Journal of Hydrology 581 (2020) 124412

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Evaluation of 23 gridded precipitation datasets across West Africa

Frédéric Satgé[a,*], Dimitri Defrance[a,b], Benjamin Sultan[a], Marie-Paule Bonnet[a], Frédérique Seyler[a], Nathalie Rouché[c], Fabrice Pierron[a], Jean-Emmanuel Paturel[c]

[a] ESPACE-DEV, Université Montpellier, IRD, Université Guyane, Université Réunion, Université Antilles, Université Avignon, Montpellier, France
[b] SYSTEM, Université Montpellier, INRA, Montpellier SupAgro, CIRAD, CIHEAM, 34000 Montpellier, France
[c] HydroSciences Montpellier, IRD, Université de Montpellier, Montpellier, CNRS, Montpellier, France

ARTICLE INFO

This manuscript was handled by Emmanouil Anagnostou, Editor-in-Chief

Keywords:
Precipitation datasets
Reliability
West Africa

ABSTRACT

This study aims reporting on 23 gridded precipitation datasets (P-datasets) reliability across West Africa through direct comparisons with rain gauges measurement at the daily and monthly time scales over a 4 years period (2000–2003). All P-datasets reliability vary in space and time. The most efficient P-dataset in term of Kling–Gupta Efficiency (KGE) changes at the local scale and the P-dataset performance is sensitive to seasonal effects. Satellite-based P-datasets performed better during the wet than the dry season whereas the opposite is observed for reanalysis P-datasets. The best overall performance was obtained for MSWEP v.2.2 and CHIRPS v.2 for daily and monthly time-step, respectively. Part of the differences in P-dataset performance at daily and monthly time step comes from the time step used to proceed the gauges adjustment (i.e day or month) and from a mismatch between gauge and satellite reporting times. In comparison to the others P-datasets, TMPA-Adj v.7 reliability is stable and reach the second highest KGE value at both daily and monthly time step. Reanalysis P-datasets (WFDEI, MERRA-2, JRA-55, ERA-Interim) present among the lowest statistical scores at the daily time step, which drastically increased at the monthly time step for WFDEI and MERRA-2. The non-adjusted P-datasets were the less efficient, but, their near-real time availability should be helpful for risk forecast studies (i.e. GSMaP-RT v.6). The results of this study give important elements to select the most adapted P-dataset for specific application across West Africa.

1. Introduction

1.1. Precipitation: A key factor subject to uncertainty

Water resources are facing unprecedented changes related to redistribution of seasonal precipitation (Saeed et al., 2018) and intensity (Fischer and Knutti, 2015; Giorgi et al., 2018) owing to climate variability. With a six-fold increase of water extraction during the 20th century in response to increases in the world population (Cosgrove and Risberman, 2000), food requirements and the economy may be particularly affected by these changes. Accurate spatiotemporal precipitation monitoring is therefore crucial for detect and quantifying ongoing changes in optimising water resource management. Traditionally, the precipitation amount is measured at the point scale from gauge measurements. However, access difficulty, political instability, and economic issues have often resulted in sparse and unevenly distributed rain gauge networks that incorrectly capture the spatial precipitation variability (Lebel et al., 1997; Li and Heap, 2014). Alternatively, weather radar stations enable precipitation monitoring with spatial distribution over larger and even remote areas. However, radar stations are expensive, and only a few are available worldwide. In addition, large amounts of radar signal interference prevent accurate estimation of precipitation over complex terrains (Tang et al., 2016; Zeng et al., 2018). Several authors have recently reported on the potential of using cellular phone signal attenuation during precipitation events to retrieve precipitation measurements (Doumounia et al., 2014; Messer et al., 2006; Overeem et al., 2011; Zinevich et al., 2008). Although these estimations are accurate, they are limited to regions with high antenna density (e.g. urban areas). Moreover, this technique faces the problem of accessing data owned by private cellular phone companies.

Regardless of the technique employed, precipitation data collection at the regional scale usually includes potential conflicts of interest in water resource management between neighbouring countries. In this context, gridded precipitation datasets (P-dataset) at an almost global scale offer an unprecedented alternative. Over remote regions, P-datasets have already shown promising perspective for water resource management by enhancing our understanding of drought (e.g. Agutu et al., 2017; Guo et al., 2017; Satgé et al., 2017a; Toté et al., 2015) and

* Corresponding author.
E-mail address: frederic.satge@ird.fr (F. Satgé).

https://doi.org/10.1016/j.jhydrol.2019.124412
Received 30 July 2019; Received in revised form 22 November 2019; Accepted 25 November 2019
Available online 30 November 2019

flood (e.g. Gao et al., 2017; Nikolopoulos et al., 2013; Toté et al., 2015) events, precipitation variability (e.g. Arvor et al., 2017; Carvalho et al., 2012), streamflow (e.g. Collischonn et al., 2008; De Paiva et al., 2013; Satgé et al., 2019; Sun et al., 2018b; Zhang et al., 2018) and snow cover dynamics (e.g. Satgé et al., 2019), and agriculture productivity (e.g. Thaler et al., 2018; De Wit et al., 2010).

1.2. State of the art for P-datasets

Three groups of gridded P-dataset can be defined depending on the input and technique used to retrieve the precipitation amounts: (1) those based on the spatial information of available gauges, (2) those based on reanalysis data derived from physical and dynamical models, and (3) those based on satellite information using passive-microwave (PMW) and infrared (IR) information. It is worth mentioning that most of the P-datasets merge aspects of these three inputs and techniques to ensure the best accuracy possible. Recently, 30 global-scale P-datasets with variable space–time coverage and resolution have been listed (Sun et al., 2018a) which present precipitation estimates discrepancy in space and time according to their different bases such as data capture, integration, and algorithms. For example, gauge-based P-dataset reliability varies in space and time according to changes in the number of available gauges used for the interpolation process (Sun et al., 2015). Similarly, satellite-based P-dataset reliability varies in space and time because the PMW and IR algorithms present limits over complex mountainous (Hussain et al., 2017; Satgé et al., 2017b) and snow-covered regions (Ferraro et al., 1998; Levizzani et al., 2002) and during short-term and slight precipitation events (Gebregiorgis and Hossain, 2013; Tian et al., 2009). Finally, reanalysis data-based P-datasets present variable reliability in space and time owing to the limited ability of the models used to represent small-scale convective cells (Beck et al., 2019). In this context, many studies assess P-dataset space–time uncertainties to evaluate their reliability (Maggioni et al., 2016; Maggioni and Massari, 2018).

A recurrent drawback of assessment studies on P-dataset reliability is the consideration of a limited number of P-datasets. A comprehensive reliability overview of available P-datasets, as listed in Table 1, can be achieved only by backcrossing the results from different P-dataset assessment studies. However, the studies are conducted over distinct regions and are based on different statistical indices, spatial and temporal scales, and periods, thus creating difficulties in intercomparing P-dataset reliability assessments. For example, when comparing TMPA, CMORPH, and PERSIANN P-datasets with reference gauge estimates, CMORPH was shown to have the most reliable P-datasets in Pakistan, China, Bali, and Indonesia (Hussain et al., 2017; Rahmawati and Lubczynski, 2017; Su et al., 2017; Zeng et al., 2018). However, TMPA was the most reliable in India, Guyana, Chile, and the South American Andean plateau (Prakash et al., 2014a; Ringard et al., 2015; Satgé et al., 2016; Zambrano-Bigiarini et al., 2017). Hence, P-dataset reliability for a given region should not be determined from results reported for other regions. In this context, it is decisive to consider the most representative P-dataset sample to insure a consistent report on P-dataset reliability across the considered region.

1.3. The need for assessing P-datasets over West Africa

Africa is particularly affected by climate changes threatening rainfed agriculture, which represents its main agricultural and economic activity (Sultan et al., 2013). However, owing to the socio-economic context, the available gauge network is limited by many spatial and temporal gaps which prevent efficient water management. According to the World Meteorological Organisation (WMO), the African continent requires uniform distribution of at least 3000 stations (ideally 10,000); however, only 744 stations are present. Moreover, only one quarter of the 744 stations conform to international standards.

Because they provide precipitation information on a regular grid at the global scale, P-datasets offer a unique opportunity for complementing traditional precipitation measurements and optimising population adaption to the ongoing changes. However, as previously mentioned, P-dataset estimates are indirect measurements with spatial and temporal uncertainties which need to be reported to evaluate their reliability. Some authors have already initiated this effort over West Africa. In 2012, seven P-datasets were tested over the basin of la Volta including CMORPH, GPROF-v6, GSMaP-MVK v5, RFE-2.0, TMPA-v6, PERSIANN, and ERA-Interim (Thiemig et al., 2012). In 2013, nine P-datasets including CMORPH, EPSAT-SG, GPCP, GSMaP-MVK, GSMaP-RT, RFE-2, TMPA-v6, TMPA-RT v6, and PERSIANN and seven P-datasets including PERSIANN, CMORPH, TMPA-RT v.6, TMPA-Adj v.6, GSMaP-MVK, GCPC-1dd, and RFE-2 were tested in Benin and Niger for hydrological (Gosset et al., 2013) and agriculture applications (Ramarohetra et al., 2013), respectively. Both studies found that their use could introduce large biases in crop or hydrological modelling framework. More recently, six P-datasets including ARC-2, CMORPH, GSMaP-MVK, PERSIANN, TAMSAT, and TMPA-v.6 were compared with gauge measurements data over the entire African continent (Awange et al., 2016).

All of the aforementioned studies focus mainly on P-datasets regularly updated by their developers to enhance the precipitation estimates. Since then, updated versions of the considered products have been made available with more accurate precipitation estimates. For example, the benefits brought by the new TMPA-v.7 in comparison to its previous version (TMPA-v.6) has been reported in many regions (e.g. Anjum et al., 2016; Prakash et al., 2014b; Satgé et al., 2016). Additionally, most of the tested P-datasets originate from the TRMM-era constellation which has limited temporal coverage from 1998 to the present. In this context, new studies have reported on recently released P-dataset versions with larger temporal coverage. For example, in 2016 over Burkina, seven P-datasets including ARC-2, CHIRPS v.2, PERSIANN-CDR, RFE v.2, TAMSAT v.2, TMPA v.7, and TMPA RT v.7 were assessed at the daily, decadal, and monthly timescales (Dembélé and Zwart, 2016). In 2017, TAMSAT v.3 was introduced and compared with its previous version (TAMSAT v.2) and with six P-datasets including ARC v.2, CHIRP v.2, CHIRPS v.2, CMORPH v.1, RFE and TMPA v.7 over West Africa, specifically Nigeria and Niger; Uganda; Zambia; and Mozambique (Maidment et al., 2017). In 2017, 10 P-datasets including CFSR, CHIRPS, CMORPH v.1 RAW and CRT, PERSIANN-CDR, RFE-2, TAMSAT v.2, TMPA v.7, TMPA-RT v.7, and GPCC were assessed over six watersheds located in Burkina, Nigeria, and Ghana (Poméon et al., 2017). Nonetheless, the reported studies indicate that the results are mostly limited in space (country or basin scale) and in terms of considered P-datasets sample. To our knowledge, only one study has reported on P-datasets at the regional West African scale with a limited sample of P-datasets including TMPA v.7, UDEL v.3.1, CRU v.3, and ARC v.2 (Akinsanola et al., 2016).

1.4. Objectives

According to the previously described context, the present study aims to compare the accuracy of 23 P-datasets in reproducing the characteristics of rain gauge measurements across West Africa, which is an unprecedented comparison. The consideration of a P-dataset sample, as large as possible, aims to provide a robust overview of P-dataset performance over West Africa. The analysis is conducted at both daily and monthly time steps. This study provides important feedback to P-dataset developers for enhancing the algorithms for next-generation P-datasets and to potential users to support their P-datasets selection.

2. Materials methods

2.1. Study area

The study area, hereafter referred to as West Africa, extends from

Table 1
Main characteristics and references of the P-datasets. In the data source column, S, R, and G stands for satellite, reanalysis, and gauge information. Spatial coverage refers to the absolute maximum and minimum latitude with precipitation information, and latency refers to the time delay for data availability. The P-datasets including gauge-based information are represented in blue, and italic font is used for P-datasets available in NRT latency of one to three days.

