

# A Process-Based Validation of GPM IMERG and Its Sources Using a Mesoscale Rain Gauge Network in the West African Forest Zone

MARLON MARANAN, ANDREAS H. FINK, AND PETER KNIPPERTZ

*Karlsruhe Institute of Technology, Karlsruhe, Germany*

LEONARD K. AMEKUDZI AND WINIFRED A. ATIAH

*Kwame Nkrumah University of Science and Technology, Kumasi, Ghana*

MARTIN STENGEL

*Deutscher Wetterdienst, Offenbach, Germany*

(Manuscript received 31 October 2019, in final form 18 February 2020)

## ABSTRACT


Using a two-year dataset (2016–17) from 17 one-minute rain gauges located in the moist forest region of Ghana, the performance of Integrated Multisatellite Retrievals for GPM, version 6b (IMERG), is evaluated based on a subdaily time scale, down to the level of the underlying passive microwave (PMW) and infrared (IR) sources. Additionally, the spaceborne cloud product Cloud Property Dataset Using SEVIRI, edition 2 (CLAAS-2), available every 15 min, is used to link IMERG rainfall to cloud-top properties. Several important issues are identified: 1) IMERG's proneness to low-intensity false alarms, accounting for more than a fifth of total rainfall; 2) IMERG's overestimation of the rainfall amount from frequently occurring weak convective events, while that of relatively rare but strong mesoscale convective systems is underestimated, resulting in an error compensation; and 3) a decrease of skill during the little dry season in July and August, known to feature enhanced low-level cloudiness and warm rain. These findings are related to 1) a general oversensitivity for clouds with low ice and liquid water path and a particular oversensitivity for low cloud optical thickness, a problem which is slightly reduced for direct PMW overpasses; 2) a pronounced negative bias for high rain intensities, strongest when IR data are included; and 3) a large fraction of missed events linked with rainfall out of warm clouds, which are inherently misinterpreted by IMERG and its sources. This paper emphasizes the potential of validating spaceborne rainfall products with high-resolution rain gauges on a subdaily time scale, particularly for the understudied West African region.

## 1. Introduction


Human activities and socioeconomic stability in developing countries within the tropics are strongly influenced by the availability and variability of precipitation (UN 2009). Droughts and torrential rainfall belong to

the risks on the extreme sides of the rainfall spectrum and have distressed West Africa in the past few decades (Nicholson 1981; Lamb and Pepler 1992; Benson and Clay 1998; L'Hôte et al. 2002; Paeth et al. 2011; Panthou et al. 2014; Sanogo et al. 2015). Historically, rain gauges have been the most reliable source for the investigation of West African rainfall characteristics and trends (e.g., Nicholson et al. 2012). In the current age of remote sensing, spaceborne rainfall information is provided almost in real time and has mitigated the dependency on often sparsely available rain gauge data in Africa, where maintenance and accessibility have frequently become subject to the lack of political will, interest, or financial means. Thus, satellite-based precipitation estimates play a key role in the ongoing development of hydrological and numerical weather models as well as

---

 Denotes content that is immediately available upon publication as open access.

---

 Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JHM-D-19-0257.s1>.

---

*Corresponding author:* Marlon Maranan, [marlon.maranan@kit.edu](mailto:marlon.maranan@kit.edu)

DOI: 10.1175/JHM-D-19-0257.1

© 2020 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#) ([www.ametsoc.org/PUBSReuseLicenses](http://www.ametsoc.org/PUBSReuseLicenses)).

water resource management, which can help preventing rainfall-related socioeconomic losses (Thiemig et al. 2012).

A recent result of continuous technical advancement is the satellite-based, globally gridded rainfall product Integrated Multi-Satellite Retrievals for Global Precipitation Measurement (GPM) (IMERG; Hou et al. 2014; Huffman et al. 2015), which went operational in 2014 and builds upon the legacy of the Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA; e.g., Kummerow et al. 1998; Huffman et al. 2007). The fundamental idea behind IMERG is a seamless blending of passive microwave (PMW) and infrared (IR) information based on a large ensemble of satellite imagers and sounders (Huffman et al. 2019a). IR retrieval methods benefit from a high data sampling rate of radiometers aboard geostationary satellites, but correlate rainfall through an indirect relationship with cloud-top temperature (e.g., Arkin et al. 1994). PMW techniques, in turn, suffer from a lower sampling rate from satellites on low-Earth orbits, but are physically more direct and rely on the interaction between upwelling MW signals and precipitation-sized hydrometeors in clouds (Petty 1995; Kidd 2001; Kidd and Levizzani 2011). The resulting high spatiotemporal resolution ( $0.1^\circ \times 0.1^\circ$  and 30 min) on a global scale makes IMERG interesting for a wide range of hydrological applications (e.g., Gaona et al. 2016; Zubieta et al. 2017; Mazzoglio et al. 2019) and the investigation of convective phenomena, particularly in the tropics (e.g., Gaona et al. 2018; Maranan et al. 2019).

Passive rainfall retrieval techniques are inherently prone to errors and biases (Islam et al. 2017), which are often region specific (McCollum et al. 2000; Petković and Kummerow 2017). The significance of IMERG as well as TMPA has led to a large number of validation efforts against ground-based rainfall observations on several time scales (e.g., Wolff et al. 2005; Nair et al. 2009; Wang and Wolff 2010; Karaseva et al. 2012; Chen et al. 2013; Mantas et al. 2015; Tan et al. 2016; Gaona et al. 2016; Xu et al. 2017), and in particular for the data-sparse African continent (e.g., Adeyewa and Nakamura 2003; Nicholson et al. 2003; Dinku et al. 2007; Roca et al. 2010; Jobard et al. 2011; Thiemig et al. 2012; Gosset et al. 2013; Pfeifroth et al. 2016; Dezfuli et al. 2017b,a; Monsieurs et al. 2018; Camberlin et al. 2019). A general conclusion that can be drawn from these studies is that IMERG and TMPA belong to the best rainfall products on monthly down to daily time scales. Much of the good performance has been credited to the monthly calibration against rain gauges, which has successfully led to an overall reduction of bias.

One ongoing challenge, however, is the question how spaceborne rainfall products perform on a subdaily time scale. Deficiencies in the observations of single rainfall

events eventually lead to erroneous rainfall amounts on larger time scales unless gauge calibration mitigates this issue. Thus, understanding the sources of errors on the shortest possible time scale is a key element in improving the overall product (Huffman et al. 2007). In the case of the densely populated West Africa, there is a general shortage of spatiotemporally high-resolution validation sources for rainfall, such as rain gauges and radars, as well as sources for environmental conditions, such as in situ weather stations and radiosondes (Fink et al. 2011), and only few studies analyzed the behavior of IMERG/TMPA for this region on a subdaily time scale. Dezfuli et al. (2017b) investigated the performance of IMERG compared to TMPA with high-resolution rain gauges from the Trans-African Hydrometeorological Observatory (TAHMO) project (van de Giesen et al. 2014) based on different rainfall types in West Africa. Owing to the higher spatiotemporal resolution, they concluded that IMERG has improved from TMPA in capturing the distributions of rainfall rates, especially during intense rainfall events, which is a known weakness of TMPA (Monsieurs et al. 2018). Furthermore, over some well-gauged West African sites, Pfeifroth et al. (2016) recently highlighted a delay in the diurnal rainfall cycle within multisatellite-based products such as TMPA, which largely originate from the underlying IR data sources. In this context of source-specific uncertainties, Tompkins and Adebisi (2012) found that TMPA reacts to deep cloud structures in the coastal area with more enhanced rainfall than products based purely on PMW data, with the latter being more sensitive to high ice content in Soudano-Saharan cloud systems than TMPA. Consequently, the works of Tan et al. (2016) and Gebregiorgis et al. (2017) recommend an individual evaluation of the underlying PMW and IR sources, ideally for different seasons, in order to detect error cancellation effects. Analyzed for North America, IR tends to produce higher magnitudes in misses and false alarms than PMW, the latter of which, however, exhibits varying error contributions between the summer and winter season.

The aim of this work is to build upon aforementioned validation strategies to identify and deduce sources of errors in IMERG at its half-hour time scale for the understudied West African forest zone. In the framework of the Dynamics-Aerosol-Chemistry-Cloud Interactions in West Africa (DACCWA) project (Knippertz et al. 2015, 2017; Flamant et al. 2018), a dense network of 17 one-minute rain gauges was established in southern Ghana in 2015, which will serve as the validation dataset. The region is a suitable testbed for the validation of IMERG because of the diversity of the rainy and dry seasons, and the occurrence

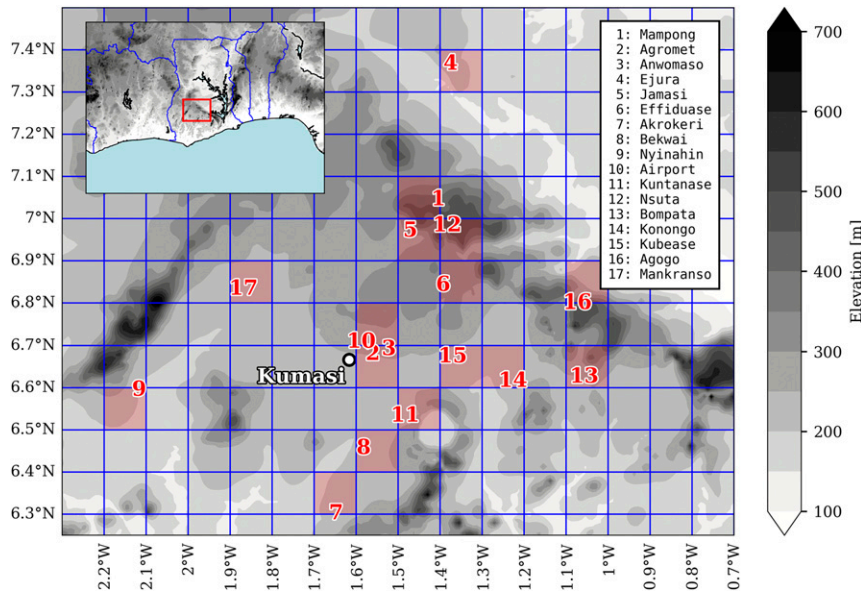


FIG. 1. Distribution of the DACCWA RG stations in the Ashanti region of Ghana (see map inset in the upper-left corner). Each red number represents an RG with names given in the top-right inset. The blue grid illustrates the IMERG pixels. RG-associated IMERG pixels that are used for the point-to-pixel comparison are denoted by the light-red shading. The gray shaded background shows the topography of the study region as provided by the Global Land One-km Base Elevation Project (GLOBE; Hastings et al. 1999).

of different rainfall types throughout the year (Hamilton et al. 1945; Eldridge 1957; Kamara 1986; Fink et al. 2006; Janiga and Thorncroft 2014; Maranan et al. 2018). In a further step, IMERG rainfall is linked to various microphysical cloud-top properties. This unique approach, that is, a subdaily, seasonal-, rainfall-type-, IMERG-source-, and cloud-property-based evaluation, can provide valuable insights into the behavior, strengths, and deficiencies of IMERG.

This study is structured as follows: After a description of the datasets and evaluation methods in sections 2 and 3, general characteristics of rainfall in the rain gauges and IMERG rainfall are given in section 4 before the performance of IMERG is evaluated in section 5. The latter is further decomposed from the perspective of different IMERG sources (section 6). Finally, the link to cloud properties is presented section in 7, before the manuscript is concluded with a discussion and summary in sections 8 and 9, respectively.

## 2. Data

### a. IMERG V6B

IMERG V6B, final version (IMERG hereafter, unless noted otherwise; Huffman et al. 2019b), is a Level 3 globally gridded precipitation product that combines data from several sources within the GPM satellite constellation.

It includes the GPM *Core Observatory* satellite with a dual-frequency precipitation radar and the 13-channel PMW imager GMI, multiple partner PMW instruments, and IR information from geostationary satellites.

Rainfall estimates in IMERG are processed on a  $0.1^\circ$  grid (blue grid in Fig. 1) every 30 min. The IMERG algorithm builds on the satellite merging techniques applied in its predecessor TMPA (Huffman et al. 2007, 2010). After an initial calibration of all partner PMW sensors toward rainfall estimates of the GPM/TRMM Combined Radar-Radiometer (CORRA), they are merged from their native spatial resolution onto the Level 3 IMERG grid at every half-hour time step. In regions without a direct PMW overpass, PMW observations are spatiotemporally “morphed” forward and backward using water vapor motion vectors from the hourly available reanalysis product Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2; Gelaro et al. 2017), similar to the principle of the Kalman filter (KF)-based Climate Prediction Center (CPC) morphing technique (CMORPH-KF; Joyce and Xie 2011). Beyond a “forecast” time of  $\pm 30$  min from the closest PMW observation, estimations from PMW-calibrated IR information based on the principles of PERSIANN-CCS (Hong et al. 2004) are additionally included (Huffman et al. 2019c). In a last step, monthly IMERG

estimates are calibrated toward rain gauge data from the Global Precipitation Climatology Centre (GPCC; Schneider et al. 2008).

In similar fashion to Tan et al. (2016), three categories of IMERG observations are considered: 1) direct PMW overpasses (PMW-direct hereafter), 2) pure PMW morphing (MORPH-only), and 3) a mixture of morphed PMW and IR (MORPH+IR). As seen later, a fourth category, IR-only, is not evaluated due to its low sample size. Within the IMERG output variable “precipitationCal” (containing the gauge-calibrated precipitation field), these categories can be discriminated using the auxiliary variables “HQprecipitation” and “IRkalmanFilterWeight.” While the former is used to identify “PMW direct” areas, the latter refers to the weight of IR observations wherever “PMW-direct” is absent. It ranges from 0% (MORPH-only) to 100% (IR-only).

### b. Rain gauge dataset

In the framework of the DACCIWA project, a total of 17 optical rain gauges (RGs hereafter) were installed within a radius of approximately 80 km around the city of Kumasi in the Ghanaian forest zone (Fig. 1) and went fully operational in December 2015. Ten RG sites coincide with rain gauge stations operated by the Ghana Meteorological Agency (GMet). The rest were placed on secured school yards.

The RG instrumentation operates on the principle that rainwater is funneled through a rain collector, forming drops equal to 0.01 mm of rainfall. These are counted by an IR sensor and stored in a logger every minute. Comparable RG networks in West Africa with such a high precision only exists in the framework of African Monsoon Multidisciplinary Analysis–Coupling the Tropical Atmosphere and the Hydrological Cycle (AMMA-CATCH; Lebel et al. 2009) and the TAHMO project (van de Giesen et al. 2014). The upper bound of measurable rainfall rate is approximately  $300 \text{ mm h}^{-1}$ , which would cause a water stream rather than the formation of drops.

