relied solely on MSG data and MTG Lightning Imager (LI) observations were not included.

14698080, 2026, 5, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70253 by UK Centre For Ecology & Hydrology, Wiley Online Library on [16/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

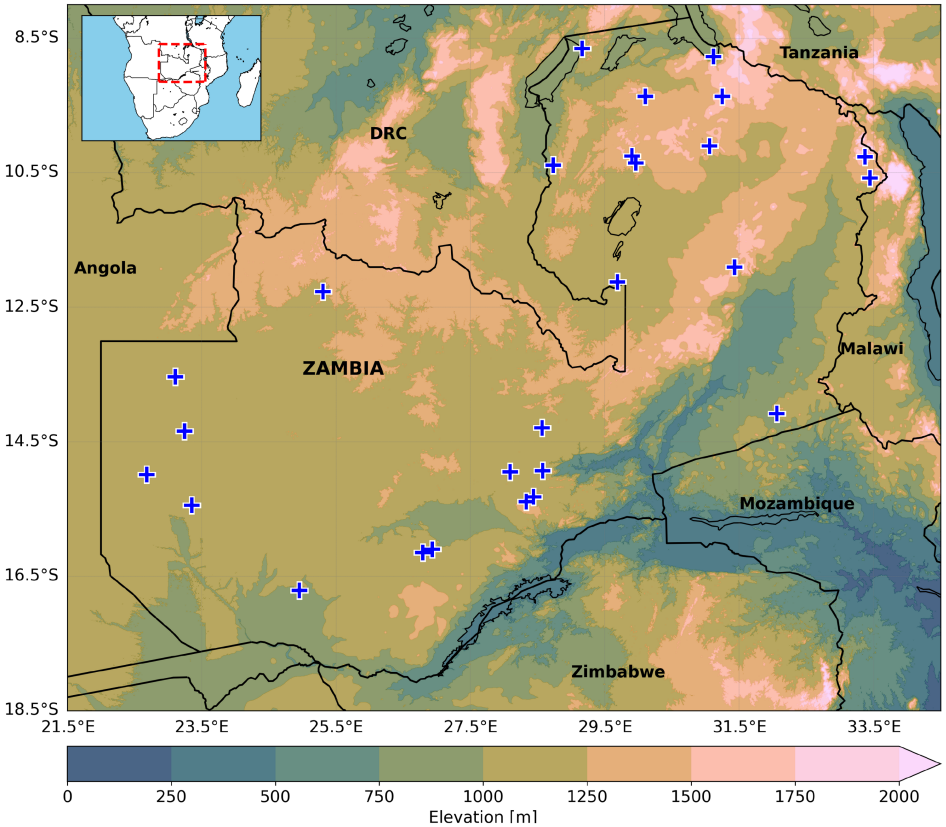

FIGURE 1 | Study domain and locations of the 27 rain gauges (blue-coloured plus signs) across Zambia that were used in the analysis. The colour shading shows the elevation based on Global Land One-km Base Elevation Project (GLOBE; GLOBE Task Team et al. 1999). The map inset highlights the location of Zambia within Southern Africa (red box).

2.1.2.2 | Cloud-Top Temperature. Most satellite-based rainfall products rely on algorithms that infer precipitation from cloud-top temperature, with colder cloud tops generally associated with higher intensity rainfall. This relationship is particularly central to IR imagery-based retrievals. In this study, we relate rainfall rates from CRR, RoA and IMERG-HQ to corresponding cloud-top temperatures. For this purpose, we use Level 1.5 Meteosat SEVIRI cloud-top temperatures derived from the 10.8 μm channel. These data are provided in a calibrated form, that is, they are geolocated to a standardised geostationary projection centred at 0° N, 0° E and corrected for geometric and radiometric effects (EUMETSAT 2017). The dataset provides relatively high temporal resolution (15-min intervals) and low latency (approximately 5 min after image acquisition), with a spatial resolution of 3 × 3 km at the sub-satellite point. However, cloud-top temperature does not constitute a strictly independent dataset for the IR-based CRR and RoA products evaluated here, as both derive rainfall estimates, at least in part, from cloud-top temperature information.

2.1.2.3 | Rain Gauges. Despite challenges such as data gaps and a sparse network, rain gauges provide the most reliable ground truth for evaluating SREs, especially when we are concerned with users who experience precipitation at the local (street/field) scale on the ground. In this study, we used data from an independent network of automatic weather stations (AWSs) across Zambia (Figure 1), provided by the Zambia Meteorological Department. These station observations were not incorporated into the calibration of the SREs evaluated in this study. Rainfall measurements are recorded at 15-min intervals and are available with a 5-min latency, which enables their use for near-real-time monitoring. Prior to use, the data was rigorously quality-checked. First, timestamps of the gauge data were verified by comparing with satellite estimated data to ensure consistency across datasets. Then, suspicious high or zero values were identified through visual inspection of time series plots and accumulation curves and compared with neighbouring stations where available; values confirmed to be wrong were dropped. The final dataset comprised 27 stations distributed across the country, covering the period from 2019 to 2024. Although some stations still contained data gaps, the overall mean data availability is relatively high for this region, at approximately 68%. Because the analysis was restricted to timesteps overlapping with the SREs, no gap filling was performed.

2.2 | Methods

The evaluation of the SREs was conducted using multiple observational datasets described in Section 2 and complementary methodological approaches to ensure a comprehensive assessment of SRE performance. The methods used in the various analyses, including an assessment of the feasibility of using SRE-derived rainfall amounts (RoA) to estimate rainfall at gauge locations, are described below.

14698080, 2026, 5, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70253 by UK Centre For Ecology & Hydrology, Wiley Online Library on [16/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2.2.1 | Evaluation of the SREs Against AWSs

2.2.1.1 | Point-To-Pixel Comparison. To facilitate comparison between the gridded satellite-based precipitation products and point-based observations from the AWSs, a point-to-pixel approach was employed. Specifically, the gauge data were compared with the values taken from the nearest grid cells of each satellite product. We acknowledge that point-based gauge data may not be fully representative of the whole area captured within a satellite product's grid cell. The potential implications of this representativeness discrepancy are taken into account when discussing our findings presented in the results section. To ensure a fair and consistent comparison, all gridded products were regridded to the coarsest native resolution among them (0.1°) using first-order conservative remapping (Jones 1999) prior to point data extraction. To assess the influence of spatial scale on the performance of the satellite-based products, we also expanded the area considered when extracting point data by taking the maximum rainfall value within increasingly larger domains (approx. 10, 20, 40, 60, 80 and 100 km diameters). After extracting the 15-min data, all datasets were aggregated to 30-min intervals to reduce noise prior to analysis. Additionally, because data gap periods varied among datasets and to maximise the amount of data points compared, the comparison was not restricted to only time steps when all the datasets overlapped. Instead, comparisons for each product were conducted using time steps for which both the AWS and the corresponding satellite-based product were available. This resulted in approximately 630,000, 215,000 and 781,000 individual 30-min observations for CRR, IMERG-HQ and RoA, respectively, aggregated across all stations for the wet seasons (October–April) from 2019 to 2024.

2.2.1.2 | Evaluation Metrics. First, the distributions of the datasets were evaluated using quantile-quantile (QQ) plots and two-dimensional (2D) kernel density estimation (KDE). In this study, QQ plots compare the empirical distributions of gauge and SRE rainfall through their corresponding quantiles, providing a direct comparison across the full range of values. The analysis was restricted to 'hits', defined as timesteps where both gauge and SRE rainfall exceeded a given threshold. Agreement is indicated by points lying along the 1:1 line, while systematic deviations reflect biases; differences in lower and upper quantiles indicate discrepancies in light and extreme rainfall, respectively, and variations in slope reflect differences in variability.

To complement this, a 2D KDE was used to examine the joint relationship between gauge and SRE rainfall. KDE provides a non-parametric estimate of the joint probability density function, with results shown as density contours representing regions of equal data density. Alignment of high-density regions along the 1:1 line indicates good agreement, while deviations reflect biases and broader contours indicate greater variability.

The ability of the satellite-based products to retrieve a rainy event were then assessed using the 2×2 contingency table (Table 1), for varying rainrates thresholds, derived from percentiles of the data sets and for the different spatio-temporal aggregation windows (i.e., 30 min, 1-, 2-, 4- and 8 h). Like in Hill et al. (2020), percentile-based thresholds are preferred over fixed thresholds because they enable fair comparisons across datasets, account for seasonal variability, reduce climate-dependent biases and provide a more robust representation of extremes across both locations and products. Given that the datasets are dominated by dry periods, only events with rain rates exceeding 0.2 mm/h are included in the percentile calculations. From the contingency table, we computed the hit rate (the proportion of rainfall events correctly detected by the satellite-based products), the probability of false alarms (POFA; the proportion of non-events incorrectly identified as rainfall) and the Heidke Skill Score (HSS; the skill of the satellite products relative to random chance, i.e., the climatological probability of rain above a threshold). These categorical metrics, defined in Table 2, allow us to assess the ability of each product to detect rainfall events across different thresholds and to rank the products against themselves. Additionally, the Receiver Operating Characteristic (ROC) curve was used to further assess the ability of the products to distinguish rainfall events from non-events across different temporal aggregation windows (30 min; 1, 2, 4 and 8 h). The ROC curve summarises

TABLE 1 | 2×2 contingency table used to compare the occurrence of rainy days in the AWS and SRE datasets.

	Gauge ≥ threshold	Gauge < threshold
SRE ≥ threshold	Hit (H)	False alarm (F)
SRE < threshold	Miss (M)	Correct rejection (R)

Note: For each dataset, a day is classified as rainy if the recorded rainfall is greater than or equal to a dataset-specific percentile-based threshold.