Acronym	Full Name	Data	Temporal Coverage	Temporal Resolution	Spatial Coverage	Spatial Resolution	Latency	Link	References
ARC-2	Africa Rainfall Climatology v.2	S, G	1983–present	Daily	Africa	0.1°	2 days	ftp://ftp.cpc.ncep.noaa.gov/fews/fewsdata/africa/arc2/	Novella and Thiaw (2012)
CHIRP v.2	Climate Hazards Group InfraRed v.2	S, R	1981–present	Daily	50°	0.05°	2 days	ftp://ftp.chg.ucsb.edu/pub/org/chg/products/	Funk et al. (2015)
CHIRPS v.2	CHIRP with Station v.2	S, R, G	1981–present	Daily	50°	0.05°	1 month	ftp://ftp.chg.ucsb.edu/pub/org/chg/products/	Funk et al. (2015)
CMORPH-Raw v.1	Climate Prediction Center MORPHing raw v.1	S	1998–present	3 h	60°	0.25°	2 days	ftp://ftp.cpc.ncep.noaa.gov/precip/CMORPH_V1.0/	Joyce et al. (2004)
CMORPH-CRT v.1	CMORPH bias corrected v.1	S, G	1998–present	3 h	60°	0.25°	6 months	ftp://ftp.cpc.ncep.noaa.gov/precip/CMORPH_V1.0/	Xie et al. (2017)
CMORPH-BLD v.1	CMORPH satellite-gauge merged v.1	S, G	1998–present	Daily	60°	0.25°	10 months	ftp://ftp.cpc.ncep.noaa.gov/precip/CMORPH_V1.0/	Xie et al. (2017)
CPC v.1	Climate Prediction Center unified v.1	G	1979–present	Daily	Global	0,5°	1 days	ftp://ftp.cpc.ncep.noaa.gov/precip/CPC_UNI_PRCP/GAUGE_GLB/	Xie et al. (2007) Chen et al. (2008)
ERA-Interim	European Centre for Medium-range Weather Forecast Re Analysis Interim	R	1979–present	3 h	60°	0.75°	3 months	https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-interim-land	Dee et al. (2011)
GSMaP-RT v.6	Global Satellite Mapping of Precipitation standard v.6	S	2000–present	Hourly	60°	0.1°	3 days	ftp://hokusai.eorc.jaxa.jp/standard/v6/	Ushio et al. (2009) Yamamoto and Shige (2014)
GSMaP-Adj v.6	GSMaP adjusted v.6	S, G	2000-–resent	Hourly	60°	0.1°	3 days	ftp://hokusai.eorc.jaxa.jp/standard/v6/	Ushio et al. (2009) Yamamoto and Shige (2014)
GPCC v.7	Global Precipitation Climatology Center	G	1901–2013	Monthly	Global	1°	Irregular	https://rda.ucar.edu/datasets/ds496.0/	Becker et al. (2013); Schneider et al. (2014)
JRA-55	Japanese 55-year Re Analysis	R	1959–present	3 h	Global	0,56°	1 Month	https://rda.ucar.edu/datasets/ds628.0/	Kobayashi et al. (2015)
JRA-55 Adj	JRA-55 Adjusted	R,G	1959–2013	3 h	Global	0,56°	Stopped	http://search.diasjp.net/en/dataset/S14FD	Izumi et al. (2017)
MERRA-2	Modern-Era Retrospective Analysis for Research and Applications 2	S, R, G	1980–present	Hourly	Global	0,5°	2 Months	https://disc.gsfc.nasa.gov/	Gelaro et al. (2017) Reichle et al. (2017)
MSWEP v.2.2	Multi-Source Weighted Ensemble Precipitation v.2.2	S, R, G	1979–present	3 h	Global	0.1°	Few months	http://www.gloh2o.org/ (Personal communication)	Beck et al. (2018) Beck et al. (2019)
PERSIANN-CDR	Precipitation Estimates from Remotely Sensed Information using Artificial Neural Network and Climate Data Record	S, G	1983–2016	Daily	60°	0.25°	6 months	https://chrsdata.eng.uci.edu/	Ashouri et al. (2015)
PERSIANN-RT	PERSIANN real time	S	2000–present	6 h	60°	0.25°	2 days	https://chrsdata.eng.uci.edu/	Hsu et al. (1997) Sorooshian et al. (2000)
PERSIANN-Adj	PERSIANN Adjusted	S, G	2000–2010	3 h	60°	0.25°	Stopped	http://fire.eng.uci.edu/PERSIANN/	Hsu et al. (1997) Sorooshian et al. (2000)
SM2Rain-CCI v.2	Soil Moisture to Rain applied on ESA Climate Change Initiative v.2	S	1998–2015	Daily	Global	0.25°	Stopped	https://zenodo.org/record/846260#.XQEZtYgzZaQ	Ciabatta et al. (2018)
TAMSAT-v.3	Tropical Applications of Meteorology using SATellite and ground-based observations v.3	S, G	1983–present	Daily	Africa	0.0375°	3 days	https://www.tamsat.org.uk/about	Maidment et al. (2017)
TMPA-RT v.7	TRMM Multi-satellite Precipitation Analysis Real Time v.7	S	1998–present	3 h	60°	0.25°	1 day	https://mirador.gsfc.nasa.gov/	Huffman et al. (2018) Huffman et al. (2010)
TMPA-Adj v.7	TMPA Adjusted v.7	S, G	2000–present	3 h	50°	0.25°	3 months	https://earthdata.nasa.gov/	Huffman et al. (2018) Huffman et al. (2010)
WFDEI	WATCH Forcing Data methodology applied to ERA-Interim	R, G	1979–2016	Daily	Land	0.5°	Stopped	ftp://ftp.iiasa.ac.at/	Weedon et al. (2014)

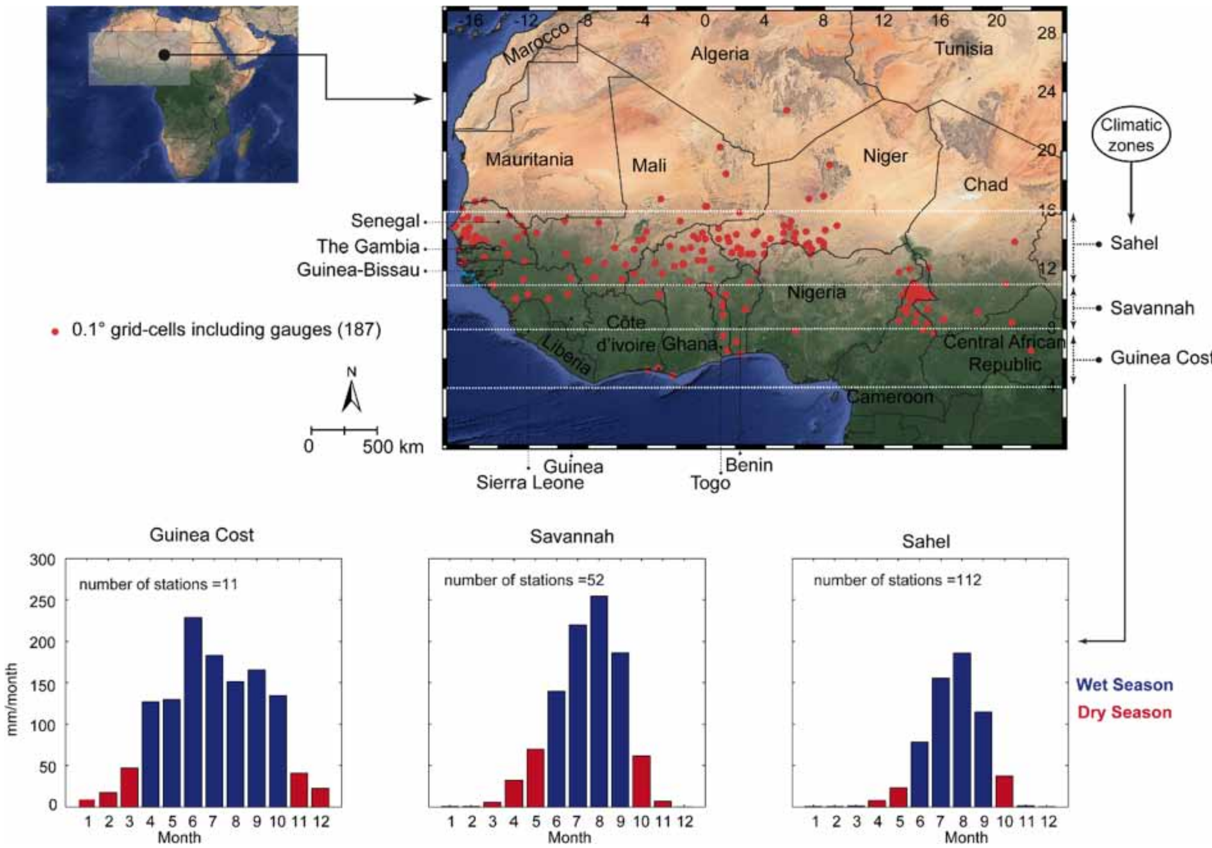

Fig. 1. Study area with the considered 0.1° grid-cell locations and the mean monthly precipitation amount given for the three climatic regions based on gauge records of 2000–2003.

the Atlantic coast of Senegal to eastern Chad and the Gulf of Guinea to north of the Sahel (18° W–25° E, 4° N–25° N) (Fig. 1). The region is characterised by a marked south–north gradient of rainfall amount ranging from 5000 mm.year^{-1} in Cameroon to less than 200 mm.year^{-1} in the northern Sahel. The West Africa region can be divided into three main climatic zones: (i) the Guinea Coast (4°–8° N), (ii) the Savannah (8°–11° N), and (iii) the Sahel (11°–16° N) (Abiodun et al., 2012; Akinsanola et al., 2016) (Fig. 1). For all zones, the year is characterised by a dry season in winter and a rainy season in summer linked to the West African Monsoon. This concentrates most of the annual rainfall amount from April to October for the Guinea Coast zone and from June to September for both the Savanah and Sahel zones (Fig. 1). The rainfall interannual temporality is great with the occurrence of drought phases (dry spell) during the rainy season and in interannual rainfall with very dry and very wet years in the 1970 s and 1950 s, respectively.

2.2. Selected P-datasets

A sample of 23 gridded P-datasets including 13 long-term P-datasets with more than 35 years of continuous observation and 10 P-datasets spanning more than 15 years was selected. Table 1 provides an overview of these P-datasets and relevant references for further information on their respective productions.

2.2.1. Comments on the selected P-datasets

Some P-datasets use gauge-based information in their respective algorithms (Table 1). Three types of gauge-based information are used: (1) punctual precipitation estimates derived from gauge records, (2) gridded precipitation estimates based on interpolation of punctual gauge records, and (3) gauge precipitation estimates (punctual or gridded) merged with different satellite datasets of precipitation, brightness, or land surface temperature.

Punctual precipitation estimates from the world meteorological organisation (WMO) Global Telecommunication System (GTS) (Novella and Thiaw, 2012) and numerous African national meteorological and hydrological centres (Maidment et al., 2014) are used for ARC-2 and TAMSAT v.3, respectively. In both cases, the gauge network is very sparse. For example, the GTS gauge network has a 1:23 000 km^2 gauge-to-area ratio across the African continent (Novella and Thiaw, 2012).

The gridded precipitation estimates are (i) the GPCC with a 1° spatial resolution (Becker et al., 2013; Schneider et al., 2014) and (ii) the daily CPC with 0.5° spatial resolution (Chen et al., 2008; Xie et al., 2007). JRA-55 Adj, TMPA-Adj v.7, and WFDEI use GPCC monthly data, whereas CMORPH-CRT v.1, CMORPH-BLD v.1, GSMaP-Adj v.6, and MERRA-2 use CPC daily data.

The gridded precipitation estimates merged with satellite precipitation estimates are (i) the CHPclim dataset with a 0.05° grid-cell size (Funk et al., 2015) (ii) the GPCP dataset with a 2.5° grid-cell size (Adler et al., 2003, 2012) and (iii) the WorldClim 2 dataset with a 1 km grid-cell size (Fick and Hijmans, 2017). CHPclim and WorldClim 2 use satellite observations as predictors to improve the interpolation from point gauge records, whereas GPCP uses the gauge record to adjust the precipitation fields derived from satellite observations. Further details are reported elsewhere (Adler et al., 2003; Funk et al., 2015). Among the considered P-datasets, CHIRPS v.2 use the CHPclim dataset, MSWEP v.2.2 use the WorldClim 2 dataset and PERSIANN-CDR uses the GPCP dataset.