For the present study, quality-controlled RG data from 2016 and 2017 are used for validation. The quality control was performed on daily rainfall and followed two steps. First, a manual removal of obviously erroneous periods, such as unrealistic values or long periods of obvious failed recordings, was performed by comparison with neighboring RGs. Second, daily RG rainfall was compared with collocated GMet data. While no specific threshold value was applied, days that exhibit a strong deviation to GMET were removed. Although valuable rainfall data exist for large parts of the two years, intermittent power outages and other issues due to electronics and environmental influences caused episodes of missing data (Fig. S1 in the online supplemental

material). Larger data gaps exist from September 2016 to May 2017, when data were temporarily obtained from only seven RGs. Therefore, RGs with longer data records may have a stronger influence in the skill measures (Monsieurs et al. 2018). Since no rainfall data from these RGs were ingested into the Global Telecommunication System, they were not part of the monthly IMERG gauge calibration and thus serve as an independent validation source. The raw rainfall data used in the present study are available under <https://doi.org/10.6096/baobab-dacciwa.1772>.

### c. CLAAS-2

To investigate cloud properties around rainy episodes, RG and IMERG rainfall is linked to cloud-top information from the Cloud Property Dataset Using SEVIRI, edition 2 (CLAAS-2) dataset (Stengel et al. 2014; Benas et al. 2017). CLAAS-2 is compiled by the Satellite Application Facility on Climate Monitoring (CM SAF), which processes data from the multichannel Spinning Enhanced Visible and Infrared Imager (SEVIRI) on board the Meteosat satellite with a spatiotemporal resolution of 3 km (at nadir) and 15 min, respectively (Aminou 2002). We make particular use of three quantities: 1) the cloud optical thickness (COT) in the visible spectrum, increasing with stronger scattering by water droplets and ice crystals (Glickman 2000); 2) the IR cloud-top brightness temperature (CTT); and 3) the cloud drop effective radius ( $R_{\text{eff}}$ ), defined as the weighted mean of the droplet size distribution (Hansen and Travis 1974). All values are taken at the nearest grid points and closest time stamps of the rainfall events.

The retrieval of the cloud properties follows the scheme described in Roebeling et al. (2008). Initially, the cloud phase at a given cloudy pixel is determined through several threshold tests with observed and simulated IR brightness temperature fields, which ultimately yields a flag (“liquid” or “ice”). Through an iterative matching algorithm similar to that described in Nakajima and King (1990),  $R_{\text{eff}}$  and COT are then estimated using lookup tables of simulated reflectances for liquid or ice phase at the wavelengths 0.6 and  $1.6 \mu\text{m}$ . While liquid droplets are assumed to be spherical with  $R_{\text{eff},l}$  ranging between 3 and  $34 \mu\text{m}$ , ice particles are considered to be monodisperse, hexagonal, and randomly orientated with  $R_{\text{eff},i}$  values from 5 to  $80 \mu\text{m}$  (CM SAF 2016). In both cases, the maximum of COT is set to 100. Beyond this value, COT becomes indistinguishable from higher values for a given  $R_{\text{eff}}$ . Combining COT and  $R_{\text{eff}}$ , the liquid and ice water path (LWP, IWP), that is, the vertically integrated amount of liquid and frozen water droplets, respectively (Glickman 2000), can be estimated via (Stephens 1978; Benas et al. 2017):

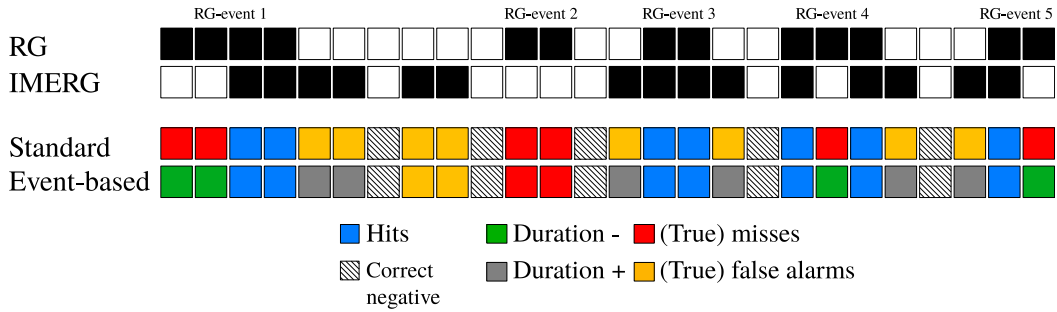


FIG. 2. Schematic showing how hits, misses, and false alarms are defined based on an exemplary half-hourly rainfall pattern in (first row) RG and (second row) IMERG with wet (black,  $\geq 0.2 \text{ mm h}^{-1}$ ) and dry time steps (white). (third row) The standard approach designates misses (false alarms) wherever a rainy time step in RG (IMERG) is associated with a dry time step in IMERG (RG). (fourth row) In the event-based approach, misses (false alarms) in adjacent time steps of hits are considered as a reduction (prolongation) of the event duration, hence termed Duration- (Duration+). ‘‘Isolated’’ errors are called true misses and false alarms, respectively.

$$\text{LWP, IWP} = \frac{2}{3} \rho_{(l,i)} \text{COTR}_{\text{eff},(l,i)}, \quad (1)$$

where  $\rho_{(l,i)}$  are the densities of water and ice, respectively. Note that since the retrieval of  $R_{\text{eff}}$  and COT require solar radiation, both can be determined only during daytime.

### 3. Methods

#### a. Measures for point-to-pixel validation

IMERG is validated on a half-hourly point-to-pixel basis by taking the closest grid cell to the respective RGs (e.g., Thieme et al. 2012). It shall be stressed that point measurements by RGs sometimes lack representativeness of the averaged rainfall in satellite pixels, which presumably becomes less severe with increasing resolution in satellites (Tan et al. 2016; Monsieus et al. 2018). In the present setting, only one IMERG pixel contains more than one RG for a potential investigation of intrapixel variabilities. Potential effects on the results are discussed in section 8. Hence, while acknowledging this caveat, no further processing such as spatial averaging of RG data is performed. Half-hour intervals with an aggregated amount of less than 0.1 mm ( $0.2 \text{ mm h}^{-1}$ ) are discarded to account for potential noise in the RG dataset. The same threshold is applied to IMERG, which corresponds to the minimum detectable rainfall rate of the GPM Ka-band radar (Tan et al. 2016).

Two groups of statistical measures are used. The first group is derived from the  $2 \times 2$  contingency table with hits  $H$  (rainfall in both RG and IMERG), misses  $M$  (rainfall in RG only), false alarms  $F$  (rainfall in IMERG only), and correct negatives  $N$  (zero rainfall in both RG and IMERG) (see Fig. 2). The probability of detection (POD), probability of false alarms (POFA), bias in

detection (BID), and the Heidke skill score (HSS) are then defined by (see Wilks 2011)

$$\text{POD} = \frac{H}{H + M}, \quad (2)$$

$$\text{POFA} = \frac{F}{H + F}, \quad (3)$$

$$\text{BID} = \frac{H + F}{H + M}, \quad (4)$$

$$\text{HSS} = \frac{2(HN - FM)}{(H + F)(F + N) + (H + M)(M + N)}. \quad (5)$$

POD quantifies the ability of IMERG to detect rainy episodes as recorded by the RGs and is perfect when  $\text{POD} = 1$ . Similarly, POFA is the fraction of false alarms relative to all rainfall occurrences in IMERG. If no false alarms are produced, then  $\text{POFA} = 0$ . BID determines whether IMERG tends to overestimate ( $\text{BID} > 1$ ) or underestimate ( $\text{BID} < 1$ ) the rainfall frequency. Finally, the HSS evaluates the performance of IMERG compared to random chance. A value of  $\text{HSS} = 1$  indicates maximum skill, a value of  $\text{HSS} = 0$  means no skill. Technically, the HSS can become negative, which would imply a lower skill of IMERG than random draws.

As in Tan et al. (2016), the second group of measures compares the rainfall rates from the subset of hits, where the mean error (ME) and mean absolute error (MAE) and their normalized counterparts, NME and NMAE, are calculated via

$$\text{ME} = \frac{1}{n} \sum_i (y_i - x_i) \quad \text{and} \quad \text{NME} = \frac{\frac{1}{n} \sum_i (y_i - x_i)}{\frac{1}{n} \sum_i x_i}, \quad (6)$$

$$\text{MAE} = \frac{1}{n} \sum_i |y_i - x_i| \quad \text{and} \quad \text{NMAE} = \frac{\frac{1}{n} \sum_i |y_i - x_i|}{\frac{1}{n} \sum_i x_i}, \quad (7)$$

where  $x_i$  and  $y_i$  denote a pair of RG and IMERG rain rates, and  $n$  the number of hits. All error measures are perfect if 0. While MAE quantifies the overall error magnitude, ME indicates the direction of the bias. Through normalization related to a background climatology of rain rates, the error magnitudes become comparable, for instance, for different rainfall rates across different seasons.

#### b. Identification and definition of rainfall types

In addition to half-hourly rainfall, IMERG's performance for different rainfall types is investigated. Here, the RG network is considered as a unit, meaning that spatiotemporally coherent signals at several RGs can be associated to the same rainfall event. The high temporal resolution of the RGs then allows an assignment to specific rainfall types.

First, the identification of rainfall events follows the correlation-regression method by Upton (2002), for which the time series of all available RGs were aggregated to 5-min data. Each rainfall event is then categorized into one of three rainfall types based on the definitions in Dezfuli et al. (2017b). Weak convective rainfall (WCR) exhibits a mean rainfall rate and duration of less than  $10 \text{ mm h}^{-1}$  and 80 min, respectively. Accordingly, strong convective rainfall (SCR) is defined for events with at least  $10 \text{ mm h}^{-1}$ . Any event exhibiting at least 80 min of uninterrupted rainfall at one RG or more is classified as a mesoscale convective system (MCS). Again, RGs affected by the same event are considered together. For instance, if the rainfall profile at one station matches the criterion for an MCS, the profiles of all other stations are collectively assigned to MCS, even if they would not fulfill the criterion individually. That way, we believe that a reasonable quantification of number and integrated rainfall of each rainfall type can be obtained.

From the perspective of rainfall events, misses and false alarms are defined slightly differently compared to single half-hour time steps (see Fig. 2). Over the length of a given rainfall event in the RGs, a "true miss" is considered when no respective IMERG time step contains any rainfall. Otherwise, the duration of the rainfall event is cut short (Duration−). The same principle applies for "true false alarms" and Duration+. Finally, we note that a half-hour RG time step is considered as rainy as soon as rainfall is detected in at least one of the 5-min periods.

## 4. General characteristics of RG and IMERG rainfall

### a. RG-based rainfall types

A total of 2552 separate rainfall events were identified within the 2-yr period. Figure 3 shows how they fall into the rainfall categories described in the previous section. The bulk of events is short lived and has low intensity (Fig. 3a) with WCRs accounting for over half of all events (see % $n$  in the legend). Roughly a tenth can be attributed to longer-lasting MCSs, but these account for over 60% of total rainfall, while WCRs contribute only 5% (see %RR). This pattern resembles the results in the satellite-based study of Maranan et al. (2018) for a broader domain in southern West Africa, where the contribution of frequent but small-scale convection is almost negligible.

The temporal evolution of rainfall rates during the passage of each rainfall type is depicted in Figs. 3b–d. It is usually marked by a sudden increase within the first 15 min followed by a more gradual weakening during the remainder of the event. We note that these profiles are highly variable as seen by the interquartile range (shaded areas). The enhanced rainfall rate in the early stages is clearly associated with the convective part of the rainfall system. It is strongest for SCRs (Fig. 3c), which likely comprise young, but vigorous convective cells. A major characteristic of MCSs is the extended trailing stratiform region, which can contribute substantially to their integrated rainfall amount (green curve in Fig. 3d). However, because of the weaker nature of this stratiform rainfall, the mean intensity of the strongest events decreases quasi-exponentially with longer event durations (Fig. 3a). Note that the intensity of the leading convective part is highly variable (cf. Dezfuli et al. 2017b), where some of the weaker events may be related to dissipating MCSs. For WCRs, a clear convective part cannot be identified in many cases, as they often last only 5–10 min. Also debris of decaying MCSs occasionally causes instances of weak and short events.

### b. Seasonal evolution of rainfall types in RGs and IMERG

The composition of rainfall types throughout the year changes depending on the stage of the West African monsoon (WAM; e.g., Fink et al. 2006; Janiga and Thorncroft 2014; Maranan et al. 2018). In Fig. 4a, the monthly evolution of both the overall number of events (green curve) and the respective fractions of the rainfall types are presented. Two number maxima are present in June and September, in line with the bimodal cycle typical of the West African forest zone

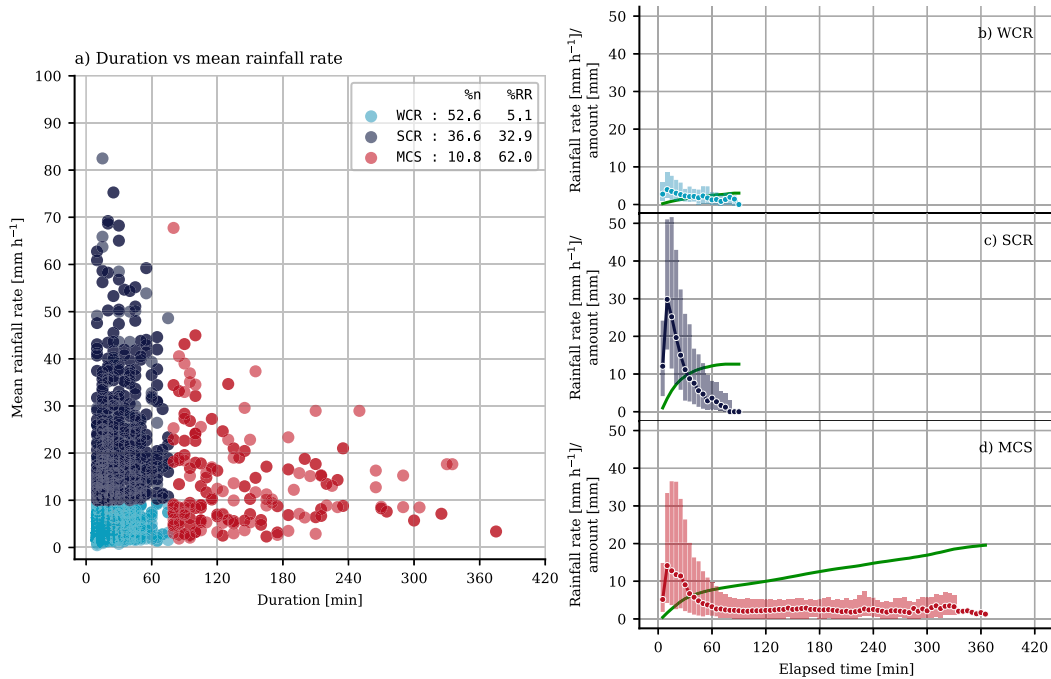


FIG. 3. (a) Scatterplot showing the duration and mean rainfall rate (i.e., integrated rainfall amount divided by the duration) of all individual rainfall events within the study period 2016–17. Different colors denote the rainfall types listed in the legend and as characterized by Dezfuli et al. (2017b). Here, %n and %RR designate the relative frequency and fractional rainfall contribution, respectively, both in %. (b)–(d) Marker plots showing composites of rainfall rates in 5-min bins against elapsed time of event for weak convective rainfall (WCR), strong convective rainfall (SCR), and mesoscale convective systems (MCSs), respectively. The bin median values are denoted by connected circles, whereas the interquartile range is given by the colored bars. The median for each bin is calculated by all available events, the number of which decreases toward longer durations (not shown). The green curve is the cumulative median rainfall.