TABLE 2 | Summary of the metrics derived from the contingency table above.

Statistic	Formula	Description
POD, [0, 1], **1**	$POD = \frac{H}{H+M}$	Proportion of observed rainy days correctly identified by SREs.
POFA, [0, 1], **0**	$POFA = \frac{F}{H+F}$	Proportion of SRE-detected rainy days not observed by gauges.
HSS, [−∞, 1], **1**	$HSS = \frac{2(HR-FM)}{(H+M)(M+R)+(H+F)(F+R)}$	Quantifies how well SREs perform compared to random chance.

Note: H, F, M and R are hits, false alarms, misses and correct rejections, respectively. The range and best score for each statistic are given in the first column in square brackets and bold numbers, respectively.

14698080, 2026, 5, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70253 by UK Centre For Ecology & Hydrology, Wiley Online Library on [16/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

performance by combining the hit rate and POFA across a range of thresholds, providing an overall measure of the discrimination capability of CRR, RoA and IMERG-HQ. Details of these evaluation metrics can be found in Wilks (2011).

2.2.2 | Inferring Rainfall Rates Based on the SRE Distribution

To infer rainfall at gauge locations from SRE-derived rainfall, a distribution-based approach was applied. For discrete RoA intensity bins, the corresponding distribution of gauge-observed rainfall was characterised using histograms and kernel density estimation (KDE; see Section 2.2.1.2). Selected percentiles (25th–99th) of the gauge rainfall distribution were then computed for each RoA bin, providing an estimate of the rainfall expected at the ground for a given SRE intensity range (e.g., 1–4, 3–6, 5–8 and 7-10 mm/day).

The framework we apply here can be interpreted in terms of the conditional distribution of gauge rainfall (Y) given SRE-derived rainfall (X). The conditional cumulative distribution function (CDF) is defined as

$$F_{Y|X}(y|x) = P(Y \leq y | X = x), \tag{1}$$

which represents the probability of observing gauge rainfall less than or equal to y given an SRE rainfall value $X = x$. In practice, this relationship was estimated empirically for discrete SRE bins B:

$$F_{Y|X\in B}(y) = P(Y \leq y | X \in B), \tag{2}$$

where B denotes a predefined RoA interval.

To summarise these conditional distributions, quantiles were computed as follows:

$$Q_p(Y | X \in B) = F^{-1}_{Y|X\in B}(p), \tag{3}$$

where Q_p denotes the p-th quantile ($p \in [0, 1]$). These conditional quantiles provide probabilistic estimates of 'gauge' rainfall for a given range of SRE intensities.

The conditional CDF was estimated empirically as follows:

$$\hat{F}_{Y|X\in B}(y) = \frac{1}{n_B} \sum_{i:X_i \in B} \mathbf{1}\left(Y_i \leq y\right), \tag{4}$$

where n_B is the number of observations within bin B and the expression represents the empirical cumulative distribution of gauge rainfall within that bin.

This approach is conceptually related to quantile-based distribution mapping techniques commonly used in precipitation studies (e.g., Gudmundsson et al. 2012; Themeßl et al. 2012). The use of percentiles enables the rainfall distribution to be characterised across its full range, which is particularly relevant given the skewed nature of precipitation data. The resulting percentile–intensity relationships were smoothed using a moving average filter to reduce sampling variability and highlight systematic trends.

The conditional quantile relationships can be interpreted as a probabilistic mapping between SRE-derived and gauge-observed rainfall. For a given RoA bin, the median (50th percentile) represents the typical rainfall expected at the gauge, while higher percentiles (e.g., 85th–99th) indicate the upper range of possible rainfall and the potential for extreme events and lower percentiles provide more conservative estimates. The spread between quantiles reflects the variability and uncertainty in the SRE–gauge relationship, with wider spreads indicating greater dispersion.

2.2.3 | Evaluation Against Proxies: Lightning and Cloud Top Temperature

Additional insight into the performance of the different datasets is provided by comparison with MTG lightning data. Specifically, we computed the probability of lightning as a function of rain rate. This comparison of CRR, RoA and IMERG-HQ with lightning data was carried out on a pixel-to-pixel basis, despite the datasets having different native spatial resolutions. Lightning data were first regridded to a 0.05° grid and pixels containing at least one lightning flash within each 15-min MSG acquisition interval were identified. Rainfall values from CRR, RoA and IMERG-HQ were then grouped into 1-mm bins (0–1, 1–2 mm, etc.) and for each bin, we assessed whether lightning occurred in the nearest 0.05° pixel. For RoA and CRR, the analysis used their native spatial grids. For IMERG-HQ, however, the lightning data were regridded to a regular 0.1° grid and lightning occurrence was checked in the closest grid cell. This approach ensures that the coarser-resolution IMERG-HQ dataset is not unduly penalised during the comparison.

To gain an understanding of the relationships between rain rate and cloud-top temperatures, we computed the fractional contribution of different cloud-top temperature ranges to total rainfall for each product. This comparison against cloud-top temperatures followed a similar approach to that used for the lightning data analysis. For each rainfall pixel in CRR, RoA and IMERG-HQ and AWS point, we extracted the minimum cloud-top temperature within a box centred on the nearest MSG pixel. The search box was set to 7×7 MSG pixels for CRR and RoA and 13×13 pixels for IMERG-HQ. The larger box sizes were chosen partly because RoA accounts for viewing effects, whereas CRR does not. For deep convective clouds, we estimate that parallax can introduce horizontal displacement errors of 1–2 pixels. Since IMERG-HQ has a coarser spatial resolution and therefore represents a larger footprint, the box size was doubled to maintain a comparable spatial context.

3 | Results

3.1 | Evaluation of Rainfall Rates and Accumulation Distribution

All products exhibit distributions broadly similar to those of the gauges, although some biases are evident, with IMERG-HQ and CRR tending to overestimate lower rainfall rates and underestimate higher rainfall rates and accumulations, whereas RoA

14698080, 2026, 5, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70253 by UK Centre For Ecology & Hydrology, Wiley Online Library on [16/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

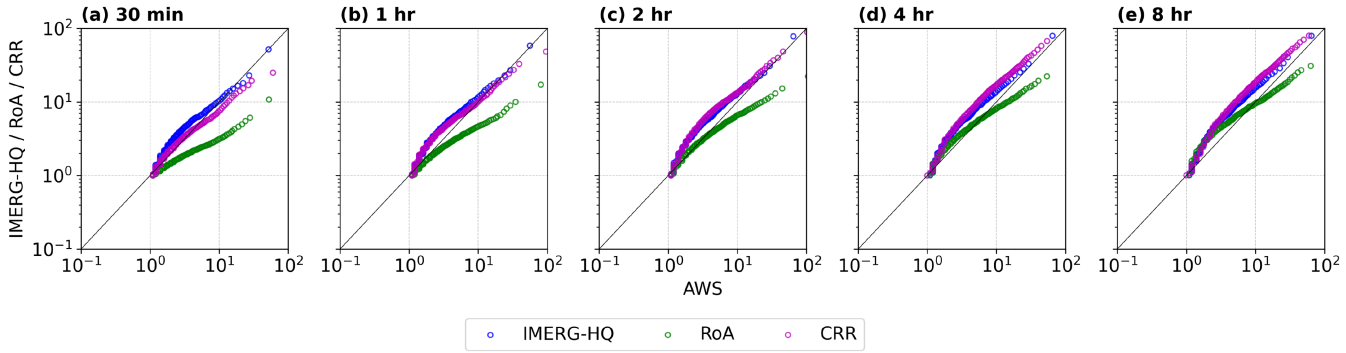

FIGURE 2 | Quantile–Quantile (Q–Q) plots comparing rainfall accumulation (mm) for the AWS gauges (*x*-axis) with those of IMERG-HQ (blue), CRR (red) and RoA (green) for different aggregation periods, that is, 30 min (a), 1h (b), 2 h (c), 4 h (d) and 8 h (e). A logarithmic scale is used to better illustrate differences at low rainfall rates.