CHIRPS v.2 also includes punctual precipitation estimates from various public data streams, private archives, and national meteorological agencies, while MSWEP v.2.2 incorporates monthly GPCC and daily gauge observations compiled from several sources (Beck et al., 2018).

Another difference between the P-datasets is the time latency for

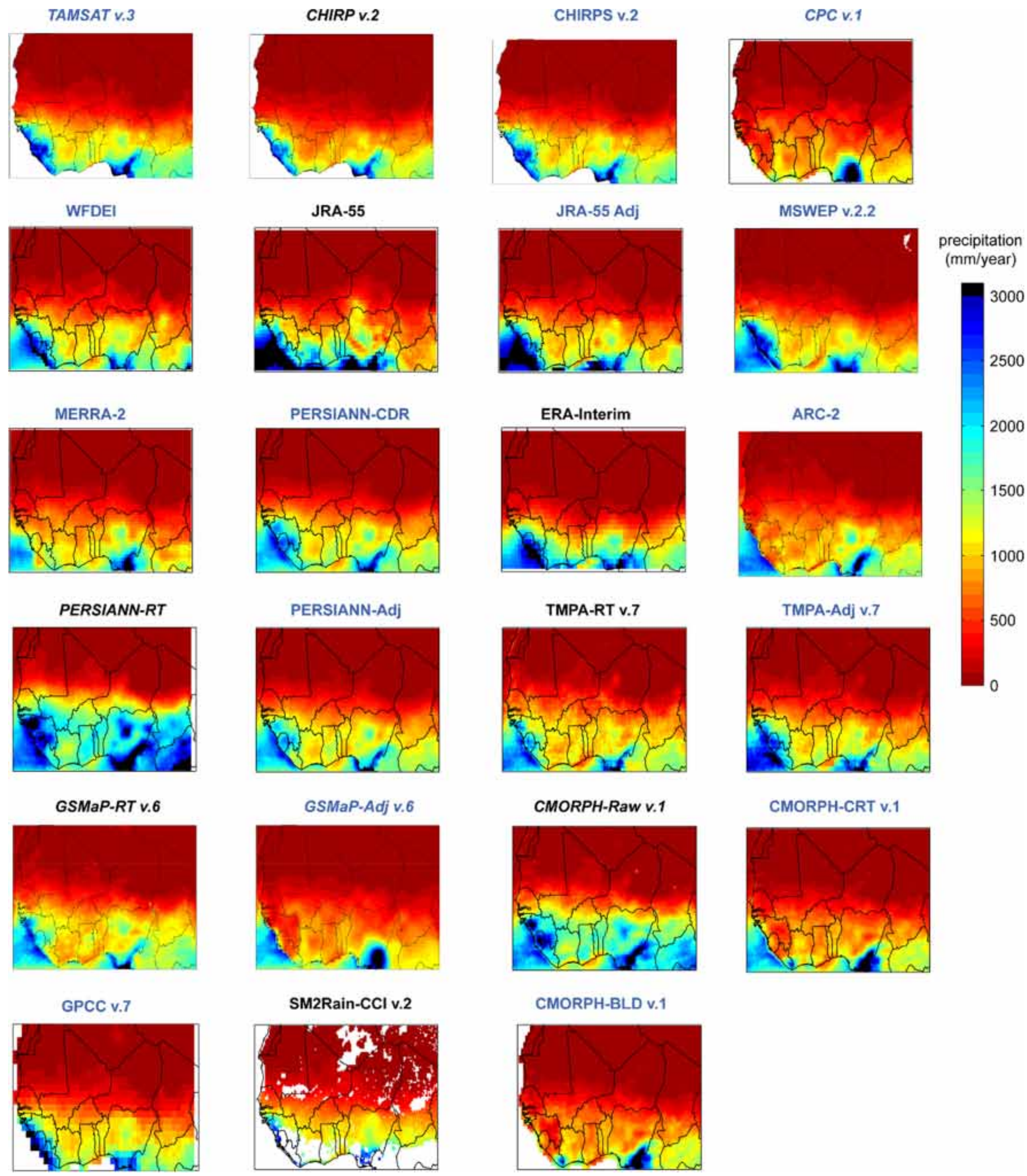

Fig. 2. Mean annual precipitation for 2000–2003 retrieved from all P-datasets at their original grid sizes. For each P-dataset, only the grid-cells with more than 80% of available daily data were retained. Blue and black colours are used to highlight P-datasets using and not using gauge-based information, respectively, and italic font is used for P-datasets available in NRT latency of one–three days. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

their availability. The P-datasets are generally available in (i) a few days or (ii) a few months after the observation (Table 1). Some are in near real time (NRT) latency of one to three days and are more adapted for flood or landslides forecasting, water resource management, and agriculture, while the others are more adapted for retrospective climatic studies.

Fig. 2 shows the mean annual precipitation patterns retrieved from all P-datasets. Except for CPC v.1 and the P-datasets, which use CPC v.1 for post adjustment processing (ARC-2, CMORPH-CRT v.1, CMORPH-BLD v.1, GSMaP-Adj v.6 and MERRA-2), all P-datasets represent the typical south–north precipitation gradient with two precipitation hotspots located over the southwest and south region. It should be noted that SM2Rain-CCI v.2 estimates are based on soil moisture estimates, which are strongly attenuated by the vegetation canopy; this results in significant gaps over areas with moderate to dense vegetation, as observed over the southern region (Fig. 2) (Dorigo et al., 2015). Additionally, a sensor failure in the ERS-2 gyroscope from January 2001 to June 2003 accentuated these gaps and explains the gaps observed over the central and northern regions (Fig. 2) (Dorigo et al., 2015).

2.2.2. P-dataset pre-processing

The P-datasets available at a sub-daily time step (Table 1) were aggregated to obtain daily time step records matching the local gauge observations (8 h to 8 h local time). It is worth mentioning that P-datasets delivered at daily time scales (Table 1) use time-windows different from those of the gauge, which can compromise the comparison

Table 2
Contingency table used to define HSS.

		Rain gauges	
		Precipitation	No precipitation
P-datasets	Precipitation	a	b
	No precipitation	c	d

at the daily scale. This point is further discussed in Section 4.1. Moreover, the P-datasets differ in terms of grid-cell size, ranging from 0.0375° for TAMSAT v.3 to 1° for GPCC v.7. To enable consistent comparison, all P-datasets were resampled to the 0.1° grid-cell size.

Bilinear averaging (interpolation) are used for P-datasets with grid-cell size $< 0.1°$ ($> 0.1°$) (Beck et al., 2019).

3. Reference precipitation dataset

A database of 1,440 gauges were made available by several African national meteorological and hydrological centres. The stations are distributed onto 952 0.1° grid-cells. For each grid-cell, a reference daily precipitation series was obtained averaging the gauges included in the grid-cell. To ensure consistent analysis, only grid-cells with more than 80% of daily records were considered. The four-year period of 2000–2003 was finally retained to consider the largest number of 0.1° grid-cells (187).

3.1. Monthly P-dataset estimate assessment

The monthly amounts were computed for only months with more than 80% of common daily records for all datasets (reference and P-datasets). The accuracy of monthly P-dataset estimates was assessed using a quantitative statistical analysis based on the modified Kling–Gupta Efficiency (KGE), an objective function combining correlation (r), bias (β), and variability (γ) components (Gupta et al., 2009; Kling et al., 2012) (Eq. (1)). We used KGE because water resource management requires reliable representation of precipitation temporal dynamics (measured by r) and volume (measured by β and γ):

$$KGE = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2} \quad (1)$$

where r represents the Pearson coefficient (Eq. (2)), β is the ratio between the mean observed and predicted precipitation (Eq. (2)), and γ is the ratio of the estimated and observed coefficients of variation (eq. (3)):

$$r = \frac{1}{n}\sum_1^n \frac{(o_n - \mu_o) * (s_n - \mu_s)}{\sigma_o * \sigma_s}, \quad (2)$$

$$\beta = \frac{\mu_s}{\mu_o} \quad (3)$$

$$\gamma = \frac{\sigma_s/\mu_s}{\sigma_o/\mu_o} \quad (4)$$

where μ and σ are the distribution mean and standard deviation, respectively; and s and o indicate the estimate and reference, respectively. KGE, β, γ, and r have their optimum at unity.

The analysis was performed considering all months of 2000–2003 and the wet and dry seasons months separately. For each grid-cell, the wet and dry seasons were selected according to their corresponding climatic zone (Fig. 1). The Sahel seasonality was applied for the grid-cells located up to latitude 16° N.

The values of KGE, r, β, and γ were computed at each grid-cell location to observe the P-dataset reliability over space, and their median values was used to observe that at the regional scale.

Considering the important gaps over space and time for SM2Rain-CCI v.2 (Fig. 2), its performance analysis was based on a reduced number of 0.1° grid-cells (79). Finally, GPCC v.7 is only available at the monthly time step (Table 1). Consistent comparison between GPCC v.7 and the reference was possible only for grid-cells and months with complete daily observation series for a total of 183 grid-cells.

3.2. Daily P-dataset estimate assessment

The daily precipitation estimates were assessed on the basis of both quantitative and categorical statistical metrics. The quantitative analysis relied on the median KGE, r, β, and γ obtained from the 187 considered grid-cells for all days and the wet and dry seasons days, separately, and from the 79 (183) grid-cells for SM2Rain-CCI v.2 (GPCC v.7).

The categorical statistics were used to measure the P-dataset capacity for detecting the daily precipitation events. Daily precipitation events are considered as discrete values with only two observable cases: rainy or not rainy days. A rainy day was considered when the precipitation amount was greater than or equal to a prescribed threshold (mm.day^{-1}). Four cases were possible (Table 2). Based on this characterisation, the Heidke Skill Score (HSS) (Eq. (4)) evaluates the P-dataset ability for detecting precipitation events in comparison with a random based prediction.

$$HSS = \frac{2 * (a * d - b * c)}{[(a + c) * (c + d) + (a + b) * (b + d)]} \quad (5)$$

The HSS values range from $-\infty$ to 1 with a perfect score of 1 and negative values indicating that random based prediction outperforms the P-dataset one.

The mean HSS value was computed from those obtained for all of the considered grid-cells for threshold values ranging from 0 to 25 mm.day^{-1} with a 1 mm.day^{-1} increment. This consideration was used to assess the P-dataset performance based on light to heavy daily precipitation events. Finally, using a 1 mm.day^{-1}, the HSS value was computed at each grid-cell location to observe the P-dataset reliability over space.

4. Results

4.1. P-dataset assessment at the monthly time step

With negative KGE values, three P-datasets (CMORPH-Raw v.1, TMPA-RT v.7, and PERSIANN-RT), were unable to represent the regional monthly precipitation (Fig. 3). Interestingly, their adjusted versions, CMORPH-BLD v.1, TMPA-Adj v.7, and PERSIANN-CDR, respectively, performed much better with KGE greater than 0.8, correlation better than 0.9, and bias and variability close to the optimum values (1). The same results were shown for CHIRP v.2, GSMaP-RT v.6, and JRA-55, which were systematically outperformed by their corresponding adjusted versions (CHIRPS v.2, GSMaP-Adj v.6, and JRA-55 Adj, respectively). In a general way, all P-datasets using gauges-based information present higher KGE than the others. The P-datasets developed for the African continent, TAMSAT v.3 and ARC-2, did not outperform the global scale P-datasets. However, the TAMSAT v.3 reliability was very close to that of the other P-datasets (KGE = 0.8).

The P-dataset performance expressed as KGE varied seasonally. TAMSAT v.3, JRA-55 Adj, PERSIANN-Adj, ARC-2, GSMaP-RT v.6, and GPCC v.7 were more effective during the wet season, and CMORPH-BLD v.1, MERRA-2, GSMaP-Adj v.6, CPC v.1, and ERA-Interim had better performance during the dry season. However, the most effective P-datasets, CHIRPS v.2, TMPA-Adj v.7, WFDEI, PERSIANN-CDR, and MSWEP v.2.2, performed similarly for both wet and dry seasons. Interestingly, all P-datasets presented higher correlation coefficient and bias values during the dry season. With respect to the variability ratio, no clear seasonal trend was observed for the different P-datasets.