(Fink et al. 2017). A local minimum in August indicates the so-called little dry season, where the relative frequency of WCRs strongly increases at the expense of SCR. The fraction of MCSs is less than those of WCRs and SCR in all months. It exhibits a distinct peak in April and an apparent decrease toward the long dry season beginning in November, but otherwise changes little throughout the year. Thus, in absolute numbers, MCSs exhibit a similar seasonal evolution as the event numbers. How the frequencies translate into the seasonal rainfall amount is depicted in Fig. 4b. First of all, the seasonal cycle of rainfall averaged over the two years and all available RGs (white curve) confirms the bimodal pattern of the event numbers. However, the pronounced intergauge spread, indicated by the standard deviation (dashed curves), emphasizes the high small-scale variability of monthly rainfall. The fractional rainfall of the individual types, indicated by the stacked bars, shows a seasonal pattern similar to the fractional number distributions, however, scaled in accordance with their respective intensity as shown in

Fig. 3. MCSs are the main contributor to rainfall, except for the long dry season where short intense rainfall events dominate. Remarkably, the high numbers of WCRs during the little dry season accounts for only little more than 10% of total rainfall.

The representation of seasonal rainfall and rainfall types in IMERG is evaluated in Fig. 4c. In general, IMERG is able to capture the fundamental characteristics well on a monthly scale (correlation coefficient  $CC = 0.98$ ). This is also true for the diurnal time scale (Fig. S2), which was already found to be well represented by IMERG in Dezfuli et al. (2017b). The high agreement in monthly rainfall is likely related to the gauge calibration, the latter of which is stronger over Ghana and Togo than elsewhere over West Africa for 2016 and 2017 (Fig. S3). During the rainy seasons, IMERG tends to underestimate monthly rainfall, causing large parts of the averaged root-mean-square error ( $RMSE = 14.05$  mm). At the same time, the interpixel variability ( $\sigma = 15.74$  mm, gray shaded area) is far less pronounced than the aforementioned intergauge variability ( $\sigma = 37.12$  mm, light-red shaded area).

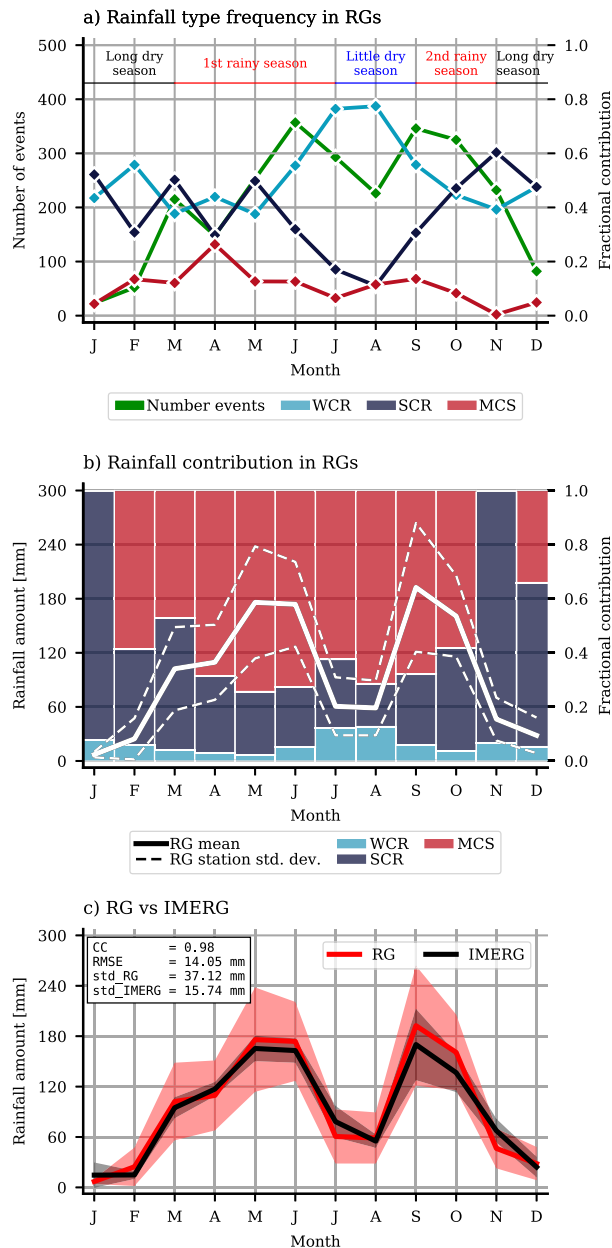


FIG. 4. Seasonal cycles of (a) the total number of individual rainfall events (green curve; left axis) and the fractional occurrence of the RG-based rainfall types (right axis). The respective WAM seasons are indicated at the top and are defined similarly to Maranan et al. (2018); (b) the fractional contribution of individual rainfall types (stacked colored bars; right axis) and the total rainfall (solid white curve; left axis) averaged over all available RGs. The dashed lines denote  $\pm 1\sigma$  of monthly rainfall within the RG network. We considered a month at a station as “available,” if at least 90% of the days contain data. The total monthly rainfall at a station was then rescaled to 100% if necessary; and (c) mean RG (red curve) and IMERG rainfall (black curve) with the rescaling applied to both RGs and IMERG. The red curve and the light red shaded area are identical to the solid and dashed white curves, respectively, in (b). Accordingly, the gray area denote the standard deviation of monthly rainfall between IMERG pixels.

It is visibly larger during the second rainy season in September and October compared to the first rainy season in May and June. Potential reasons for this behavior as well as the overall skill of IMERG are investigated in the next section.

## 5. Evaluation of IMERG

Building upon the previous paragraph, the skill of IMERG on a half-hourly and a point-to-pixel basis is evaluated for different categories listed as sections in Table 1. In the following, the results in each section of Table 1 is discussed and further analyzed. Unless noted otherwise, the standard approach of the contingency table is considered (Fig. 2).

### a. Rainfall occurrence

The occurrence frequency of the standard contingency table elements based on all available half-hour time steps ( $n = 419147$ ) is presented in Fig. 5a. First of all, less than 10% of the time steps in either RGs or IMERG contain rainfall and a total of 1.2% are hits. The errors, in turn, are clearly dominated by false alarms with a fraction of 6.2%. However, the decomposition of these false alarms and misses using the event-based approach of the contingency table reveals that not all errors emerge from a misinterpretation of isolated rainfall events (Figs. 5b,c). Almost 40% of falsely detected rainy time steps occur in association with rainfall events observed by the RGs, tantamount to an overestimation of the event duration in IMERG (Duration+, gray bar). The underestimation of the event duration (Duration-) comprises roughly a quarter of all misses. However, given the low percentage of misses in general (0.4%), Duration- rarely occurs.

Section A of Table 1 summarizes the results in Fig. 5a as skill measures introduced in section 3. As expected, an eye-catching result is the high POFA with 0.83, meaning that 83% of all rainy IMERG time steps are false alarms. At the same time, 23% of all rainy RG time steps are missed by IMERG (POD = 0.77). This preponderance of false alarms compared to misses is reflected in a BID of 4.61. With an HSS of 0.25, however, IMERG statistically performs better than observations based on pure chance. It shall be stressed again that these metrics are based on a simple rain-no rain condition without any information about rainfall rates. Applying the error measures, IMERG rainfall exhibits a mean absolute error of  $7.22 \text{ mm h}^{-1}$  and is negatively biased on average ( $\text{ME} = -4.53 \text{ mm h}^{-1}$ ).

### b. Rainfall rates

An important aspect to consider about the rain rate error measures is that they refer to the same RG and



TABLE 1. Performance evaluation of IMERG based on the measures introduced in section 3 for different categories, namely, all available time steps (section A), rain rate intervals (section B), rainfall types (section C), and seasons (section D). Note again that only hits are considered for calculation of the rain rate error measures ME, MAE, NME, and NMAE. In section B, rain rates are based on the RGs. Therefore, false alarms and correct negatives are not quantified. In section C, rainy time steps in the RGs and IMERG not associated with the respective rainfall type, i.e., all true false alarms and rainfall of the other rainfall types, are set to zero. Thus, misses include both true misses and Duration $-$ , whereas false alarms contain only Duration $+$ .

Section	Description	Contingency measures				Rain rate error measures (hits only)			
		POD	POFA	BID	HSS	ME (mm h $^{-1}$ )	MAE (mm h $^{-1}$ )	NME	NMAE
A	All time steps	0.77	0.83	4.61	0.25	-4.53	7.22	-0.55	0.88
B	<1 mm h $^{-1}$	0.69	—	—	—	2.47	2.54	4.39	4.50
	1–5 mm h $^{-1}$	0.77	—	—	—	1.01	2.46	0.40	0.97
	5–10 mm h $^{-1}$	0.77	—	—	—	-3.07	4.42	-0.44	0.63
	10–20 mm h $^{-1}$	0.82	—	—	—	-10.25	10.67	-0.72	0.75
	>20 mm h $^{-1}$	0.87	—	—	—	-31.90	31.95	-0.87	0.87
C	WCR	0.33	0.75	1.33	0.28	-0.34	1.99	-0.15	0.86
	SCR	0.63	0.76	2.62	0.35	-7.71	9.08	-0.75	0.88
	MCS	0.92	0.63	2.48	0.52	-4.22	7.45	-0.49	0.87
D	First rainy season	0.83	0.80	4.26	0.29	-4.22	7.41	-0.51	0.90
	Little dry season	0.57	0.88	4.70	0.18	-3.88	6.67	-0.49	0.85
	Second rainy season	0.79	0.81	4.11	0.27	-4.98	7.00	-0.60	0.84
	Long dry season	0.66	0.92	8.06	0.14	-5.85	7.59	-0.70	0.91

IMERG time steps. The scatterplot in Fig. 6a illustrates how the half-hourly rain rate pairs are distributed. Note that only hits are considered here. The bulk of data points comprise rainfall rates in the range of 1–10 mm h $^{-1}$  and is located close to the 1:1 line. However, the overall variability is high, suggesting issues in rain rate estimation and/or timing. The latter was found to affect the skill of PMW retrievals (You et al. 2019, see section 8 for a brief discussion). The regression line, determined with the error model in Tian et al. (2013), further indicates a positive and negative bias for low and high rain rates, respectively. Ignoring corresponding time steps and arranging this data subset in a quantile–quantile (Q–Q) plot, differences in the underlying distribution of rainfall rates between the RGs and IMERG as well as biases can be made visible in a more comprehensive manner (Fig. 6b). While rain rates are almost evenly distributed up to 2 mm h $^{-1}$ , the negative bias in IMERG at higher rain rates becomes increasingly evident. Overall, IMERG is unable to resolve the most extreme rainfall rates. Expressed as cumulative distributions, rainfall rates for hits and other elements from the event-based contingency table are compared in Fig. 6c. Around 70% of time steps containing true false alarms are equal or less than 1 mm h $^{-1}$  with a median of 0.55 mm h $^{-1}$  (short vertical orange line at the bottom). This is also true for roughly 50% of all Duration $+$  time steps. This hints toward a generally flawed formulation for very low-intensity rainfall in IMERG. At the same time, the subsets of true misses as well as Duration $-$  comprise markedly higher rain rates.

The dependence of IMERG's performance on certain rain rate intervals observed by the RGs is evaluated in section B of Table 1. Here, only POD can be quantified out of the contingency measures. Increasing rain rates as measured by the RGs are associated with an increase in POD. However, the rain rate intervals are differently biased. As seen in ME, the positive bias at low RG intensities turns strongly negative at high rain rates. Interestingly, for the weakest and most intense intervals, the absolute value of ME is nearly the same as MAE. Hence, at simultaneous RG and IMERG time steps, low- and high-intensity RG values are almost exclusively over and underestimated, respectively.