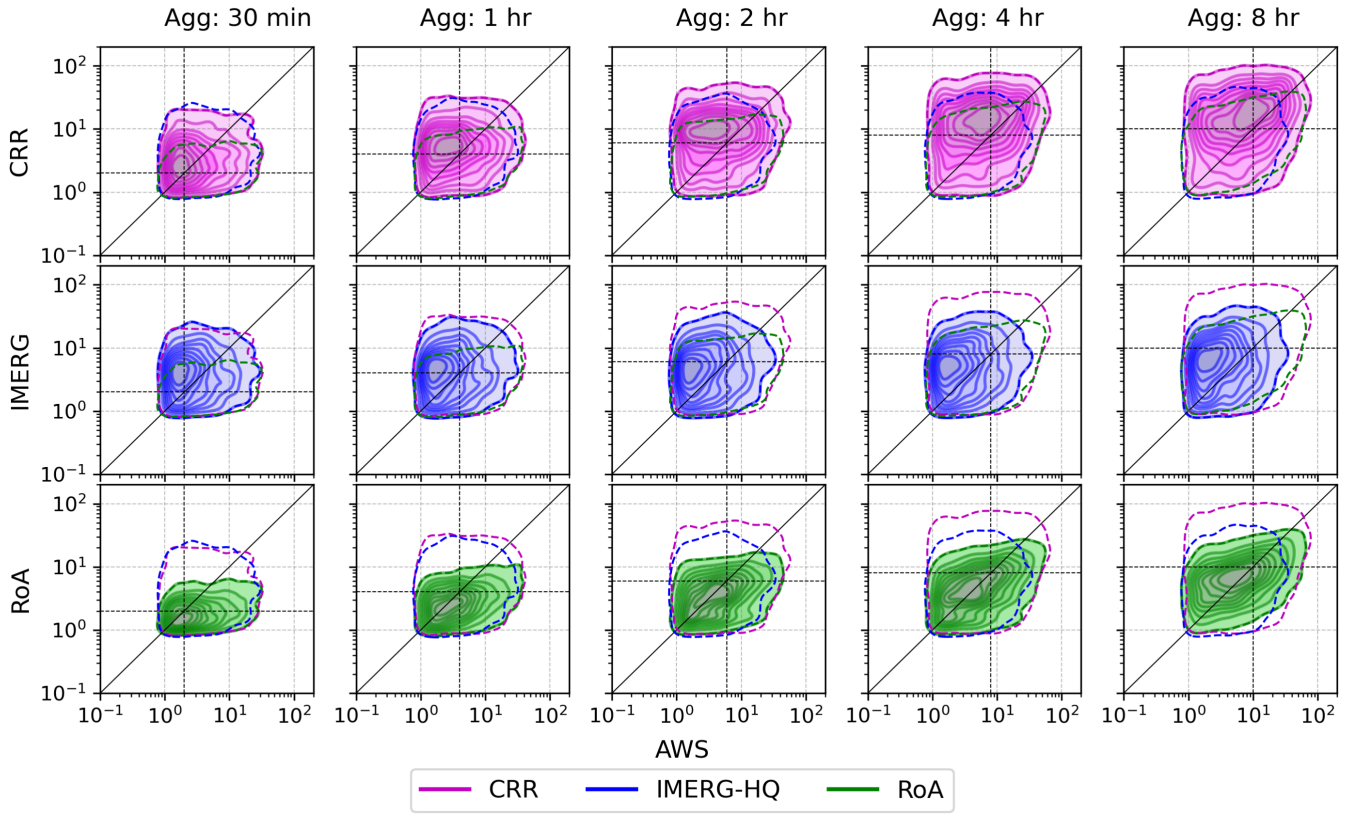

FIGURE 3 | Kernel Density Estimation (KDE) plots illustrating the distribution rainfall accumulation (mm) of CRR (red colour), IMERG (blue colour) and RoA (green colour) compared to the gauge observations from the AWS. The maximum contours (dotted line) for each of the three gridded datasets are overlaid on the plots for easy comparison. The five aggregation periods used are indicated at the top of each columns. Vertical and horizontal lines are equal to 2, 4, 6, 8 and 10 mm for the 30 min, 1, 3, 4 and 8 h aggregation windows.

generally underestimates rainfall across all rainfall intensities and accumulations (Figures 2 and 3). IMERG-HQ and CRR overestimate at lower rainfall rates, as evidenced by points lying above the 1:1 diagonal line (Figure 2). This tendency becomes more pronounced with longer aggregation periods and remains consistent across these extended timescales. The top row of Figure 3 shows that CRR consistently overestimates rainfall, as indicated by the highest density of points lying above the 1:1 line. This overestimation becomes more pronounced as the aggregation window increases. Similarly, IMERG-HQ (middle row) also tends to overestimate rainfall, though to a lesser extent than CRR. Notably, IMERG-HQ's density shifts towards lower rainfall intensities, suggesting it has a higher frequency of low rain rates.

In contrast, for RoA the points lie below the 1:1 line suggesting underestimation of rainfall rates, particularly during shorter aggregation periods (green-coloured points in 2). This bias diminishes over longer accumulation periods, with points progressively approaching the diagonal and even slightly overestimating low rainfall intensities starting from the 2-h aggregation period. A similar behaviour is confirmed in the 2D KDE plots (bottom row of 3), points mostly clustering below the diagonal, especially at shorter aggregation windows. RoA shows

14698080, 2026, 5, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70253 by UK Centre For Ecology & Hydrology, Wiley Online Library on [16/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

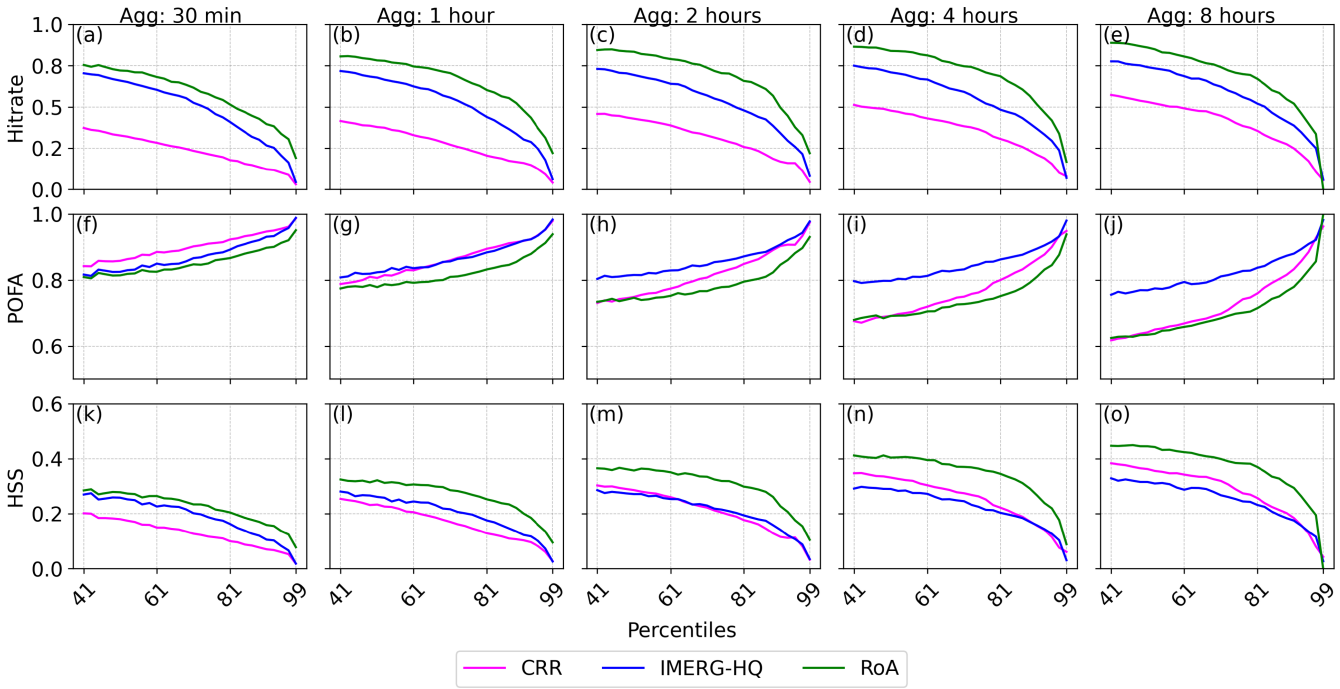

FIGURE 4 | Comparison of hit rates (a–e), POFA (f–j) and HSS (k–o) for CRR, IMERG-HQ and RoA against gauge observations over Zambia, across varying rainfall-accumulation percentile thresholds and accumulation periods. The percentiles are computed independently for AWS, CRR, IMERG-HQ and RoA and considering only values greater than 0.2 mm/h. This implies that the rain rate thresholds used to compute the metrics for the AWS and satellite-based products are not necessarily identical. This approach is adopted to mitigate the influence of systematic dry or wet biases among the datasets.

clear improvement as the aggregation window increases, with the density progressively moving towards the 1:1 line, indicating better agreement with gauge measurements. Based on the dotted lines in Figure 3, we observe that by the 4-h aggregation window, RoA provides the closest alignment with the gauge observations, reflecting the best overall agreement among the three products.

RoA demonstrates the best performance and highest skill in detecting rainfall events across all aggregation windows and percentile thresholds, followed by IMERG-HQ and CRR performs the least effectively (Figure 4). The performance differences are substantial, for example, at the 41st percentile and a 1-h aggregation window, RoA's hit rate exceeds those of CRR and IMERG-HQ by nearly 0.5 and 0.1, respectively (Figure 4a). Although all products exhibit high false alarm rates (POFA >0.8) at the 30-min aggregation window, RoA consistently records the lowest POFA across all aggregation windows. CRR exhibits the highest POFA at shorter aggregation windows, whereas IMERG-HQ records the highest values at longer aggregation windows. Consequently, RoA achieves the highest HSS values, exceeding 0.4 at the 4- and 8-h aggregation windows, while HSS values for IMERG-HQ and CRR do not surpass 0.4. At shorter aggregation windows, CRR has the lowest HSS, followed by IMERG-HQ. As the aggregation window increases, however, IMERG-HQ becomes the lowest-performing product, while CRR shows moderate improvement and RoA continues to outperform both products (bottom row of Figure 4). This shift is largely attributable to the influence of POFA on HSS, as increasing false alarm rates reduce overall skill scores. Across all products, hit rate, POFA and HSS deteriorate with increasing rainfall percentile, indicating greater difficulty in detecting more extreme rainfall events. Conversely, increasing the aggregation window improves all three metrics, likely because longer accumulation periods reduce temporal mismatches between satellite-based estimates and gauge observations (e.g., a 30-min timing offset is effectively absorbed within hourly or longer aggregations). Because the percentile thresholds are calculated independently for each SRE, differences in hit rate, POFA and HSS among the products are expected. CRR exhibits higher percentile values and relative to gauge observations, shows a systematic positive bias across all percentiles and aggregation windows, whereas IMERG-HQ and RoA generally exhibit negative biases (Figure A1).