Adjustment of CHIRP v.2, JRA-55, PERSIANN-RT, TMPA-RT v.7, GSMaP-RT v.6, and CMORPH-Raw v.1 increased the KGE values

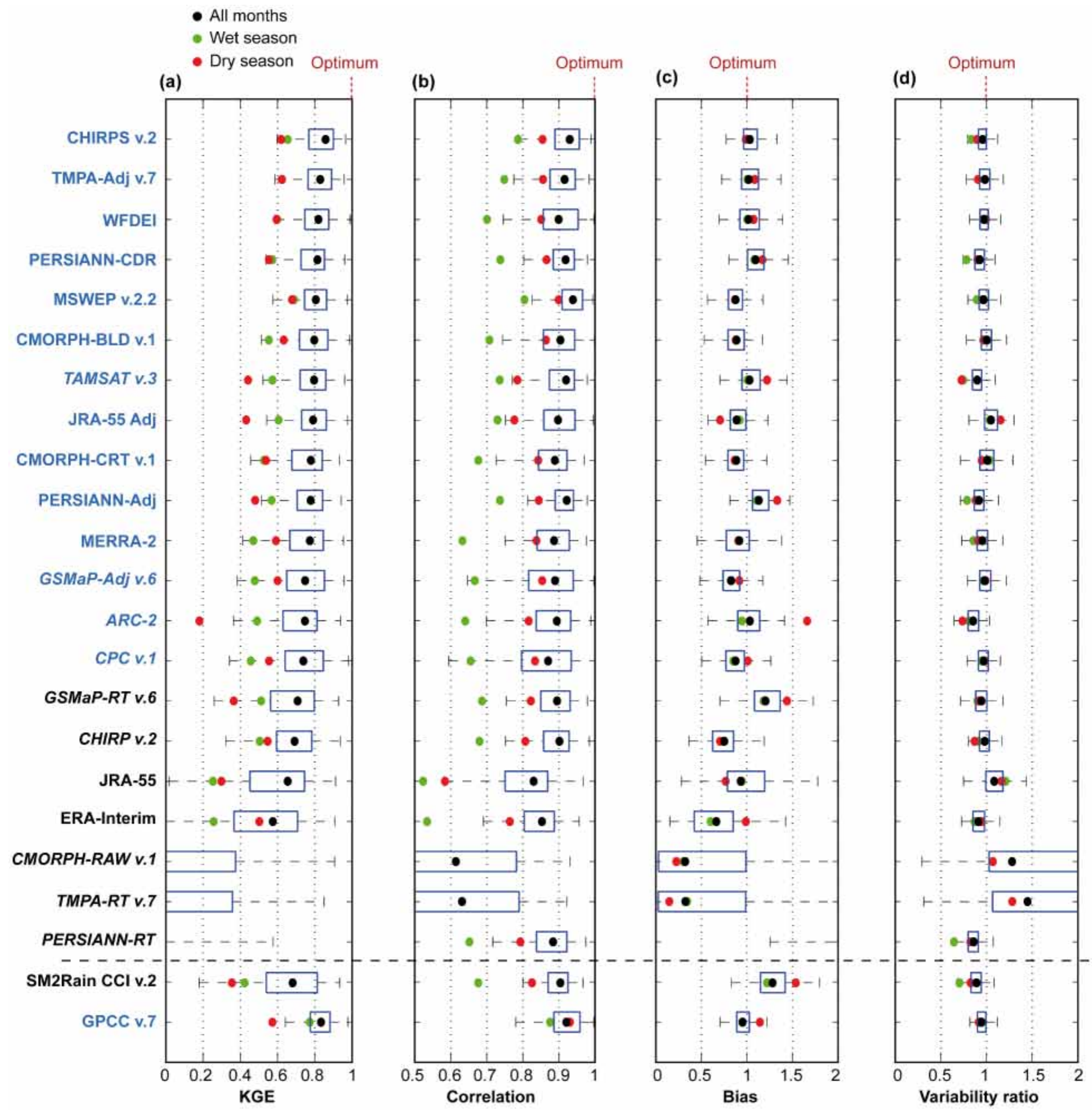

Fig. 3. P-dataset reliability at the regional. The right and left edges of the boxes represent the 25th and 75th percentile values, respectively. The P-datasets are sorted from the most (top) to the least (bottom) efficient in term of KGE. SM2Rain-CCI v.2 and GPCC v.7 are at the bottom because their analyses are based on a different number of 0.1° grid-cells, at 79 and 183, respectively. Blue and black colours are used to highlight P-datasets using or not using gauge-based information, respectively, and italic font is used for P-datasets available in NRT latency of one to three days. The graphics were inspired by Beck et al. (2019).

considerably at most of the grid-cell locations. The adjustment applied to GSMaP-Adj v.6 was not effective over the western region, where KGE values decreased in comparison to GSMaP-RT v.6, its non-adjusted version. Similarly, the CMORPH adjusted versions (CMORPH-CRT v.1 and BLD v.1) presented the lowest registered KGE values over the western region. CPC v.1, MERRA-2, and ARC-2 also presented the lowest KGE value over this region. Regarding the most effective P-datasets, CHIRPS v.2, GPCC v.7, WFDEI, and TMPA-Adj v.7 presented similar KGE distributions.

Most of the P-datasets were well correlated to the reference, with correlation better than 0.8 (Fig. 5). The adjusted version systematically presented higher correlation values, with MSWEP v.2.2 presenting the highest number of grid-cells with correlation better than 0.9 and only one grid-cell with correlation worse than 0.7. Interestingly, CHIRPS presented the lowest correlation score over the northern very arid region, with correlation worse than 0.7 (Fig. 5).

The P-datasets without gauge-based information presented higher bias (Fig. 6). PERSIANN-RT highly overestimate precipitation throughout the region (bias greater than 1.55). The bias decreased in the post-adjusted version (PERSIANN-Adj) with acceptable bias estimates ($1.15 < \text{bias} < 1.35$) over many grid-cells.

Similar results were observed for PERSIANN-CDR. CMORPH-Raw v.1 and TMPA-RT v.7 presented similar bias distributions, from overestimation to underestimation in the northern arid and southern humid regions, respectively. TMPA-RT gauge adjustment was highly successful, with most of the TMPA-Adj v.7 grid-cells presenting acceptable bias values at $0.85 < \text{bias} < 1.15$. CPC v.1 strongly underestimates precipitation over the western region. This bias spread for all P-datasets

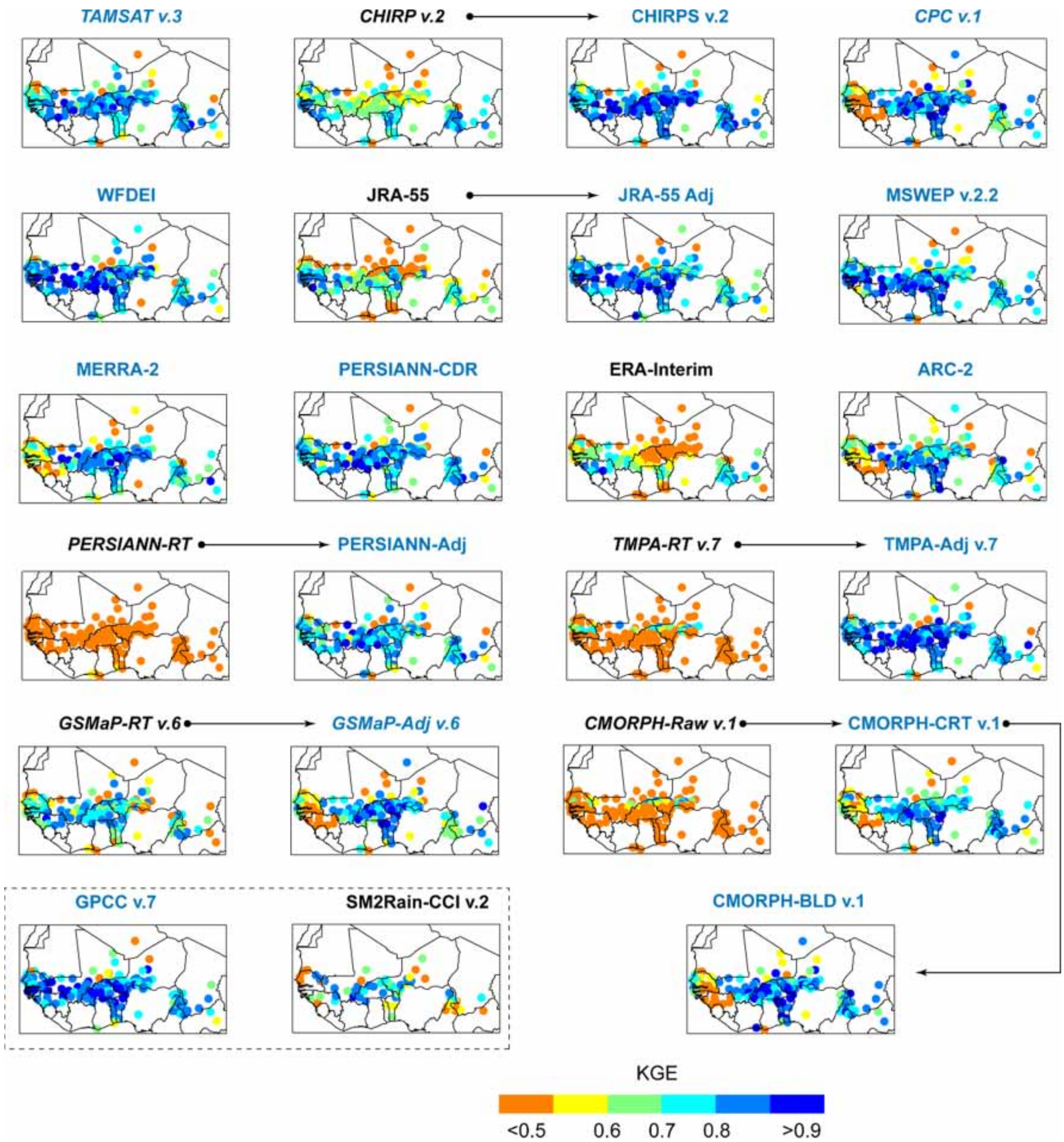

Fig. 4. P-dataset reliability at the grid-cell level expressed in the form of KGE considering all months in 2000–2003. Arrows are used to highlight the potential benefit of using gauge-based information. Blue and black colours are used to highlight P-datasets using or not using gauge-based information, respectively, and italic font is used for P-datasets available in NRT latency of one the three days. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

using CPC v.1 for their adjustement process (CMORPH-CRT and BLD v.1, ARC-2, GSMaP-Adj v.6 and MERRA-2).

Interestingly, the precipitation adjustment applied on GSMaP-RT v.6 increased the bias on GSMaP-Adj v.6. WFDEI, TMPA-Adj v.7, and CHIRPS v.2 presented less-biased precipitation estimates with reasonable bias values of 0.85 < bias < 1.15 in most of the considered grid-cells.

Regarding the variability ratio distribution, the efficiency of using gauge-based information to retrieve the precipitation estimates was obvious when comparing PERSIANN-RT, TMPA-RT v.7, and CMORPH-Raw v.1 with their post-adjusted versions (Fig. 7). The non-adjusted products CMORPH-RAW v.1, and TMPA-RT v.7 strongly overestimated the precipitation variability in the majority of grid-cells with variability ratios better than 1.25. To the contrary, PERSIANN-RT strongly underestimated the precipitation variability in most grid-cells, with a variability ratio worse than 0.85. However, when considering JRA-55 (JRA-55 Adj) and CHIRP v.2 (CHIRPS v.2), the use of gauge-based information did not significantly enhance the variability ratio. Finally, the two African P-datasets underestimated the precipitation variability, over most of the grid-cells (variability ratios < 0.90).