### c. Rainfall types

Using the analysis techniques from the previous paragraphs, the ability of IMERG in capturing RG-based rainfall types is shown in Fig. 7. Here, rainfall in the RGs and IMERG, which are not associated with the respective rainfall type, is set to zero. This also involves true false alarms in IMERG. Therefore, misses are represented by both true misses and Duration $-$ , false alarms solely by Duration $+$ . Evidently, more hits and less true misses are observed going from WCRs to MCSs (Fig. 7a). Thus, the degree of convective organization is an important factor in IMERG's detection ability. However, the overestimation of the event duration is an issue for all rainfall types. Over half of all rainfall-type-related time steps in IMERG are Duration $+$  (dark gray bars in Fig. 7a). By contrast, Duration $-$  plays an inferior role in detection errors. The Q–Q plot for each rainfall type highlights remarkable

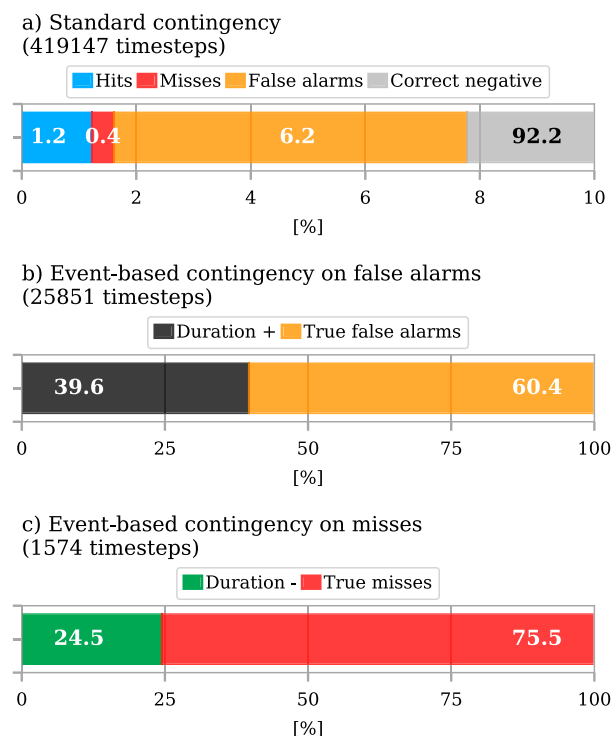


FIG. 5. (a) Standard approach of the contingency table on all available IMERG time steps. Note that the axis is truncated at 10% for more clarity. The fraction of correct negatives extends further to 100%. (b),(c) Event-based approach of the contingency table applied on the false alarm and miss subsets, respectively, of IMERG. Refer to Fig. 2 for both definitions of the contingency table. The numbers in the bars denote the percentages.

differences in the rain rate distributions between the RGs and IMERG (Fig. 7b). Rain rate pairs around WCR events are well aligned to the 1:1 line, whereas those of SCRs and MCSs indicate a strong underestimation of high-intensity rain rates in IMERG, which was seen already in Fig. 6b. The most notable difference between SCR and MCS distributions is found for lower rain rates. Low-intensity SCR rainfall is clearly too weak in IMERG as seen by the early deviation of the SCR curve from the 1:1 line. Since the curve never approaches the 1:1 line again at higher rain rates, the integrated SCR rainfall within the subset of hits is almost exclusively underestimated. Conversely, low-intensity MCS rainfall, largely occurring during the overpass of the stratiform part, is slightly too strong in IMERG. However, IMERG generally fails to adequately capture rain rates above  $5 \text{ mm h}^{-1}$ , from where the curve deviates strongly from the 1:1 line.

The skill is summarized in section C of Table 1. Considering POD, less than half of WCR time steps are identified by IMERG but confidence in detection is strongly increased around MCS events (0.92). Some

cases of true misses do occur even for MCSs. These are confined to cases where stations were located at the periphery of MCS passages (not shown). HSS increases from WCRs toward MCSs, again indicating a higher detection skill as well as better POFA for organized convection. It is interesting to note that the values for BID are still larger than 1. This means that time steps containing false alarms due to Duration+ outnumber the sum of time steps with true misses and Duration-. In other words, the net event duration of all rainfall types is considerably overestimated by IMERG, which became clear already in Fig. 7a. This is supported by the fact that rain rate distribution for WCRs within the subset of hits is even slightly negatively biased ( $\text{ME} = -0.34 \text{ mm h}^{-1}$ ). Consequently, the integrated WCR rainfall is generally overestimated by IMERG, whereas there are compensational effects between longer event duration and a mean underestimation of rain rates for SCR and MCS cases ( $-7.71$  and  $-4.22 \text{ mm h}^{-1}$ , respectively).

#### d. Seasonality

Projecting the previous results onto a seasonal perspective, Fig. 8a shows the averaged, monthly accumulated rainfall difference associated with the occurrence of the rainfall types by considering the event-based contingency (Fig. 2), where Duration+ contributes to a positive bias, Duration- as well as true misses to a negative bias. Confirming previous findings, monthly rainfall amounts associated with WCR events are overestimated and those linked to SCRs and MCSs are underestimated. However, September stands out exhibiting by far the largest negative and positive number biases for MCSs and WCRs, respectively. Pronounced underestimation of MCS rainfall is also visible in October and higher than both in May and June. Monthly IMERG rainfall obviously consists of substantial error compensations between the different rainfall types. Decomposing the seasonal cycle of IMERG into the contributions of rainfall types in the same manner as for the RGs (see Fig. 4b) yields remarkable discrepancies (Fig. 8c). More than a fifth of IMERG's total rainfall can be attributed to true false alarms (light orange bars). This potentially has important implications for the monthly gauge-calibration process in IMERG where rainfall estimates in the case of hits may be scaled in the wrong direction. At the same time, true misses are observed as well (Fig. 8c). Both SCRs and MCSs dominate the fractional rainfall of misses. As mentioned previously, true misses of MCSs occurred at stations located at the periphery of MCS passages, but still account for over half of missed rainfall in some months. WCRs exhibit a marked peak during

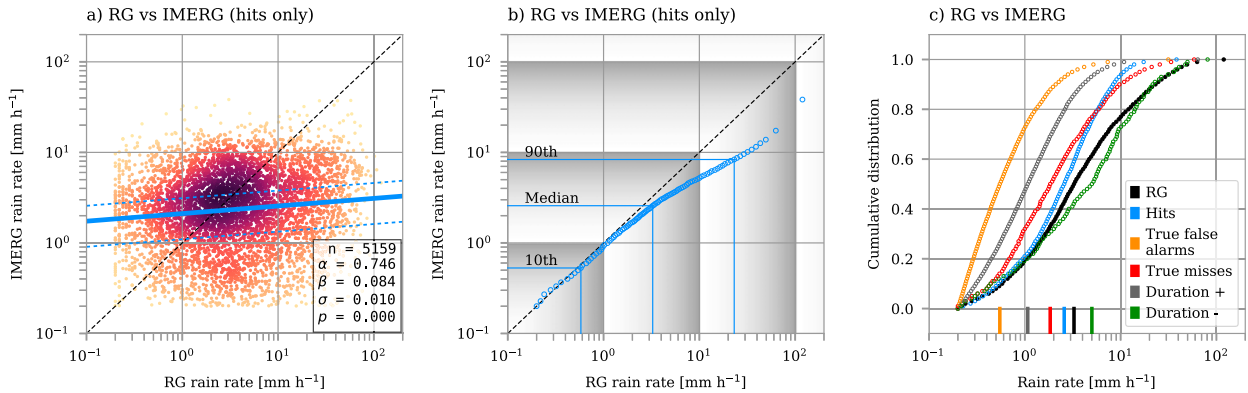


FIG. 6. Comparison of rain rates between the RGs and IMERG. (a) Scatterplot with RG rain rates on the  $x$  axis and IMERG rain rates on the  $y$  axis. Only hits are considered. Darker colors indicate a higher density of points. The regression line was calculated using the multiplicative error model of Tian et al. (2013). The parameters on the bottom right are total number of hits  $n$ , the  $y$  intercept  $\alpha$ , slope  $\beta$ , and standard deviation  $\sigma$  of the regression in the natural logarithm space of the error model, and the  $p$  value of the regression. (b) Quantile–quantile (Q–Q) plot of RG ( $x$  axis) and IMERG ( $y$  axis) rain rates. The positions of the 10th, 50th, and 90th percentiles are highlighted. (c) Cumulative rain rate distribution of the event-based contingency table (Fig. 2). The distribution of the RG rain rates are plotted in black as a reference. The colored lines at the bottom denote the respective median values.

the little dry season. The increased frequency of WCRs during this time of the year suggests season-specific difficulties in IMERG in detecting low-intensity events.

More generally, the event-based contingency table highlights a pronounced difference between dry and rainy seasons (Fig. 9a). Both dry seasons are dominated by a much higher frequency of true false alarms (>60%) compared to the rainy seasons. However, these (low-intensity) true false alarms appear to be a general issue within IMERG. In contrast, true misses are far less frequent overall, but are maximized during the little dry season. The latter is in large part due to IMERG’s inability to capture WCR events during this period (cf. Fig. 8c). Considering the subset of hits, the respective distributions in the Q–Q plot in Fig. 9b largely reflect the dominating rainfall type in the respective seasons. While the curves of both rainy seasons resemble that of MCSs, the long dry season exhibit a pattern similar to SCRs (cf. Fig. 7b). However, IMERG underestimates high rain rates stronger during the second rainy season compared to the first. The quality of rain rate estimation during the little dry season is distinctively better compared to the long dry season, but exhibits a similar weak negative bias at the lowest rain rates. Overall, the obvious commonality in all seasons is the negative bias at high rainfall intensities.

Summarizing the seasonal dependence in section D of Table 1, the skill of IMERG, although still better than random chance, is markedly lower during both dry seasons compared to the rainy seasons due to both decreased detection ability and frequent false alarms. During the little dry season, the skill of

IMERG particularly suffers from frequent misses of WCRs. Interestingly, all error measures are worst for the long dry season, which is, in some parts, related to SCRs being the dominant rainfall type during this period.

## 6. Source-based evaluation of IMERG

### a. Rainfall occurrence and rates

As described in section 2, rainfall observation in IMERG is composed of estimates based on direct PMW overpasses (PMW-direct), spatiotemporally advected PMW information (MORPH-only), and the combination of MORPH and IR (MORPH+IR). As seen in Fig. 10a, MORPH-only is the most frequently used source (37.2%) over the study area, followed by MORPH+IR (35.1%) and PMW-direct (27.6%). Only a small fraction is represented by IR-only and is therefore not subject to further study. While the fraction of misses hardly changes among the sources, it becomes evident that both PMW morphing and the inclusion of IR information increase the frequency of false alarms. IR retrievals are known to misjudge cold cloud features as rainy, for example, nonprecipitating anvils (Liu et al. 2007). However, the prevalence of false alarms in comparison to misses exists in all sources and suggests a general deficiency of overestimating rain occurrences.

Focusing again on hits, Fig. 10b shows the Q–Q plot for all sources. Most notably, the curves shift toward the right going from PMW-direct to MORPH+IR, indicating an increasing underestimation of rain intensities. Both PMW-direct and MORPH-only are closely aligned

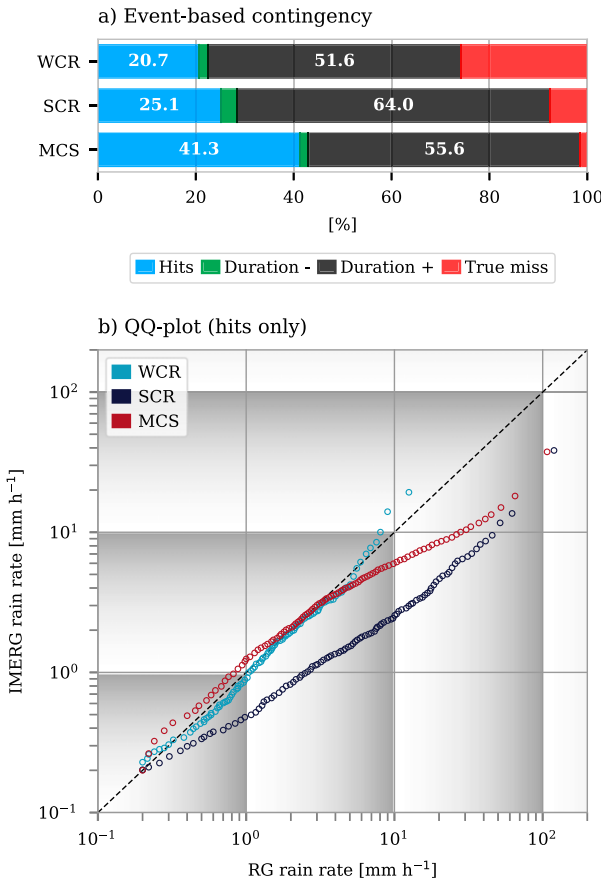


FIG. 7. (a) Fractional distribution of the event-based contingency table for WCRs, SCR, and MCS. True false alarms are excluded since only time steps associated with rainfall types are considered. (b) Q-Q plots of rain rates for time steps with hits for WCRs, SCR, and MCS.

to the 1:1 line for very low rain rates, but MORPH-only deviates from it earlier. Thus, while the underestimation of high rain rates in IMERG results from every source, the negative bias is weakest for PMW-direct. However, this bias appears to be an inherent problem in the PMW algorithm, which is amplified by morphing and IR data.

*b. Rainfall types and seasonality*

Figures 11a–c decomposes the results in Fig. 7a by source. Again, more hits and less frequent true misses are detected going from WCRs to MCSs across all sources. However, the duration of rainfall events is drastically increased within MORPH-only and/or MORPH+IR, and is strongest for SCR events. On the other hand, source-dependent tendencies for true misses and Duration– are less obvious, in large parts due to their low frequency (Fig. 10a). The general pattern of the source-based Q–Q plots all resemble

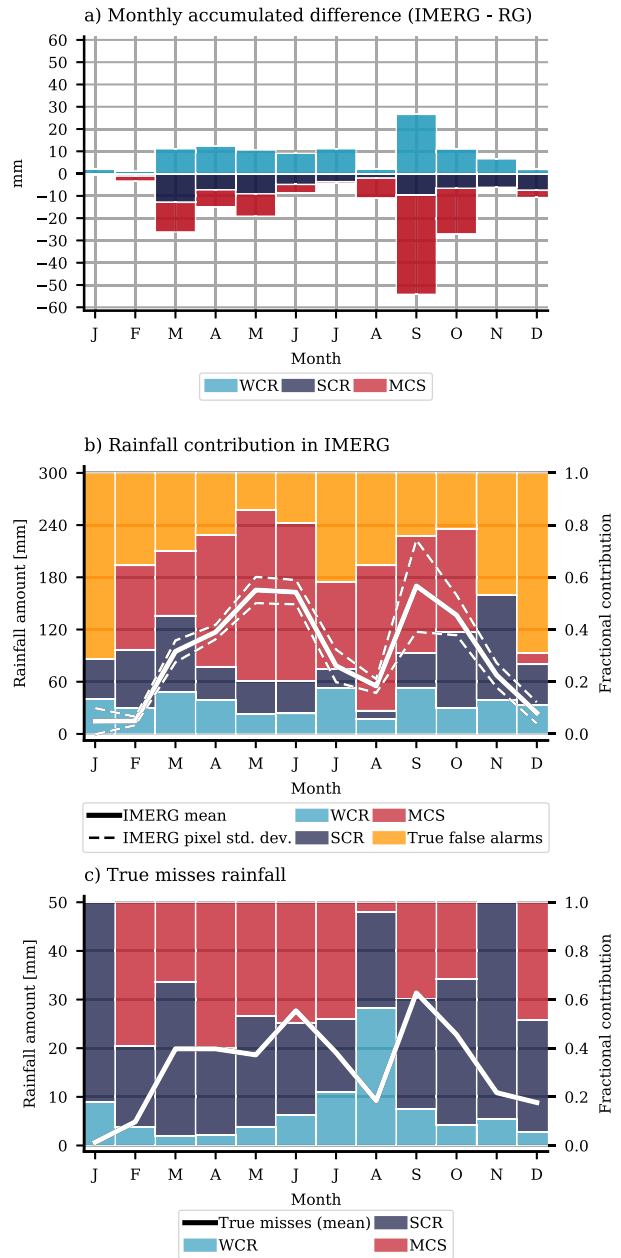


FIG. 8. Comparison between RG and IMERG with respect to total rainfall and representation of rainfall types using the event-based approach of the contingency table (Fig. 2), all averaged over collocated RGs and IMERG pixels. (a) Monthly integrated rainfall difference for rainfall types. (b) As in Fig. 4b, but for IMERG. The fractional rainfall contribution of true false alarms is denoted as orange bars. The solid and dashed white curves are identical to the black curve and gray area, respectively, in Fig. 4c. (c) As in (b), but for true misses. The left axis refers to the white curve and the right axis to the stacked bars.