Spatial scale influences the skill of the satellite-based precipitation products. All products achieve their highest Heidke Skill Score (HSS) at the smallest spatial scale (10 km), while the lowest HSS values occur at the largest scale (80 km; bottom row of Figure 5). This decline is primarily driven by the Probability of False Alarm (POFA), which increases with spatial scale despite improvements in hit rate. For CRR, the hit rate increases steadily with spatial scale, indicating progressively better agreement with gauge observations as the spatial domain expands, presumably as spatial mismatches are reduced. In contrast, both RoA and IMERG-HQ achieve higher hit rates than CRR across all spatial scales, with RoA consistently performing best. However, for these two products, the hit rate improves only up to a spatial scale of approximately 40 km, beyond which it reaches a plateau. As expected, POFA values (second row) generally increase with spatial scale for all products, reflecting the greater likelihood of detecting rainfall somewhere within a larger area that does

14698080, 2026, 5, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70253 by UK Centre For Ecology & Hydrology, Wiley Online Library on [16/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

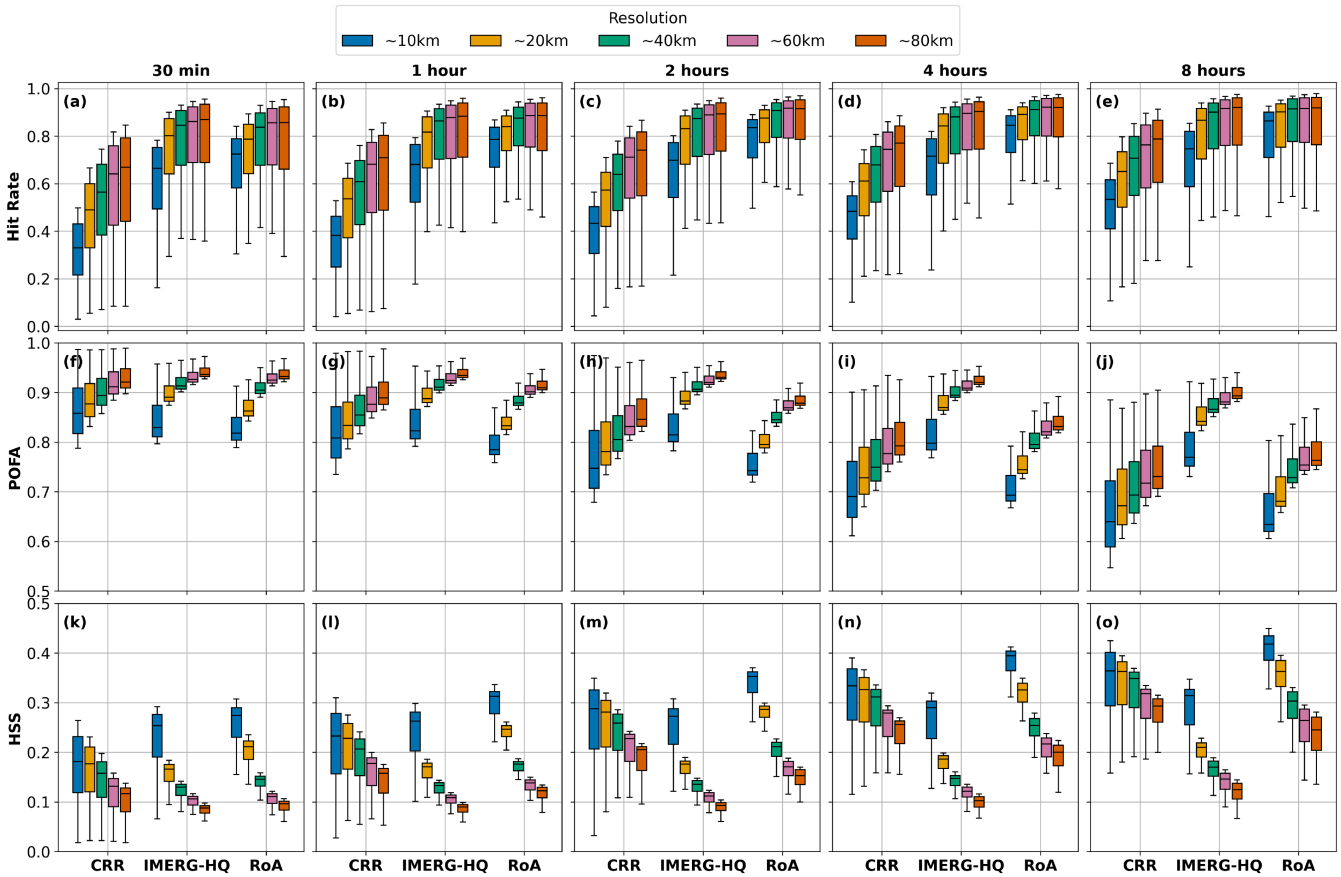

FIGURE 5 | Comparison of hit rates (a–e), POFA (f–j) and HSS (k–o) for CRR, IMERG-HQ and RoA against gauge observations over Zambia, across various rainfall intensity thresholds and accumulation periods and increasing spatial sizes.

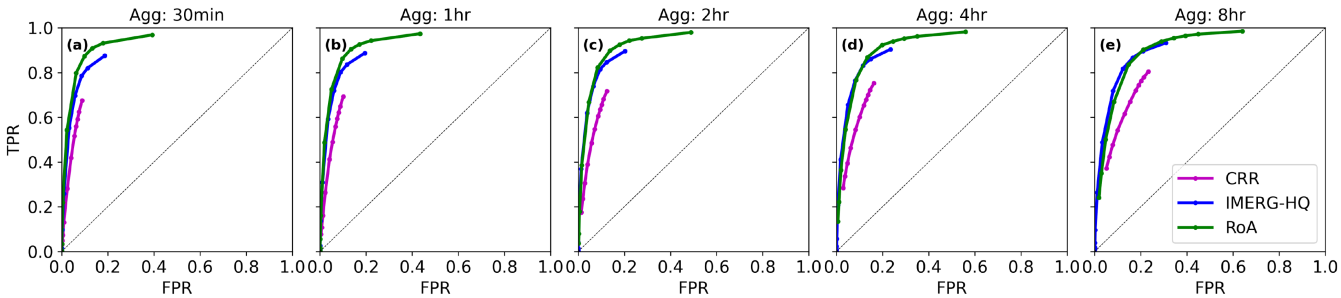

FIGURE 6 | Receiver Operating Characteristic (ROC) curves for CRR, IMERG-HQ and RoA, evaluated across multiple rainfall accumulation periods; 30 min (a), 1 h (b), 2 h (c), 4 h (d) and 8 h (e). The curves are generated by plotting the true positive rate (TPR) against the false positive rate (FPR) for varying decision thresholds. The dashed black 1:1 diagonal represents the performance of a random classifier. Curves lying above the diagonal are preferred as they indicate skillful discrimination.

not coincide with rainfall at the gauge location. For CRR, however, the degradation in POFA is partially compensated by its continuously increasing hit rate. Consequently, although CRR performs less well at smaller spatial scales and low aggregation windows, it achieves the highest HSS among the products at the largest spatial domains.

Results in Figure 6 further support the earlier findings: RoA consistently outperforms both IMERG-HQ and CRR across accumulation periods ranging from 30 min to 8 h in discriminating between rain and no-rain events. Its ROC curves lie closest to the top-left corner, indicating higher true positive rates at lower false positive rates. In contrast, CRR, which also relies on IR brightness temperatures to infer rainfall, performs less effectively, particularly at shorter accumulation periods, owing to its higher false alarm rate and lower sensitivity to short-duration rainfall events. IMERG-HQ, leveraging passive microwave observations, shows intermediate performance, benefiting from its physical basis. The overall trend of improved detection with longer accumulation periods indicates the benefit of temporal integration in reducing uncertainties and enhancing rain detection. This analysis not only confirms the better performance of the RoA retrieval, likely due to its training against GPM microwave data, but also highlights the trade-offs in each method, emphasising the need for careful product selection in operational forecasting and hydrological modelling.

14698080, 2026, 5, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70253 by UK Centre For Ecology & Hydrology, Wiley Online Library on [16/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

3.2 | Inferring Rainfall Rates on the Ground Based SRE Distribution

The distributions of gauge rainfall conditioned on RoA rainfall ranges (i.e., 1–4, 3–6, 5–8 and 7–10 mm/h) differ considerably, with the largest discrepancies occurring at the higher rainfall ranges (Figure 7a–d). For each RoA rainfall threshold, the gauge observations consistently exhibit peak values shifted to the left of the corresponding RoA distributions, indicating generally lower recorded rainfall amounts. For the lower RoA ranges (1–4 and 3–6 mm/h), the gauge distributions are strongly weighted towards low rainfall values, despite the corresponding RoA estimates indicating moderate rainfall intensities. As the RoA rainfall range increases, the divergence between the two distributions becomes more pronounced. In the 7–10 mm/h category (Figure 7d), gauge observations span a very wide range of rainfall intensities, extending from near-zero values to more than 35 mm/h, while the RoA estimates remain comparatively concentrated. This demonstrates that high RoA rainfall intensities can correspond to substantially different ground-observed rainfall amounts. Another notable feature is the increasing spread and variability of gauge rainfall at higher RoA ranges, suggesting growing uncertainty in the relationship between satellite-derived and ground-observed rainfall for intense events. However, we also note that at the higher rainfall categories contain fewer samples, reflecting the rarity of extreme rainfall events at the 30-min aggregation scale and the limited temporal coverage of the available observations.

