Reliability Assessment of 15 Gridded Rainfall Datasets for the Construction of a Daily High-Resolution Reanalysis across Senegal for Agroclimatic Applications

A. MBENGUE,[a,b] B. SULTAN,[b] R. FANIRIANTSOA,[c] O. NDIAYE,[a] A. DIONGUE-NIANG,[a] F. SATGÉ,[b] M. L. DIOP,[a] D. NDIAYE,[a] AND O. KONTE[a]

[a] *Meteorological Exploitation Directorate, ANACIM, Dakar, Senegal*
[b] *University of Montpellier, ESPACE-DEV, IRD, France*
[c] *International Research Institute for Climate and Society, Columbia University, Palisades, New York*

(Manuscript received 31 December 2024, in final form 20 July 2025, accepted 1 August 2025)

ABSTRACT: This study focuses on developing a new high-resolution gridded rainfall dataset for Senegal, essential for supporting rainfed agriculture, which is sensitive to climate variability. Given the limited number of rain gauges, the research evaluates 15 publicly available gridded rainfall datasets (P datasets) against data from 21 stations of the Senegalese National Meteorological Service (ANACIM) over a 17-yr period (2005–21). The evaluation employs several agroclimatic indices, including the onset and cessation of rain, duration of the rainy season, and extreme events. The findings reveal that the reliability of P datasets varies significantly based on the metrics used. For total rainfall, African Rainfall Climatology, version 2 (ARC2), Climate Hazards Infrared Precipitation with Station (CHIRPS), ERA5, and Rainfall Estimation Algorithm, version 2 (RFEv2) emerged as the most reliable datasets, with ERA5 achieving the highest Kling–Gupta efficiency (KGE) value of 0.81 at daily scale. In terms of agroclimatic parameters, ARC2, CHIRPS, and RFEv2 excelled in accurately representing the start (KGE $\geq$ 0.45) and end (KGE $\geq$ 0.39) dates of the rainy season. However, P datasets generally overestimate rainfall events and struggle with identifying dry spells. The newly constructed merged dataset (M dataset) demonstrated over 100% improvement in correlation for daily estimates and significant bias reductions: 99.19% for ARC2, 80% for CHIRPS, and 90.57% for RFEv2. This research provides critical insights for selecting appropriate datasets to enhance climate information for agricultural decision-making in Senegal.

KEYWORDS: Climatology; Satellite observations; Surface observations; Data assimilation; Reanalysis data

1. Introduction

Rainfall is a critical component of the hydrological cycle and plays an important role in rainfed agriculture, hydropower generation, socioeconomic development, and human activities (Sultan et al. 2013; Degefu et al. 2022). The pronounced variability of rainfall and the resurgence of extreme hydrometeorological events such as floods, droughts, and heavy rains pose a threat to people and ecosystems. This is particularly true throughout West Africa, and especially for the agricultural sector, which is critical to the survival of the population and highly dependent on rainfall (Houngnibo et al. 2023). Variations in total rainfall could have adverse effects on rainfed agriculture and food security.

In general, rainfall in West Africa exhibits high interannual and multidecadal variability (Nicholson and Webster 2007; Owusu and Waylen 2013). Several studies have shown a declining trend in rainfall for the period 1970–2000 with observed signs of recovery in the 2000s (Owusu and Waylen 2013; Manzanas et al. 2014). High-quality rainfall data are essential to monitor changes in rainfall patterns over time, to evaluate and improve climate and numerical weather models, and to support decision-making in the most climate-vulnerable sectors. Historical data assess climate risks or long-term trends, while near-real-time data are critical for monitoring weather-related risks and enabling the development of timely and actionable early warning systems by governments and agencies (Maidment et al. 2017). Weather radars monitor rainfall with spatial distribution over large, even remote, areas. However, radar stations are expensive, and few are available worldwide. Furthermore, radar signal interference hinders accurate rainfall estimation in complex areas (Tang et al. 2016; Zeng et al. 2018). Recently, several authors have highlighted the potential use of mobile phone signal attenuation during rain events to collect rainfall amounts (Messer et al. 2006; Zinevich et al. 2008; Overeem et al. 2011; Doumounia et al. 2014). While these estimates are accurate, they are limited to densely populated regions (e.g., urban areas). In addition, this technique faces challenges in accessing data held by private mobile phone companies.

In most cases, records from networks of in situ meteorological stations are the most reliable means of obtaining accurate local information on weather and climate (Yonaba et al. 2022). However, although Africa is one of the most vulnerable regions to climate variability and change, there are spatial and temporal gaps in the available measurements. In the Global Climate Observing System Surface Network (GCOS–GSN), the African continent stands out in terms of the quality of its measurement network. In 2019, only 26% of stations met the

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/JAMC-D-24-0238.s1.

Corresponding author: A. Mbengue, asse.mbengue@anacim.sn

DOI: 10.1175/JAMC-D-24-0238.1

World Meteorological Organization (WMO) standards, and 35% of stations were not operational (WMO 2020). Difficulties in access, political instability, and economic issues have often resulted in sparsely and unevenly distributed rain gauge networks that inadequately capture the spatial variability of rainfall (Lebel et al. 1997). Since the advent of remote sensing, numerous gridded rainfall datasets (P datasets) have been developed as an alternative and complement to rain gauge networks to provide more comprehensive global coverage of rainfall. P datasets often combine data from (i) in situ rain gauge networks, (ii) satellite remote sensing data from geostationary and polar orbiting platforms, and (iii) climate reanalyses, which are datasets that integrate historical weather observations with modern meteorological models to provide a consistent and comprehensive record of past atmospheric conditions. Examples include ECMWF Re-Analysis (ERA) and Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA2). In remote regions, P datasets have proven effective for water resource management, improving our understanding of droughts (Toté et al. 2015; Agutu et al. 2017; Guo et al. 2017; Satgé et al. 2017), floods (Nikolopoulos et al. 2013; Toté et al. 2015; Gao et al. 2017), rainfall variability (Carvalho et al. 2012; Arvor et al. 2017), river discharge (Collischonn et al. 2008; Sun et al. 2018; Zhang et al. 2018; Satgé et al. 2019), snow-cover dynamics (Satgé et al. 2019), and agricultural productivity (de Wit et al. 2010; Thaler et al. 2018). Meteorological observation data are the backbone of national meteorological services. For Senegal, long-term rainfall records are available for several meteorological stations throughout the country in the database of the Senegalese National Meteorological Service (ANACIM) since the 1800s. However, some stations have long periods of missing data and are unevenly distributed in certain regions, particularly in the north and central-east (agropastoral zone). As a result, large areas of Senegal are not well covered (in space and time), which limits climate monitoring at the local level and necessitates the use of rainfall estimates that are representative of the area surrounding the station over several square kilometers. P datasets, which are available at a global scale and on a regular grid, provide an unprecedented opportunity to overcome this problem.

However, P datasets are subject to significant biases that need to be assessed to appropriately weigh the use of P-dataset estimates. In this context, a handful of studies have focused on the entire West African region, including Senegal (Awange et al. 2016; Akinsanola et al. 2017, and more recently, Satgé et al. 2020; Houngnibo et al. 2023; Kouakou et al. 2023), while the majority of assessment studies focus on reduced spatial scales (countries or watersheds) of different climatic zones and are based on different statistical indices, spatial and temporal scales, and time periods, leading to difficulties in comparing reliability assessments of P datasets. For example, when comparing TMPA, CMORPH, and PERSIANN P datasets with reference gauge estimates, CMORPH was shown to be the most reliable P dataset in Pakistan, China, Bali, and Indonesia (Hussain et al. 2018; Rahmawati and Lubczynski 2018; Su et al. 2017; Zeng et al. 2018). Thiemig et al. (2012) evaluated seven P datasets including CMORPH, GPROFv6, GSMaP moving vector with Kalman filter, version 5 (GSMaP-MVKv5), Rainfall Estimation Algorithm, version 2 (RFEv2.0), TMPAv6, PERSIANN, and ERA-Interim in the Volta basin; Gosset et al. (2013) and Ramarohetra et al. (2013) investigated the reliability of nine [CMORPH, Estimation of Precipitation by Satellites-second generation (EPSAT-SG), GPCP, GSMaP-MVK, GSMaP-real time (GSMaP-RT), RFEv2, TMPAv6, TMPA-real time, version 6 (TMPA-RTv6), and PERSIANN] and seven (PERSIANN, CMORPH, TMPA-RTv6, TMPA-adjusted, version 6 (TMPA-Adjv6), GSMaP-MVK, GPCP-one-degree daily (GPCP-1dd), and RFEv2] P datasets in Benin and Niger, respectively, for agricultural applications.

These authors demonstrated that their use could introduce significant biases in crop or hydrological modeling. More recently, in Burkina Faso, Dembélé and Zwart (2016) evaluated seven P datasets and showed that African Rainfall Climatology, version 2 (ARC2), RFE, and Tropical Applications of Meteorology using Satellite Data and Ground-Based Observations (TAMSAT) African Rainfall Climatology and Time Series (TARCAT) were more reliable for drought monitoring, while PERSIANN, Climate Hazards Infrared Precipitation with Station (CHIRPS), and TMPA were more suitable for flood monitoring. Maidment et al. (2017) evaluated two versions of TAMSAT (TAMSATv3 and TAMSATv2) and six other P datasets, including ARCv2, CHIRPv2, CHIRPSv2, CMORPHv1, RFE, and TMPAv7 over West Africa, including Nigeria, Niger, Uganda, Zambia, and Mozambique. The results indicate that both versions of TAMSAT daily estimates reliably detect rainy days but are less competent in capturing rainfall amounts. In addition, Poméon et al. (2017) evaluated ten P datasets, including CFSR, CHIRPS, CMORPH, version 1, raw and corrected (CMORPHv1 RAW and CRT), PERSIANN-CDR, RFEv2, TAMSATv2, TMPAv7, TMPA-RTv7, and GPCP, over six watersheds in Burkina, Nigeria, and Ghana. While performance varies, most P datasets accurately predict observed runoff in each watershed. The best results in this study were obtained with GPCP data, which uses a variety of inputs, including infrared and microwave satellite data, as well as rain gauge observations. Nwachukwu et al. (2020) evaluated 16 satellite-based rainfall products [Satellite Precipitation Products (SPPs)] using rainfall data from 11 rain gauge stations across Nigeria from 2000 to 2012. Although some SPPs replicated the north–south rainfall pattern, significant discrepancies were found, with Multisource Weighted-Ensemble Precipitation, version 2.2 (MSWEP v2.2) and IMERG final run, version 6 (IMERG-F v6) performing best overall. Nevertheless, the reported studies show that the results are mostly limited in space (country or basin scale) and in terms of the sample of P datasets considered. However, to our knowledge, none of these evaluations has been carried out in Senegal.

In general, the most realistic P datasets are those that combine satellite estimates and rain gauge observations (Sun et al. 2014; Houngnibo et al. 2023; Kouakou et al. 2023). These calibrated P datasets use rain gauge measurements from the Global Telecommunication System (GTS) network to reduce the monthly bias relative to the gauges. For example, Gosset et al. (2013) and Ramarohetra et al. (2013) found that such calibrated satellite products are best suited for hydrological and agricultural applications because they minimize bias in both daily rainfall distribution and total annual rainfall. However,

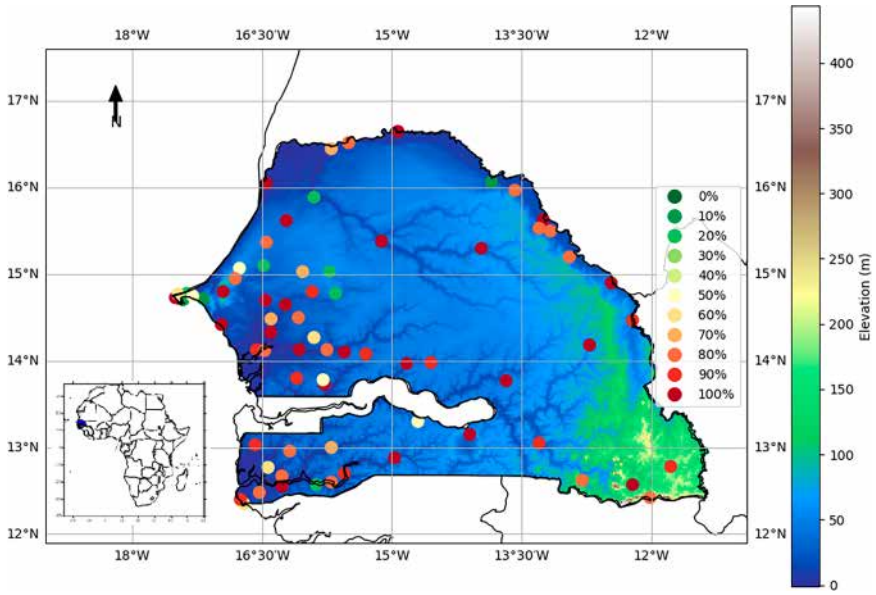

FIG. 1. Elevation map of Senegal showing the location of the country, the positions of weather stations, and the total availability of data used in this study over the period 1981–2021. Height obtained from the shuttle radar topography model.

the reliability of these hybrid datasets varies in space and time with the number of available gauges used in the interpolation process (Sun et al. 2014). This is an important issue in Africa, where the number of gauges is lower than in other regions of the world (WMO 2020). Furthermore, there has been a significant decline in the number of stations across the African continent, particularly in Senegal, that are used to calibrate datasets such as CHIRPS, which could limit their performance. There have been initiatives across the African continent to address the uneven distribution of meteorological stations by assimilating ground-based rain observations with gridpoint estimates. These efforts have demonstrated their importance in producing accurate and reliable climate information. By integrating satellite data with sparse ground observations, these initiatives improve the spatial coverage and quality of precipitation data, which is critical for various hydrological and agricultural applications. Such advances help mitigate the challenges posed by the limited number of weather stations and ensure better climate monitoring and forecasting capabilities (Ayehu et al. 2018; Dinku 2018; Siebert et al. 2019).

Given the context and issues outlined above, a holistic assessment of available P datasets is required prior to their use in the production of climate information to support decision-making. The overall objective of this study is to make best use of the national network of meteorological station data to identify the most accurate, publicly available P datasets in Senegal. A set of 15 P datasets of daily rainfall estimates over Senegal was compared with rain gauge measurements over a 17-yr period (2005–21). As a first step, we investigated the reliability of the 15 P datasets in reproducing rainfall variability at daily time scales and in estimating important agroclimatic indices such as the onset of rain, the length of the rainy season, and dry spells. In a second step, we merged the ANACIM meteorological dataset with the three best P datasets to create a new high-resolution merged datasets (M datasets) for Senegal. These two steps provide important insights for selecting the most appropriate datasets for a specific application in Senegal and for proposing a customized dataset based on the combination of rain gauge and P-dataset estimates.

2. Materials and methods

a. Study area

Senegal is located between 12°8′–16°09′N latitude and 11°–18°W longitude. It is bordered by the Atlantic Ocean to the west, Mauritania to the north-northeast, Mali to the east-southeast, Guinea to the southeast, and Guinea-Bissau to the southwest (Fig. 1). Senegal's latitude allows winds of different origins and characteristics to alternate throughout the year. Thus, two main seasons characterize the climatic regime: a dry season (from November to April–May), characterized by the predominance of maritime trade winds (from the west) and continental trade winds (inland), and a rainy season, from May–June to October, dominated by the monsoon flow of the Saint Helena anticyclone. Rainfall peaks in August. The main climatic feature is the significant spatial variability of rainfall, which averages over 1000 mm in the south and less than 300 mm in the north. According to Sagna (2007), the spatial distribution of rainfall divides the country into two major climatic regions on either side of the 500-mm isohyet (Fig. S1 in the online supplemental material):

- The Sahelian region north of this isohyet comprises two rainfall regimes: (i) the northern Sahelian (N Sahelian) regime with rainfall less than 300 mm and (ii) the southern Sahelian (S Sahelian) regime with rainfall between 300 and 500 mm.