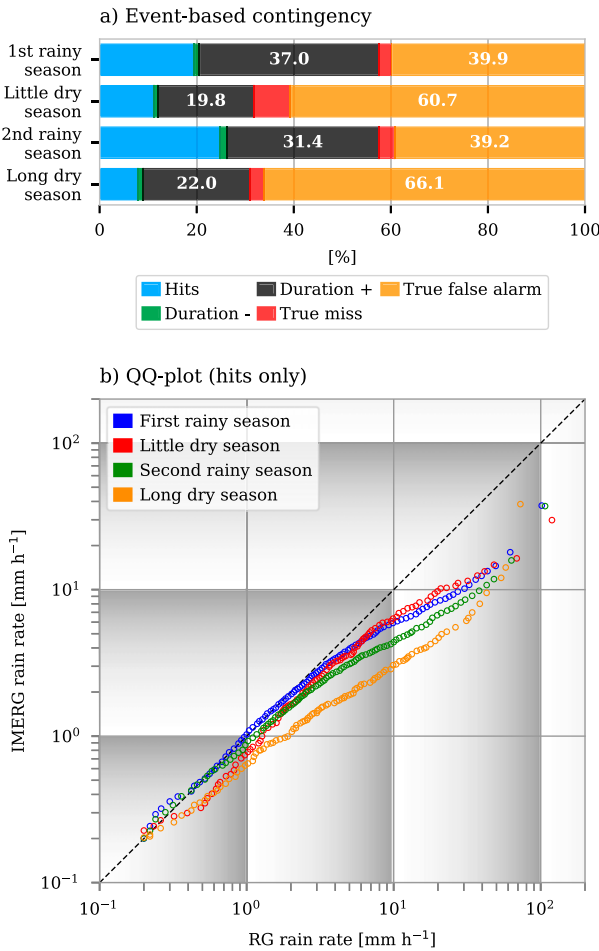


FIG. 9. As in Fig. 7, but for the different seasons of the West African monsoon. Refer to Fig. 4a for their definition.

the curves for the respective rainfall types seen in Figs. 7d–f. While PMW-direct again exhibit the weakest negative bias at high rain rates, the variation among the sources with respect to rainfall types is otherwise relatively low.

In the same manner, the different seasons are analyzed in Fig. 12. First of all, the dominance of true false alarms is again evident in all seasons (Figs. 12a–d). Interestingly, their fractions barely show a dependency on the source and rather exhibit similar values. With the exception of the long dry season, it is Duration+ that increases going from PMW-direct to MORPH+IR, which eventually causes the increase in false alarms in the standard approach of the contingency table (Fig. 10a). Source-based variations in the Q–Q plots (Figs. 12e–h) are most apparent for the little dry season, where rain rates below  $10 \text{ mm h}^{-1}$  are stronger negatively biased in MORPH+IR than in the other sources. Unlike the other seasons, which are dominated by deep convection, the larger fraction of shallow precipitating

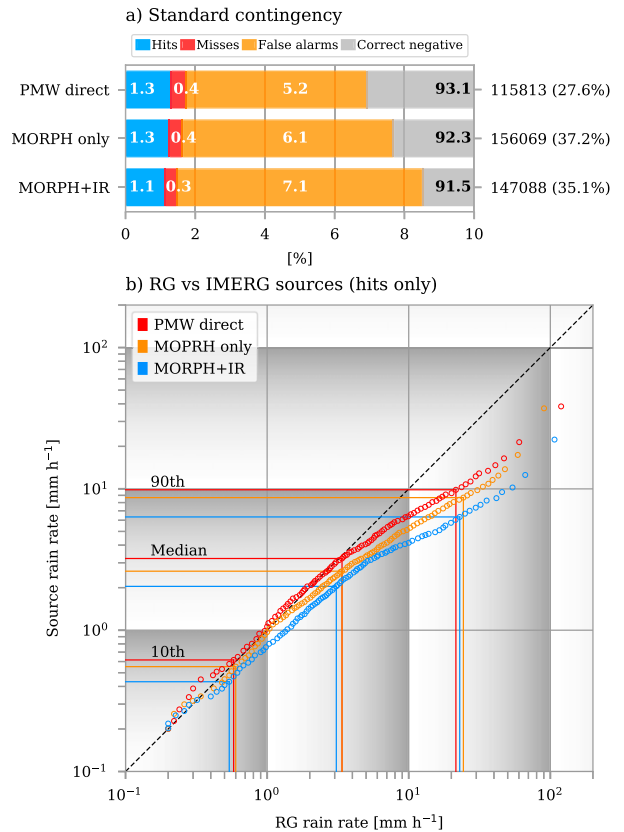


FIG. 10. (a) Standard approach of the contingency table based on all available time steps in the IMERG sources. The absolute numbers of time steps as well as their fraction relative to all IMERG time steps are given on the right axis. Note that the axis is truncated at 10% for more clarity. The fraction of correct negatives extends further to 100%. (b) Q–Q plot of RG rain rates ( $x$  axis) and those of the IMERG sources ( $y$  axis). The respective positions of the 10th, 50th, and 90th percentiles are highlighted.

clouds during the little dry season (not shown) likely imposes bigger challenges for the CTT-based IR rainfall estimation.

In summary, the clear benefits of filling data gaps in IMERG through morphing and inclusion of IR information come at the expense of amplifying the weaknesses of the PMW algorithm, that is, longer event durations and a stronger negative bias of intense rain rates.

### 7. Link to cloud-top properties

The high temporal resolution of the CLAAS-2 dataset allows us to break down the behavior of IMERG based upon the presence of different cloud-top properties and to compare it with the observations from the RGs. As CLAAS-2 contains cloud-top information only, this analysis can contribute additional

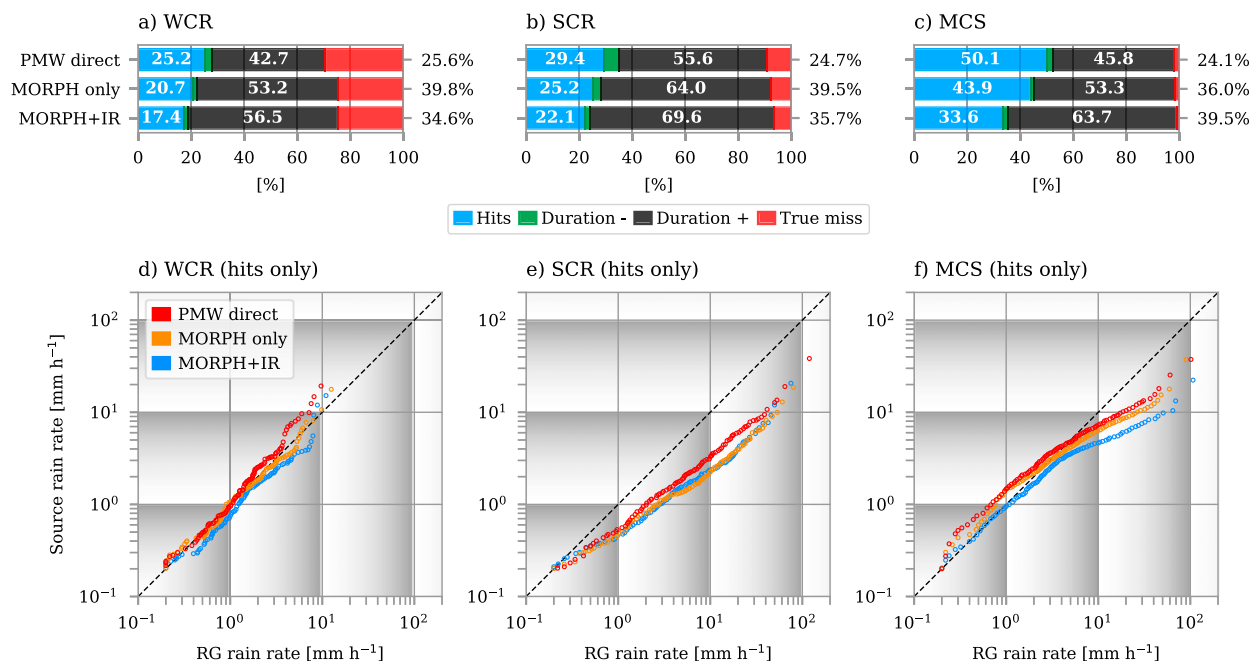


FIG. 11. As in Fig. 7, but based on the IMERG sources.

and independent information about error sources, particularly in PMW measurements, which contain information about the precipitation depth within clouds.

#### a. Cloud characteristics around rainy episodes and skill of IMERG

Figure 13 compares the probability distribution of cloud-top properties described in section 2 around all

rainy time steps within the different sources (colored lines) with those of the RGs (gray shade). Here, we distinguish between cloud tops in ice (Figs. 13a–d) and liquid phases (Figs. 13e–h), the latter of which is associated with warm rain. Note again that the sample consists of daytime rainfall only since the retrieval of COT and  $R_{\text{eff}}$  requires sunlight as input (see section 2). Although available at all times, the

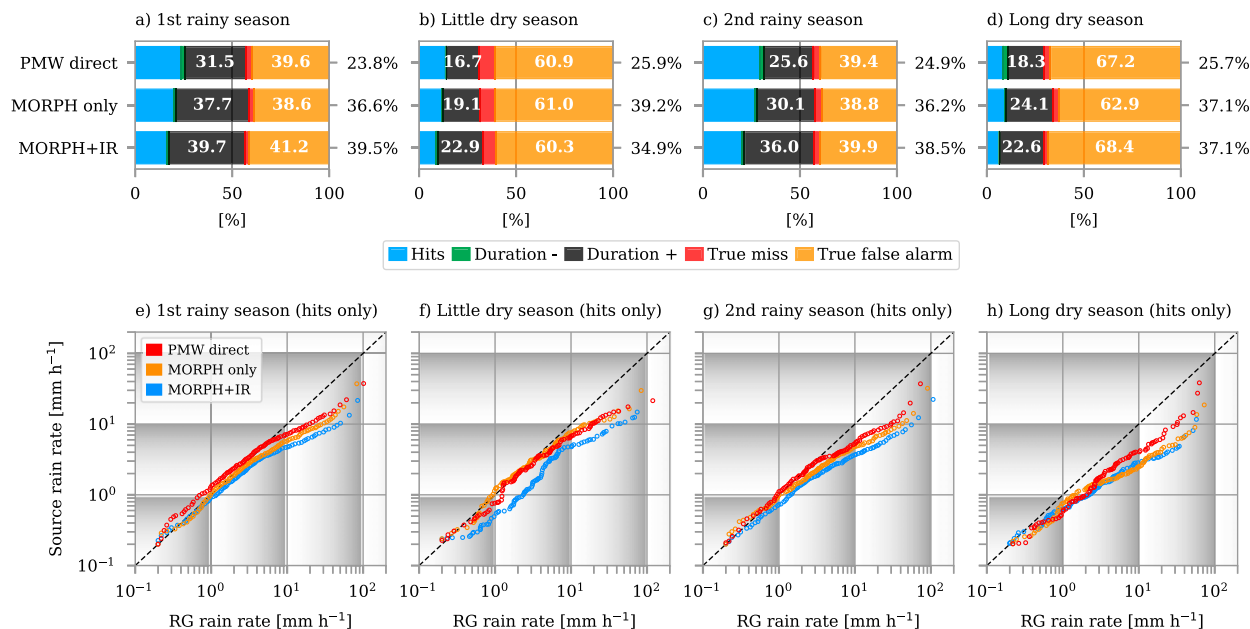


FIG. 12. As in Fig. 9, but based on the IMERG sources.

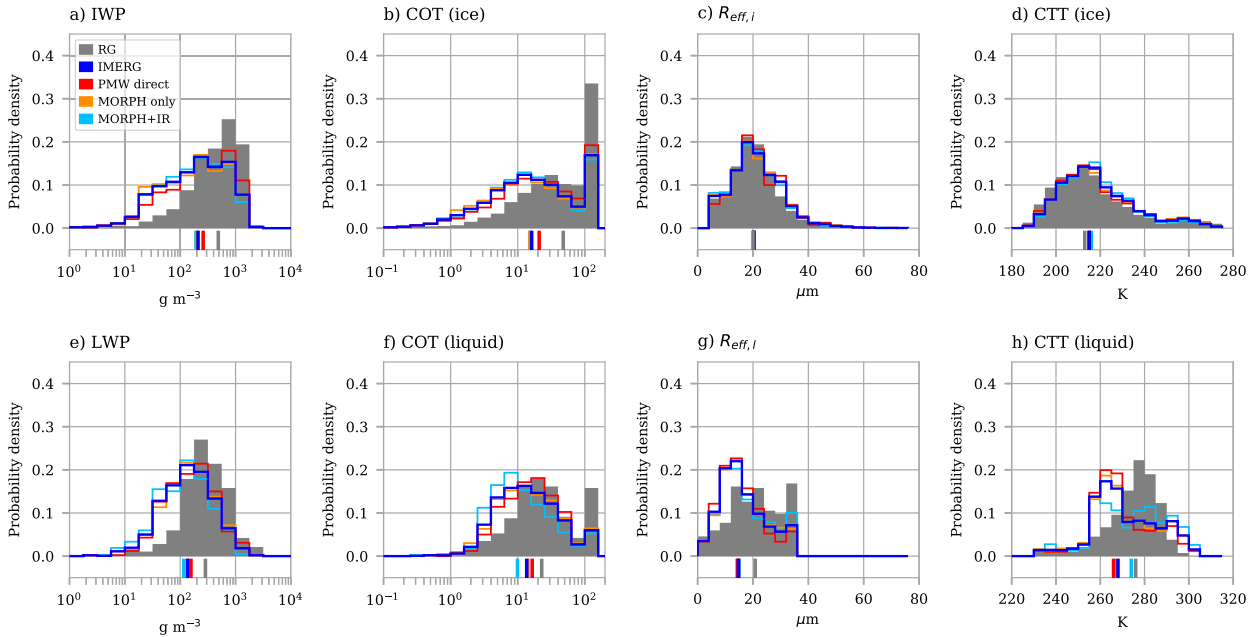


FIG. 13. Probability distributions of cloud properties described in section 2 based on all rainy time steps in the RGs (gray shaded), IMERG, and its sources (colored lines), separated into (a)–(d) glaciated cloud tops and (e)–(h) warm/liquid cloud tops. Note the different scales for CTT between glaciated and liquid cloud tops. The colored vertical lines at the bottom of each plot denote the median values of the respective distribution.

same temporal subset is taken for CTT in order to retain consistency.