4.2. P-dataset assessment at the daily time step

At the regional scale, the ability of the P-datasets to quantify the daily precipitation amount was relatively low, with most having median KGE values worse than 0.4 (Fig. 8). Only MSWEP v.2.2, TMPA-Adj v.7, CMORPH-BLD v.1, CMORPH-CRT v.1, GSMaP-RT v.6, and PERSIANN-Adj had KGE scores superior to 0.4, with the best score achieved by MSWEP v.2.2 (KGE = 0.52). Generally, non-adjusted P-datasets presented the lowest KGE values except for GSMaP v.6. The KGE decreased from 0.44 (GSMaP-RT v.6) to 0.35 for (GSMaP-Adj v.6). Interestingly, PERSIANN-RT presented a negative KGE value but one of

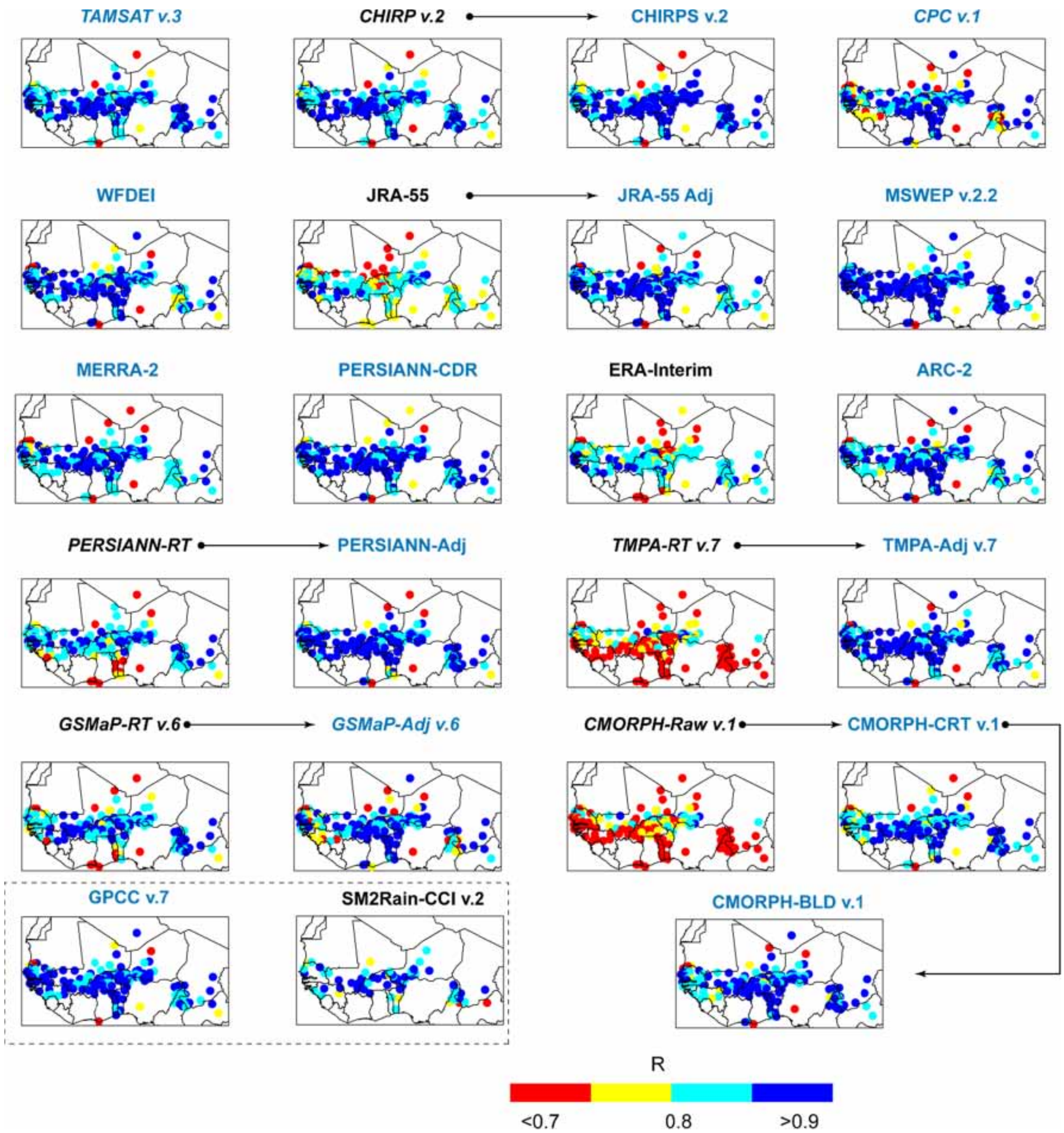

Fig. 5. P-dataset reliability at the grid-cell scale expressed in the form of correlation considering all months in 2000–2003. Arrows are used to highlight the potential benefit of using gauge-based information. Blue and black colours are used to highlight P-datasets using or not using gauge-based information, respectively, and italic font is used for P-datasets available in NRT latency of one to three days.

the highest correlation score, at 0.5. Therefore, its low KGE score appears to be influenced by its very high positive bias value of 2.5. This is in line with observation made at the monthly time step and the dominant influence of the bias values on the KGE score.

In term of KGE, the P-dataset accuracy was higher during the wet than that in the dry season. Interestingly, MERRA-2, WFDEI, ERA-Interim, and JRA-55 performed better during the dry season, which is line with the results obtained over the Continental United States (CONUS) (Beck et al., 2019). However, the performances of MERRA-2, WFDEI, ERA-Interim, and JRA-55 were very low, with KGE < 0.2.

Most of the P-datasets presented the highest HSS scores using a threshold value of 1 mm.day^{-1} (Fig. 9). In particular, the HSS values of CHIRP v.2 and MERRA-2 were close to 0 when considering a 0 mm.day^{-1} threshold value; the values jumped to 0.3 and 0.36, respectively, when considering a 1 mm.day^{-1} threshold value. Actually, the P-datasets detected many precipitation events with less than 1 mm.day^{-1} which were not detected by the gauges. This can be explained by different factors: (1) The gauges are not sensitive enough to such precipitation amounts; (2) difference in the spatial scale between point (gauge) and average area (P-dataset grid-cell) measurements; (3) the P-dataset algorithm is deficient. Because these precipitation events are insignificant ($<$ 1 mm.day-1), they should be considered as no-precipitation events.

The highest HSS score was achieved by CMORPH-BLD v.1 (HSS = 0.58) and MSWEP v.2.2 (HSS = 0.55). The P-dataset ability in reproducing daily precipitation amounts decreased for increasing intensity. Two P-dataset groups measured differently events of more than 15 mm.day^{-1}. The first group (CMORPH-CRT and BLD v.1, GSMaP-RT and Adj v.6, MSWEP v.2.2, PERSIANN-RT and Adj, ARC-2, and CPC and TMPA-Adj v.7) was much more suited for reproducing high-intensity precipitation events than the second group (CHIRP v.2, CHIRPS v.2, CMORPH-RAW v.1, JRA-55, JRA-55 Adj, PERSIANN-CDR, TAMSAT v.3, TMPA-RT v.7, WFDEI, MERRA-2, ERA-Interim, and SM2Rain-CCI v.2) (Fig. 9). It is worth mentioning that the first group includes (i) P-datasets with gauge-based calibration using daily data and (ii) P-datasets available at the sub-daily time step. The second group includes (i)

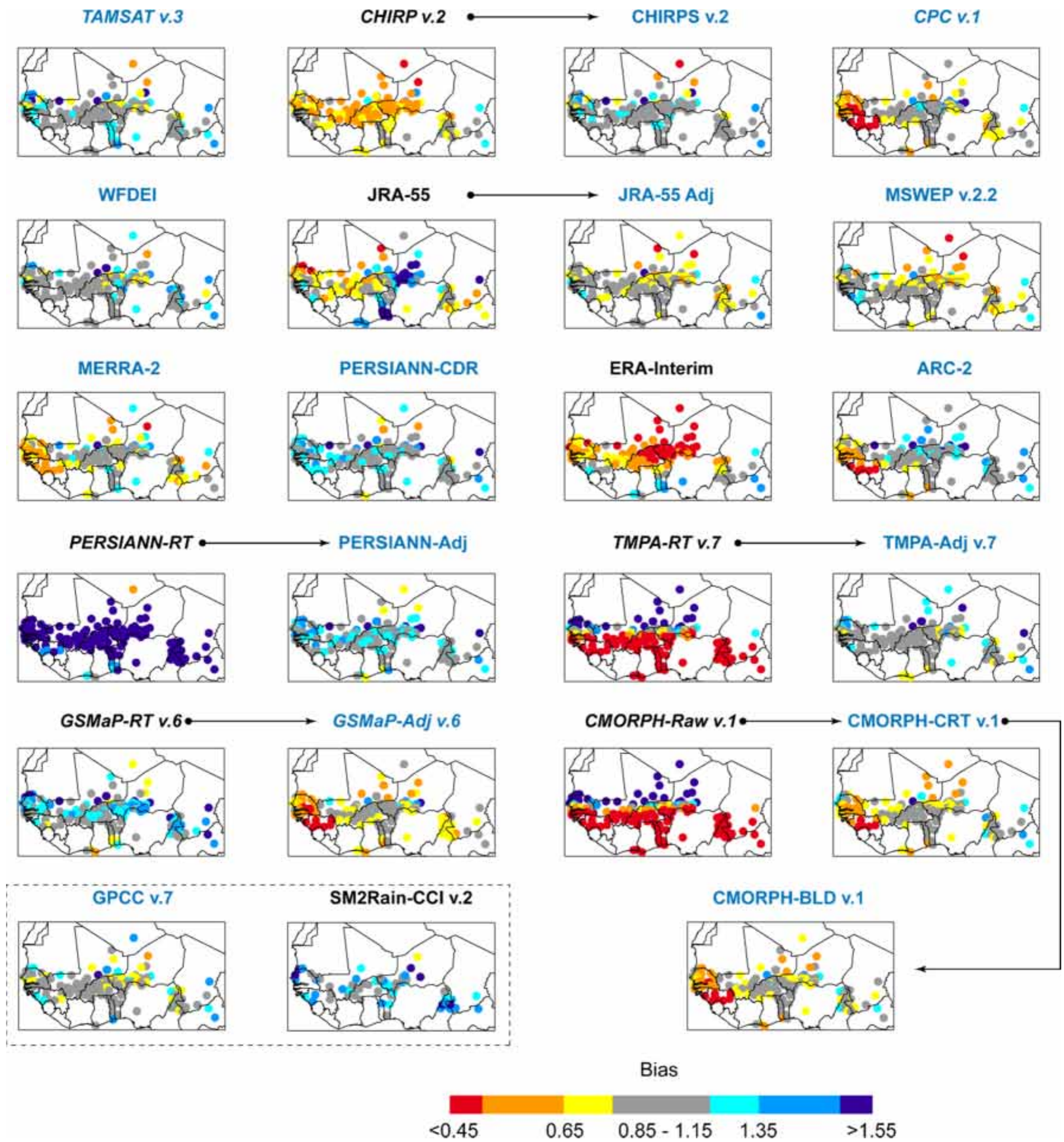

Fig. 6. P-dataset reliability at the grid-cell scale expressed in the form of bias considering all months for in 2000–2003. Arrows are used to highlight the potential benefit of using gauge-based information. Blue and black colours are used to highlight P-datasets using or not using gauge-based information, respectively, and italic font is used for P-datasets available in NRT latency of one to three days. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

non-adjusted P-datasets or (ii) those adjusted with monthly gauges-based data, (iii) P-datasets delivered at the daily time step, and (iv) reanalysis P-datasets which generally have the largest discrepancies when compared with other P-datasets (Sun et al., 2018a). Therefore, the gauge-based information used for P-datasets and the delivered time step (daily or sub-daily) considerably influence the P-dataset reliability at the daily time scale. This point is further discussed in Section 4 in the Discussion.

Using gauge-based information improved the HSS score over space. For instance, TMPA-RT v.7, CMORPH-Raw v.1, CHIRP v.2 and PERSIANN-RT adjusted versions provided much better HSS scores throughout the region (Fig. 10). The adjusted versions of JRA-55, and GSMaP-RT v.6 did not show significant enhancement. Overall, the first group identified in Fig. 9 (CMORPH-CRT and BLD v.1, GSMaP-RT and Adj v.6, MSWEP v.2.2, PERSIANN-RT and Adj, ARC-2, CPC and TMPA-Adj v.7) presents the highest HSS all over West Africa (Fig. 10).

5. Discussion

5.1. Monthly versus daily P-dataset reliability

Interestingly, the P-dataset performance ranking differed at the monthly and daily timescale (Fig. 11). We identified two main factors to explain these discrepancies.