- The Sudanese region south of the 500-mm isohyet includes (i) the north Sudanese (N Sudanese) regime between the 500- and 1000-mm isohyets and (ii) the south Sudanese (S Sudanese) regime beyond 1000 mm.

Beyond the climatic characterization described above, seasonal rainfall in Senegal plays a fundamental role in several crucial aspects of development, particularly in the areas of water resources, agriculture, and livestock. This annual rainfall is an important source for recharging groundwater, thereby contributing to the availability of water resources essential for daily life, agricultural irrigation, and the preservation of local ecosystems. The importance of rainfall to food security in the region is undeniable. Farmers rely heavily on seasonal rainfall to ensure sufficient agricultural yields, and variations in the amount and distribution of rainfall can have significant impacts on food production, the availability of pasture for livestock, and ultimately the food security of local populations.

b. *Reference rainfall dataset*

Daily rainfall from 72 stations has been made available by ANACIM for the period 1981–2021 (Fig. 1). The percentage of available (nonmissing) data over this period varies from 10% to 100%, with strong spatial and temporal variability. Due to their recent establishment, several stations provided less than 40% of the data over the period considered. For the reliability assessment of the P dataset, a subset of 21 stations with more than 95% of complete observations during the rainy season in Senegal (May–October) for the period 2005–21 was retained. The period 2005–21 was chosen to match the availability of satellite data, as many satellite missions started in 2000. To ensure an objective comparison between the P datasets, synoptic and climatological stations that were generally assimilated in some of the considered P datasets were systematically excluded. On the other hand, the 72 stations covering the period from 1981 to 2021 were used to produce the merged data.

These data were subjected to a quality control consisting of the following steps: (i) verification of the geographical coordinates of the in situ station; (ii) verification of false zeros (not reported) for a given month. A false zero is defined as an abnormal zero value observed throughout the month during the rainy season. The main source of false zeros is the ambiguity in the coding of days without observations (missing values) and days without rain (zero values) during the digitization of the records; (iii) checking for the presence of outliers in the time series. Outliers can be checked for temporal and spatial cases; (iv) a temporal check is performed each month to ensure that each observed value is consistent with the climatology of each station. Any values detected by the test are flagged as suspicious or anomalous and require careful verification. Spatial checking is done by comparing each value from a given station with the values from neighboring stations on the same date. Incorrect and anomalous values were thus identified. However, even with objective methods, undetected observation errors are possible. Erroneous values are replaced by missing values, unless the exact value is found after consulting the documentation. All days with more than 300 mm day^{-1} rainfall, which are considered as anomalous values, are marked as missing.

c. *Selected P datasets*

Fifteen P datasets were selected for the study. Two of these are reanalyses (ERA5, MERRA2) and the other 13 are satellite-based rainfall estimates. Climate reanalysis datasets are produced by combining observations from different sources, such as ground weather stations, radiosondes, satellites, and other meteorological instruments. These data are then used in numerical models to reconstruct the past state of the atmosphere over a given period. Satellite precipitation estimates, on the other hand, are methods used to estimate the amount of precipitation falling on Earth's surface using data from meteorological satellites. Meteorological satellites can be equipped with specific sensors designed to detect various atmospheric features, including clouds and precipitation. Rainfall sensors, such as microwave radiometers, radars, or lidars, measure various properties of water droplets or ice crystals in the atmosphere. The data collected by these sensors are then used in complex algorithms to estimate the amount of rainfall (see the references cited in Table 1 for more details). These algorithms consider several factors such as atmospheric temperature, cloud reflectivity, cloud-top temperature, and other characteristics to estimate rainfall. These rainfall estimates are either quasi real time or calibrated to ground observations. The CHIRP dataset is a global, uncalibrated satellite-based rainfall estimate derived from infrared (IR) observations at 0.05° spatial resolution. It provides daily precipitation data from 1981 to the present and serves as the foundation for the bias-corrected CHIRPS product (Funk et al. 2014). Soil moisture to rain ASCAT (SM2RainASCAT) is slightly different, as it uses a "bottom–up" approach that uses satellite observations of soil moisture to estimate rainfall using the SM2RAIN algorithm (Brocca et al. 2014).

P datasets have different spatial and temporal resolutions. The mean annual precipitation over Senegal (2005–21) may differ slightly from one P dataset to another (Fig. 2). While the north–south rainfall gradient is generally well reproduced by P datasets compared to observations, the amplitude of annual rainfall over the country may be over- or underestimated depending on the P dataset considered. For example, PERSIANN estimates show significantly higher rainfall amounts in Senegal, with a cumulative total of at least 600 mm in the northern half of the country, as this area is framed by the 200–400-mm isohyet. This is also observed with CMORPH-CRT and PERSIANN-CCS-CDR data, which tend to overestimate rainfall in Senegal but to a lesser extent. The ERA5, CHIRP, MERRA2, and SM2RainASCAT data simulate less rainfall, especially in the northern half of the country.

d. *Evaluation of rainfall estimates over Senegal*

1) P-DATASET PREPROCESSING

To ensure a consistent comparison, all P datasets are interpolated onto the TAMSATv3.1 product grid (0.0375° × 0.0375°), which is the finest resolution of all datasets, using bilinear transformation. This technique estimates values at the target grid cells by taking a distance-weighted average of the four surrounding points. It is commonly used for downscaling or remapping gridded climate data due to its simplicity and smooth spatial

TABLE 1. Characteristics of rainfall estimation datasets.

Data	Category	Data sources	Grid	Spatial coverage	Temporal resolution	Availability	Study period	Reference
ARC2	Regional calibration	IR and gauges	0.1°	Africa, 20.0°W–55.0°E and 40.0°S–40.0°N	Daily	1983–the present	2005–21	Novella and Thiaw (2013)
CHIRP	No calibrated	IR	0.05°	Global	Daily	1984–the present	2005–21	Funk et al. (2014)
CHIRPS	Global calibration	IR and gauges	0.05°	Global	Daily	1985–the present	2005–21	Funk et al. (2015)
CMORPH-BLD	Near–real time	IR and PMW	0.25°	Global	Daily	1998–2022	2005–21	Xie et al. (2019)
CMORPH-CRT	Near–real time	IR, PMW, and gauges	0.25°	Global	Daily	1998–2019	2005–19	Xie et al. (2019)
ERA5	Global calibration	Combines vast amounts of historical observations into global estimates using advanced modeling and data assimilation systems	0.25°	Global	Hourly	1940–the present	2005–21	Copernicus Climate Change Service (2017)
IMERG v06 F	Global calibration	IR, PMW, radars, and gauges	0.1°	Global	Daily	2000–21	2005–21	Huffman et al. (2019)
MERRA2		Modern-Era Retrospective Analysis for Research and Applications, version 2	0.5° × 0.625°	Global	Hourly	1980–the present	2005–21	GMAO (2014)
PERSIANN-CCS-CDR	Near–real time	Modified PERSIANN algorithm, the National Centers for Environmental Prediction (NCEP) Stage IV hourly rainfall to train the ANN model	0.04°	60°S–60°N	Daily	1983–2022	2005–21	Sadeghi et al. (2021)
PERSIANN-CDR	Near–real time	IR, performing segmentation of the cloud image under different temperature thresholds	0.25°	60°S–60°N	Daily	1983–2022	2005–21	Nguyen et al. (2019)
PERSIANN	Near–real time	Artificial neural network and climate data record	1.25°	60°S–60°N	Daily	2000–the present	2005–21	Nguyen et al. (2019), Hsu et al. (1997)
PRISM	Regional calibration	IR from soil moisture	0.25°	Africa	Daily	2010–20	2010–20	Pellarin (2019)
RFEV2	Regional calibration	IR, PMW, and gauges	0.1°	Africa	Daily	2000–the present	2005–21	Love (2002)
SM2RainASCAT	No calibrated	Soil moisture to rain applied on ESA climate change initiative v2	0.1°	Global	Daily	2007–21	2007–21	Brocca et al. (2019)
TAMSATv3.1	Regional calibration	IR and gauges	0.0375°	African continent, including Madagascar (N: 38.025°, S: −35.9625°, W: −19.012°, E: 51.975°)	Daily	1983–the present	2005–21	Maidment et al. (2017), Tarnavsky et al. (2014), Maidment et al. (2014)

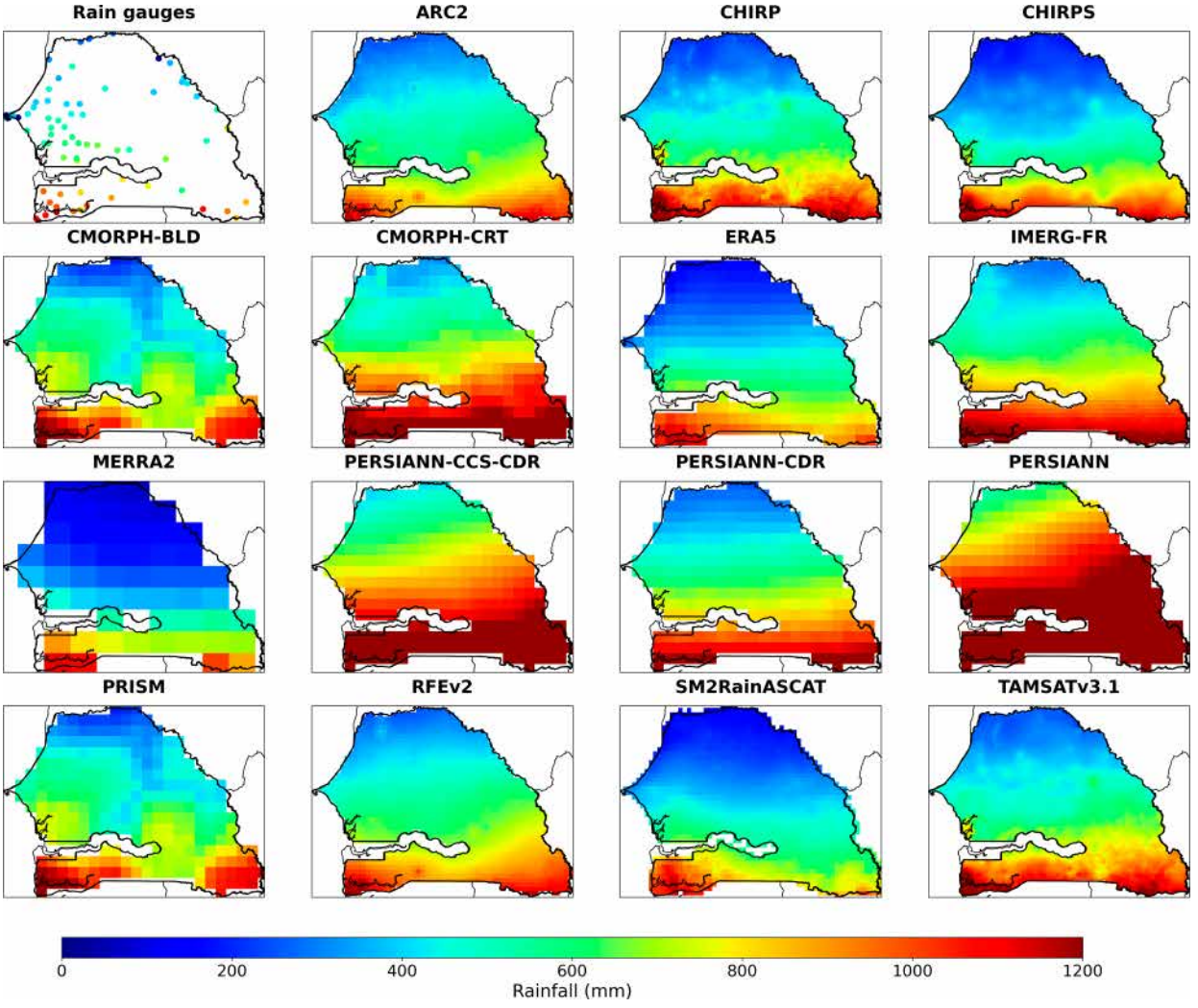

FIG. 2. Annual cumulative rainfall of observed and P-dataset estimates for the period 2005–21.

transitions (Beck et al. 2019). Daily accumulation was obtained by aggregating the P-dataset estimates from 6 to 6 h for those produced at the hourly time step (ERA5, MERRA2); the others available per day were considered as such. Only gridded products were interpolated, while station data remained unchanged. For comparison, gridcell value matching station locations were extracted.

2) P-DATASET ASSESSMENT BASED ON QUANTITATIVE METRICS

The evaluation is done by comparing the estimates of the 21 rain gauges and their nearest grid points for the 15 P datasets at the daily time step. The reliability of the P datasets is assessed by calculating a number of skill scores, including the Pearson correlation (CORR) coefficient [Eq. (1)], the multiplicative bias [BIAS; Eq. (2)], the mean absolute error [MAE; Eq. (3)], the root-mean-square error [RMSE; Eq. (4)], and the Kling–Gupta efficiency (KGE) index [Eq. (5)]. The KGE is a robust measure that assesses the quality of a model simulation by comparing simulated time series with observations. It combines several components to provide an overall assessment of model performance. It was developed by Kling et al. (2012) and has already been used in several P-dataset evaluation studies (Satgé et al. 2019; Houngnibo et al. 2023; Kouakou et al. 2023):

$$\mathrm{CORR} = \frac{\sum_{i=1}^{N}(F_i - \overline{F})(O_i - \overline{O})}{\sqrt{\sum_{i=1}^{N}(F_i - \overline{F})^2}\sqrt{\sum_{i=1}^{N}(O_i - \overline{O})^2}} \tag{1}$$

$$\mathrm{BIAS} = \frac{1/N\sum_{i=1}^{N} F_i}{1/N\sum_{i=1}^{N} O_i}, \tag{2}$$

$$\mathrm{MAE} = 1/N\sum_{i=1}^{N}|F_i - O_i|, \tag{3}$$

TABLE 2. Contingency table used for the categorical statistical analysis of rainfall. A threshold value of 1.0 mm is used to separate rainfall to no rainfall events.

		Gauge observations	
		Rainfall	No rainfall
Estimations products	Rainfall	Hits "a"	False alarms "b"
	No rainfall	Misses "c"	Correct negatives "d"

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(F_i - O_i)^2}, \tag{4}$$

$$\mathrm{KGE} = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2}. \tag{5}$$

Here, O_i = rainfall observation; $\overline{O}$ = average rainfall observation; F_i = estimates of P dataset; $\overline{F}$ = average estimates of P dataset; r = same as CORR, the Pearson correlation coefficient between observations and simulations; $\beta = (\mu F_i/\mu O_i)$ is the ratio between the mean of simulations μF_i and that of observations μO_i, an indicator of mean flow bias; and $\gamma = (\sigma_{F_i}/\sigma_{O_i})$ is the ratio between the standard deviation of the simulations σ_{F_i} and that of the observations σ_{O_i}, an indicator of variability bias.