Figure 7e summarises the expected gauge rainfall associated with different RoA rainfall ranges using conditional percentiles of the gauge distributions. As the RoA rainfall intensity increases, the expected rainfall recorded by the gauges also increases across all percentiles. However, the spread between percentiles becomes progressively larger at higher RoA ranges, indicating increasing variability and uncertainty in the corresponding ground-observed rainfall amounts. For lower RoA intensities (1–4 mm/h), the gauge rainfall distribution remains relatively compact, with most percentiles clustered within a narrower range. In contrast, for higher RoA intensities (7–10 mm/h), the gauge rainfall estimates span a much broader range, reflecting the substantial variability observed in the conditional distributions shown in Figure 7d. For example, while the median expected gauge rainfall remains comparatively moderate, the upper percentiles increase sharply, suggesting that high RoA values can be associated with both moderate and very intense local rainfall events.

3.3 | Comparison of CRR, IMERG and RoA and Gauge Rainfall Rates Against Lightning and Cloud-Top Temperature

The probability of lightning occurrence generally increases with rainfall intensity, although notable differences exist among the products, with CRR exhibiting no clear relationship between rainfall intensity and lightning occurrence (Figure 8a). RoA and IMERG-HQ show a strong positive relationship between rainfall intensity and lightning probability, with probabilities exceeding 0.6 for rainfall rates above approximately 20 mm hr^{-1}, after which the relationship begins to level off. The oscillations at higher rainfall rates for both products are likely due to the small number of events at these intensities, particularly for IMERG-HQ. In contrast, CRR shows almost no correspondence with lightning occurrence across all rainfall rates, with probabilities remaining close to zero (approximately 0.15 even at the highest rainfall rates; red line in Figure 8a). The lack of correspondence between CRR and lightning probability, even at higher rainfall rates (> 10 mm hr^{-1}), is concerning, particularly given that CRR is one of the primary nowcasting resources used in the region.

As expected, the rainfall distributions conditioned on cloud-top temperature show that the largest contribution to total rainfall

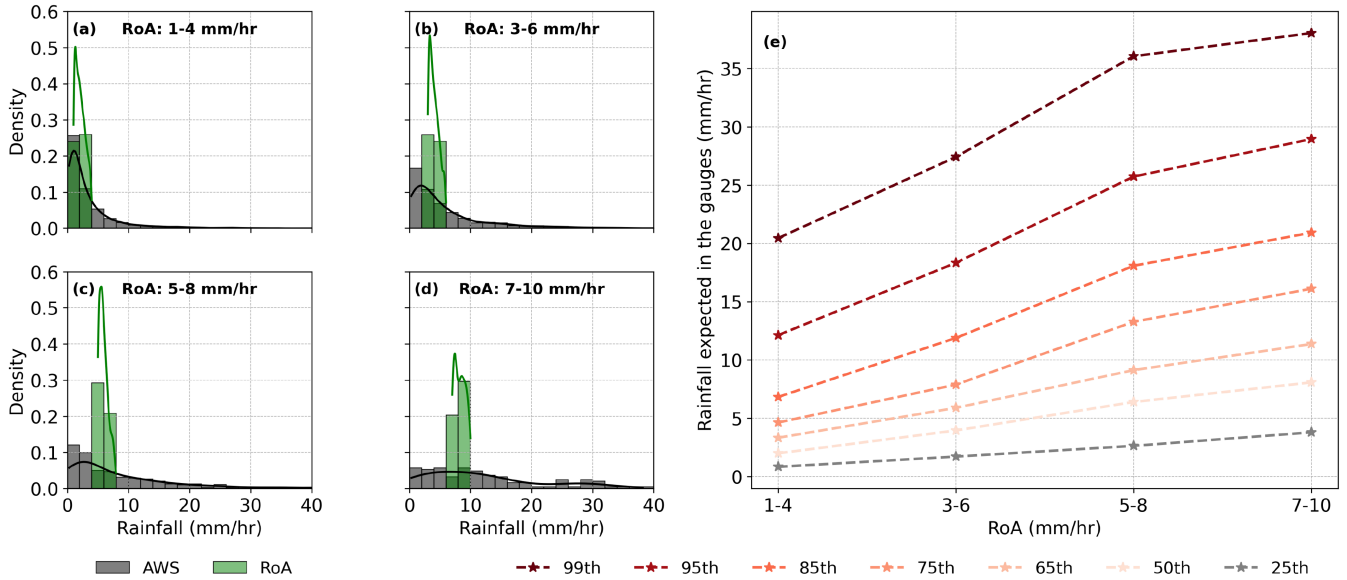

FIGURE 7 | Probability density functions of the gauge and RoA rainfall intensities for different RoA rain rate classes (a–d), computed using fixed 2 mm/h histogram bins over the range 0–51 mm/h, with overlaid kernel density estimates. The right-hand panel (e) shows the corresponding gauge-expected rainfall intensity percentiles (25th–99th) as a function of RoA rain rate class. Only coincident rainfall events are considered.

14698080, 2026, 5, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70253 by UK Centre For Ecology & Hydrology, Wiley Online Library on [16/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

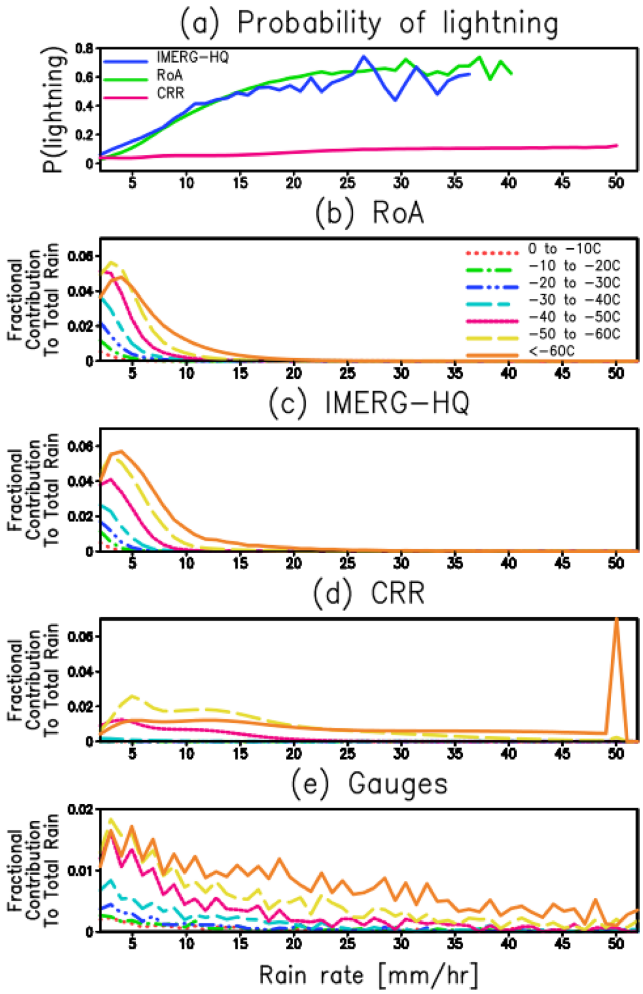

FIGURE 8 | Evaluation of satellite-based products against MTG lightning observations (a) and cloud-top temperature fields (b–d). (e) is a comparison of rain rates measured by gauges against cloud-top temperature. For the lightning comparison, the probability of lightning occurrence is shown as a function of SREs rain rates. Rain rates are grouped into 1 mm/h bins and lightning occurrence is evaluated within specified nearest-neighbour pixels. For the cloud-top analysis, the minimum cloud-top temperature within a defined neighbourhood around each nearest pixel is extracted. Each line represents the fractional contribution to total rainfall for each 10°C cloud-top temperature bin and 1 mm/h rainfall bin.

generally originates from the coldest cloud-top temperature ranges, although CRR exhibits a distinctly different distribution from the other two products (Figure 8b–e). RoA and IMERG-HQ retrieve rainfall across a broad range of cloud-top temperatures, including relatively warm cloud tops (> −30°C), although colder cloud tops contribute increasingly at higher rainfall intensities. In contrast, CRR retrieves little to no rainfall from cloud tops warmer than −30°C and is strongly dominated by colder cloud-top classes, particularly below −60°C. Gauge observations, however, show substantial rainfall contributions across all cloud-top temperature ranges, including warmer cloud tops, with flatter and more gradually decaying rainfall distributions than those of the satellite products (Figure 8e). However, caution should be exercised when visually comparing the gauge distributions with those of the SREs because the gauge panel uses a different *y*-axis scale. Overall, the gauge observations indicate greater rainfall variability than any of the SRE products, particularly at higher rainfall intensities.

4 | Discussion

4.1 | Distribution the Rain Rate and Accumulation

In agreement with Hill et al. (2020), all satellite-based products demonstrate useful skill in reproducing rainfall distributions similar to those derived from gauge data, however, systematic biases remain. CRR and IMERG-HQ tend to overestimate low-to-medium rainfall rates while underestimating higher intensities. In contrast, RoA generally underestimates rainfall across all intensity ranges and aggregation periods considered. Biases in satellite-based products in the region have been shown by past studies (e.g., Cattani et al. 2021; de Waal et al. 2025). Some of the products, rely on infra-red imagery based on cloud-top temperatures and hence, are prone to reporting rain for cold non-precipitating clouds like cirrus, leading to high POFA values, hence, overestimation rainfall and lowering the skill of the products. Hill et al. (2020) showed that although both IMERG and CRR detect rainfall more frequently at night, CRR displays distinct day-night rain rate distributions, with higher rain rates detected more often during the day due to the use of the 0.6 μm channel. Algorithms based primarily on cloud-top temperature cannot capture the full vertical structure of precipitating clouds and therefore miss important microphysical and dynamical information. As demonstrated for CRR in Section 3.3, such IR-only approaches show limited sensitivity and struggle to detect rainfall associated with moderately deep or complex convective systems. Biases also arise from missed events, particularly during warm-rainfall conditions. All products, especially CRR, are susceptible to missing warm-rain events because IR algorithms typically depend on cold-cloud thresholds to infer precipitation. As shown earlier, CRR does not detect any rainfall from cloud-top temperatures warmer than −30°C. Although warm-rain events are not usually as hazardous as severe convective thunderstorms, they can still generate impactful flash floods, especially when synoptic conditions lead to rainfall over many hours or where the ground is already saturated. The latter case was evident in the WISER EWSA Testbed 1 in Lusaka, where warm-rain episodes contributed to flooding in urban areas with inadequate drainage infrastructure (Ageet et al. 2024). The underestimation observed in RoA may be linked to its probabilistic algorithm design. Instead of applying fixed cold-cloud thresholds, RoA estimates a range of possible rainfall outcomes and assigns a high probability of zero rainfall when satellite signals are ambiguous, reducing final rainfall estimates (Amell et al. 2025). This smoothing approach can weaken extreme rainfall signals, resulting in high precipitation detection rates but lower gauge-measured accumulations.