The first is related to the gauge-based information time step used to adjust P-dataset estimates. Indeed, the five most efficient P-datasets at the monthly time step, CHIRPS v.2, TMPA-Adj v.7, WFDEI, PERSIANN-CDR, and MSWEP v2.2, were adjusted using monthly gauge-based information, whereas three of the five most efficient P-datasets at the daily time step, MSWEP v.2.2, CMORPH-BLD v.1, and CMORPH-CRT v.1, were adjusted using daily gauge-based information (Fig. 11). Additionally, the reliability of the gauge-based information can also influence the P-dataset accuracy. Accordingly, over the Sahel region, the

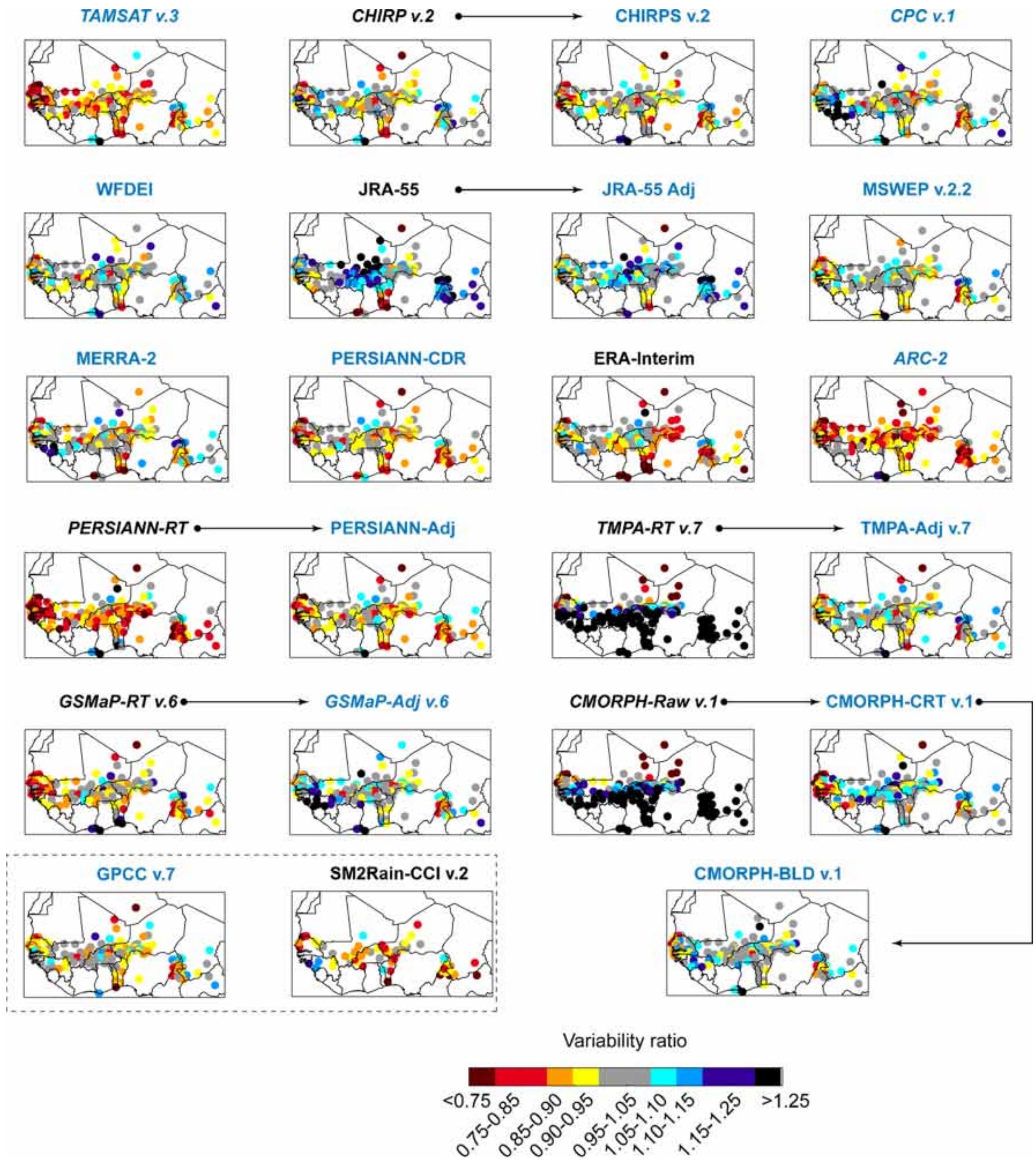

Fig. 7. P-dataset reliability at the grid-cell scale expressed in the form of variability ratio considering all months for in 2000–2003. Arrows are used to highlight the potential benefit of using gauge-based information. Blue and black colours are used to highlight P-datasets using or not using gauge-based information, respectively, and italic font is used for P-datasets available in NRT latency of one to three days. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

CHPclim monthly dataset reliability was higher than that of GPCC (Funk et al., 2015). Because most of the considered grid-cells used to assess P-dataset reliability are in the Sahel region (Fig. 1), CHIRPS, which uses CHPclim, provides more realistic monthly precipitation estimates, at KGE = 0.86, than WFDEI and PERSIANN-CDR, which uses GPCC and GPCP, respectively (Fig. 3). This demonstrates the importance of maintaining reliable gauge networks to insure accurate P-dataset estimates.

The second factor is the P-datasets delivered time step. Some P-datasets are delivered at the daily aggregation level (Table 1), which is based on different time windows than those used for local records. For example, PERSIANN-CDR daily estimates correspond to a given 0 h to 0 h UTC aggregation time period, whereas the gauges used in this study register daily amount from 8 h to 8 h UTC. Such temporal inconsistency can introduce large differences between the P-datasets and the gauge measurements (Ashouri et al., 2015; Satgé et al., 2019). Therefore, only one of the P-datasets, CMORPH-BLD v.1, delivered at the daily time step ranked in the top five most efficient P-datasets. On the contrary, four of the five most efficient P-datasets at the daily time step were delivered at the sub-daily time step (3-hourly to hourly) (Table 1 and Fig. 11). The sub-daily time step enables matching of the computed daily estimates with the local record time windows to ensure consistent comparison between gauges and P-dataset estimates.

Our results demonstrate the importance of considering both daily and monthly time steps when assessing P-dataset reliability because the latter is influenced by the gauge-based adjustment process and the delivered time step.

Generally, P-datasets using gauge-based information achieved

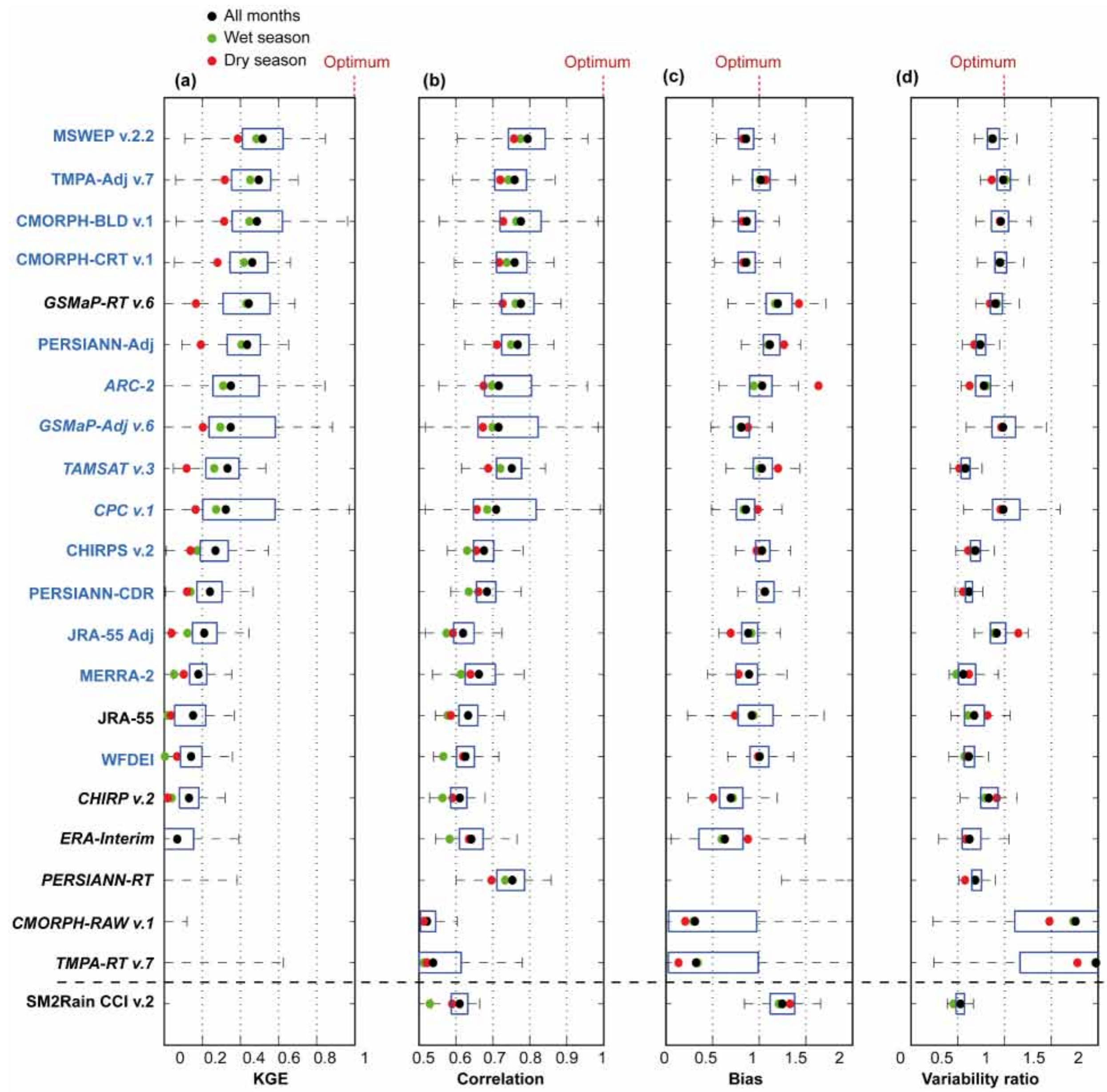

Fig. 8. P-dataset reliability at the regional scale. The right and left edges of the box represent the 25th and 75th percentile values, respectively. The P-datasets are sorted from the most (top) to the least (bottom) efficient in terms of KGE. SM2Rain-CCI v.2 and GPCC v.7 are at the bottom because their analyses are based on a different number of 0.1° grid-cells, at 79 and 183, respectively. Blue and black colours are used to highlight P-datasets using or not using gauge-based information, respectively, and italic font is used for P-datasets available in NRT latency of one to three days. The graphics were inspired by Beck et al. (2019).

highest KGE scores at both monthly and daily time step (Figs. 3 and 8), which supports previous results obtained over West African regions (Casse et al., 2015; Gosset et al., 2013; Poméon et al., 2017) and elsewhere (Beck et al., 2019; Dinku et al., 2007; Satgé et al., 2016). However, the use of gauge-based information for P-dataset adjustment is not always as effective. Indeed, GSMaP-RT v.6 outperformed its adjusted version GSMaP-Adj v.6 at the daily time step. This result is consistent with previous observation over the CONUS, (Beck et al., 2019) and illustrates the potential P-dataset algorithm limit to consider the best of gauge data. GSMaP-RT v.6 is the only P-dataset with no gauge-based information of the top-five daily ranking. Therefore, GSMaP-Adj v.6 should be highly effective if using the gauge-based information in the optimal form.

5.2. *P-dataset reliability in space and time*

CMORPH-BLD v.1, CMORPH-CRT v.1, GSMaP-Adj v.6, and MERRA-2 presented weaker performance over the western region in comparison with other P-datasets (Fig. 3). We identified one factor to explain this spatial inconsistency.

Different from other P-datasets which use GPCC or GPCP data, CMORPH-BLD v.1, GSMaP-Adj v.6, and MERRA-2 use CPC data. The gauge number used to retrieve CPC is lower than that used to retrieve GPCC (Fig. 12a). Over the Senegal, Gambia, Guinea-Bissau, and Guinea regions only two CPC grid-cells counted with more than one gauge against seven for GPCC (Fig. 12b and c). As a result, compared with GPCC v.7, CPC v.1 presents the lowest efficiency over the western region (Fig. 4), which propagates into the use of CPC by the P-datasets to adjust their estimates.