3) P-DATASET ASSESSMENT BASED ON THE CATEGORICAL INDEX

At the daily time step, the reliability of the P datasets is also analyzed using categorical indices based on a contingency table (Table 2). Two categorical indices are considered: the probability of detection [POD; Eq. (6)] and the false alarm rate [FAR; Eq. (7)]. POD indicates the proportion of precipitation events (i.e., rainfall occurrence > 1 mm) from station observations that are correctly detected by the estimates, ranging from 0 (no detection) to 1 (perfect detection). FAR corresponds to the proportion of events identified by estimates but not confirmed by station observations, with a value of 0 representing a perfect result:

$$\mathrm{POD} = \frac{a}{a+c}, \tag{6}$$

$$\mathrm{FAR} = \frac{b}{a+b}. \tag{7}$$

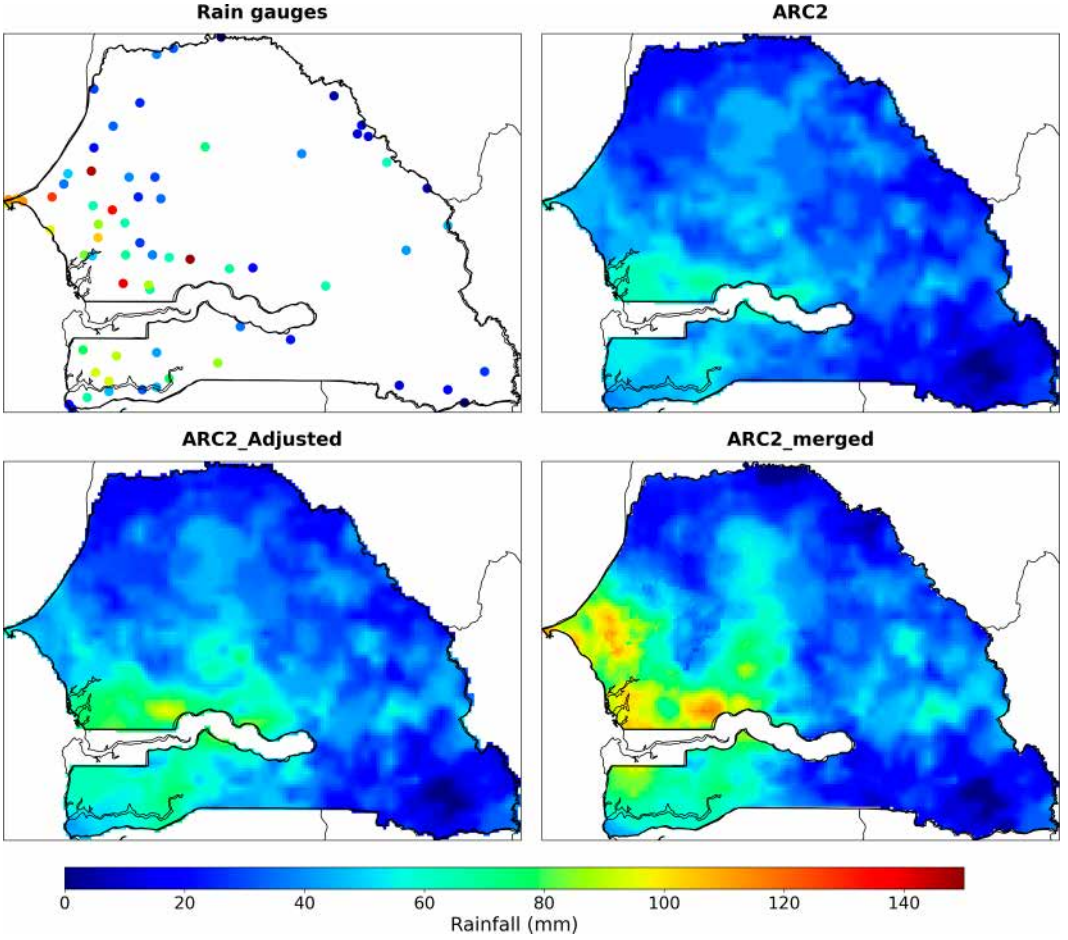

FIG. 3. Steps of the rainfall data merging process: (1) station data, (2) P-dataset rainfall estimates, (3) bias-corrected P-dataset estimation, and (4) M dataset: a combination of (1) and (3).

TABLE 3. Validation scores for daily rainfall estimates of P dataset and merged data.

Dataset	CORR	BIAS	MAE (mm day^{-1})	RMSE (mm day^{-1})	POD	FAR
ARC2	0.19	0.88	4.76	11.93	0.53	0.59
CHIRP	0.35	0.83	4.12	10.35	0.75	0.59
CHIRPS	0.32	0.98	4.40	11.05	0.68	0.53
CMORPH-BLD	0.20	0.91	4.77	12.03	0.60	0.59
CMORPH-CRT	0.28	1.26	5.12	13.10	0.65	0.52
ERA5	0.42	0.81	3.86	10.21	0.71	0.51
IMERG-FR	0.35	1.11	4.57	12.18	0.70	0.52
MERRA2	0.31	1.05	4.75	12.59	0.88	0.59
PERSIANN-CCS-CDR	0.31	1.38	5.25	12.21	0.82	0.57
PERSIANN-CDR	0.35	1.12	4.53	10.72	0.84	0.59
PERSIANN	0.27	1.63	6.19	14.68	0.73	0.57
PRISM	0.26	0.96	4.58	12.21	0.64	0.51
RFEv2	0.21	0.89	4.72	11.93	0.60	0.59
SM2RainASCAT	0.35	0.64	3.86	10.19	0.86	0.61
TAMSATv3.1	0.12	0.85	4.90	11.49	0.57	0.65
ARC2-merged	0.68	1.00	1.59	6.02	0.83	0.37
CHIRPS-merged	0.67	1.01	1.72	6.16	0.83	0.41
RFEv2-merged	0.69	0.99	1.58	5.96	0.86	0.37

4) P-DATASET ASSESSMENT BASED ON AGROCLIMATIC INDICATORS

First, the date of season start (DSS) and the end date of the rainy season (DES) are identified for each year, taking into account rainy (dry) days, days with more (less) than 1 mm of rainfall. The DSS is determined based on the Sivakumar criterion (1991) with different variants depending on the region considered. For the northern zone (>14°) and the southern zone (<14°), the DSS is determined when 15 and 20 mm of rainfall, respectively, are recorded in one or 3 consecutive days within a 30-day period starting from 1 May (122nd Julian day). The DSS is only valid if there is no rainfall gap of more than 20 days during this 30-day period. The DES is determined when 20 consecutive dry days are observed from 15 September (259th Julian day).

Second, several agroclimatic indices are calculated for each year based on DSS and DES:

- The length of the rainy season (SL) as the number of days between DSS and DES.
- The number of rainy days (NRainDay1) as the number of rainy days (rainfall > 1 mm) between DSS and DES.
- Number of days with precipitation above the 95th percentile (NRainDay95) between DSS and DES.
- Longest dry spells during the rainy season (DS) as the longest sequence of dry days and occurrence of dry spells longer than 5, 7, 10, and 15 days between DSS and DES.
- Seasonal cumulative rainfall (CUMS) as cumulative rainfall between DSS and DES.

All these indices are calculated from rain gauge observations and corresponding 15 P-dataset observations to assess the reliability of P datasets for agricultural applications.

e. Merging of stations with gridded data

The average KGE obtained from all the quantitative and agroclimatic indicators considered is used to rank the 15 P datasets from least to most reliable. The three best P datasets were individually merged with the rainfall data station to construct three M datasets following a two-step procedure.

First, the multiplicative bias time step variable (MBTSV) method is applied to minimize the P-dataset bias (Dinku et al. 2014, 2018; Siebert et al. 2019). MBTSV involves calculating a bias coefficient [Eq. (8)] of P-dataset estimates compared to gauge data for each station and each day over a climatological period (in this case, 1991–2020). The bias coefficients obtained at each gauge location are then interpolated at the spatial resolution of the P dataset using inverse distance weighting interpolation. A minimum of three and a maximum of nine stations within a 1° influence radius are considered. Thus, for all 365 or 366 days of the year, we constructed a gridpoint bias mask for each day by taking the entire climatological data series from 1991 to 2020 for the day in question. Finally, the P-dataset estimates are multiplied by the interpolated bias coefficients for the day in question. For example, to correct the ARC2 estimate on 15 August of any year, we multiplied it by the bias coefficient calculated between the climatological series of estimates (1991–2020) from ARC2 and the observations at each grid point for that date:

$$\mathrm{CB}_{tj} = \frac{\sum_{(t \text{ for each year } \pm 5 \text{ days})}^{n} G}{\sum_{(t \text{ for each year } \pm 5 \text{ days})}^{n} P}, \tag{8}$$

where G are the gauge data, P are the observations from the P dataset, and n is the number of years; we apply a smoothing window of ±5 days when calculating the daily estimate bias to better capture temporal dynamics.

Second, the P-dataset estimates are merged with the gauge data using the simple bias adjustment (SBA) method. SBA involves interpolating the differences between the gauge observations and

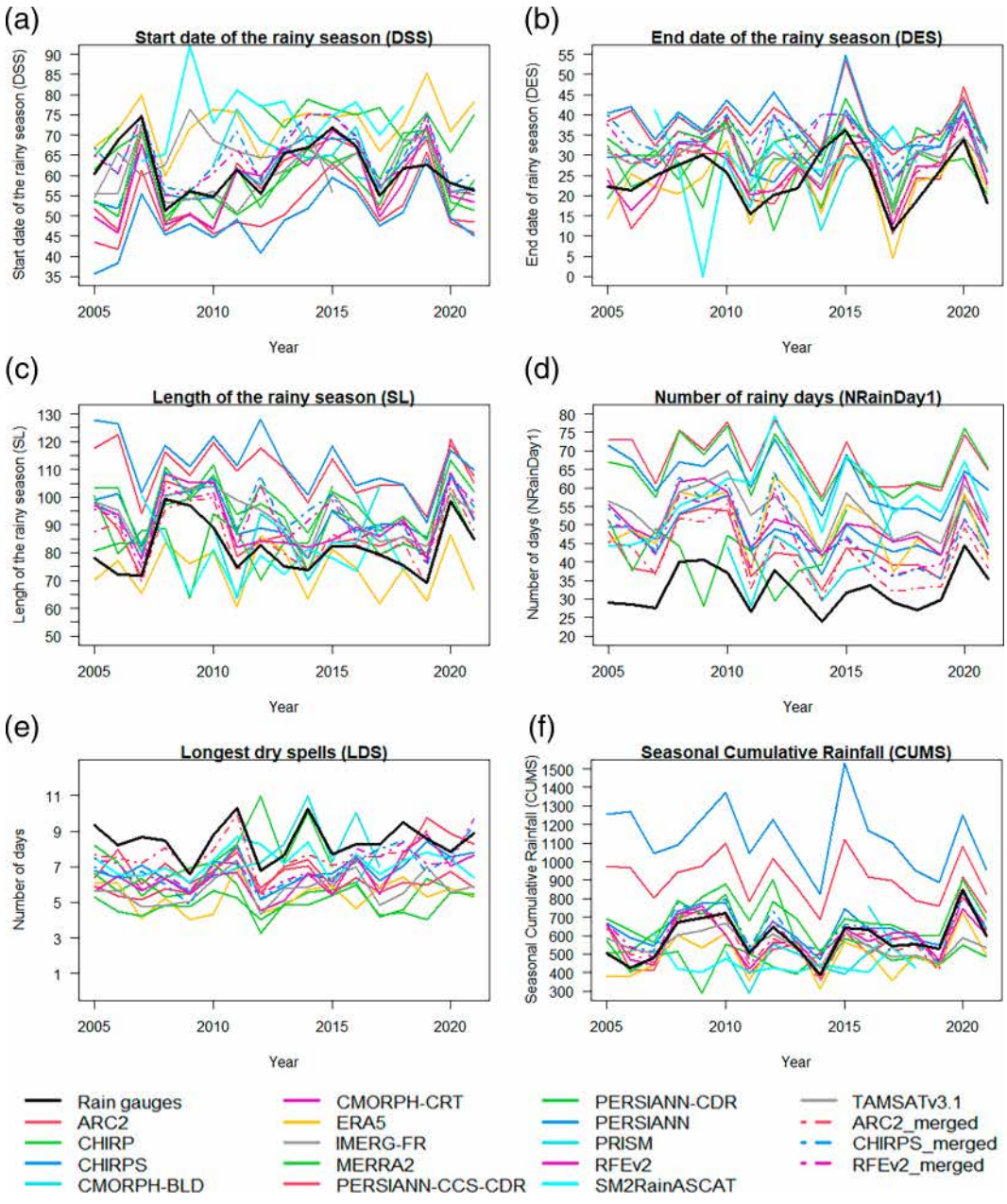

FIG. 4. Interannual variability of agroclimatic parameters in average across Senegal using rain gauge data, rainfall estimates from P dataset, and merged data.

the bias-corrected P-dataset estimates to further adjust the P-dataset observations (Dinku et al. 2014, 2018).

Each step in this process is performed using the Climate Data Tools (CDT), an open-source R package with functions for quality control, homogenization, and merging of meteorological data from stations with satellite or other proxy data such as reanalysis (Dinku et al. 2014, 2018). The final merged product [M dataset; Eq. (9)] is a function of

$$\begin{aligned} \text{M dataset} &= \text{F1(trend spatial)} + \text{F2(random)} \\ &= \text{F1}(P) + \text{F2}(G - P). \end{aligned} \tag{9}$$

Figure 3 summarizes the steps of the rainfall data fusion process with the example of 5 September 2020, for ARC2 rainfall estimates.

To evaluate the efficiency of the proposed method, the leave-one-out cross validation (LOOCV) is considered. In LOOCV, one of the available gauging stations is left out and one of them is set aside for validation, while the merging is performed with the remaining gauging stations. This process is repeated n times (so that each available station can be used for validation). At each iteration, the values estimated at the validation stations are compared with the

TABLE 4. KGE scores of P dataset and merged rainfall estimates for daily data and agroclimatic indicators.

Dataset	Daily rainfall	DSS	DES	SL	DS5	DS7	DS10	DS15	LDS	NRainDay1	NRainDay95	CUMS	KGE mean
ARC2	0.63	0.45	0.46	0.52	0.21	0.07	0.13	−0.24	0.28	0.60	0.08	0.70	0.32
CHIRPS	0.61	0.52	0.39	0.53	0.10	0.12	0.13	−0.50	0.24	0.55	0.17	0.79	0.30
RFEv2	0.67	0.47	0.48	0.55	0.10	0.04	0.26	−0.40	0.23	0.39	0.08	0.74	0.30
IMERG-FR	0.52	−0.24	0.25	0.47	0.21	0.24	−0.40	0.12	0.42	0.48	0.28	0.77	0.26
TAMSATv3.1	0.59	0.41	0.38	0.49	−0.06	−0.18	−0.33	NA	0.14	0.34	−0.03	0.69	0.22
ERA5	0.81	0.13	0.51	0.52	−0.02	−0.13	−0.26	−0.51	0.20	0.27	0.19	0.72	0.20
CMORPH-BLD	0.50	0.21	0.23	0.28	0.02	−0.04	0.20	−0.18	0.14	0.43	−0.20	0.66	0.19
PRISM	0.56	0.15	0.21	0.38	0.14	0.10	0.13	−0.05	−0.13	0.43	−0.10	0.37	0.18
CHIRP	0.63	0.46	0.43	0.55	−0.08	−0.29	−0.38	NA	0.12	0.10	−0.27	0.73	0.18
PERSIANN	0.49	0.70	0.41	0.38	0.18	0.03	−0.19	−0.24	0.09	0.09	−0.10	0.05	0.16
CMORPH-CRT	0.27	0.07	0.23	0.28	0.02	−0.04	0.20	−0.18	0.14	0.43	−0.20	0.66	0.16
PERSIANN-CCS-CDR	0.52	0.60	0.43	0.41	−0.05	−0.16	−0.40	−0.51	0.08	−0.09	−0.26	0.43	0.08
PERSIANN-CDR	0.52	0.08	0.44	0.52	−0.25	−0.45	−0.58	−0.51	−0.03	−0.03	−0.36	0.72	0.01
MERRA2	0.40	0.28	0.35	0.51	−0.40	−0.51	−0.45	NA	−0.17	−0.63	−0.49	0.71	−0.04
SM2RainASCAT	0.58	0.23	0.32	0.40	−0.35	−0.37	−0.37	−0.15	−0.69	−0.68	−0.72	0.44	−0.11
KGE mean	0.55	0.30	0.37	0.45	−0.02	−0.10	−0.15	−0.28	0.07	0.18	−0.13	0.61	0.16
Correlation value (CV) (%)	22.16	81.54	26.79	20.31	−1251.30	−211.61	−189.74	−74.40	366.85	221.66	−211.56	33.13	79.11
ARC2-merged	0.61	0.50	0.66	0.64	0.30	0.17	0.11	0.11	0.31	0.61	0.29	0.78	0.42
RFEv2-merged	0.69	0.57	0.63	0.63	0.33	0.27	0.05	0.11	0.28	0.50	0.24	0.77	0.42
CHIRPS-merged	0.61	0.52	0.40	0.57	0.30	0.17	0.02	−0.01	0.31	0.47	0.28	0.76	0.37
KGE mean	0.64	0.53	0.56	0.61	0.31	0.20	0.06	0.07	0.30	0.53	0.27	0.77	0.41
CV (%)	7.59	7.06	25.61	6.01	6.27	29.86	72.84	99.14	5.46	14.25	10.17	1.28	8.21

actual gauge observations, considering the previous indices described in sections 2d(3) and 2d(4).