A fundamental issue with IMERG and its sources is an overestimation of rainfall related to low LWP and IWP ( $<100 \text{ g m}^{-3}$ ), leading to substantially lower median values compared to the RGs (Figs. 13a,e). One major reason for this is the oversensitivity for low COT (Figs. 13b,f), which is more pronounced for ice. Overall, this issue slightly reduces as soon as direct PMW observations come into play. With respect to  $R_{\text{eff}}$ , IMERG performs well for ice clouds, but is oversensitive for liquid (i.e., warm) clouds, particularly below  $15 \mu\text{m}$  (Figs. 13c,g). In fact, warm rain becomes considerably likely for  $R_{\text{eff}} > 14 \mu\text{m}$  (e.g., Lensky and Rosenfeld 1997; Freud and Rosenfeld 2012), which already represents the median value of IMERG and its sources. Thus, while uncertainties in the rainfall occurrence associated with glaciated clouds are mostly related to COT, it is the combination of COT and  $R_{\text{eff},l}$  for warm clouds. With respect to CTT, differences between IMERG and RGs for frozen cloud tops are subtle but become more apparent for warm clouds (Figs. 13d,h). Here, IMERG strongly overestimates the occurrence frequency of rain around CTTs of 260–270 K but underestimates it for cloud tops above this temperature range. Interestingly, the fractional number of rainy time steps for  $\text{CTT} > 270 \text{ K}$  is highest in MORPH+IR. Overall, IMERG predicts more rain

occurrences from supercooled clouds than recorded by the RGs.

A look into the skill measures separated for warm and glaciated clouds as well as all sources reveals a considerable discrepancy in skill between and warm and cold cloud rainfall (Table 2). Around warm clouds, POD is substantially lower, POFA is even higher despite being already around 0.8 for ice clouds, and BID is overall higher (5.0 versus 4.34 for IMERG). Consequently, HSS is lower, but still indicates a slightly better skill than random chance for all sources. ME and MAE are higher for ice clouds, which is unsurprising due to heavy rainfall being mostly associated with deep convection. In fact, the NMAE values are very similar between cold and warm clouds, the latter of which, however, are associated with a stronger negative bias (see NME).

#### b. Origin of hits, false alarms, and misses

Focusing on IMERG only, Fig. 14 illustrates the distribution of the standard contingency table elements (Fig. 2) based on all rainy time steps. For IWP (Fig. 14a), the distributions for hits (blue curve) and the RGs (gray shade, same as in Fig. 13) are nearly identical. In other words, with IWP as a reference, IMERG is generally able to detect rain occurrences as measured by the RGs. As seen previously, however, IMERG is tuned such that it produces too many low-intensity false alarms. In the case of ice clouds, this stems from the aforementioned

TABLE 2. As in Table 1, but for glaciated cloud tops (section A) and warm/liquid cloud tops (section B), and further decomposed into IMERG and its sources.

Section	Description	Source	Contingency measures				Rain rate error measures (hits only)			
			POD	POFA	BID	HSS	ME (mm h <sup>-1</sup> )	MAE (mm h <sup>-1</sup> )	NME	NMAE
A	Glaciated clouds	IMERG	0.83	0.81	4.34	0.26	-4.57	7.33	-0.55	0.88
		PMW-direct	0.81	0.77	3.53	0.31	-2.50	7.54	-0.31	0.93
		MORPH-only	0.83	0.80	4.08	0.28	-4.73	7.54	-0.55	0.87
		MORPH+IR	0.85	0.85	5.54	0.21	-5.37	7.20	-0.65	0.87
B	Warm clouds	IMERG	0.29	0.94	5.00	0.09	-3.14	3.61	-0.74	0.85
		PMW-direct	0.28	0.94	5.15	0.08	-3.10	3.88	-0.71	0.90
		MORPH-only	0.34	0.95	6.34	0.08	-3.78	4.19	-0.76	0.84
		MORPH+IR	0.26	0.93	3.70	0.10	-2.17	2.67	-0.67	0.83

oversensitivity toward low IWP values (orange), which can be traced back to a flawed relationship with COT (Fig. 14b). In contrast, the high similarity between the contingency elements and RGs with respect to  $R_{eff,i}$  indicates that hits, false alarms and misses can hardly be predicted with  $R_{eff,i}$ . Considering CTT, the distribution of misses is shifted towards higher values compared to that of the RG observation (Fig. 14d), that is, clouds at lower altitudes.

With respect to warm clouds, IMERG behaves differently. As expected from the low POD in Table 2, IMERG's current relationship to LWP predominantly yields misses. In addition, the similarity between hits and false alarms indicates that warm clouds are frequently misinterpreted. This deficiency is partly related to a combined oversensitivity for low COT and  $R_{eff,i}$  (Figs. 14f,g). Furthermore, the uncertainty is enhanced

by CTT, where pronounced differences between the RGs and hits are apparent as well (Fig. 14h). Overall, this is the reason for frequent misses of WCR events during the little dry season, which are predominantly produced by warm clouds (not shown). At the same time, the aforementioned overestimation of rain occurrences for CTTs around 260 K are typically false alarms.

## 8. Discussion

As mentioned in section 2, all results presented in this study must first be understood from the perspective of a point-to-pixel validation. While IMERG contains area-averaged rainfall information within its 0.1° grid (Huffman et al. 2019c), almost every pixel is compared with only a single RG. This discrepancy in spatial representativeness may affect some of the error measures

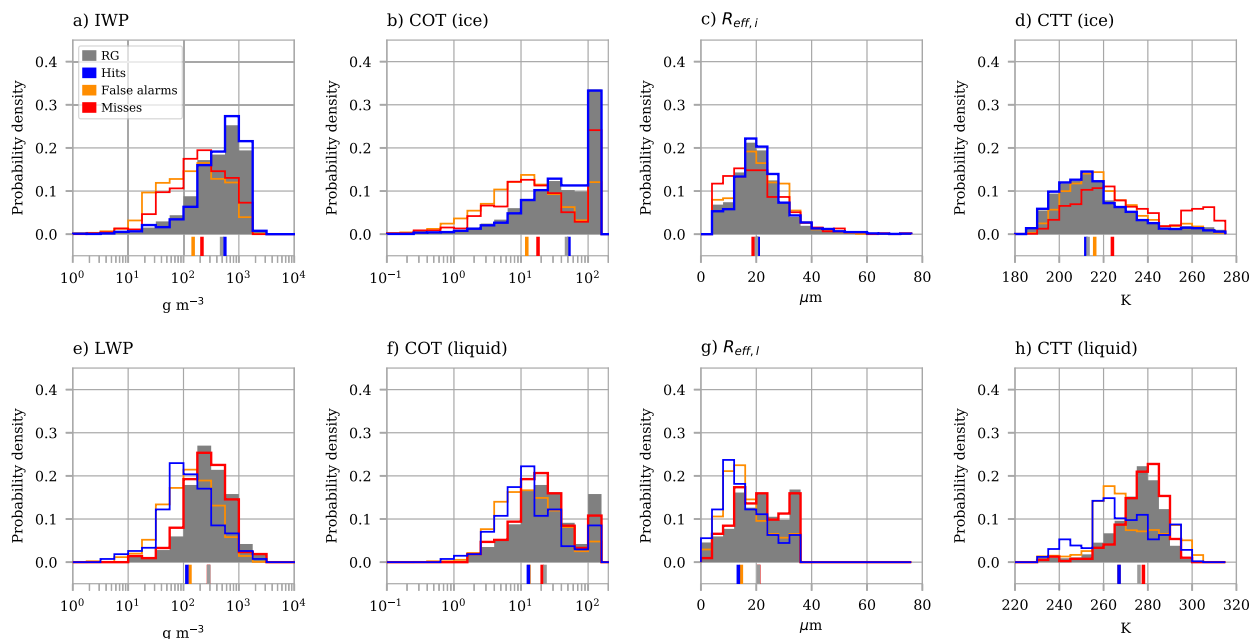


FIG. 14. As in Fig. 13, but for the elements of the standard contingency table. Results are shown for IMERG only.



and may hamper the comparability of the most extreme rain rates between the RGs and IMERG. In fact, further investigations have shown that a successively increased number of RGs within coarse-grained IMERG grid boxes improves skill (according to HSS) and mitigates bias issues (at high rain rates), although at the expense of the detection ability (see Table S1 and Fig. 4). This is consistent with [Monsieurs et al. \(2018\)](#) and [Tian et al. \(2018\)](#), with the latter authors arguing that the POD is particularly reduced at low rain rates. By contrast, gradual coarse-graining of the IMERG grid around a single RG tends to improve rainfall detection, but lowers HSS as a result of more false alarms (see Tables S2–S5). Given this general point-to-pixel uncertainty, we conclude that some of the error magnitudes in [Tables 1 and 2](#), particularly MAE, are overestimated.

A particular aspect in IMERG that has rarely been documented in other studies to the best of the authors' knowledge is the high occurrence frequency of low-intensity false alarms. First, we note that the RGs potentially underestimate the frequency of light rain due to wind and dirt-filter-related undercatch. For reasons outlined above, we expect that the POFA values are overestimated. Although examined for a different climatic zone, an additional aspect stressed by [You et al. \(2019\)](#) is a time lag effect in PMW observations, in which false alarms (correlation coefficients) were reduced (increased) through a temporal shift of PMW estimations relative to surface observations. Indeed, shifting IMERG backward by one time step (i.e.,  $-30$  min) results in the highest correlation coefficient as well as an increased POD (Fig. S5a and Tables S6 and S7). This implies that strong rainfall in IMERG tends to lag its counterpart in the observation, which becomes evident in all but the little dry season (Fig. S5b). It is speculated that this lag appears (i) due to limitations of the morphing technique when, for example during the first rainy season, particularly fast moving convective systems are observed within a highly sheared environment (e.g., [Maranan et al. 2018](#)); or (ii) due to the time needed during the cumulonimbus development until a critical level of ice water path is reached for rain detection while the convective cell has already started precipitating (e.g., [Pfeifroth et al. 2016](#)). Nonetheless, frequent false alarms remain a distinct issue in the present study domain despite a slight improvement in POFA after the temporal shift. They are source independent, but become more pronounced as soon as spatiotemporal morphing of PMW data and inclusion of IR data come into play. For TRMM, the latter was found to be largely associated with nonprecipitating anvils in convective situations ([Liu et al. 2007](#)). As these false alarms constitute more than a fifth of

monthly IMERG rainfall, they promote a reduction of daily and subdaily rainfall (and in particular high rainfall rates) around hits if monthly rainfall is reduced through gauge calibration. In fact, the early run of IMERG, which contains data prior to the gauge correction, exhibits a decreased negative bias at high rain rates (Fig. S6). Thus, it can be argued that the early run is more suitable for the evaluation as well as statistics of extreme rainfall. Either way, this pronounced negative bias at high rain rates in IMERG must be considered in future rainfall studies. Quantile mapping techniques, usually applied for bias corrections in climate models, are a potent way to address this issue ([Lafon et al. 2013](#); [Cannon et al. 2015](#)), and here particularly with respect to the different IMERG sources.

While false alarms are a particular challenge for glaciated clouds, especially with low COT, misses are frequently related to warm rain (cf. [Young et al. 2018](#)), which in turn highlights issues of PMW rainfall retrievals in the absence of frozen precipitation-sized hydrometeors. Thus, the overall detection skill of IMERG may depend on the moment in which precipitation-sized ice particles are eventually formed within a convective cloud. Over West Africa, rainfall processes and the timing of ice formation are likely influenced by the high aerosol load documented in a number of studies (e.g., [Knippertz et al. 2015](#); [Deroubaix et al. 2019](#); [Deetz et al. 2018](#); [Taylor et al. 2019](#); [Haslett et al. 2019](#)). In general, [Rosenfeld et al. \(2011\)](#) found that under heavy aerosol load conditions, clouds glaciate at warmer temperatures and the activity of ice nuclei, for example, Saharan dust in the case of West Africa, becomes dominant for precipitation formation. At the same time, [McCollum et al. \(2000\)](#) argued that these conditions may explain the substantial overestimation of monthly rainfall over central Africa in the Global Precipitation Climatology Project (GPCP; [Huffman et al. 1997](#)) due to the reduction of drop size and thus precipitation efficiency in deep convection. Eventually, these opposing rainfall processes may significantly affect the performance of IMERG with regard to rainfall detection and rain rate estimation.

The WAM dynamics determine the occurrence of the different rainfall types presented in this study, and thus the event rainfall amount. In this regard, [Hamada et al. \(2015\)](#) interestingly noticed a weak relationship between deep intense radar echoes and extreme near surface rainfall for many moist tropical regions. They argued that extreme rain rates are rather controlled by abundant low-level moisture, leading to low cloud bases and thus a deep warm cloud layer where collision-coalescence processes are enhanced. This, however, was in absence of high radar reflectivities in the upper-level portion of the convective clouds, which are

usually caused by large, precipitation-sized ice particles. In contrast, intense convection featuring these high upper-level reflectivities were associated to a lesser extent with the most extreme near-surface rain rates. This weak linkage between cloud ice content and rain rates potentially adds to the uncertainty in rainfall estimation in IMERG given the fact that scattering of microwave signals by ice is its key principle over land. The described situation is frequent during the second rainy season in September and October, which has been found to exhibit the strongest integrated underestimation of SCR and MCS rainfall in this study.

We stress that this study is representative primarily for the West African forest zone and other regions with comparable conditions regarding climate and aerosols. Petković and Kummerow (2017) and McCollum et al. (2000) already emphasized how region-dependent rainfall biases are. They can partly be understood through the complex interplay between underlying dynamics, aerosol load, and their influence on the evolution and characteristics of clouds. Thus, together with further regional-scale validation efforts, we anticipate that the consideration of such additional information can help to improve IMERG.