The available gauge information for retrieving CPC and GPCC

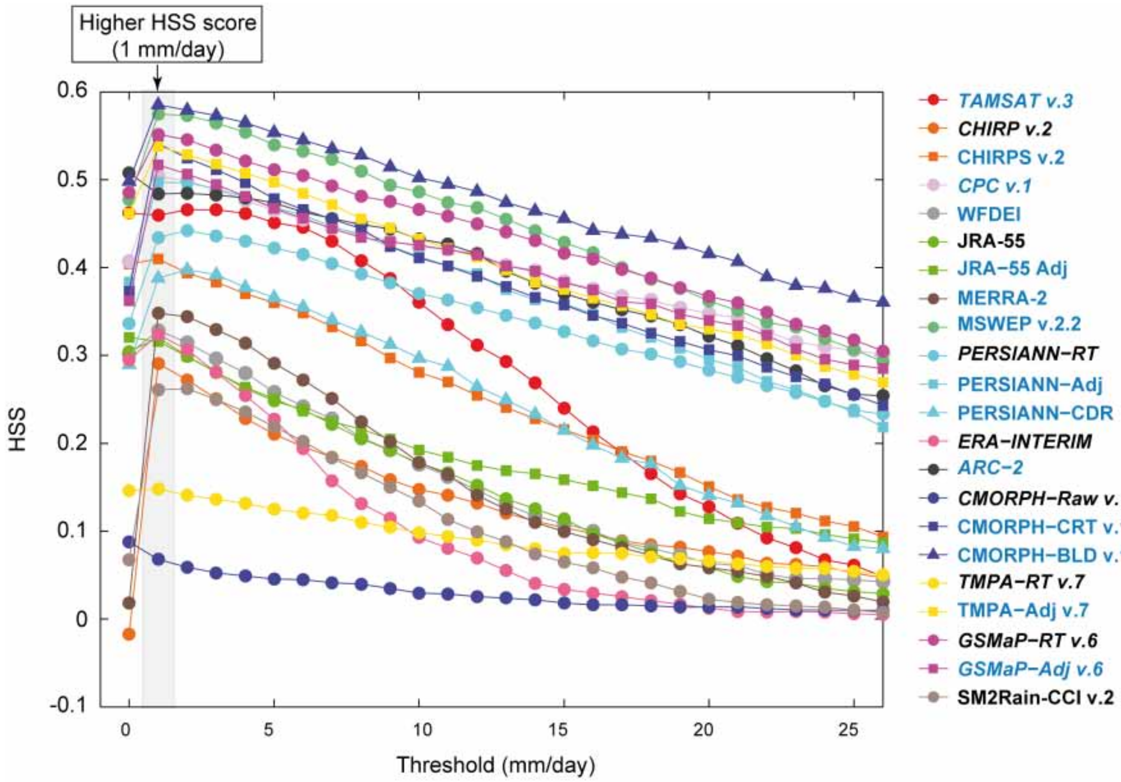

Fig. 9. P-dataset ability in reproducing daily precipitation events of different intensities expressed in the form of HSS. Blue and black colours are used to highlight P-datasets using or not using gauge-based information, respectively, and italic font is used for P-dataset available in NRT latency of one to three days.

datasets also varies with time (Fig. 12a). Therefore, the P-dataset reliability could be better (worse) if considering a period with more (fewer) available gauges for retrieving GPCC or CPC datasets. In this context, TAMSAT v.3 uses consistent gauge-based information in space and time rather than a continuously updated information to avoid adding any space or time discrepancy (Maidment et al., 2017).

Actually, P-datasets present space and time inconsistencies which cannot be reported by using single temporal windows (Satgé et al., 2019). The authors assessed P-dataset reliability over three different four years period and one twelve years period across the Lake Titicaca region. Results show that the P-dataset reliability conclusion vary according to the considered period. Therefore, the analysis should be conducted over different temporal windows to adequately evaluate the P-dataset space and time reliability. Such a consideration is challenging over West Africa owing to the scarcity of gauge networks and the important temporal gaps present. To overcome this issue, an alternative method could use satellite-based soil moisture estimates rather than traditional rain gauges measurements as a reference benchmark (Massari et al., 2017).

It is worth mentioning that GPCC and CPC share common gauges with the reference network used in this study as highlighted by many overlapping between both network (Fig. 12b and c). Similar observation should be done if considering the others gauges based datasets used for P-dataset calibration and presented in Section 2.2.1. Therefore, the gauges network used for the assessment is not totally independent of the considered P-datasets and could influence P-dataset reliability conclusions. The P-dataset reliability conclusion could have been less optimist if only based on independent gauges network. In this context, future studies should try to consider totally independent gauges network to provide more consistent feedback on actual P-dataset reliability. However, information on the shared information between national gauges networks and gauges based dataset (i.e. CPC and GPCC) is hard to obtain and compromise this kind of consideration.

5.3. P-datasets sensitivity to seasonal variation

Reanalysis P-datasets, ERA-Interim, MERRA-2, JRA-55 and WFDEI, performed better during the dry than during the wet season (Fig. 3). This agrees with previous results obtained over the CONUS (Beck et al., 2019). The authors explained that reanalysis P-datasets are better adapted to detecting large-scale stratiform systems, which are typical in the dry season, than unpredictable small-scale convective cells, which are typical in the wet season. On the contrary, only satellite-based P-datasets performed better during the wet than the dry season (Beck et al., 2019; Salles et al., 2019; Satgé et al., 2017b). Actually, the irregular sampling of the low earth orbiting satellites and the limited number of overpasses hardly captures short precipitation events which are typical during the dry season (Gebregiorgis and Hossain, 2013; Tian et al., 2009). Therefore, GSMaP-RT v.6 presented a better KGE value during the wet than that during the dry season (Fig. 3). The seasonality sensitivity of the other P-datasets incorporating satellite, reanalysis, or gauge-based information shows a greater contrast because they consider the different inputs.

Despite the seasonal variation in KGE value, the P-datasets presented significantly higher coefficient correlation during the dry season (Fig. 3). This difference could be related to the higher monthly precipitation variability during the dry season (Fig. 1) tending to increase the correlation coefficient. Accordingly, all P-datasets presented higher correlation coefficients considering the entire period because the precipitation variability is even more marked than at the seasonal scale. At the contrary, the P-datasets were more biased during the dry season (Fig. 3) except for CPC v.1, GSMaP-Adj v.6, and ERA-Interim. The P-datasets with higher bias values during the dry season (TAMSAT v.3, PERSIANN-Adj, ARC-2, GSMaP-RT v.6, SM2Rain-CCI v.2, JRA-55 Adj, and GPCC v.7) or the wet season (CPC v.1, GSMaP-Adj v.6, and ERA-Interim) presented lower KGE scores during this specific season whereas the P-datasets with close bias values for both wet and dry seasons (CHIRPS v.2, TMPA-Adj v.7, WFDEI, PERSIANN-CDR, MSWEP

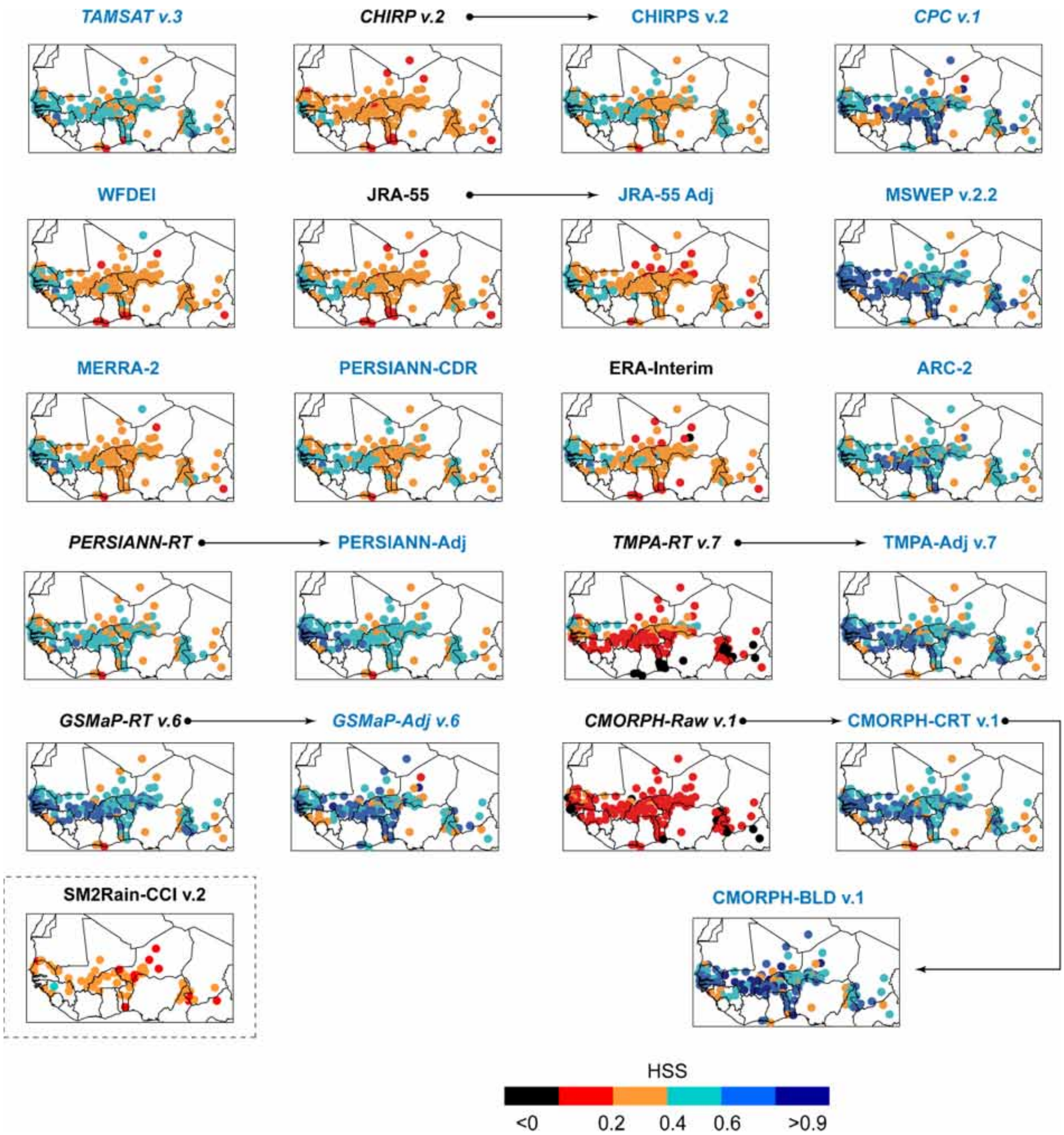

Fig. 10. Daily P-dataset reliability expressed in the form of HSS considering all days for 2000–2003 period. The HSS was obtained for a threshold value of 1 mm/month. Arrows are used to highlight the potential benefit of using gauge-based information. Blue and black colours are used to highlight P-datasets using or not using gauge-based information, respectively, and italic font is used for P-datasets available in NRT latency of one to three days.

v2.2, CMORPH-CRT v.1, CHIRP v.2) presented similar KGE values for both seasons. Considering the similar seasonal trend observed for both KGE and bias values, the bias appears to have a dominant influence on the KGE score.

Interestingly, if considering the P-dataset bias in millimetres, all P-datasets had systematically higher bias values during the wet season (Fig. 13). Because lower monthly precipitation occurred during the dry season, the same volumetric error (millimetres) expressed in ratio (Eq. 2) corresponds to higher bias during the dry in comparison to that during the wet season. For most of the P-datasets, the reported bias value during the dry season was < 5 mm.month^{-1} (Fig. 13), which should have an insignificant influence on the water budget.

Therefore, despite the low KGE value during the dry season, P-datasets still provide valuable additional information to follow both temporal and volume monthly precipitation dynamics over West Africa.

5.4. P-dataset time latency

Fig. 14 shows the KGE scores of the NRT P-datasets in comparison with the most accurate P-dataset at the daily time step.

At the daily time step, GSMaP-RT v.6 was the most reliable NRT P-dataset. With three days of time latency, the GSMaP-RT KGE, at 0.44, was close to the most effective P-dataset (MSWEP v.2.2) with KGE = 0.52, which is available with a few months of latency.