Finally, the improved performance at longer rainfall accumulation periods is encouraging, as it benefits applications focused on cumulative rainfall amounts rather than precise event timing. Longer accumulation intervals are expected to reduce temporal mismatches between satellite estimates and gauge observations, which have been identified as a source

14698080, 2026, 5, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70253 by UK Centre For Ecology & Hydrology, Wiley Online Library on [16/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

of uncertainty in previous comparisons (e.g., Li et al. 2023; Mekonnen et al. 2021), particularly for products with short sampling intervals (e.g., 15 min). This is consistent with our results, as longer accumulation periods generally produce rainfall distributions that more closely resemble those derived from gauge observations.

4.2 | Ability of the SREs to Capture Extreme Rainfall

An important result of this study is that all three satellite-based products show limitations in retrieving extreme rainfall events, that are often the most hazardous. This under-detection of extremes has been documented in previous validation studies in the region (e.g., Ageet et al. 2022; de Waal et al. 2025) and can be attributed to several factors. First, IR-based products are unable to reliably detect the coldest cloud-top temperatures associated with the deepest convective cores (Maggioni et al. 2016). As shown for CRR, the proportion of rainfall retrieved from cloud tops colder than −60°C is only about half of that retrieved by IMERG-HQ, which exclusively makes use of microwave signals or RoA, which more effectively leverages its microwave-informed training. Secondly, extreme rainfall often originates from localised convective bursts with steep vertical gradients in hydrometeors (Minjarez-Sosa and Waissman 2017). These structures are difficult to resolve with the current spatiotemporal sampling of satellite sensors, particularly passive microwave instruments aboard sun-synchronous satellites, which provide only a limited number of overpasses per day. We also acknowledge that our point-to-pixel comparison may influence the results; satellite products are gridded and inherently smoothed (Risser et al. 2019), while point rain gauges can capture sharp, short-lived peaks characteristic of extreme events. Together, these factors contribute to the systematic under-detection of the most intense rainfall and underscore the need for region-specific calibration, higher-resolution retrievals and stronger integration of lightning and radar observations in future products.

4.3 | Comparative Performance of SRE Products for Nowcasting in the Region

Another key result of our study is that the new IR-only and ML-based product, RoA, mostly outperforms not only the near-real-time CRR, but also the post-processed and optimised IMERG-HQ, which makes use of microwave data to optimise its retrievals. RoA showed consistently strong performance across all validation metrics: RoA provided the highest event detection rates across all percentile thresholds and, together with IMERG-HQ, demonstrated greater sensitivity to a wider range of cloud-top temperatures and correlation to lightning data, indicating a more robust ability to capture diverse precipitating systems. Given that RoA is a fast ML algorithm, it therefore offers a substantially improved input for nowcasting applications. We attribute the superior performance of RoA to three primary factors: (a) RoA is trained using GPM-based radar and microwave observations, which also underpin the strong performance of IMERG-HQ, (b) the application of machine-learning techniques reduces systematic errors and optimises physical consistency and (c) RoA derives rainfall from inputs of multi-channel data over two-dimensional maps, meaning that the spatial structure of rainfall systems contributes to the computation of pointwise rainfall. In contrast, IMERG-HQ only uses local-pixel information to estimate the rainfall at that point. Similar results, in which ML-based SRE products outperform more traditional retrieval algorithms, have also been reported by (Brempong et al. 2025; Wang et al. 2026).

In this comparison, CRR demonstrated useful skill for nowcasting but performed less successfully than the other products. The CRR algorithm was originally calibrated using European radar networks and was specifically designed to exclude warm stratiform rainfall (Hernanz and Lahuerta 2019). Applying CRR to African climatic systems therefore represents an extrapolation beyond the conditions for which the algorithm was originally developed and calibrated. The weaker correlation of CRR with lightning activity, its comparatively large biases and its pronounced difficulty in detecting warm-rain events collectively suggest that the existing calibration of CRR is less well suited to the convective regimes typical of southern Africa than RoA, but we acknowledge that future recalibration of CRR with African radars may significantly improve its performance. Furthermore, CRR will soon be updated to use MTG data inputs, which should enhance its skill. Overall, our results underscore the need for regional calibration, as cloud and precipitation characteristics differ markedly across the globe and retrieval algorithms must be adapted accordingly to support effective early warning and nowcasting systems.

4.4 | Capability of SREs to Represent Ground-Based Rainfall Rates

The conditional rainfall distributions and percentile relationships demonstrate that, despite systematic biases, RoA retains meaningful information about the relative intensity and variability of rainfall events. Similar biases in satellite rainfall estimates (SREs), particularly for extreme rainfall events, have previously been reported across parts of tropical Africa (e.g., Ageet et al. 2022; de Waal et al. 2025). Although RoA is generally capable of distinguishing between lower- and higher-intensity rainfall regimes, the relationship between satellite-derived and gauge-observed rainfall becomes increasingly uncertain at higher rainfall intensities. This is reflected in the broader spread of gauge rainfall distributions and the increasing separation between percentile estimates for higher RoA ranges. The results further show that a single RoA rainfall range can correspond to a wide spectrum of ground-observed rainfall amounts, particularly for extreme events Figure 7a–d. These discrepancies likely arise from a combination of retrieval uncertainties, spatial-scale mismatches between satellite pixels and point-based gauge observations and the high spatial variability of convective rainfall. In addition, gridded precipitation products are known to smooth localised rainfall extremes, thereby reducing their ability to fully capture the magnitude and variability of intense rainfall events (Risser et al. 2019).

Consequently, deterministic interpretations of SRE rainfall magnitudes may be inappropriate for extreme rainfall applications. Instead, the results presented in Figure 7e support the use of probabilistic or distribution-based frameworks to translate SRE estimates into expected ground rainfall conditions. Such approaches may provide a more realistic characterisation of uncertainty for

14698080, 2026, 5, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70253 by UK Centre For Ecology & Hydrology, Wiley Online Library on [16/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

hydrological modelling, flood forecasting and early warning systems. However, the robustness of the higher-intensity relationships remains limited by the relatively small sample sizes associated with extreme rainfall events, owing both to the rarity of intense rainfall at the 30-min timescale and to the limited observational record available in this study. Extending the temporal record and increasing gauge network density would therefore help improve confidence in the estimated conditional rainfall relationships.

4.5 | Comparison Against Precipitation Proxies: Lightning and Cloud-Top Temperature Data

The relationships between rainfall intensity, lightning occurrence and cloud-top temperature shown in Figure 8 provide insight into how the different satellite rainfall products represent convective rainfall processes, since lightning activity is commonly associated with deep convective systems and intense precipitation (e.g., Minjarez-Sosa and Waissman 2017). The increasing probability of lightning occurrence with rainfall intensity for RoA and IMERG-HQ contrasts with CRR, which shows almost no correspondence between rainfall intensity and lightning activity.

The cloud-top temperature analysis further highlights important differences in retrieval behaviour between the products. RoA and IMERG-HQ retrieve rainfall across a broad range of cloud-top temperatures, including relatively warm cloud tops, whereas CRR retrieves little to no rainfall from cloud tops warmer than −30°C. This behaviour is largely consistent with the design of the CRR algorithm, in which warm stratiform clouds are suppressed (Hernanz and Lahuerta 2019). However, substantial rainfall can also occur from warmer cloud tops, particularly in tropical environments where warm-rain processes and shallow convection contribute significantly to precipitation, as observed during the WISER-EWSA Testbed I (Ageet et al. 2024). The gauge observations in Figure 8e support this, showing rainfall contributions across all cloud-top temperature classes, including warmer clouds. The flatter and more gradually decaying rainfall distributions in the gauge observations further suggest that surface rainfall variability is considerably larger than represented by the satellite products. Similar behaviour has been documented in previous evaluations of satellite rainfall estimates (SREs) and, besides algorithm flaws, has been attributed to spatial averaging within satellite pixels tending to smooth localised rainfall extremes and reduce rainfall variability relative to point-based gauge observations (Risser et al. 2019). Additional discrepancies may also arise from scale mismatches between gridded satellite products and point-based gauges, which are particularly important for convective rainfall at sub-daily timescales where rainfall variability is highly localised (Maggioni et al. 2016).

5 | Conclusions

During the WISER EWSA nowcasting testbeds, substantial positive feedback was received, highlighting the value of nowcasting services in southern Africa. However, this feedback needed to be complemented by a systematic, objective evaluation of the issued products to ensure that decisions are supported by accurate and reliable information. This study therefore evaluated the quality of a widely used nowcasting product, CRR, using in situ observations from a network of Automatic Weather Stations (AWS) in Zambia. In addition, two satellite rainfall estimates, the microwave-only product of IMERG and RoA, a newly developed ML-based dataset, are assessed to determine whether they offer a more reliable starting point for nowcasting applications than CRR. In particular, RoA is a product that can be computed and released in near real time and therefore offers a possible future operational nowcasting function.