## 9. Summary

The present work evaluated the performance of IMERG V6B (final run) with respect to different rainfall types, WAM seasons, its sources (PMW-direct, MORPH-only, MORPH+IR), and cloud-top characteristics on a subdaily time scale. Two years of data from a dense network of 17 high-resolution rain gauges deployed in the forest zone of Ghana in southern West Africa served as the reference. We found the following:

- Very frequent but low-intensity false alarms contribute more than a fifth to total IMERG rainfall. They occur in every IMERG source.
- The duration of rainfall events is generally overestimated, but increasingly more pronounced going from PMW-direct to MORPH+IR. Overall, we find a systematic overestimation in integrated rainfall for weak and short convective events.
- High rainfall intensities are negatively biased in every IMERG source, leading to an underestimation in integrated rainfall for SCRs as well as MCSs and ultimately to an error compensation with WCRs. This particularly applies to the second rainy season in September and October.
- IMERG and its sources are too sensitive toward low values in IWP and LWP, accounting for the majority of false alarms. For ice clouds, it is mainly the over-sensitivity toward a low COT, whereas for warm

clouds, it is the combination of both low COT and  $R_{\text{eff},i}$ . IMERG performs drastically better in the presence of ice clouds than warm clouds, the latter of which is subject to a lot of missed events.

This study has emphasized the potential of regional- and subdaily-scale validations of spaceborne rainfall products in combination with high-resolution rain gauges, particularly for data-sparse regions such as West Africa.

*Acknowledgments.* The DACCIWA project has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under Grant Agreement 603502 (EU project DACCIWA: Dynamics-aerosol-chemistry-cloud interactions in West Africa). The research leading to these results has in parts been done within the subproject “C2-Statistical-dynamical forecasts of tropical rainfall” of the Transregional Collaborative Research Center SFB/TRR 165 “Waves to Weather” ([www.wavestoweather.de](http://www.wavestoweather.de)) funded by the German Research Foundation (DFG).

## REFERENCES

- Adeyewa, Z. D., and K. Nakamura, 2003: Validation of TRMM radar rainfall data over major climatic regions in Africa. *J. Appl. Meteor.*, **42**, 331–347, [https://doi.org/10.1175/1520-0450\(2003\)042<0331:VOTRRD>2.0.CO;2](https://doi.org/10.1175/1520-0450(2003)042<0331:VOTRRD>2.0.CO;2).
- Aminou, D., 2002: MSG’s SEVIRI instrument. *ESA Bull.*, **111**, 15–17.
- Arkin, P. A., R. Joyce, and J. E. Janowiak, 1994: The estimation of global monthly mean rainfall using infrared satellite data: The GOES Precipitation Index (GPI). *Remote Sens. Rev.*, **11**, 107–124, <https://doi.org/10.1080/02757259409532261>.
- Benas, N., S. Finkensieper, M. Stengel, G.-J. van Zadelhoff, T. Hanschmann, R. Hollmann, and J. F. Meirink, 2017: The MSG-SEVIRI-based cloud property data record CLAAS-2. *Earth Syst. Sci. Data*, **9**, 415–434, <https://doi.org/10.5194/essd-9-415-2017>.
- Benson, C., and E. Clay, 1998: *The Impact of Drought on Sub-Saharan African Economies: A Preliminary Examination*. The World Bank, 80 pp.
- Camberlin, P., and Coauthors, 2019: Evaluation of remotely sensed rainfall products over Central Africa. *Quart. J. Roy. Meteor. Soc.*, **145**, 2115–2138, <https://doi.org/10.1002/qj.3547>.
- Cannon, A. J., S. R. Sobie, and T. Q. Murdock, 2015: Bias correction of GCM precipitation by quantile mapping: How well do methods preserve changes in quantiles and extremes? *J. Climate*, **28**, 6938–6959, <https://doi.org/10.1175/JCLI-D-14-00754.1>.
- Chen, Y., E. E. Ebert, K. J. Walsh, and N. E. Davidson, 2013: Evaluation of TRMM 3B42 precipitation estimates of tropical cyclone rainfall using PACRAIN data. *J. Geophys. Res. Atmos.*, **118**, 2184–2196, <https://doi.org/10.1002/JGRD.50250>.
- CM SAF, 2016: SEVIRI cloud physical products, CLAAS edition 2. Algorithm Theoretical Basis Doc., 35 pp., [https://doi.org/10.5676/EUM\\_SAF\\_CM/CLAAS/V002](https://doi.org/10.5676/EUM_SAF_CM/CLAAS/V002).
- Deetz, K., and Coauthors, 2018: Numerical simulations of aerosol radiative effects and their impact on clouds and atmospheric

- dynamics over southern West Africa. *Atmos. Chem. Phys.*, **18**, 9767–9788, <https://doi.org/10.5194/acp-18-9767-2018>.
- Deroubaix, A., and Coauthors, 2019: Diurnal cycle of coastal anthropogenic pollutant transport over southern West Africa during the DACCWA campaign. *Atmos. Chem. Phys.*, **19**, 473–497, <https://doi.org/10.5194/acp-19-473-2019>.
- Dezfuli, A. K., C. M. Ichoku, G. J. Huffman, K. I. Mohr, J. S. Selker, N. Van De Giesen, R. Hochreutener, and F. O. Annor, 2017a: Validation of IMERG precipitation in Africa. *J. Hydrometeorol.*, **18**, 2817–2825, <https://doi.org/10.1175/JHM-D-17-0139.1>.
- , —, K. I. Mohr, and G. J. Huffman, 2017b: Precipitation characteristics in West and East Africa from satellite and in situ observations. *J. Hydrometeorol.*, **18**, 1799–1805, <https://doi.org/10.1175/JHM-D-17-0068.1>.
- Dinku, T., P. Ceccato, E. Grover-Kopec, M. Lemma, S. Connor, and C. Ropelewski, 2007: Validation of satellite rainfall products over East Africa's complex topography. *Int. J. Remote Sens.*, **28**, 1503–1526, <https://doi.org/10.1080/01431160600954688>.
- Eldridge, R., 1957: A synoptic study of West African disturbance lines. *Quart. J. Roy. Meteor. Soc.*, **83**, 303–314, <https://doi.org/10.1002/qj.49708335704>.
- Fink, A. H., D. Vincent, and V. Ermert, 2006: Rainfall types in the West African sudanian zone during the summer monsoon 2002. *Mon. Wea. Rev.*, **134**, 2143–2164, <https://doi.org/10.1175/MWR3182.1>.
- , and Coauthors, 2011: Operational meteorology in West Africa: Observational networks, weather analysis and forecasting. *Atmos. Sci. Lett.*, **12**, 135–141, <https://doi.org/10.1002/asl.324>.
- , and Coauthors, 2017: Mean climate and seasonal cycle. *Meteorology of Tropical West Africa: The Forecasters' Handbook*, D. J. Parker and M. Diop-Kane, Eds., John Wiley and Sons, 1–39.
- Flamant, C., and Coauthors, 2018: The dynamics–aerosol–chemistry–cloud interactions in West Africa field campaign: Overview and research highlights. *Bull. Amer. Meteor. Soc.*, **99**, 83–104, <https://doi.org/10.1175/BAMS-D-16-0256.1>.
- Freud, E., and D. Rosenfeld, 2012: Linear relation between convective cloud drop number concentration and depth for rain initiation. *J. Geophys. Res.*, **117**, D02207, <https://doi.org/10.1029/2011JD016457>.
- Gaona, M. F. R., G. Villarini, W. Zhang, and G. A. Vecchi, 2018: The added value of IMERG in characterizing rainfall in tropical cyclones. *Atmos. Res.*, **209**, 95–102, <https://doi.org/10.1016/j.atmosres.2018.03.008>.
- , A. Overeem, H. Leijnse, and R. Uijlenhoet, 2016: First-year evaluation of GPM rainfall over The Netherlands: IMERG day 1 final run (V03D). *J. Hydrometeorol.*, **17**, 2799–2814, <https://doi.org/10.1175/JHM-D-16-0087.1>.
- Gebregiorgis, A. S., P.-E. Kirstetter, Y. E. Hong, N. J. Carr, J. J. Gourley, W. Petersen, and Y. Zheng, 2017: Understanding overland multisensor satellite precipitation error in TMPA-RT products. *J. Hydrometeorol.*, **18**, 285–306, <https://doi.org/10.1175/JHM-D-15-0207.1>.
- Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2). *J. Climate*, **30**, 5419–5454, <https://doi.org/10.1175/JCLI-D-16-0758.1>.
- Glickman, T. S., Ed., 2000: *Glossary of Meteorology*. 2nd ed. Amer. Meteor. Soc., 855 pp., <http://glossary.ametsoc.org/>.
- Gosset, M., J. Viarre, G. Quantin, and M. Alcoba, 2013: Evaluation of several rainfall products used for hydrological applications over West Africa using two high-resolution gauge networks. *Quart. J. Roy. Meteor. Soc.*, **139**, 923–940, <https://doi.org/10.1002/qj.2130>.
- Hamada, A., Y. N. Takayabu, C. Liu, and E. J. Zipser, 2015: Weak linkage between the heaviest rainfall and tallest storms. *Nat. Commun.*, **6**, 6213, <https://doi.org/10.1038/ncomms7213>.
- Hamilton, R., J. Archbold, and C. Douglas, 1945: Meteorology of Nigeria and adjacent territory. *Quart. J. Roy. Meteor. Soc.*, **71**, 231–264, <https://doi.org/10.1002/qj.49707130905>.
- Hansen, J. E., and L. D. Travis, 1974: Light scattering in planetary atmospheres. *Space Sci. Rev.*, **16**, 527–610, <https://doi.org/10.1007/BF00168069>.
- Haslett, S. L., and Coauthors, 2019: Remote biomass burning dominates southern West African air pollution during the monsoon. *Atmos. Chem. Phys.*, **19**, 15 217–15 234, <https://doi.org/10.5194/ACP-19-15217-2019>.
- Hastings, D. A., and Coauthors, 1999: The Global Land One-Kilometer Base Elevation (GLOBE) digital elevation model, version 1.0. NOAA National Geophysical Data Center, <https://www.ngdc.noaa.gov/mgg/topo/globe.html>.
- Hong, Y., K.-L. Hsu, S. Sorooshian, and X. Gao, 2004: Precipitation Estimation from Remotely Sensed Imagery Using an Artificial Neural Network Cloud Classification System. *J. Appl. Meteor.*, **43**, 1834–1853, <https://doi.org/10.1175/JAM2173.1>.
- Hou, A. Y., and Coauthors, 2014: The Global Precipitation Measurement mission. *Bull. Amer. Meteor. Soc.*, **95**, 701–722, <https://doi.org/10.1175/BAMS-D-13-00164.1>.
- Huffman, G. J., and Coauthors, 1997: The Global Precipitation Climatology Project (GPCP) combined precipitation dataset. *Bull. Amer. Meteor. Soc.*, **78**, 5–20, [https://doi.org/10.1175/1520-0477\(1997\)078<0005:TGPCPG>2.0.CO;2](https://doi.org/10.1175/1520-0477(1997)078<0005:TGPCPG>2.0.CO;2).
- , and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.*, **8**, 38–55, <https://doi.org/10.1175/JHM560.1>.
- , R. F. Adler, D. T. Bolvin, and E. J. Nelkin, 2010: The TRMM Multi-Satellite Precipitation Analysis (TMPA). *Satellite Rainfall Applications for Surface Hydrology*, M. Gebremichael and F. Hossain, Eds., Springer, 3–22.
- , D. T. Bolvin, D. Braithwaite, K. Hsu, R. Joyce, C. Kidd, E. J. Nelkin, and P. Xie, 2015: NASA Global Precipitation Measurement Integrated Multi-satellitE Retrievals for GPM (IMERG). Algorithm Theoretical Basis Doc., version 4.5, 30 pp., [http://pmm.nasa.gov/sites/default/files/document\\_files/IMERG\\_ATBD\\_V4.5.pdf](http://pmm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V4.5.pdf).
- , D. Bolvin, E. Nelkin, and J. Tan, 2019a: Integrated multi-satellite retrievals for GPM (IMERG) technical documentation. NASA Tech. Doc., 77 pp., accessed 5 August 2019, [https://docserver.gesdisc.eosdis.nasa.gov/public/project/GPM/IMERG\\_doc.06.pdf](https://docserver.gesdisc.eosdis.nasa.gov/public/project/GPM/IMERG_doc.06.pdf).
- , E. Stocker, D. Bolvin, E. Nelkin, and J. Tan, 2019b: GPM IMERG final precipitation 13 half hourly 0.1 degree × 0.1 degree v06. Goddard Earth Sciences Data and Information Services Center, accessed 30 June 2019, <https://doi.org/10.5067/GPM/IMERG/3B-HH/06>.
- , and Coauthors, 2019c: NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG). Algorithm Theoretical Basis Doc., version 06, 32 pp., accessed 5 August 2019, [https://docserver.gesdisc.eosdis.nasa.gov/public/project/GPM/IMERG\\_ATBD\\_V06.pdf](https://docserver.gesdisc.eosdis.nasa.gov/public/project/GPM/IMERG_ATBD_V06.pdf).