Finally, the skill of the M-dataset indices obtained with the LOOCV is compared with that obtained with the corresponding P dataset [Eq. (10)]:

$$\text{Improvement} = \frac{\text{skill score M\_Dataset} - \text{skill score P\_Dataset}}{\text{skill score P\_Dataset}} \times 100. \quad (10)$$

The improvement is computed for the skill scores described in section 2d(2) (CORR, BIAS, MAE, RMSE, and KGE). Positive (negative) values for CORR and KGE indicate that the M dataset improves (decreases) the performance of the corresponding P dataset. Negative (positive) values for BIAS, MAE, and RMSE indicate that the M dataset improves (decreases) the performance of the corresponding P dataset. It is worth mentioning here that for the bias, the absolute values of the difference are compared to 1.

3. Results

a. *P-dataset reliability at the regional scale*

Table 3 shows the average statistics over Senegal of daily rainfall estimates from the P dataset to represent observations. Overall, the daily rainfall estimates show a low correlation with the observations. The ERA5 data show the best performance with a CORR of 0.42; however, they underestimate the daily rainfall in Senegal (BIAS = 0.81 with a RMSE of 10.21 mm). The PERSIANN data show the largest bias (BIAS = 1.63) in the simulation of daily rainfall in Senegal. All evaluated datasets have a FAR of at least 0.5 in Senegal.

Table S1 summarizes the intraseasonal performance (during the agricultural activity period) of the daily rainfall estimates from the P dataset. The estimates show good consistency, especially during the core of the rainy season, with an average KGE between July and October of between 0.56 and 0.58. In contrast, at the beginning of the rainy season, in May, the P dataset shows lower performance in estimating daily rainfall, with an average KGE of only 0.04.

Except for CUMS, the ability of rainfall estimates to reproduce agroclimatic variables is generally low. The characteristics of onset (DSS), cessation (DES), longest dry spells (LDS), and NRainDay1 are not accurately reproduced by most rainfall estimates, even at the Senegal scale (Fig. 4). P-dataset estimates of the end of the rainy season are generally too late, resulting in an overestimation of the length (SL) of the rainy season. In addition, the P dataset significantly overestimates the average number of rainy events per year, which can increase by 30–35 days on average compared to observations for PERSIANN and PERSIANN-CDR.

The P-dataset performance evaluation scores at the regional level for agroclimatic parameters are also presented in supplemental material Table 2.

Table 4 shows the average KGE values of the P datasets for each climatic and agroclimatic area considered, to classify all the P datasets in relation to the observations (supplemental material Table 3). The three best products ARC2, CHIRPS, and RFEv2 are those that gave the three highest average KGE increases compared to the assessment indicators. It is also very variable from one dataset to another. Although the

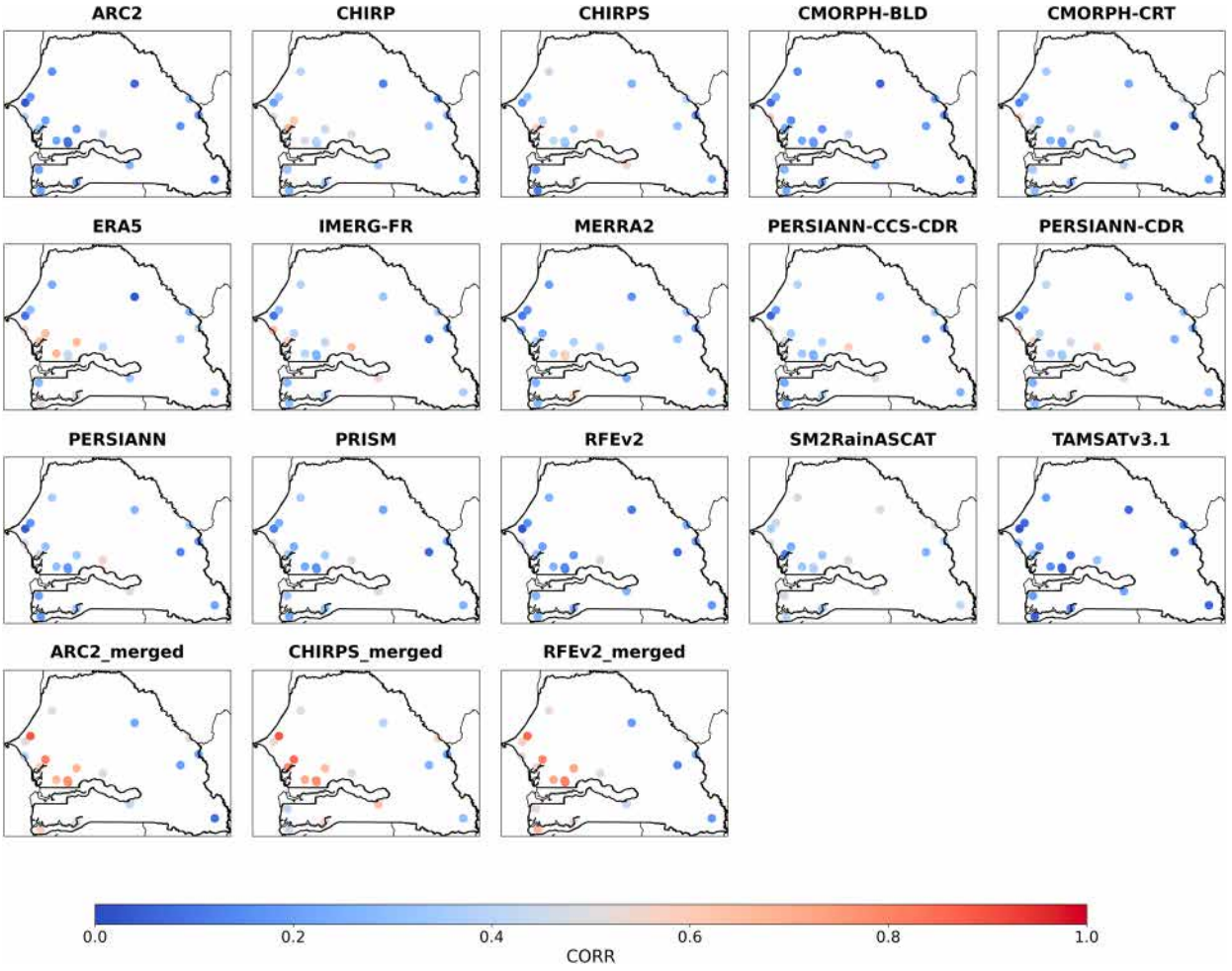

FIG. 5. Correlation of observed daily rainfall at the station level compared to P-dataset estimates and merged data.

ERA5 data show the highest values for the daily rainfall estimation (KGE = 0.81), this product is less effective in simulating the agroclimatic characteristics of the season compared to the ARC2, CHIRPS, and RFEv2 data. This table shows that for the determination of the DSS, the estimates of ARC2, RFEv2, TRMM, CHIRP, CHIRPS, TAMSATv3.1, PERSIANN, and PERSIANN-CCS-CDR gave the best performances with a KGE higher than the average of the P datasets (0.30). These same estimates remain the most efficient for simulating the end of season date with a KGE higher than the average of the P datasets (0.37). Furthermore, the P-dataset estimates showed poor performance in representing dry breaks and the occurrence of dry breaks of 5, 7, 10, and 15 days with large variability. The IMERG-FR estimates showed the best performance for determining large amounts of rainfall in Senegal with a KGE of 0.28; however, the P-dataset estimates all showed a good performance for representing the seasonal cumulative rainfall, except SM2RainASCAT, PRISM, PERSIANN, and PERSIANN-CDR which showed a KGE lower than the P-dataset average (0.61). From our ranking result against agroclimatic applications in Senegal, ARC2 (KGE = 0.32), RFEv2 (KGE = 0.30), and CHIRPS (KGE = 0.30) are the best rainfall estimation products.

b. *P-dataset reliability in space*

At the station level, daily P-dataset estimates show a weak correlation with observations, ranging from 0.10 to 0.60 depending on the station (Fig. 5). The CHIRP, CHIRPS, ERA5, IMERG-FR, and PERSIANN data present a better spatial distribution of the correlation, with values of at least 0.30 over the western half of the country.

Figure 6 shows the distribution of the bias of the P-dataset estimates compared to the ground observations. We can see that the CMORPH-CRT, IMERG-FR, and the three versions of PERSIANN generally overestimate daily precipitation in Senegal with a bias > 2.5 at a few stations in the central-west of the country.

P-dataset estimates show a higher MAE in the south and far east of the country with values between 6 and 10 mm day^{-1} (Fig. 7). From the rest in the north of the country, they present less bias with an MAE of less than 5 mm for daily rainfall estimates.

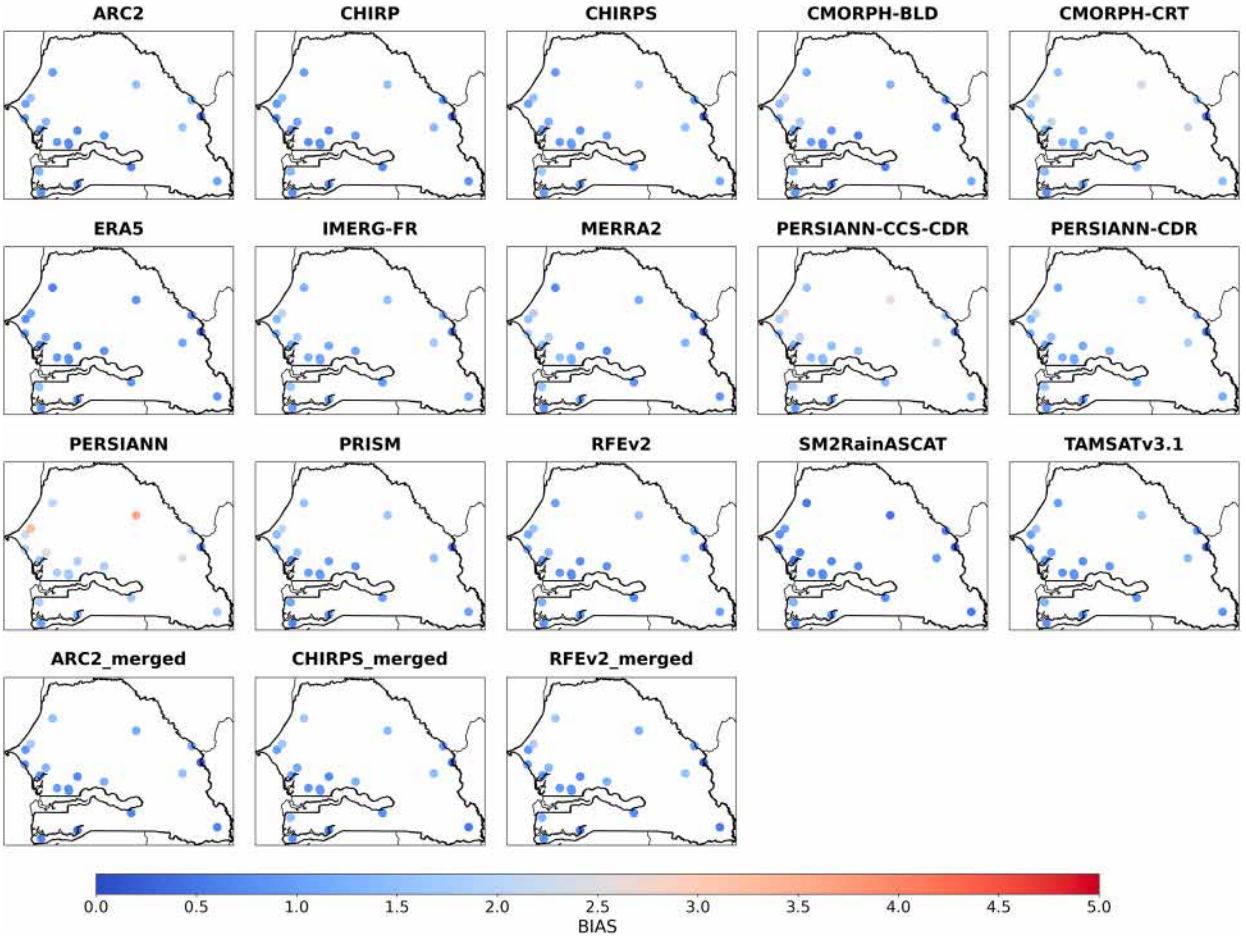

FIG. 6. Bias of observed daily rainfall at the station level compared to P-dataset estimates and merged data.

We note that the estimates from MERRA2 and PERSIANN present high values of RMSE (Fig. 8) on the western facade of the country which can deviate by 18 mm day^{-1} on average at certain stations compared to ground observations. Moreover, in the northern half of the country, the P-dataset estimates except these two previously cited present the same distribution of RMSE with values varying between 2 and 8 mm day^{-1} on average. As for the MAE, it is in the south of the country that the differences in estimates are greater with an RMSE of more than 12 mm day^{-1}.

Furthermore, for the detection of rainy events in Senegal, the P-dataset estimates are very variable depending on the station with a POD varying from 0.40 to 1 (Fig. 9). The estimates of ERA5, MERRA2, IMERG-FR, the three versions of PERSIANN, and SM2RainASCAT show the best spatial configuration of the POD with values of 0.60–0.80 depending on the station. Moreover, the other P datasets show an average rain event detection power of 0.50.

However, even if the estimates from MERRA2, IMERG-FR, the three versions of PERSIANN, and SM2RainASCAT show good POD performance, this indicator can hide a lot of information. As shown in Fig. 10, the latter with TAMSATv3.1 generate more false alarms in Senegal, particularly on the western front influenced by the sea current, with an FAR of 0.60 on some stations.

Evaluation of P dataset for agroclimatic parameters at the rain gauge level is presented as the supplemental material (Fig. S2 for DDS, Fig. S3 for DES, Fig. S4 for SL, Fig. S5 for NRainDay1, Fig. S6 for NRainDay95, Fig. S7 for DS5, Fig. S8 for DS7, Fig. S9 for DS10, Fig. S10 for DS15, Fig. S11 for LDS, Fig. S12 for CUMS).

c. *Evaluation of the* M *dataset*

1) EVALUATION OF DAILY RAINFALL ESTIMATES

The three new M datasets are very close to each other and show a reduced dispersion when compared to the P dataset. A significant improvement in correlation to represent observed daily rainfall is notable when compared to the P-dataset estimates (Fig. 5). Correlations range from 0.3 to 0.9 depending on the stations. From a spatial perspective, although CHIRPS-merged data show a lower correlation compared to the other two merged data, the merging has improved the correlation of daily rainfall, especially over the western half of the country for all M datasets.