In summary, we find that:

- Although CRR, IMERG-HQ and RoA demonstrate useful skill in reproducing rain-rate and rainfall-accumulation distributions similar to those observed by the gauges, all products exhibit systematic biases. In general, CRR and IMERG-HQ tend to overestimate rainfall at lower intensities or accumulations and underestimate it at moderate to high rain rates or accumulations. In contrast, RoA generally underestimates rainfall across all aggregation periods considered.
- Worryingly, all three satellite-based products show limitations in retrieving extreme rainfall events and exhibit substantially lower variability than the gauge observations, despite these events often being the most hazardous.
- Overall we find that the new IR-only and ML-based product, RoA, mostly outperforms not only the near-real-time CRR but also the post-processed and optimised IMERG-HQ, which makes use of microwave data to optimise its retrievals.
- Despite their biases, SREs still provide useful information that can be leveraged to estimate probabilistic (percentile-based) rainfall at the ground level, as demonstrated for RoA. Given these biases, this approach is akin to post-processing or downscaling of rainfall estimates, providing valuable information for applications such as early warning systems.
- IMERG-HQ and RoA retrieve rainfall across a broader range of cloud-top temperatures, spanning from warmer than −30°C to colder than −60°C. In contrast, CRR operates over a more limited temperature range and is unable to retrieve rainfall from clouds warmer than −30°C (partly by the algorithm design), which are often associated with intense rainfall events, particularly in tropical Africa.

Nowcasting represents a major opportunity for African weather services to provide the timely and high-quality short-range forecasts, up to 6 h ahead, that are urgently needed by vulnerable communities. The WISER EWSA project demonstrated both the demand for such information across Southern Africa and its alignment with the United Nations Early Warnings for All (EW4All) initiative. While nowcasts already outperform the classical daily forecasts typically issued by many NMHSs, especially in terms of spatio-temporal precision, our results underline the importance of continued validation and calibration to ensure their reliability in operational settings. Advances in technology, particularly in machine learning, highlight the need for continued reassessment and

14698080, 2026, 5, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70253 by UK Centre For Ecology & Hydrology, Wiley Online Library on [16/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

improvement of rainfall estimation and nowcasting products. Future developments using ML algorithms trained to exploit new-generation satellite observations, including MTG observations, offer further opportunities to enhance performance. Continued evaluation through regional initiatives such as the WISER EWSA testbeds will be essential for refining these capabilities, demonstrating operational value and building confidence among national meteorological services.

Author Contributions

Simon Ageet: conceptualization, methodology, writing – review and editing, writing – original draft, formal analysis, software. **Christopher M. Taylor:** conceptualization, methodology, writing – review and editing, formal analysis. **Adrià Amell:** methodology, software, data curation, writing – review and editing. **Edson Nkonde:** data curation, writing – review and editing. **John H. Marsham:** conceptualization, methodology, supervision, funding acquisition, writing – review and editing. **James Groves:** software, data curation. **Douglas J. Parker:** conceptualization, methodology, writing – review and editing, supervision, funding acquisition. **Abhilash Singh:** methodology, software, writing – review and editing. **Alexander J. Roberts:** methodology, software, writing – review and editing, data curation.

Acknowledgements

S.A. acknowledges support from the German Federal Ministry of Research, Technology and Space Grant 02WAS1713A 'Co-Design of a Hydrometeorological Information System for Sustainable Water Resources Management in Southern Africa (CO-HYDIM-SA)'. The work of A.A. has been supported in part by the European Union's HORIZON Research and Innovation Programme under Grant 101120657, project ENFIELD (European Lighthouse to Manifest Trustworthy and Green AI). Open Access funding enabled and organized by Projekt DEAL.

Funding

This work has been done under the WISER Early Warnings for Southern Africa (WISER-EWSA) project, funded by the Met Office as part of the Weather and Climate Information Services (WISER) programme on behalf of the UK government's Foreign, Commonwealth and Development Office (FCDO) and the NAIAR Project (NERC NE/Y000420/1).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Ageet, S., A. H. Fink, M. Maranan, et al. 2022. "Validation of Satellite Rainfall Estimates Over Equatorial East Africa." *Journal of Hydrometeorology* 23, no. 2: 129–151. https://doi.org/10.1175/JHM-D-21-0145.1.

Ageet, S., A. H. Fink, M. Maranan, and B. Schulz. 2023. "Predictability of Rainfall Over Equatorial East Africa in the Ecmwf Ensemble Reforecasts on Short-To Medium-Range Time Scales." *Weather and Forecasting* 38, no. 12: 2613–2630.

Ageet, S., D. Parker, H. Msemo, et al. 2024. *Weather and Climate Information Services Early Warning for Southern Africa (wiser-ewsa) Testbed 1*. University of Leeds. https://doi.org/10.48785/100/267.

Amell, A., L. Hee, S. Pfreundschuh, and P. Eriksson. 2025. "Probabilistic Near-Real-Time Retrievals of Rain Over Africa Using Deep Learning." *Journal of Geophysical Research: Atmospheres* 130, no. 20: e2025JD044595. https://doi.org/10.1029/2025JD044595.

Anderson, S. R., S. J. Cole, C. Klein, C. M. Taylor, C. A. Diop, and M. Kamara. 2024. "Nowcasting Convective Activity for the Sahel: A Simple Probabilistic Approach Using Real-Time and Historical Satellite Data on Cloud-Top Temperature." *Quarterly Journal of the Royal Meteorological Society* 150, no. 759: 597–617. https://doi.org/10.1002/qj.4607.

Ayzel, G., and M. Heistermann. 2025. "Brief Communication: Training of Ai-Based Nowcasting Models for Rainfall Early Warning Should Take Into Account User Requirements." *Natural Hazards and Earth System Sciences* 25, no. 1: 41–47. https://doi.org/10.5194/nhess-25-41-2025.

Bližňák, V., and Z. Sokol. 2026. "First Validation of the Lightning Imager Aboard Meteosat Third Generation Satellite With Earth Networks Total Lightning Network." *International Journal of Applied Earth Observation and Geoinformation* 147: 105205.

Bližňák, V., Z. Sokol, and P. Zacharov. 2017. "Nowcasting of Deep Convective Clouds and Heavy Precipitation: Comparison Study Between nwp Model Simulation and Extrapolation." *Atmospheric Research* 184: 24–34. https://doi.org/10.1016/j.atmosres.2016.10.003.

Brempong, E. A., M. A. Hassen, M. MohamedKhair, et al. 2025. Oya: Deep Learning for Accurate Global Precipitation Estimation. https://arxiv.org/abs/2511.10562.

Cattani, E., O. Ferguglia, A. Merino, and V. Levizzani. 2021. "Precipitation Products' Inter–Comparison Over East and Southern Africa 1983–2017." *Remote Sensing* 13, no. 21: 4419. https://doi.org/10.3390/rs13214419.

de Coning, E., M. Gijben, B. Maseko, and L. van Hemert. 2015. "Using Satellite Data to Identify and Track Intense Thunderstorms in South and Southern Africa." *South African Journal of Science* 111, no. 7–8: 1–5.

de Waal, J., M. Maranan, A. H. Fink, A. Watson, A. Bennett, and J. Helmschrot. 2025. "Validation of Satellite Precipitation Estimates Over the Western Cape Region, South Africa." *Journal of Hydrometeorology* 26, no. 6: 759–779. https://doi.org/10.1175/JHM-D-24-0123.1.

Erdmann, F., and D. R. Poelman. 2023. "Automated Lightning Jump (Lj) Detection From Geostationary Satellite Data." *Journal of Applied Meteorology and Climatology* 62, no. 11: 1573–1590. https://doi.org/10.1175/JAMC-D-22-0144.1.

EUMETSAT. 2017. Level 1.5 Image Data Format Description. Image, 127.

EUMETSAT. 2025a. *Mtg Li Level 2 Data Guide*. https://user.eumetsat.int/resources/user-guides/mtg-li-level-2-data-guide.

EUMETSAT. 2025b. *Using mtg Lightning Imager (li) to Track Convective Storms*. https://user.eumetsat.int/resources/user-guides/using-mtg-lightning-imager-li-to-track-convective-storms#ID-LI-data-in-convective-storms-nowcasting.

Funk, C., P. Peterson, M. Landsfeld, et al. 2015. "The Climate Hazards Infrared Precipitation With Stations—A New Environmental Record for Monitoring Extremes." *Scientific Data* 2: 150066. https://doi.org/10.1038/sdata.2015.66.

Gijben, M., and E. De Coning. 2017. "Using Satellite and Lightning Data to Track Rapidly Developing Thunderstorms in Data Sparse Regions." *Atmosphere* 8, no. 4: 67. https://doi.org/10.3390/atmos8040067.

GLOBE Task Team, D. A. Hastings, P. K. Dunbar, et al. 1999. *The global land one-kilometer base elevation (globe) digital elevation model, version 1.0*. Digital Database. http://www.ngdc.noaa.gov/mgg/topo/globe.html.

Gudmundsson, L., J. B. Bremnes, J. E. Haugen, and T. Engen-Skaugen. 2012. "Technical Note: Downscaling Rcm Precipitation to the Station Scale Using Statistical Transformations Ndash; a Comparison of

14698080, 2026, 5, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70253 by UK Centre For Ecology & Hydrology, Wiley Online Library on [16/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Methods." *Hydrology and Earth System Sciences* 16, no. 9: 3383–3390. https://doi.org/10.5194/hess-16-3383-2012.

Hermes, K., J. H. Marsham, M. Bollasina, M. Brooks, M. Klose, and F. Marenco. 2025. "Nowcasting of Dust and Convective Storms via Diffusion-Model Predictions of Seviri Rgb Imagery." *Weather and Climate Extremes* 50: 100828. https://doi.org/10.1016/j.wace.2025.100828.