- Islam, T., Y. Hu, A. A. Kokhanovsky, and J. Wang, 2017: *Remote Sensing of Aerosols, Clouds, and Precipitation*. Elsevier, 364 pp.
- Janiga, M. A., and C. D. Thorncroft, 2014: Convection over tropical Africa and the east Atlantic during the West African monsoon: Regional and diurnal variability. *J. Climate*, **27**, 4159–4188, <https://doi.org/10.1175/JCLI-D-13-00449.1>.
- Jobard, I., F. Chopin, J. C. Bergès, and R. Roca, 2011: An inter-comparison of 10-day satellite precipitation products during West African monsoon. *Int. J. Remote Sens.*, **32**, 2353–2376, <https://doi.org/10.1080/01431161003698286>.
- Joyce, R. J., and P. Xie, 2011: Kalman filter-based CMORPH. *J. Hydrometeorol.*, **12**, 1547–1563, <https://doi.org/10.1175/JHM-D-11-022.1>.
- Kamara, S. I., 1986: The origins and types of rainfall in West Africa. *Weather*, **41**, 48–56, <https://doi.org/10.1002/j.1477-8696.1986.tb03787.x>.
- Karaseva, M. O., S. Prakash, and R. Gairola, 2012: Validation of high-resolution TRMM-3B43 precipitation product using rain gauge measurements over Kyrgyzstan. *Theor. Appl. Climatol.*, **108**, 147–157, <https://doi.org/10.1007/s00704-011-0509-6>.
- Kidd, C., 2001: Satellite rainfall climatology: A review. *Int. J. Climatol.*, **21**, 1041–1066, <https://doi.org/10.1002/joc.635>.
- , and V. Levizzani, 2011: Status of satellite precipitation retrievals. *Hydrol. Earth Syst. Sci.*, **15**, 1109–1116, <https://doi.org/10.5194/hess-15-1109-2011>.
- Knippertz, P., and Coauthors, 2015: The dacciya project: Dynamics–aerosol–chemistry–cloud interactions in West Africa. *Bull. Amer. Meteor. Soc.*, **96**, 1451–1460, <https://doi.org/10.1175/BAMS-D-14-00108.1>.
- , and Coauthors, 2017: A meteorological and chemical overview of the DACCIIWA field campaign in West Africa in June–July 2016. *Atmos. Chem. Phys.*, **17**, 10893–10918, <https://doi.org/10.5194/ACP-17-10893-2017>.
- Kummerow, C., W. Barnes, T. Kozu, J. Shiue, and J. Simpson, 1998: The tropical rainfall measuring mission (TRMM) sensor package. *J. Atmos. Oceanic Technol.*, **15**, 809–817, [https://doi.org/10.1175/1520-0426\(1998\)015<0809:TTRMMT>2.0.CO;2](https://doi.org/10.1175/1520-0426(1998)015<0809:TTRMMT>2.0.CO;2).
- Lafon, T., S. Dadson, G. Buys, and C. Prudhomme, 2013: Bias correction of daily precipitation simulated by a regional climate model: A comparison of methods. *Int. J. Climatol.*, **33**, 1367–1381, <https://doi.org/10.1002/joc.3518>.
- Lamb, P. J., and R. A. Pepler, 1992: Further case studies of tropical Atlantic surface atmospheric and oceanic patterns associated with sub-Saharan drought. *J. Climate*, **5**, 476–488, [https://doi.org/10.1175/1520-0442\(1992\)005<0476:FCSOTA>2.0.CO;2](https://doi.org/10.1175/1520-0442(1992)005<0476:FCSOTA>2.0.CO;2).
- Lebel, T., and Coauthors, 2009: AMMA-CATCH studies in the Sahelian region of West-Africa: An overview. *J. Hydrol.*, **375**, 3–13, <https://doi.org/10.1016/j.jhydrol.2009.03.020>.
- Lensky, I. M., and D. Rosenfeld, 1997: Estimation of precipitation area and rain intensity based on the microphysical properties retrieved from NOAA AVHRR data. *J. Appl. Meteor.*, **36**, 234–242, [https://doi.org/10.1175/1520-0450\(1997\)036<0234:EOPAAR>2.0.CO;2](https://doi.org/10.1175/1520-0450(1997)036<0234:EOPAAR>2.0.CO;2).
- L'Hôte, Y., G. Mahé, B. Somé, and J. P. Triboulet, 2002: Analysis of a Sahelian annual rainfall index from 1896 to 2000; The drought continues. *Hydrol. Sci. J.*, **47**, 563–572, <https://doi.org/10.1080/02626660209492960>.
- Liu, C., E. J. Zipser, and S. W. Nesbitt, 2007: Global distribution of tropical deep convection: Different perspectives from TRMM infrared and radar data. *J. Climate*, **20**, 489–503, <https://doi.org/10.1175/JCLI4023.1>.
- Mantas, V. M., Z. Liu, C. Caro, and A. Pereira, 2015: Validation of TRMM Multi-Satellite Precipitation Analysis (TMPA) products in the Peruvian Andes. *Atmos. Res.*, **163**, 132–145, <https://doi.org/10.1016/j.atmosres.2014.11.012>.
- Maranan, M., A. Fink, and P. Knippertz, 2018: Rainfall types over southern West Africa: Objective identification, climatology and synoptic environment. *Quart. J. Roy. Meteor. Soc.*, **144**, 1628–1648, <https://doi.org/10.1002/qj.3345>.
- , A. H. Fink, P. Knippertz, S. D. Francis, A. B. Akpo, G. Jegede, and C. Yorke, 2019: Interactions between convection and a moist vortex associated with an extreme rainfall event over southern west africa. *Mon. Wea. Rev.*, **147**, 2309–2328, <https://doi.org/10.1175/MWR-D-18-0396.1>.
- Mazzoglio, P., F. Laio, S. Balbo, P. Boccoardo, and F. Disabato, 2019: Improving an extreme rainfall detection system with GPM IMERG data. *Remote Sens.*, **11**, 677, <https://doi.org/10.3390/rs11060677>.
- McCollum, J. R., A. Gruber, and M. B. Ba, 2000: Discrepancy between gauges and satellite estimates of rainfall in equatorial Africa. *J. Appl. Meteor.*, **39**, 666–679, <https://doi.org/10.1175/1520-0450-39.5.666>.
- Monsieurs, E., and Coauthors, 2018: Evaluating TMPA rainfall over the sparsely gauged East African Rift. *J. Hydrometeorol.*, **19**, 1507–1528, <https://doi.org/10.1175/JHM-D-18-0103.1>.
- Nair, S., G. Srinivasan, and R. Nemani, 2009: Evaluation of multi-satellite TRMM derived rainfall estimates over a western state of India. *J. Meteor. Soc. Japan*, **87**, 927–939, <https://doi.org/10.2151/JMSJ.87.927>.
- Nakajima, T., and M. D. King, 1990: Determination of the optical thickness and effective particle radius of clouds from reflected solar radiation measurements. Part I: Theory. *J. Atmos. Sci.*, **47**, 1878–1893, [https://doi.org/10.1175/1520-0469\(1990\)047<1878:DOTOTA>2.0.CO;2](https://doi.org/10.1175/1520-0469(1990)047<1878:DOTOTA>2.0.CO;2).
- Nicholson, S. E., 1981: Rainfall and atmospheric circulation during drought periods and wetter years in West Africa. *Mon. Wea. Rev.*, **109**, 2191–2208, [https://doi.org/10.1175/1520-0493\(1981\)109<2191:RAACDD>2.0.CO;2](https://doi.org/10.1175/1520-0493(1981)109<2191:RAACDD>2.0.CO;2).
- , and Coauthors, 2003: Validation of TRMM and other rainfall estimates with a high-density gauge dataset for West Africa. Part II: Validation of TRMM rainfall products. *J. Appl. Meteor.*, **42**, 1355–1368, [https://doi.org/10.1175/1520-0450\(2003\)042<1355:VOTAOR>2.0.CO;2](https://doi.org/10.1175/1520-0450(2003)042<1355:VOTAOR>2.0.CO;2).
- , A. K. Dezfuli, and D. Klotter, 2012: A two-century precipitation dataset for the continent of Africa. *Bull. Amer. Meteor. Soc.*, **93**, 1219–1231, <https://doi.org/10.1175/BAMS-D-11-00212.1>.
- Paeth, H., A. H. Fink, S. Pohle, F. Keis, H. Mächel, and C. Samimi, 2011: Meteorological characteristics and potential causes of the 2007 flood in sub-Saharan Africa. *Int. J. Climatol.*, **31**, 1908–1926, <https://doi.org/10.1002/joc.2199>.
- Panthou, G., T. Vischel, and T. Lebel, 2014: Recent trends in the regime of extreme rainfall in the central Sahel. *Int. J. Climatol.*, **34**, 3998–4006, <https://doi.org/10.1002/joc.3984>.
- Petković, V., and C. D. Kummerow, 2017: Understanding the sources of satellite passive microwave rainfall retrieval systematic errors over land. *J. Appl. Meteor. Climatol.*, **56**, 597–614, <https://doi.org/10.1175/JAMC-D-16-0174.1>.
- Petty, G. W., 1995: The status of satellite-based rainfall estimation over land. *Remote Sens. Environ.*, **51**, 125–137, [https://doi.org/10.1016/0034-4257\(94\)00070-4](https://doi.org/10.1016/0034-4257(94)00070-4).
- Pfeifroth, U., J. Trentmann, A. H. Fink, and B. Ahrens, 2016: Evaluating satellite-based diurnal cycles of precipitation in the African tropics. *J. Appl. Meteor. Climatol.*, **55**, 23–39, <https://doi.org/10.1175/JAMC-D-15-0065.1>.
- Roca, R., P. Chambon, I. Jobard, P.-E. Kirstetter, M. Gosset, and J. C. Bergès, 2010: Comparing satellite and surface rainfall

- products over West Africa at meteorologically relevant scales during the AMMA campaign using error estimates. *J. Appl. Meteor. Climatol.*, **49**, 715–731, <https://doi.org/10.1175/2009JAMC2318.1>.
- Roebeling, R., H. Deneke, and A. Feijt, 2008: Validation of cloud liquid water path retrievals from SEVIRI using one year of cloudnet observations. *J. Appl. Meteor. Climatol.*, **47**, 206–222, <https://doi.org/10.1175/2007JAMC1661.1>.
- Rosenfeld, D., and Coauthors, 2011: Glaciation temperatures of convective clouds ingesting desert dust, air pollution and smoke from forest fires. *Geophys. Res. Lett.*, **38**, L21804, <https://doi.org/10.1029/2011GL049423>.
- Sanogo, S., A. H. Fink, J. A. Omotosho, A. Ba, R. Redl, and V. Ermert, 2015: Spatio-temporal characteristics of the recent rainfall recovery in West Africa. *Int. J. Climatol.*, **35**, 4589–4605, <https://doi.org/10.1002/joc.4309>.
- Schneider, U., T. Fuchs, A. Meyer-Christoffer, and B. Rudolf, 2008: Global precipitation analysis products of the GPCP. Global Precipitation Climatology Centre (GPCP), DWD, 12 pp.
- Stengel, M., A. Kniffka, J. Meirink, M. Lockhoff, J. Tan, and R. Hollmann, 2014: CLAAS: The CM SAF cloud property data set using SEVIRI. *Atmos. Chem. Phys.*, **14**, 4297–4311, <https://doi.org/10.5194/acp-14-4297-2014>.
- Stephens, G., 1978: Radiation profiles in extended water clouds. II: Parameterization schemes. *J. Atmos. Sci.*, **35**, 2123–2132, [https://doi.org/10.1175/1520-0469\(1978\)035<2123:RPIEWC>2.0.CO;2](https://doi.org/10.1175/1520-0469(1978)035<2123:RPIEWC>2.0.CO;2).
- Tan, J., W. A. Petersen, and A. Tokay, 2016: A novel approach to identify sources of errors in IMERG for GPM ground validation. *J. Hydrometeorol.*, **17**, 2477–2491, <https://doi.org/10.1175/JHM-D-16-0079.1>.
- Taylor, J. W., and Coauthors, 2019: Aerosol influences on low-level clouds in the West African monsoon. *Atmos. Chem. Phys.*, **19**, 8503–8522, <https://doi.org/10.5194/acp-19-8503-2019>.
- Thiemig, V., R. Rojas, M. Zambrano-Bigiarini, V. Levizzani, and A. De Roo, 2012: Validation of satellite-based precipitation products over sparsely gauged African river basins. *J. Hydrometeorol.*, **13**, 1760–1783, <https://doi.org/10.1175/JHM-D-12-032.1>.
- Tian, F., S. Hou, L. Yang, H. Hu, and A. Hou, 2018: How does the evaluation of the GPM IMERG rainfall product depend on gauge density and rainfall intensity? *J. Hydrometeorol.*, **19**, 339–349, <https://doi.org/10.1175/JHM-D-17-0161.1>.
- Tian, Y., G. J. Huffman, R. F. Adler, L. Tang, M. Sapiiano, V. Maggioni, and H. Wu, 2013: Modeling errors in daily precipitation measurements: Additive or multiplicative? *Geophys. Res. Lett.*, **40**, 2060–2065, <https://doi.org/10.1002/grl.50320>.
- Tompkins, A. M., and A. A. Adebisi, 2012: Using cloudsat cloud retrievals to differentiate satellite-derived rainfall products over West Africa. *J. Hydrometeorol.*, **13**, 1810–1816, <https://doi.org/10.1175/JHM-D-12-039.1>.
- UN, 2009: *Water in a Changing World*. United Nations World Water Assessment Programme Development Rep. 3, Vol. 1, Earthscan, 318 pp.
- Upton, G. J., 2002: A correlation–regression method for tracking rainstorms using rain-gauge data. *J. Hydrol.*, **261**, 60–73, [https://doi.org/10.1016/S0022-1694\(01\)00618-7](https://doi.org/10.1016/S0022-1694(01)00618-7).
- van de Giesen, N., R. Hut, and J. Selker, 2014: The trans-African hydro-meteorological observatory (TAHMO). *Wiley Interdiscip. Rev.: Water*, **1**, 341–348, <https://doi.org/10.1002/wat2.1034>.
- Wang, J., and D. B. Wolff, 2010: Evaluation of TRMM ground-validation radar-rain errors using rain gauge measurements. *J. Appl. Meteor. Climatol.*, **49**, 310–324, <https://doi.org/10.1175/2009JAMC2264.1>.
- Wilks, D. S., 2011: *Statistical Methods in the Atmospheric Sciences*. 3rd ed. International Geophysics Series, Vol. 100, Academic Press, 704 pp.
- Wolff, D. B., D. Marks, E. Amitai, D. Silberstein, B. Fisher, A. Tokay, J. Wang, and J. Pippitt, 2005: Ground validation for the Tropical Rainfall Measuring Mission (TRMM). *J. Atmos. Oceanic Technol.*, **22**, 365–380, <https://doi.org/10.1175/JTECH1700.1>.
- Xu, R., F. Tian, L. Yang, H. Hu, H. Lu, and A. Hou, 2017: Ground validation of GPM IMERG and TRMM 3B42V7 rainfall products over southern Tibetan Plateau based on a high-density rain gauge network. *J. Geophys. Res. Atmos.*, **122**, 910–924, <https://doi.org/10.1002/2016JD025418>.
- You, Y., H. Meng, J. Dong, and S. Rudlosky, 2019: Time-lag correlation between passive microwave measurements and surface precipitation and its impact on precipitation retrieval evaluation. *Geophys. Res. Lett.*, **46**, 8415–8423, <https://doi.org/10.1029/2019GL083426>.
- Young, M. P., J. C. Chiu, C. J. Williams, T. H. Stein, M. Stengel, M. D. Fielding, and E. Black, 2018: Spatio-temporal variability of warm rain events over southern West Africa from geostationary satellite observations for climate monitoring and model evaluation. *Quart. J. Roy. Meteor. Soc.*, **144**, 2311–2330, <https://doi.org/10.1002/qj.3372>.
- Zubieta, R., A. Getirana, J. C. Espinoza, W. Lavado-Casimiro, and L. Aragon, 2017: Hydrological modeling of the Peruvian-Ecuadorian amazon basin using GPM-IMERG satellite-based precipitation dataset. *Hydrol. Earth Syst. Sci.*, **21**, 3543–3555, <https://doi.org/10.5194/hess-21-3543-2017>.