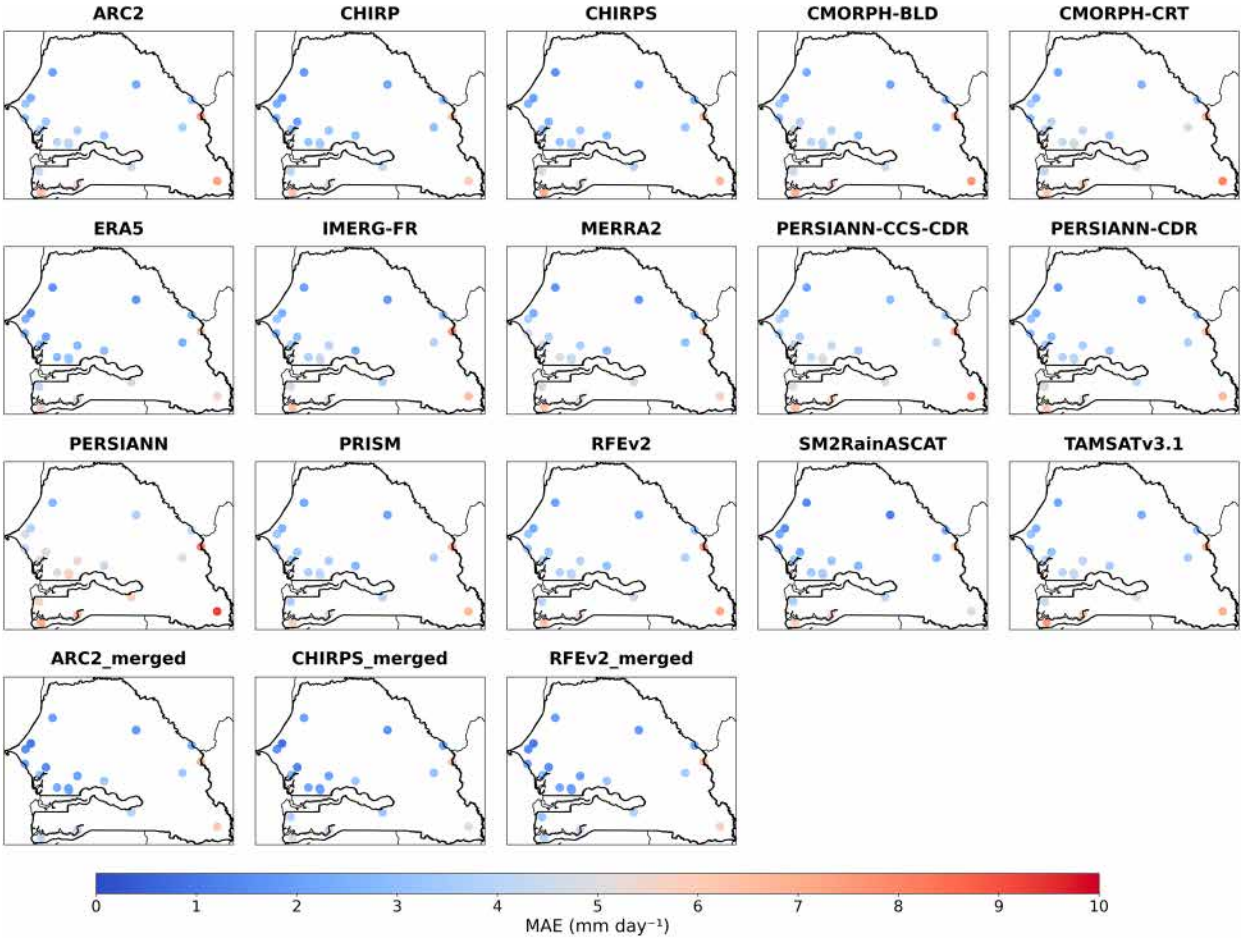

FIG. 7. MAE of observed daily rainfall at the station level compared to P-dataset estimates and merged data.

Furthermore, for bias (Fig. 6), ARC2 merged and RFEv2 merged show the best performances (bias close to 1) on practically all stations, and CHIRPS merged overestimates daily rainfall in the north of Senegal. Regarding MAE (Fig. 7) and RMSE (Fig. 8), the M datasets are less biased than the P-dataset estimates. For example, in the southern half of the country, the RMSE is less than 10 mm day^{-1}, compared with 20 mm day^{-1} with P-dataset estimates for certain stations. The merging of P-dataset estimates with station data has improved the POD with a score ranging between 0.74 and 0.9 depending on the station (Fig. 9) and reduced the FAR. However, the merged data show an average false alarm rate of 50% in the northeastern part of the country, where the meteorological observation network is still sparse (Fig. 10).

The scores of daily M datasets with ground-based data compared to observations are presented at the end of Table 3. RFEv2 data, generated by fusion between RFEv2 and in situ observations, showed the highest correlation (CORR = 0.69) among the three merged products with a very low bias (BIAS = 0.99). All three merged products have a POD greater than 0.8. The RFEv2-merged product proved to be the most proficient in representing observed daily accumulations with a KGE of 0.69 (Table 4).

The merging process improved the performance of the daily precipitation estimates from the P dataset throughout the rainy season, from May to October, with monthly KGE values above 0.6 (Fig. 11). The best performance is observed during the peak of the rainy season, from June to October, where KGE values range between 0.73 and 0.78.

The ability of the different datasets to reproduce a single extreme event is illustrated through a case study of a dramatic event occurred on 2020 (Fig. 12). On 5 September 2020, cumulative rainfall above the 95th percentile was observed at many stations along the Senegalese coast, causing flooding almost everywhere, with fatalities and damage to infrastructure. The ARC2 and CHIRPS rainfall estimates for that day in this part of the country do not exceed 40 mm, while many stations received more than 90 mm or even 150 mm of rain. The RFEv2 estimates best reproduced the observed signal but with a bias in the center of the country. Merging, as seen here, made it possible to correct the P-dataset estimates to reproduce the extreme events recorded and to reduce the intensity of the RFEv2 estimate bias in the center.

Table 5 provides an insight into the added value of merging the rainfall dataset estimates with ground observations. Overall, the merged data are better correlated with observations,

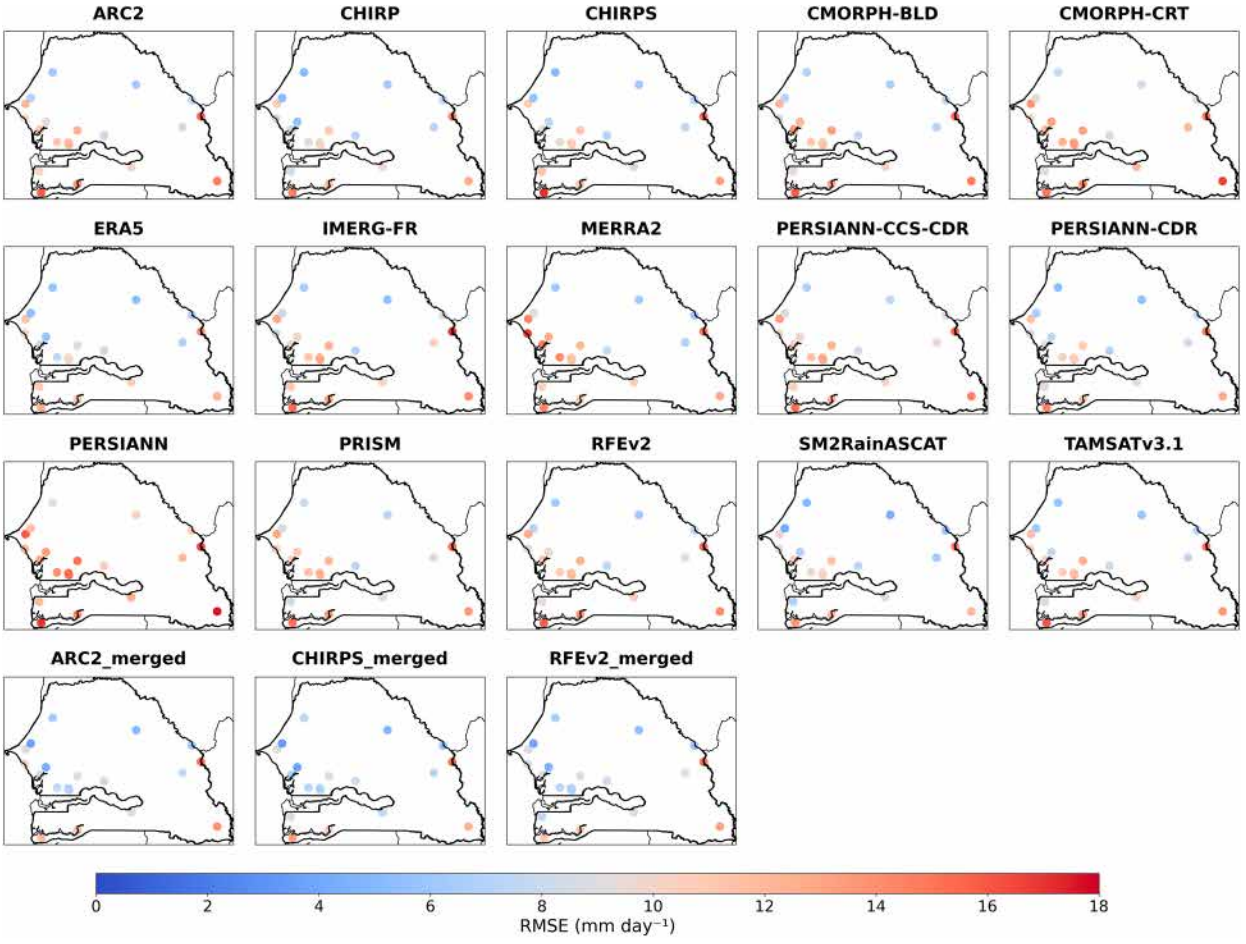

FIG. 8. RMSE of observed daily rainfall at the station level compared to P-dataset estimates and merged data.

especially for daily estimates where we observed a significant increase in correlation of over 100%. The merging technique reduced the bias by at least 80%. The MAE and RMSE of daily precipitation estimates were reduced by merging, which also increased the detection power of ARC2 by 56.82%, CHIRPS by 23.23%, and RFEv2 by 44.22%, while the FAR of these different P datasets decreased by 37.27%, 22.48%, and 38.05%, respectively.

2) EVALUATION OF AGROCLIMATIC PARAMETERS

The merging has significantly improved the estimates of the P datasets to reproduce the intrasequential characteristics of the agricultural season, but the data still lengthen the season and overestimate the long dry spells and the number of rain events (Fig. 4). The evaluation of M dataset for agroclimatic parameters at the rain gauge level is provided in the supplemental material. Specifically, Figs. S2–S12 illustrate the results for each indicator: DDS (Fig. S2), DES (Fig. S3), SL (Fig. S4), NRainDay1 (Fig. S5), NRainDay95 (Fig. S6), DS5 (Fig. S7), DS7 (Fig. S8), DS10 (Fig. S9), DS15 (Fig. S10), LDS (Fig. S11), and CUMS (Fig. S12). The average statistics of the merged data to reproduce the agroclimatic variables are presented in supplemental material Table 2. These values, when compared to the values of the P dataset, showed that the merging improved the correlation coefficients and the KGE of the three datasets with respect to the beginning of the season and reduced the bias by 46.15% for ARC2, 10.35% for CHIRPS, and 84.91% for RFEv2, respectively (Table 5). The merged RFEv2 data were the most competent in determining this parameter with a KGE of 0.57 (Table 4). For the end of the season, the CHIRPS-merged dataset performed the worst (KGE = 0.4) compared to the other merged products, although the values generally increased in the right direction. Despite the improvement in daily rainfall estimates from the P dataset, the merging only slightly or moderately improves the detection of drought events at different thresholds, the number of rainy days, and rainfall extremes.

Furthermore, we have seen that these parameters are less well captured by the P-dataset estimates, as shown in the classification table in supplemental material Table 3.

4. Discussion

Reliable and accurate rainfall estimates are not only critical for the study of climate variability but also important for

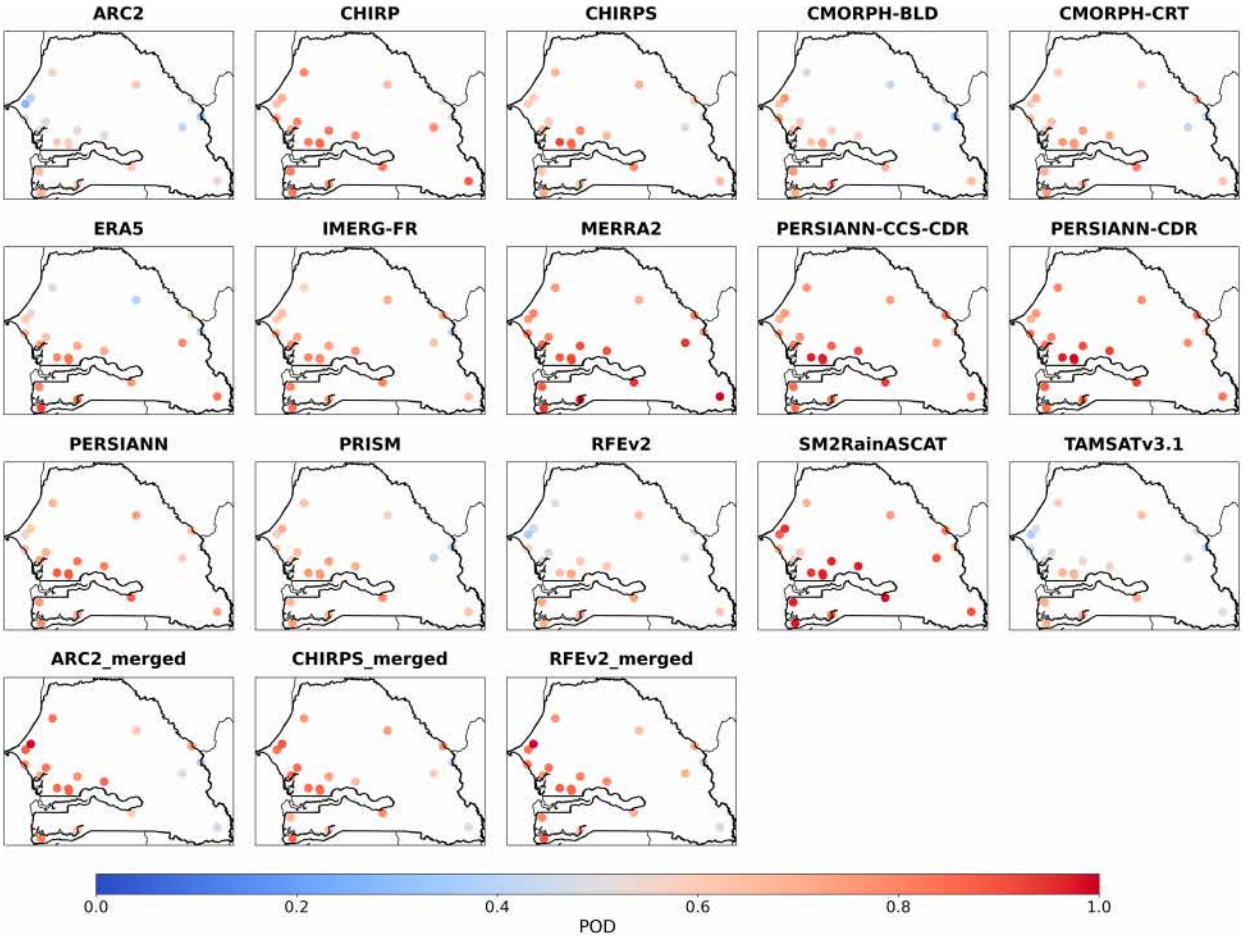

FIG. 9. Observed rainfall event detection power at the station level compared to P-dataset estimates and merged data.

water resource management, agriculture, and weather, climate, and hydrological forecasting (Sarojini et al. 2016). This study provides a comprehensive overview of 15 existing rainfall products and quantifies biases in daily rainfall estimates and agroclimatic parameters. The 15 rainfall products evaluated had spatial resolutions ranging from 0.0375° to 0.625° and included gauge-based products (ARC2, CHIRPS, IMERG-FR, PRISM, RFEv2, TAMSATv31), satellite-based products (CHIRP, PERSIANN-CCS-CDR, PERSIANN-CDR, PERSIANN, both CMORPH versions mentioned in Table 1), rainfall estimation products from soil moisture (SM2RainASCAT), and reanalyses (ERA5, MERRA2).