Hernanz, A., and J. A. Lahuerta. 2019. *Algorithm Theoretical Basis Document for the Precipitation Product Processors of the nwc/geo*. EUMETSAT NWCSAF. https://www.nwcsaf.org/Downloads/GEO/2018/Documents/Scientific_Docs/NWC-CDOP2-GEO-AEMET-SCI-ATBD-Precipitation_v2.1.pdf.

Hill, P. G., T. H. Stein, A. J. Roberts, J. K. Fletcher, J. H. Marsham, and J. Groves. 2020. "How Skilful Are Nowcasting Satellite Applications Facility Products for Tropical Africa?" *Meteorological Applications* 27, no. 6: e1966.

Hou, A. Y., R. K. Kakar, S. Neeck, et al. 2014. "The Global Precipitation Measurement Mission." *Bulletin of the American Meteorological Society* 95, no. 5: 701–722. https://doi.org/10.1175/BAMS-D-13-00164.1.

Huffman, G. J., D. T. Bolvin, D. Braithwaite, et al. 2020. *Nasa Global Precipitation Measurement (gpm) Integrated Multi-Satellite Retrievals for gpm (imerg): Algorithm Theoretical Basis Document, Version 06*. NASA. https://gpm.nasa.gov/sites/default/files/2020-05/IMERG_ATBD_V06.3.pdf.

Huffman, G. J., E. F. Stocker, D. T. Bolvin, E. J. Nelkin, and J. Tan. 2023. Gpm Imerg Final Precipitation l3 Half Hourly 0.1 Degree x0.1 Degree v07. Accessed: 10 December 2025.

Jones, P. W. 1999. "First- and Second-Order Conservative Remapping Schemes for Grids in Spherical Coordinates." *Monthly Weather Review* 127, no. 9: 2204–2210. https://doi.org/10.1175/1520-0493(1999)127<2204:FASOCR>2.0.CO;2.

Kendon, E. J., R. A. Stratton, S. Tucker, et al. 2019. "Enhanced Future Changes in Wet and Dry Extremes Over Africa at Convection-Permitting Scale." *Nature Communications* 10: 1794. https://doi.org/10.1038/s41467-019-09776-9.

Klein, C., D. Belušić, and C. M. Taylor. 2018. "Wavelet Scale Analysis of Mesoscale Convective Systems for Detecting Deep Convection From Infrared Imagery." *Journal of Geophysical Research: Atmospheres* 123, no. 6: 3035–3050. https://doi.org/10.1002/2017JD027432.

Kummerow, C. D., D. L. Randel, M. Kulie, et al. 2015. "The Evolution of the Goddard Profiling Algorithm to a Fully Parametric Scheme." *Journal of Atmospheric and Oceanic Technology* 32, no. 12: 2265–2280. https://doi.org/10.1175/JTECH-D-15-0039.1.

Lamptey, B., S. Sahabi Abed, M. Gudoshava, et al. 2024. "Challenges and Ways Forward for Sustainable Weather and Climate Services in Africa." *Nature Communications* 15: 2664. https://doi.org/10.1038/s41467-024-46742-6.

Li, R., C. Guilloteau, P. E. Kirstetter, and E. Foufoula-Georgiou. 2023. "How Well Does the Imerg Satellite Precipitation Product Capture the Timing of Precipitation Events?" *Journal of Hydrology* 620: 129563.

Maggioni, V., P. C. Meyers, and M. D. Robinson. 2016. "A Review of Merged High-Resolution Satellite Precipitation Product Accuracy During the Tropical Rainfall Measuring Mission (trmm) Era." *Journal of Hydrometeorology* 17, no. 4: 1101–1117. https://doi.org/10.1175/JHM-D-15-0190.1.

Mekonnen, K., A. M. Melesse, and T. A. Woldesenbet. 2021. "Effect of Temporal Sampling Mismatches Between Satellite Rainfall Estimates and Rain Gauge Observations on Modelling Extreme Rainfall in the Upper Awash Basin, Ethiopia." *Journal of Hydrology* 598: 126467.

Milton, S., A. Diongue-Niang, B. Lamptey, C. Bain, C. Birch, and P. Bougeault. 2017. "Numerical Weather Prediction Over Africa." In *Meteorology of Tropical West Africa* edited by D.J. Parker and M. DiopKane, 380–422. Wiley. https://doi.org/10.1002/9781118391297.ch10.

Minjarez-Sosa, C. M., and J. Waissman. 2017. "A Survey and Perspectives on Mathematical Models for Quantitative Precipitation Estimation Using Lightning." *Advances in Meteorology* 2017, no. 1: 1351308. https://doi.org/10.1155/2017/1351308.

Parker, D. J., A. M. Blyth, S. J. Woolnough, et al. 2022. "The African Swift Project: Growing Science Capability to Bring About a Revolution in Weather Prediction." *Bulletin of the American Meteorological Society* 103, no. 2: E349–E369. https://doi.org/10.1175/BAMS-D-20-0047.1.

Pfreundschuh, S., P. Eriksson, D. Duncan, B. Rydberg, N. Håkansson, and A. Thoss. 2018. "A Neural Network Approach to Estimating a Posteriori Distributions of Bayesian Retrieval Problems." *Atmospheric Measurement Techniques* 11, no. 8: 4627–4643. https://doi.org/10.5194/amt-11-4627-2018.

Risser, M. D., C. J. Paciorek, M. F. Wehner, T. A. O'Brien, and W. D. Collins. 2019. "A Probabilistic Gridded Product for Daily Precipitation Extremes Over the United States." *Climate Dynamics* 53, no. 5: 2517–2538. https://doi.org/10.1007/s00382-019-04636-0.

Roberts, A. J., J. K. Fletcher, J. Groves, et al. 2022. "Nowcasting for Africa: Advances, Potential and Value." *Weather* 77, no. 7: 250–256.

San Roman Pineda, I., K. Vincent, W. Colenbrander, et al. 2025. *Wiser Ewsa King-Size Testbed Report*, University of Leeds. https://doi.org/10.48785/100/394.

Tan, J., W. A. Petersen, and A. Tokay. 2016. "A Novel Approach to Identify Sources of Errors in IMERG for GPM Ground Validation." *Journal of Hydrometeorology* 17: 2477–2491. https://doi.org/10.1175/JHM-D-16-0079.1.

Taylor, C. M., D. Belušić, F. Guichard, et al. 2017. "Frequency of Extreme Sahelian Storms Tripled Since 1982 in Satellite Observations." *Nature* 544, no. 7651: 475–478.

Themeßl, M., A. Gobiet, and G. Heinrich. 2012. "Empirical-Statistical Downscaling and Error Correction of Regional Climate Models and Its Impact on the Climate Change Signal." *Climatic Change* 112: 449–468. https://doi.org/10.1007/s10584-011-0224-4.

UN DESA. 2018. *World Urbanization Prospects: The 2018 Revision*. https://population.un.org/wup/.

Viticchié, B., S. E. Enno, D. Navia, J. Avbelj, and J. Grandell. 2026. "Meteosat Third Generation Lightning Imager Detection Performances." *EGUsphere* 2026: 1–35.

Vogel, P., P. Knippertz, A. H. Fink, A. Schlueter, and T. Gneiting. 2018. "Skill of Global Raw and Postprocessed Ensemble Predictions of Rainfall Over Northern Tropical Africa." *Weather and Forecasting* 33, no. 2: 369–388.

Wang, X., R. S. Alharbi, O. M. Baez-Villanueva, et al. 2026. Mswep v3: Machine Learning-Powered Global Precipitation Estimates at 0.1° Hourly Resolution (1979-Present). https://arxiv.org/abs/2602.01436.

White, C. J., H. Carlsen, A. W. Robertson, et al. 2017. "Potential Applications of Subseasonal-To-Seasonal (s2s) Predictions." *Meteorological Applications* 24, no. 3: 315–325. https://doi.org/10.1002/met.1654.

Wilks, D. S., ed. 2011. *Statistical Methods in the Atmospheric Sciences*. Vol. 100. 3rd ed. Academic Press.

WMO. 2017. *Guidelines for Nowcasting Techniques*.

WMO. 2025. *Wmo Radar Database (wrd)*. https://wrd.mgm.gov.tr/Home/Wrd.

Woodhams, B. J., C. E. Birch, J. H. Marsham, C. L. Bain, N. M. Roberts, and D. F. Boyd. 2018. "What Is the Added Value of a Convection-Permitting Model for Forecasting Extreme Rainfall Over Tropical East Africa?" *Monthly Weather Review* 146, no. 9: 2757–2780.

14698080, 2026, 5, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70253 by UK Centre For Ecology & Hydrology, Wiley Online Library on [16/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

World Bank. 2021. *A Regional Analysis of Weather, Climate, Water and Early Warning Services in Southern Africa: Status Quo and Proposed Actions*, University of Leeds. https://documents1.worldbank.org/curated/en/974411636364188920/pdf/Regional-Analysis-of-Weather-Climate-Water-and-Early-Warning-Services-in-Southern-Africa-Status-Quo-and-Proposed-Actions.pdf.

Žagar, N. 2017. "A Global Perspective of the Limits of Prediction Skill of Nwp Models." *Tellus A: Dynamic Meteorology and Oceanography* 69, no. 1: 1–13. https://doi.org/10.1080/16000870.2017.1282159.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Figure A1.** Percentile values for the three SREs along those of the AWSs for the five-aggregation window.

14698080, 2026, 5, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70253 by UK Centre For Ecology & Hydrology, Wiley Online Library on [16/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License