It is noteworthy that the performance of the evaluated rainfall datasets differed between daily rainfall estimates and agroclimatic parameters. Daily estimates were found to have significant uncertainties in the magnitude and variability of rainfall, even between products within the same category. The results highlight the low correlation between satellite rainfall estimates and daily observations, which is consistent with several previous studies (Roca et al. 2010; Satgé et al. 2020; Houngnibo et al. 2023; Nicholson et al. 2003), which also reported limitations in the ability of satellite data to accurately reproduce daily rainfall events.

In addition, some authors have reported the underperformance of IR- or passive microwave (PMW)-based products in detecting cloud heat over coastal and orographic regions. According to Dinku et al. (2007, 2008, 2011), this is mainly due to the threshold they use to discriminate rain clouds that are too cold. PMW-based algorithms are primarily scattered by high altitude ice, but orographic rainfall is a warm cloud process that does not necessarily produce much ice, leading to an underestimation of daily estimates (Petković et al. 2019).

Furthermore, our results highlighted a marked seasonal variability in the performance of daily rainfall estimates from P-dataset products. Their performance is significantly better during the core of the rainy season (July–October), with an average KGE index between 0.56 and 0.58. In contrast, at the beginning of the rainy season in May, performance is significantly lower, with an average KGE of only 0.04. This drop in performance at the start of the season can be explained by several factors. Fusion algorithms, such as those used in CHIRPS or RFEv2 products, partly rely on regression or spatial interpolation methods, which tend to perform better when rainfall gradients are well defined, which is generally the case at the peak of the rainy season but not necessarily at the beginning of the season (Toté et al. 2015). Furthermore,

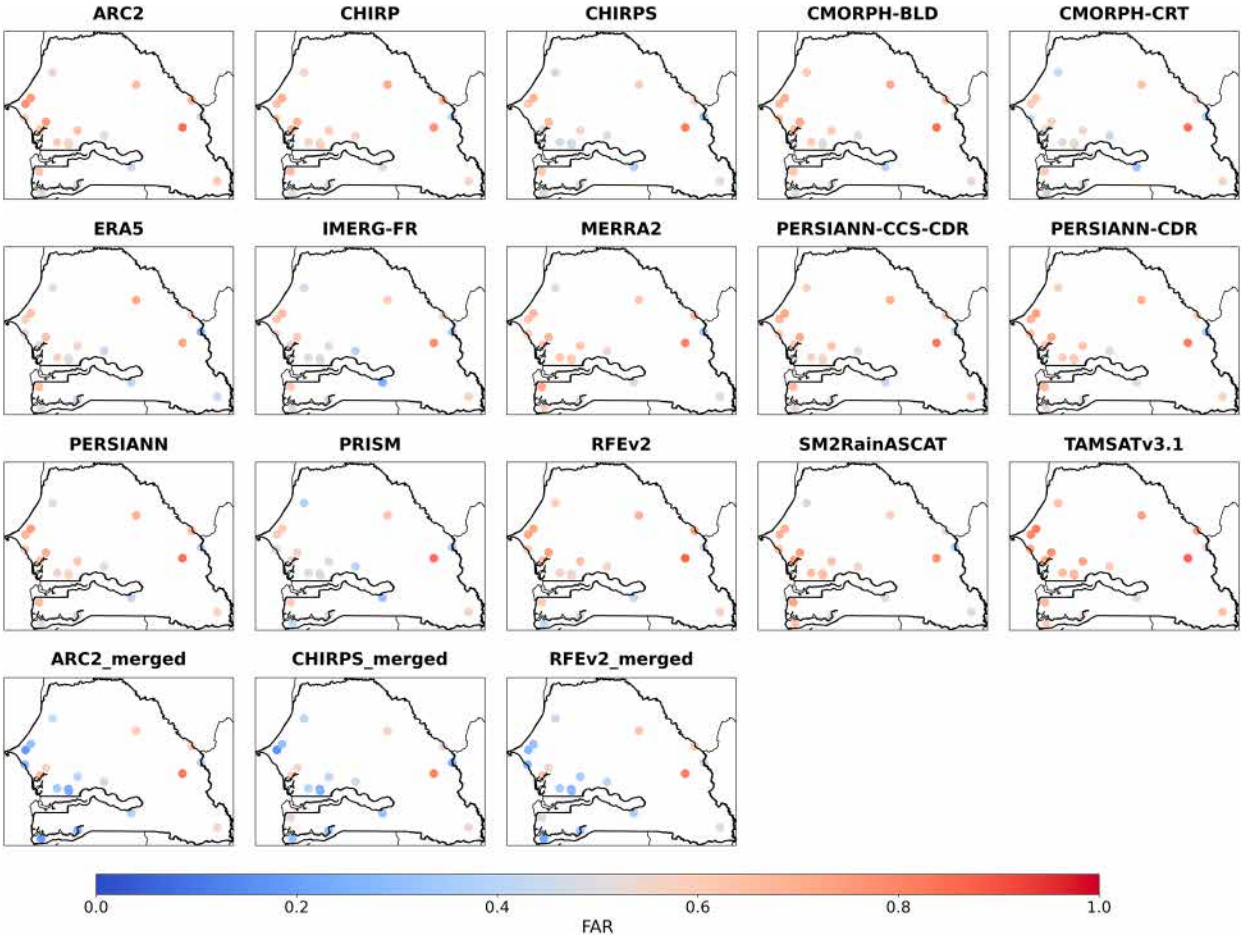

FIG. 10. Estimation performance of dataset P and merged data for false alarm.

some products use climatological averages or fixed approaches to detect the onset of rainfall, which limits their ability to adapt to interannual variability in the actual onset of the season. These results highlight a major challenge for agricultural applications: The accurate detection of early season rainfall is crucial for the planting calendar. Underestimating this rainfall can lead to errors in the estimation of the start of the growing season and thus distort recommendations to farmers or yield projections in agroclimatic models.

In general, estimation datasets using gauge-based information (ARC2, CHIRPS, ERA5, RFEv2) are the best at representing daily observations from stations with KGE > 0.6. This confirms previous results obtained in the West African region (Gosset et al. 2013; Casse et al. 2015; Poméon et al. 2017). The good performance of these datasets can be attributed to calibration strategy, which gives them better accuracy in areas with a relatively large observing network. We have also shown that their performance can be significantly improved by merging with a much denser network of ground stations than the GTS network. To limit the autocorrelation biases associated with using the same data for calibration and evaluation, we have applied a rigorous leave-one-out approach. Nevertheless, we recognize that some form of spatial dependence remains. Indeed, neighboring stations, which are often highly spatially correlated, can still indirectly influence the performance of the merged products in their immediate vicinity. This covariation is particularly pronounced in areas with high station density, which promotes a certain homogeneity of rainfall fields and mechanically improves the adjustment of satellite products (Maidment et al. 2017; Dinku et al. 2018). Therefore, some of the "high skill" observed in our results may partly reflect this geographical proximity effect. These findings highlight the need to place the assessment of fused products in their spatial context, taking into account the distribution of available stations. It is important to consider this limitation in operational applications, especially in areas where ground-based observations are rare or absent, as performance is likely to be lower. Furthermore, their performance can also be attributed to the prevalence of convective rainfall in West Africa. Indeed, IR- and PMW-based estimation methods work well with convective rainfall (Ebert 2007), hence the overall competence in these regions.

The main difficulty in observing rainfall is the intermittency of these phenomena, combined with limited sampling in time and space (Hegerl et al. 2015; Trenberth et al. 2017). Although rain gauges are essential for direct measurement of rainfall,

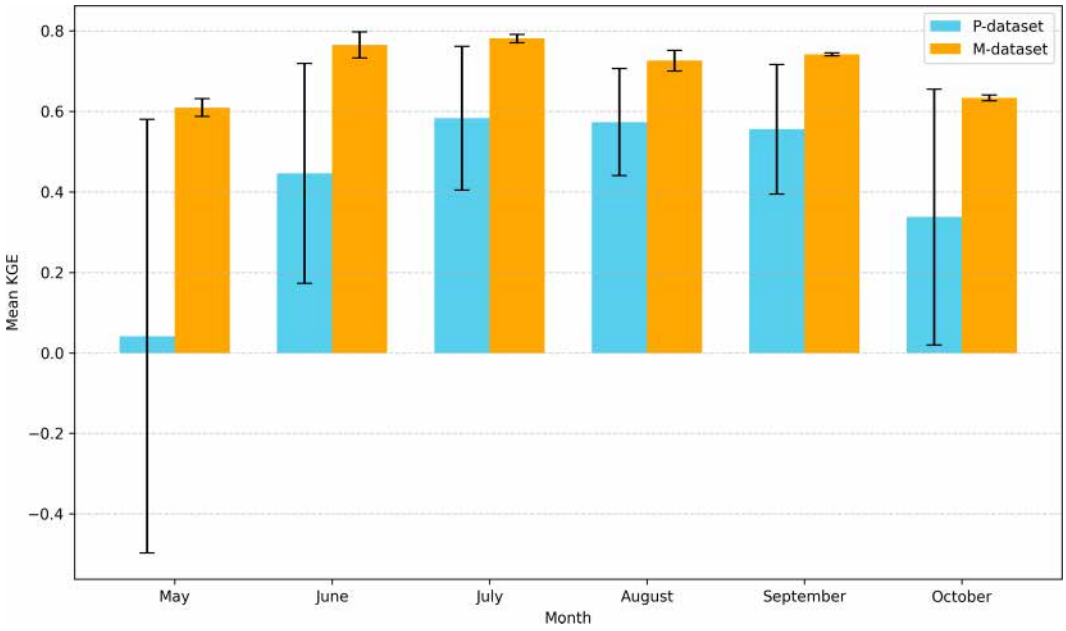

FIG. 11. Comparison of daily rainfall estimation performance between P dataset and M dataset during the rainy season (May–October). Bars represent the average of monthly KGE of the different P datasets and M datasets.

their global density is limited and their distribution is uneven, particularly with a limited number of gauges in the oceans (Kidd et al. 2017). Worryingly, the number of these available rain gauges has been declining, which may compromise our ability to monitor rainfall variability in the future. Temporally, gauge-based rainfall products are limited to monthly samples. Delays in data availability and limited temporal resolution make it difficult to use these products for real-time research or to track short-term variations in rainfall. In addition, gauge observations are interpolated to produce gridded datasets covering the globe, which tends to attenuate extremes and affect long-term trends, especially in regions where gauges are scarce. Gauge-based rainfall estimates show inconsistencies between different products and vary considerably globally depending on the number of stations used, their homogeneity, analytical methods, quality control procedures, and management of data coverage over time (Hegerl et al. 2015; Sun et al. 2014).

An improvement in the performance of satellite rainfall estimates could be achieved by aggregating data over longer time scales, such as decade, month, and season, as shown by Fenta et al. (2018) and Hailesilassie et al. (2021). This could be due to the elimination of errors as the integration period increases and the imprecision caused by the disaggregation of daily values, which is smaller at the monthly scale, or errors at smaller time scales that cancel out when aggregated over a longer time scale (Belay et al. 2019; Gebremicael et al. 2019).

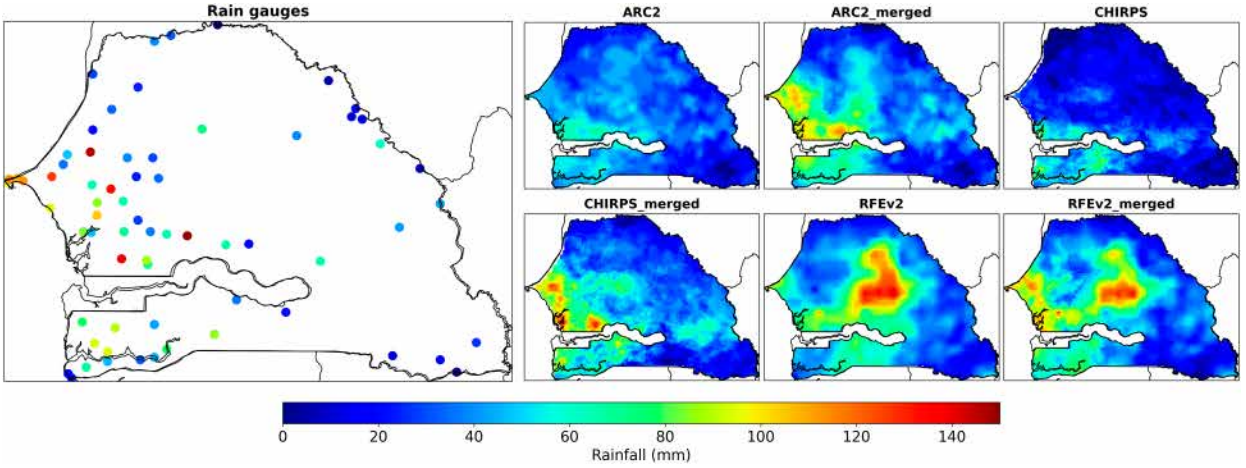

FIG. 12. The extreme event of the 5 Sep 2020 in Senegal as measured by rain gauges, P datasets, and M datasets.

TABLE 5. Evaluation of the improvement of the M dataset compared to the original M datasets. Improvement (%) is computed as defined in method section [Eq. (10)]. For example, the value of 100% for correlation would mean that the M dataset double the CV obtained using the corresponding P dataset. The asterisk (*) indicates significant increase or decrease in score at 95% confidence level.

Variable	Dataset	CORR improvement (%)	BIAS improvement (%)	MAE improvement (%)	RMSE improvement (%)	KGE improvement (%)	POD improvement (%)	FAR improvement (%)
Daily rainfall	ARC2-merged	260.96*	−99.19*	−66.48*	−49.53	−3.71	56.82*	−37.27
	CHIRPS-merged	112.70*	−80.00*	−60.88*	−44.25	0.20	23.22	−22.48
	RFEv2-merged	226.42*	−90.57*	−66.63*	−50.05*	2.65	44.22*	−38.05
DSS	ARC2-merged	2.11	−46.15	2.65	2.98	11.54		
	CHIRPS-merged	23.49	10.34	−8.57	−5.26	0.85		
	RFEv2-merged	11.01	−89.61*	−4.20	−4.07	21.49		
DES	ARC2-merged	45.10*	468.57*	−21.52	−14.64	45.19		
	CHIRPS-merged	17.50	115.19	29.48	33.05	2.99		
	RFEv2-merged	30.04*	467.95*	30.96	15.76	26.30		
SL	ARC2-merged	24.81*	−60.19*	−24.68	−21.20	24.66*		
	CHIRPS-merged	8.10	16.31	0.51	−0.46	7.23		
	RFEv2-merged	14.81	15.25	−1.31	−5.69	14.16		
NRainday1	ARC2-merged	1.76	−4.67	−8.52	−5.35	2.77		
	CHIRPS-merged	−7.54	8.77	10.32	12.93	−13.89		
	RFEv2-merged	1.79	−19.46	−18.47	−15.20	29.41*		
NRainday95	ARC2-merged	161.76*	−2.74	−17.59	−17.59	257.24*		
	CHIRPS-merged	46.18	6.01	−8.73	−10.33	65.39		
	RFEv2-merged	40.36	−19.96	−12.92	−12.79	219.13*		
DS5	ARC2-merged	44.21	13.22	−7.53	−3.02	44.65		
	CHIRPS-merged	79.79	−19.81	−0.33	−3.04	198.15*		
	RFEv2-merged	58.65	−47.82	−16.65	−15.07	224.63*		
DS7	ARC2-merged	92.86*	9.20	−7.55	−4.53	166.15*		
	CHIRPS-merged	11.11	−1.98	−4.46	1.38	35.25		
	RFEv2-merged	38.78	46.63	−5.75	−1.78	602.18*		
DS10	ARC2-merged	35.89	36.81	−7.55	−5.09	−15.10		
	CHIRPS-merged	−30.39	−22.22	6.46	7.55	−145.90		
	RFEv2-merged	−14.50	−18.15	10.17	7.01	−264.41*		
DS15	ARC2-merged	4.55	0.00	4.17	2.75	−137.68		
	CHIRPS-merged	1063.64*	−50.06	19.44	9.47	−98.02		
	RFEv2-merged	1187.50*	−50.00	17.07	8.42	−126.14		
LDS	ARC2-merged	16.44	24.63	−5.85	−2.29	10.90		
	CHIRPS-merged	19.18	−2.26	−4.58	−0.35	31.05		
	RFEv2-merged	14.96	−29.78	−10.63	−9.65	21.82		
CUMS	ARC2-merged	12.10	−83.33	−22.66	−20.93	11.39		
	CHIRPS-merged	1.38	1.75	−8.32	−5.52	−2.96		
	RFEv2-merged	5.37	25.00	−8.02	−8.93	5.01		

Dembélé and Zwart (2016) validated ARC2, RFEv2, TARCAT, and other estimation datasets for Burkina Faso and found that all estimates showed very good correlation at monthly time scales. In terms of overall performance, they found that RFEv2 and ARC2 performed better than TARCAT. Other studies on the Sahel, such as Jobard et al. (2011) and Pierre et al. (2011), have also shown the good performance of RFEv2 in the subregion. At the monthly scale, Abdourahamane (2021) found that CHIRPS and ARC2 have good potential for hydroclimatic applications in Niger. However, his study used four estimation datasets (CHIRPSv2.0, ARC2, CMORPH, and TAMSAT). Our results demonstrate the importance of considering daily time steps when evaluating the reliability of estimation datasets, as it is influenced by the gauge-based adjustment and the time step provided.

To our knowledge, no study to date has evaluated the performance of satellite estimates in representing the intrinsic characteristics of the agricultural season (DSS, DES, SL, DS, NRainDay1, CUMS). Dunning et al. (2016) developed a novel approach to calculate the start and end dates of the season across Africa and compared it with other traditional criteria using only satellite rainfall estimates. However, this study did not compare results with ground stations. Most publications focusing on the evaluation of satellite estimates have examined the performance of daily or aggregated estimates at decadal and monthly time steps. Our study, although limited to Senegal, is a pioneering effort focusing on the performance of rainfall estimates in estimating key parameters of the agricultural season. Our results have revealed significant biases in the rainfall products in their estimation of agroclimatic indices in Senegal. In general, they detect too long a rainy season with too early onset and too late an end with too short dry spells. These biases could result from the high sensitivity of satellite sensors to low-intensity events with sporadic rainfall and the consequent overestimation of the number of rainy days in Senegal compared to ground observations, especially when using a threshold of 1 mm to detect a rainy day. This overestimation of rainy days by satellite sensors has also been documented in previous studies, such as that by Maidment et al. (2017), and highlights the need for caution when interpreting satellite data, especially in specific regional contexts, to avoid significant biases in climatological and hydrological analyses.

5. Conclusions

Accurate rainfall estimates at fine spatiotemporal scales are crucial for hydrological resource assessment. This study evaluates and compares fifteen rainfall estimates using rain gauge data at different spatial and temporal scales and for 11 key parameters of the season from 2005 to 2021. The results indicate that datasets using information based on ground observations achieve average scores in representing daily rainfall from stations. In addition, the estimation products tend to overestimate the length of the season (too early start and too late end) and to underestimate the occurrence of dry spells, implying a notable presence of false alarms and a significant bias in the overestimation of rainfall events. The best overall performances are achieved by ARC2 (KGE = 0.32), RFEv2 (KGE = 0.30), and CHIRPSv2 (KGE = 0.30). Their performance could be improved by merging the ANACIM observational rain gauge dataset with these new rainfall estimates, M datasets at a spatial resolution of 0.0375° × 0.0375° from 1983 to 2021 for ARC2 merged (KGE = 0.42), from 1985 to 2021 for CHIRPS merged (KGE = 0.37), and from 2000 to 2021 for RFEv2 merged (KGE = 0.42). Merging these estimates with the ANACIM observational rain gauge dataset improved the correlation with daily estimates by over 113%–261%, depending on the product, and reduced the bias by 99.187% for ARC2, 80% for CHIRPS, and 90.566% for RFEv2.

Finally, although this study has identified the most efficient satellite products for agroclimatic applications in Senegal, it has some limitations that should be highlighted. In particular, the density of stations in certain areas may indirectly influence the local performance of the merged products due to the high spatial covariability of the observations. To increase the robustness of future evaluations, it would be relevant to consider complementary approaches, such as analyzing the performance in areas independent of the calibration network or integrating recently installed stations that were not used for the initial calibration. These strategies would make it possible to avoid an indirect influence of neighboring stations on the evaluation of the merged products and to better assess their real ability to represent precipitation in more limited observation contexts.

Such a rainfall product, publicly available on the ANACIM website, is an invaluable resource for the scientific community, academics, and decision-makers. These netCDF formatted data provide a comprehensive and consistent view of rainfall over a significant period, enabling in-depth analysis in hydrology, monitoring of climate change and climatological studies. The availability of these datasets promotes collaboration and information sharing within the community, paving the way for multiple applications in research and decision-making related to meteorological and climatic phenomena.

Acknowledgments. The activities described in this article were funded as part of my thesis by the Water Cycle and Climate Change (CECC) project. We also thank the IRD and the administration of the ESPACE-DEV Laboratory for all the logistics necessary to carry out this work. This is an opportunity to thank the Senegalese National Meteorological Service (ANACIM) for allowing us to carry out this research and especially for making the field data available.

Data availability statement. 1) Gridded rainfall datasets: This study evaluated 17 publicly available gridded rainfall datasets (P datasets) that are freely accessible from various sources. Detailed information about these datasets, including their sources and references, is provided in Table 1 of the manuscript. 2) Merged high-resolution dataset (M dataset): The high-resolution (4 km) merged rainfall dataset developed in this study, which incorporates gridded rainfall products and in situ observations from 21 meteorological stations in Senegal, is available upon reasonable request. Interested researchers may

contact the corresponding author at mbengueass91@gmail.com for access under appropriate research collaboration agreements. 3) Software: (i) Bias correction and agroclimatic index computation scripts were developed in R and are publicly accessible. The merging of gridded datasets with ground station data was performed using the CDT software package, available at https://rdrr.io/github/rijaf-iri/CDT/. (ii) For further details regarding the methodology, please refer to the materials and methods section of the manuscript. For any additional questions or requests concerning the datasets or software, please contact the corresponding author.

REFERENCES

Abdourahamane, Z. S., 2021: Evaluation of fine resolution gridded rainfall datasets over a dense network of rain gauges in Niger. *Atmos. Res.*, **252**, 105459, https://doi.org/10.1016/j.atmosres.2021.105459.

Agutu, N. O., J. L. Awange, A. Zerihun, C. E. Ndehedehe, M. Kuhn, and Y. Fukuda, 2017: Assessing multi-satellite remote sensing, reanalysis, and land surface models' products in characterizing agricultural drought in East Africa. *Remote Sens. Environ.*, **194**, 287–302, https://doi.org/10.1016/j.rse.2017.03.041.

Akinsanola, A. A., K. O. Ogunjobi, V. O. Ajayi, E. A. Adefisan, J. A. Omotosho, and S. Sanogo, 2017: Comparison of five gridded precipitation products at climatological scales over West Africa. *Meteor. Atmos. Phys.*, **129**, 669–689, https://doi.org/10.1007/s00703-016-0493-6.

Arvor, D., B. M. Funatsu, V. Michot, and V. Dubreuil, 2017: Monitoring rainfall patterns in the southern Amazon with PERSIANN-CDR data: Long-term characteristics and trends. *Remote Sens.*, **9**, 889, https://doi.org/10.3390/rs9090889.

Awange, J. L., V. G. Ferreira, E. Forootan, S. A. Andam-Akorful, N. O. Agutu, X. F. He, and Khandu, 2016: Uncertainties in remotely sensed precipitation data over Africa. *Int. J. Climatol.*, **36**, 303–323, https://doi.org/10.1002/joc.4346.

Ayehu, G. T., T. Tadesse, B. Gessesse, and T. Dinku, 2018: Validation of new satellite rainfall products over the upper Blue Nile basin, Ethiopia. *Atmos. Meas. Tech.*, **11**, 1921–1936, https://doi.org/10.5194/amt-11-1921-2018.

Beck, H. E., E. F. Wood, M. Pan, C. K. Fisher, D. G. Miralles, A. I. J. M. van Dijk, T. R. Mcvicar, and R. F. Adler, 2019: MSWEP V2 global 3-hourly 0.1° precipitation: Methodology and quantitative assessment. *Bull. Amer. Meteor. Soc.*, **100**, 473–500, https://doi.org/10.1175/BAMS-D-17-0138.1.

Belay, A. S., and Coauthors, 2019: Evaluation and application of multi-source satellite rainfall product CHIRPS to assess spatio-temporal rainfall variability on data-sparse western margins of Ethiopian highlands. *Remote Sens.*, **11**, 2688, https://doi.org/10.3390/rs11222688.

Brocca, L., and Coauthors, 2014: Soil as a natural rain gauge: Estimating global rainfall from satellite soil moisture data. *J. Geophys. Res. Atmos.*, **119**, 5128–5141, https://doi.org/10.1002/2014JD021489.

——, and Coauthors, 2019: SM2RAIN-ASCAT (2007–2018): Global daily satellite rainfall from ASCAT soil moisture. *Earth Syst. Sci. Data*, **11**, 1583–1601, https://doi.org/10.5194/essd-11-1583-2019.

Carvalho, L. M. V., C. Jones, A. N. D. Posadas, R. Quiroz, B. Bookhagen, and B. Liebmann, 2012: Precipitation characteristics of the South American monsoon system derived from multiple datasets. *J. Climate*, **25**, 4600–4620, https://doi.org/10.1175/JCLI-D-11-00335.1.

Casse, C., M. Gosset, C. Peugeot, V. Pedinotti, A. Boone, B. A. Tanimoun, and B. Decharme, 2015: Potential of satellite rainfall products to predict Niger river flood events in Niamey. *Atmos. Res.*, **163**, 162–176, https://doi.org/10.1016/j.atmosres.2015.01.010.

Collischonn, B., W. Collischonn, and C. E. M. Tucci, 2008: Daily hydrological modeling in the Amazon basin using TRMM rainfall estimates. *J. Hydrol.*, **360**, 207–216, https://doi.org/10.1016/j.jhydrol.2008.07.032.

Copernicus Climate Change Service, 2017: ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service Climate Data Store (CDS), accessed 15 March 2024, https://cds.climate.copernicus.eu/datasets/reanalysis-era5-complete?tab=overview.

Degefu, M. A., W. Bewket, and Y. Amha, 2022: Evaluating performance of 20 global and quasi-global precipitation products in representing drought events in Ethiopia I: Visual and correlation analysis. *Wea. Climate Extremes*, **35**, 100416, https://doi.org/10.1016/j.wace.2022.100416.

Dembélé, M., and S. J. Zwart, 2016: Evaluation and comparison of satellite-based rainfall products in Burkina Faso, West Africa. *Int. J. Remote Sens.*, **37**, 3995–4014, https://doi.org/10.1080/01431161.2016.1207258.

de Wit, A., and Coauthors, 2010: Using ERA-INTERIM for regional crop yield forecasting in Europe. *Climate Res.*, **44**, 41–53, https://doi.org/10.3354/cr00872.

Dinku, T., 2018: The ENACTS approach: Transforming climate services in Africa, one country at a time. *2018 AGU Fall Meeting*, Washington, D.C., Amer. Geophys. Union, Abstract PA11D-0820, https://ui.adsabs.harvard.edu/abs/2018AGUFMPA11D0820D/abstract.

——, P. Ceccato, E. Grover-Kopec, M. Lemma, S. J. Connor, and C. F. 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.

——, S. J. Connor, P. Ceccato, and C. F. Ropelewski, 2008: Comparison of global gridded precipitation products over a mountainous region of Africa. *Int. J. Climatol.*, **28**, 1627–1638, https://doi.org/10.1002/joc.1669.

——, P. Ceccato, and S. J. Connor, 2011: Challenges of satellite rainfall estimation over mountainous and arid parts of East Africa. *Int. J. Remote Sens.*, **32**, 5965–5979, https://doi.org/10.1080/01431161.2010.499381.

——, K. Hailemariam, R. Maidment, E. Tarnavsky, and S. Connor, 2014: Combined use of satellite estimates and rain gauge observations to generate high-quality historical rainfall time series over Ethiopia. *Int. J. Climatol.*, **34**, 2489–2504, https://doi.org/10.1002/joc.3855.

——, M. C. Thomson, R. Cousin, J. del Corral, P. Ceccato, J. Hansen, and S. J. Connor, 2018: Enhancing National Climate Services (ENACTS) for development in Africa. *Climate Dev.*, **10**, 664–672, https://doi.org/10.1080/17565529.2017.1405784.

Doumounia, A., M. Gosset, F. Cazenave, M. Kacou, and F. Zougmore, 2014: Rainfall monitoring based on microwave links from cellular telecommunication networks: First results from a West African test bed. *Geophys. Res. Lett.*, **41**, 6016–6022, https://doi.org/10.1002/2014GL060724.

Dunning, C. M., E. C. L. Black, and R. P. Allan, 2016: The onset and cessation of seasonal rainfall over Africa. *J. Geophys. Res. Atmos.*, **121**, 11 405–11 424, https://doi.org/10.1002/2016JD025428.

Ebert, E. E., 2007: Methods for verifying satellite rainfall estimates. *Measuring Rainfall from Space: EURAINSAT and the Future*, V. Levizzani et al., Eds., Springer, 345–356, https://doi.org/10.1007/978-1-4020-5835-6_22.

Fenta, A. A., and Coauthors, 2018: Evaluation of satellite rainfall estimates over the Lake Tana basin at the source region of the Blue Nile River. *Atmos. Res.*, **212**, 43–53, https://doi.org/10.1016/j.atmosres.2018.05.009.

Funk, C., and Coauthors, 2015: The climate hazards infrared precipitation with stations-a new environmental record for monitoring extremes. *Sci. Data*, **2**, 150066, https://doi.org/10.1038/sdata.2015.66.

Funk, C. C., and Coauthors, 2014: A quasi-global rainfall time series for drought monitoring. USGS Data Series 832, 4 pp., https://pubs.usgs.gov/ds/832/pdf/ds832.pdf.

Gao, Z., D. Long, G. Tang, C. Zeng, J. Huang, and Y. Hong, 2017: Assessing the potential of satellite-based precipitation estimates for flood frequency analysis in ungauged or poorly gauged tributaries of China's Yangtze River basin. *J. Hydrol.*, **550**, 478–496, https://doi.org/10.1016/j.jhydrol.2017.05.025.

Gebremicael, T. G., Y. A. Mohamed, and P. Van der Zaag, 2019: Attributing the hydrological impact of different land use types and their long-term dynamics through combining parsimonious hydrological modelling, alteration analysis and PLSR analysis. *Sci. Total Environ.*, **660**, 1155–1167, https://doi.org/10.1016/j.scitotenv.2019.01.085.

GMAO, 2014: MERRA-2 inst1_2d_asm_Nx: 2d,1-hourly,instantaneous,single-level,assimilation,single-level diagnostics V5.12.4. Goddard Earth Sciences Data and Information Services Center (GES DISC), accessed 9 April 2024, https://cmr.earthdata.nasa.gov/search/concepts/C1276812820-GES_DISC.html.

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.

Guo, H., A. Bao, T. Liu, F. Ndayisaba, D. He, A. Kurban, and P. De Maeyer, 2017: Meteorological drought analysis in the lower Mekong basin using satellite-based long-term CHIRPS product. *Sustainability*, **9**, 901, https://doi.org/10.3390/su9060901.

Hailesilassie, W. T., T. Ayenew, and S. Tekleab, 2021: Analysing trends and spatio-temporal variability of precipitation in the main central rift valley Lakes basin, Ethiopia. *Environ. Earth Sci. Res. J.*, **8**, 37–47, https://doi.org/10.18280/eesrj.080104.

Hegerl, G. C., and Coauthors, 2015: Challenges in quantifying changes in the global water cycle. *Bull. Amer. Meteor. Soc.*, **96**, 1097–1115, https://doi.org/10.1175/BAMS-D-13-00212.1.

Houngnibo, M. C. M., B. Minoungou, S. B. Traore, R. I. Maidment, A. Alhassane, and A. Ali, 2023: Validation of high-resolution satellite precipitation products over West Africa for rainfall monitoring and early warning. *Front. Climate*, **5**, 1185754, https://doi.org/10.3389/fclim.2023.1185754.

Hsu, K.-L., X. Gao, S. Sorooshian, and H. V. Gupta, 1997: Precipitation estimation from remotely sensed information using artificial neural networks. *J. Appl. Meteor.*, **36**, 1176–1190, https://doi.org/10.1175/1520-0450(1997)036<1176:PEFRSI>2.0.CO;2.

Huffman, G. J., E. F. Stocker, D. T. Bolvin, E. J. Nelkin, and J. Tan, 2019: IMERG-FR IMERG final rainfall L3 1 day 0.1 degree × 0.1 degree V06. Goddard Earth Sciences Data and Information Services Center (GES DISC), accessed 15 March 2024, https://doi.org/10.5067/GPM/IMERG/3B-HH/06.

Hussain, Y., F. Satgé, M. B. Hussain, H. Martinez-Carvajal, M.-P. Bonnet, M. Cárdenas-Soto, H. L. Roig, and G. Akhter, 2018: Performance of CMORPH, TMPA, and PERSIANN rainfall datasets over plain, mountainous, and glacial regions of Pakistan. *Theor. Appl. Climatol.*, **131**, 1119–1132, https://doi.org/10.1007/s00704-016-2027-z.

Jobard, I., F. Chopin, J. C. Bergès, and R. Roca, 2011: An intercomparison of 10-day satellite precipitation products during West African monsoon. *Int. J. Remote Sens.*, **32**, 2353–2376, https://doi.org/10.1080/01431161003698286.

Kidd, C., A. Becker, G. J. Huffman, C. L. Muller, P. Joe, G. Skofronick-Jackson, and D. B. Kirschbaum, 2017: So, how much of the Earth's surface is covered by rain gauges? *Bull. Amer. Meteor. Soc.*, **98**, 69–78, https://doi.org/10.1175/BAMS-D-14-00283.1.

Kling, H., M. Fuchs, and M. Paulin, 2012: Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *J. Hydrol.*, **424–425**, 264–277, https://doi.org/10.1016/j.jhydrol.2012.01.011.

Kouakou, C., J.-E. Paturel, F. Satgé, Y. Tramblay, D. Defrance, and N. Rouché, 2023: Comparison of gridded precipitation estimates for regional hydrological modeling in west and central Africa. *J. Hydrol.*, **47**, 101409, https://doi.org/10.1016/j.ejrh.2023.101409.

Lebel, T., J.-D. Taupin, and N. D'Amato, 1997: Rainfall monitoring during HAPEX-Sahel. 1. General rainfall conditions and climatology. *J. Hydrol.*, **188–189**, 74–96, https://doi.org/10.1016/S0022-1694(96)03155-1.

Love, T., 2002: The climate prediction center rainfall algorithm version 2. NOAA/NCEP Climate Prediction Center Tech. Rep. 2.0, 28 pp., https://www.cpc.ncep.noaa.gov/products/fews/RFE2.ppt.

Maidment, R. I., D. Grimes, R. P. Allan, E. Tarnavsky, M. Stringer, T. Hewison, R. Roebeling, and E. Black, 2014: The 30 year TAMSAT African Rainfall Climatology And Time series (TARCAT) data set. *J. Geophys. Res. Atmos.*, **119**, 10 619–10 644, https://doi.org/10.1002/2014JD021927.

——, and Coauthors, 2017: A new, long-term daily satellite-based rainfall dataset for operational monitoring in Africa. *Sci. Data*, **4**, 170063, https://doi.org/10.1038/sdata.2017.63.

Manzanas, R., L. K. Amekudzi, K. Preko, S. Herrera, and J. M. Gutierrez, 2014: Precipitation variability and trends in Ghana: An intercomparison of observational and reanalysis products. *Climatic Change*, **124**, 805–819, https://doi.org/10.1007/s10584-014-1100-9.

Messer, H., A. Zinevich, and P. Alpert, 2006: Environmental monitoring by wireless communication networks. *Science*, **312**, 713, https://doi.org/10.1126/science.1120034.

Nguyen, P., and Coauthors, 2019: The CHRS data portal, an easily accessible public repository for PERSIANN global satellite precipitation data. *Sci. Data*, **6**, 180296, https://doi.org/10.1038/sdata.2018.296.

Nicholson, S. E., and Coauthors, 2003: Validation of TRMM and other rainfall estimates with a high-density gauge dataset for West Africa. Part I: Validation of GPCC rainfall product and pre-TRMM satellite and blended products. *J. Appl. Meteor.*, **42**, 1337–1354, https://doi.org/10.1175/1520-0450(2003)042<1337:VOTAOR>2.0.CO;2.

——, and P. J. Webster, 2007: A physical basis for the interannual variability of rainfall in the Sahel. *Quart. J. Roy. Meteor. Soc.*, **133**, 2065–2084, https://doi.org/10.1002/qj.104.

Nikolopoulos, E. I., E. N. Anagnostou, and M. Borga, 2013: Using high-resolution satellite rainfall products to simulate a major

flash flood event in northern Italy. *J. Hydrometeor*, **14**, 171–185, https://doi.org/10.1175/JHM-D-12-09.1.

Novella, N. S., and W. M. Thiaw, 2013: African rainfall climatology version 2 for famine early warning systems. *J. Appl. Meteor. Climatol.*, **52**, 588–606, https://doi.org/10.1175/JAMC-D-11-0238.1.

Nwachukwu, P. N., F. Satgé, S. E. Yacoubi, S. Pinel, and M.-P. Bonnet, 2020: From TRMM to GPM: How reliable are satellite-based precipitation data across Nigeria? *Remote Sens.*, **12**, 3964, https://doi.org/10.3390/rs12233964.

Overeem, A., H. Leijnse, and R. Uijlenhoet, 2011: Measuring urban rainfall using microwave links from commercial cellular communication networks. *Water Resour. Res.*, **47**, W12505, https://doi.org/10.1029/2010WR010350.

Owusu, K., and P. Waylen, 2013: The changing rainy season climatology of mid-Ghana. *Theor. Appl. Climatol.*, **112**, 419–430, https://doi.org/10.1007/s00704-012-0736-5.

Pellarin, T., 2019: PrISM satellite rainfall product based on SMOS soil moisture measurements in Africa (3h, 0.25°). Zenodo, accessed 25 March 2024, https://doi.org/10.5281/zenodo.3565610.

Petković, V., M. Orescanin, P. Kirstetter, C. Kummerow, and R. Ferraro, 2019: Enhancing PMW satellite precipitation estimation: Detecting convective class. *J. Atmos. Oceanic Technol.*, **36**, 2349–2363, https://doi.org/10.1175/JTECH-D-19-0008.1.

Pierre, C., G. Bergametti, B. Marticorena, E. Mougin, T. Lebel, and A. Ali, 2011: Pluriannual comparisons of satellite-based rainfall products over the Sahelian belt for seasonal vegetation modeling. *J. Geophys. Res.*, **116**, D18201, https://doi.org/10.1029/2011JD016115.

Poméon, T., D. Jackisch, and B. Diekkrüger, 2017: Evaluating the performance of remotely sensed and reanalysed precipitation data over West Africa using HBV light. *J. Hydrol.*, **547**, 222–235, https://doi.org/10.1016/j.jhydrol.2017.01.055.

Rahmawati, N., and M. W. Lubczynski, 2018: Validation of satellite daily rainfall estimates in complex terrain of Bali Island, Indonesia. *Theor. Appl. Climatol.*, **134**, 513–532, https://doi.org/10.1007/s00704-017-2290-7.

Ramarohetra, J., B. Sultan, C. Baron, T. Gaiser, and M. Gosset, 2013: How satellite rainfall estimate errors may impact rainfed cereal yield simulation in West Africa. *Agric. For. Meteor.*, **180**, 118–131, https://doi.org/10.1016/j.agrformet.2013.05.010.

Roca, R., P. Chambon, I. Jobard, P.-E. Kirstetter, M. Gosset, and J. C. Berges, 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.

Sadeghi, M., P. Nguyen, M. R. Naeini, K. Hsu, D. Braithwaite, and S. Sorooshian, 2021: PERSIANN-CCS-CDR, a 3-hourly 0.04° global precipitation climate data record for heavy precipitation studies. *Sci. Data*, **8**, 157, https://doi.org/10.1038/s41597-021-00940-9.

Sagna, P., 2007: Caractéristiques climatiques. *Atlas du Sénégal*, Paris, Les éditions J.A., 66–69.

Sarojini, B. B., P. A. Stott, and E. Black, 2016: Detection and attribution of human influence on regional precipitation. *Nat. Climate Change*, **6**, 669–675, https://doi.org/10.1038/nclimate2976.

Satgé, F., A. Xavier, R. P. Zolá, Y. Hussain, F. Timouk, J. Garnier, and M.-P. Bonnet, 2017: Comparative assessments of the latest GPM mission's spatially enhanced satellite rainfall products over the main Bolivian watersheds. *Remote Sens.*, **9**, 369, https://doi.org/10.3390/rs9040369.

——, D. Ruelland, M.-P. Bonnet, J. Molina, and R. Pillco, 2019: Consistency of satellite-based precipitation products in space and over time compared with gauge observations and snow-hydrological modelling in the Lake Titicaca region. *Hydrol. Earth Syst. Sci.*, **23**, 595–619, https://doi.org/10.5194/hess-23-595-2019.

——, D. Defrance, B. Sultan, M.-P. Bonnet, F. Seyler, N. Rouché, F. Pierron, and J.-E. Paturel, 2020: Evaluation of 23 gridded precipitation datasets across West Africa. *J. Hydrol.*, **581**, 124412, https://doi.org/10.1016/j.jhydrol.2019.124412.

Siebert, A., T. Dinku, F. Vuguziga, A. Twahirwa, D. M. Kagabo, J. delCorral, and A. W. Robertson, 2019: Evaluation of ENACTS-Rwanda: A new multi-decade, high-resolution rainfall and temperature data set—Climatology. *Int. J. Climatol.*, **39**, 3104–3120, https://doi.org/10.1002/joc.6010.

Su, J., H. Lü, J. Wang, A. M. Sadeghi, and Y. Zhu, 2017: Evaluating the applicability of four latest satellite–gauge combined precipitation estimates for extreme precipitation and streamflow predictions over the upper Yellow River basins in China. *Remote Sens.*, **9**, 1176, https://doi.org/10.3390/rs9111176.

Sultan, B., and Coauthors, 2013: Assessing climate change impacts on sorghum and millet yields in the Sudanian and Sahelian savannas of West Africa. *Environ. Res. Lett.*, **8**, 014040, https://doi.org/10.1088/1748-9326/8/1/014040.

Sun, Q., C. Miao, Q. Duan, D. Kong, A. Ye, Z. Di, and W. Gong, 2014: Would the 'real' observed dataset stand up? A critical examination of eight observed gridded climate datasets for China. *Environ. Res. Lett.*, **9**, 015001, https://doi.org/10.1088/1748-9326/9/1/015001.

Sun, W., J. Ma, G. Yang, and W. Li, 2018: Statistical and hydrological evaluations of multi-satellite precipitation products over Fujiang River basin in humid southeast China. *Remote Sens.*, **10**, 1898, https://doi.org/10.3390/rs10121898.

Tang, G., Y. Ma, D. Long, L. Zhong, and Y. Hong, 2016: Evaluation of GPM day[1] IMERG and TMPA version-7 legacy products over mainland China at multiple spatiotemporal scales. *J. Hydrol.*, **533**, 152–167, https://doi.org/10.1016/j.jhydrol.2015.12.008.

Tarnavsky, E., D. Grimes, R. Maidment, E. Black, R. P. Allan, M. Stringer, R. Chadwick, and F. Kayitakire, 2014: Extension of the TAMSAT satellite-based rainfall monitoring over Africa from 1983 to present. *J. Appl. Meteor. Climatol.*, 53, 2805–2822, https://doi.org/10.1175/JAMC-D-14-0016.1.

Thaler, S., L. Brocca, L. Ciabatta, J. Eitzinger, S. Hahn, and W. Wagner, 2018: Effects of different spatial precipitation input data on crop model outputs under a central European climate. *Atmosphere*, **9**, 290, https://doi.org/10.3390/atmos9080290.

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. Hydrometeor.*, **13**, 1760–1783, https://doi.org/10.1175/JHM-D-12-032.1.

Toté, C., D. Patricio, H. Boogaard, R. Van Der Wijngaart, E. Tarnavsky, and C. Funk, 2015: Evaluation of satellite rainfall estimates for drought and flood monitoring in Mozambique. *Remote Sens.*, **7**, 1758–1776, https://doi.org/10.3390/rs70201758.

Trenberth, K. E., Y. X. Zhang, and M. Gehne, 2017: Intermittency in precipitation: Duration, frequency, intensity, and amounts using hourly data. *J. Hydrometeor.*, **18**, 1393–1412, https://doi.org/10.1175/JHM-D-16-0263.1.

WMO, 2020: 2020 State of Climate Services: Risk information and early warning systems. WMO-1252, 25 pp., https://library.wmo.int/idurl/4/57191.

Xie, P., R. Joyce, S. Wu, S.-H. Yoo, Y. Yarosh, F. Sun, and R. Lin, 2019: NOAA Climate Data Record (CDR) of CPC Morphing Technique (CMORPH) high resolution global rainfall estimates, Version 1. NOAA NCEI, accessed 25 March 2024, https://doi.org/10.25921/w9va-q159.

Yonaba, R., A. Belemtoogri, F. Tazen, L. A. Mounirou, M. Koïta, H. Karambiri, and H. Yacouba, 2022: Assessing the accuracy of SM2RAIN (Soil Moisture to Rainfall) products in poorly gauged countries: The case of Burkina Faso in the West African Sahel. Copernicus Meetings No. IAHS2022-263, 21 pp., https://www.researchgate.net/profile/Roland-Yonaba/publication/361074000_Assessing_the_accuracy_of_SM2RAIN_Soil_Moisture_to_Rainfall_products_in_poorly_gauged_countries_the_case_of_Burkina_Faso_in_the_West_African_Sahel/links/629a91b5416ec50bdb09536b/Assessing-the-accuracy-of-SM2RAIN-Soil-Moisture-to-Rainfall-products-in-poorly-gauged-countries-the-case-of-Burkina-Faso-in-the-West-African-Sahel.pdf.

Zeng, Q., Y. Wang, L. Chen, Z. Wang, H. Zhu, and B. Li, 2018: Inter-comparison and evaluation of remote sensing precipitation products over China from 2005 to 2013. *Remote Sens.*, **10**, 168, https://doi.org/10.3390/rs10020168.

Zhang, Y., Y. Li, X. Ji, X. Luo, and X. Li, 2018: Evaluation and hydrologic validation of three satellite-based precipitation products in the upper catchment of the Red River basin, China. *Remote Sens.*, **10**, 1881, https://doi.org/10.3390/rs10121881.

Zinevich, A., P. Alpert, and H. Messer, 2008: Estimation of rainfall fields using commercial microwave communication networks of variable density. *Adv. Water Resour.*, **31**, 1470–1480, https://doi.org/10.1016/j.advwatres.2008.03.003.