It is worth mentioning that the low score achieved by the P-datasets at the daily time step is partly related to the difference between spatially averaged (P-dataset grid-cell) and point (reference gauges) measurements (Salles et al., 2019; Satgé et al., 2019; Tang et al., 2018). The P-dataset reliability increased with the number of gauges used to represent the spatially average grid-cell measurement (Salles et al., 2019; Tang et al., 2018). In this study, most of the considered 0.1° grid-cells

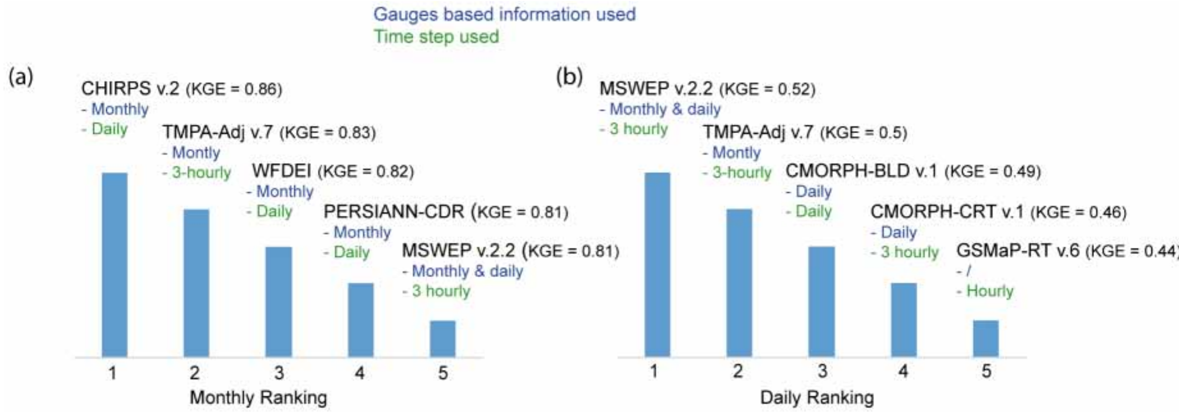

Fig. 11. P-dataset top-five ranking for the (a) monthly and (b) daily precipitation estimates based on their median KGE value.

were counted with only one gauge. Therefore, the presented KGE score may underestimate the actual P-dataset reliability. Testing the sensitivity of streamflow modelling to P-datasets at basin outlets overcome the influence of scarce and unevenly distributed gauge networks. Indeed, aggregation of precipitation at the basin scale eliminates the difference in spatial representation between point (gauge) and areal (P-datasets) measurements because both gauge and P-datasets represent precipitation at the basin spatial scale. Therefore, the reliability of P-datasets varies significantly when used to reproduce gauge precipitation estimates or streamflow observations (Satgé et al. 2019).

In this context, NRT P-dataset coverage and latency can fit the needs of an early warning system across sparsely gauged or ungauged regions. Recent studies have successfully used NRT P-datasets to follow flood events in terms of streamflow (Yuan et al., 2019) and flood extent (Belabib et al., 2019) or for landslide occurrence estimations (Brunetti et al., 2018). Future studies should assess NRT P-datasets in the scope of early warning studies to consistently evaluate NRT dataset reliability in this specific context.

5.5. Towards an enhanced P-dataset over West Africa

This study considers an unprecedented sample of 23 P-datasets over the West African region to provide a consistent guideline for potential users. The results suggest that during 2000–2003, CHIRPS v.2 and MSWEP v.2.2 showed the best estimates of monthly and daily precipitation, respectively. The most reliable P-dataset can change at the local scale. As an example, Fig. 15 shows the most suitable P-datasets for representing both monthly and daily precipitation at the grid-cell level. Interestingly, at the daily time step, MSWEP v.2.2 was more consistent for the western region, whereas CMORPH P-datasets provided more accurate estimates over the central and southern regions (Fig. 15). At the monthly time step, even if CHIRP(S) P-datasets are counted with the highest number of grid-cells, large spatial heterogeneity is observed with many grid-cells where WFDEI, JRA-55 Adj, CMORPH, and TMPA outperformed CHIRP(S) (Fig. 15). To take advantage of all available P-datasets, merging all P-datasets to produce an enhanced P-dataset over the region is a good option. Previous studies have reported on the benefit of such an approach to retrieve a more realistic P-dataset over Pakistan (Muhammad et al., 2018; Rahman et al., 2018), Tibet (Ma et al., 2018) and different tropical complex terrain (Bhuiyan et al., 2019). These ensemble precipitation datasets enhance the regional precipitation representation and should be used as guideline over West Africa.

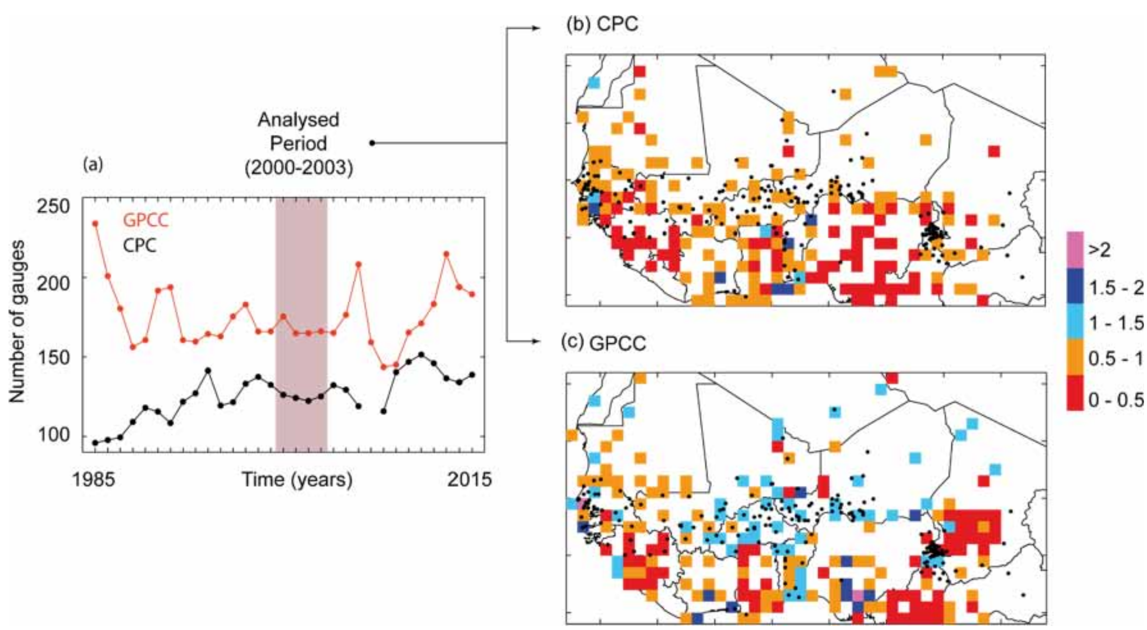

Fig. 12. (a) Mean numbers of available gauges used to retrieve GPCC and CPC for 1985–2015 and their spatial distribution for the analysed period 2000–2003 (b, c). The black points in (b) and (c) represent the centroid of the 0.1° grid-cell considered in this study to assess P-dataset reliability.

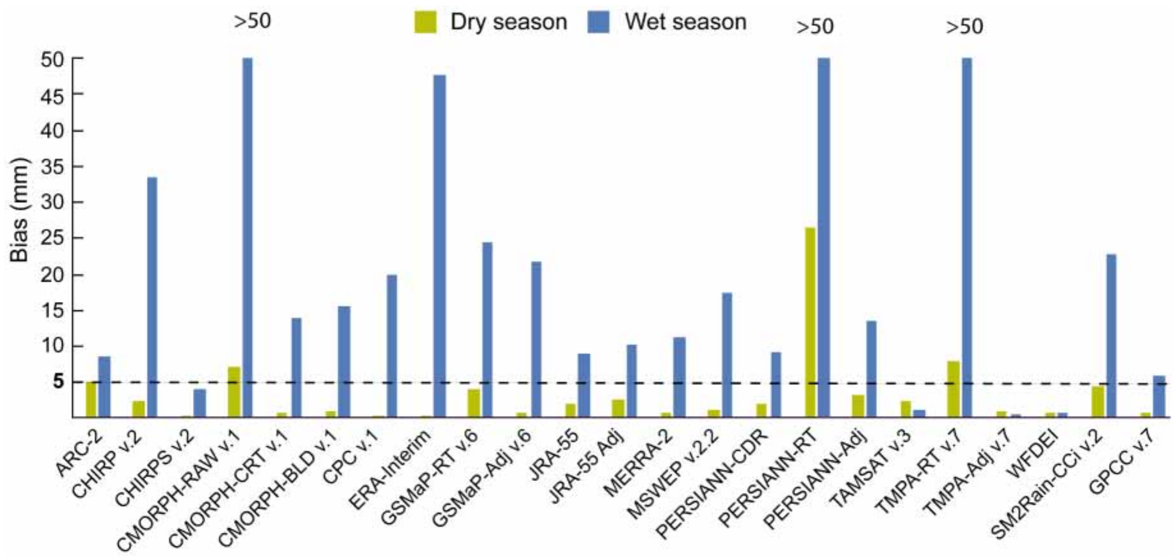

Fig. 13. Monthly bias value expressed in millimetres for both dry and wet seasons. The values are expressed in terms of absolute values. To facilitate the analysis, bias values greater than 50 mm are not shown.

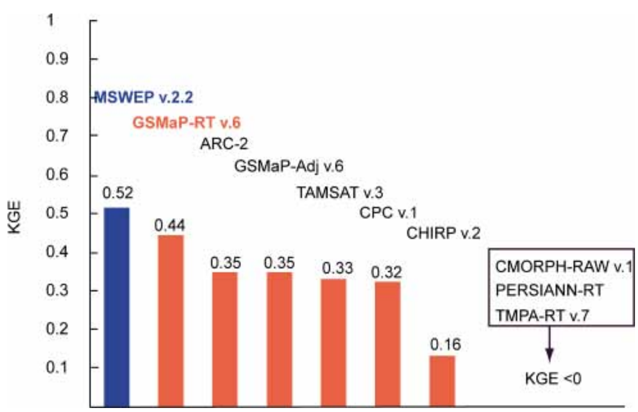

Fig. 14. NRT P-dataset reliability at both daily time steps in comparison to the most effective P-datasets, represented in blue.

6. Conclusions

The present study evaluates the accuracy of 23 gridded P-datasets over the West African region at both monthly and daily time step for the 2000–2003 period. Despite the limited coverage and scarcity of the ground reference points, some consistent features emerged from the analysis:

- The P-dataset performance ranking differs at the monthly and daily timescale. P-datasets using sub-daily (monthly) gauge information perform better at the daily (monthly) time step. Additionally, for the P-datasets released at the daily time step, the temporal mismatch between gauge and satellite reporting times decrease their reliability at the daily time step. In this line, MSWEP v.2.2 and CHIRPS v.2 provide the most reliable daily and monthly precipitation estimates, respectively whereas TMPA-Adj v.7 performance is very good for both daily and monthly estimates.
- The only satellite based P-datasets (CMORPH-RT v.1, TMPA-RT v.7, PERSIANN-RT, GSMaP v.6-RT) performance is very low at both monthly and daily time scale. Their reliability drastically increase for their adjusted versions (CMORPH-CRT and BLD v.1, TMPA-Adj v.7, PERSIANN-Adj) excepted for GSMaP v.6 at the daily time step.
- All the considered reanalysis P-datasets (WFDEI, JRA-55, JRA-55 Adj, ERA-Interim) are unreliable at the daily time step. The use of monthly GPCC P-dataset to adjust their estimates considerably increase their reliability at the monthly time step (WFDEI, JRA-55 Adj).
- The two African P-datasets (TAMSAT v.3 and ARC-2) present an overall lower performance in comparison to the almost global scale P-datasets at both daily and monthly time-step. Despite good performance in some parts of the region, SM2Rain-CCI v.2 still suffers too many gaps in space and time across West African.
- All P-datasets present spatial discrepancies in their statistical score suggesting the use of a spatial P-datasets' merging approach to take advantage from all available P-datasets across West Africa.

It should be reminded that most of the considered 0.1° grid-cells count with only one gauge to represent the observed precipitation. Because of spatial inconsistency between point (gauges) and spatially average (P-datasets) measurement, different conclusion regarding the P-datasets reliability, could have been drawn if more gauges had been available per grid-cells or if using P-datasets as forcing data for hydrological modelling. Additionally, the study is based on a single four years temporal window. However, P-dataset reliability vary in time and the results could have been different if considering another four years temporal window or a larger one. Therefore, this study aims more at compare the P-dataset reliability between them rather than to provide definitive conclusion on their respective accuracy.

CRediT authorship contribution statement

Frédéric Satgé: Conceptualization, Methodology, Software, Formal analysis, Investigation, Project administration, Writing - original draft. **Dimitri Defrance:** Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - review & editing. **Benjamin Sultan:** Investigation, Writing - review & editing. **Marie-Paule Bonnet:** Funding acquisition, Investigation, Writing - review & editing. **Frédérique Seyler:** Funding acquisition, Investigation, Writing - review & editing. **Nathalie Rouché:** Resources, Writing - review & editing. **Fabrice Pierron:** Resources, Writing - review & editing. **Jean-Emmanuel Paturel:** Funding acquisition, Resources, Investigation, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial

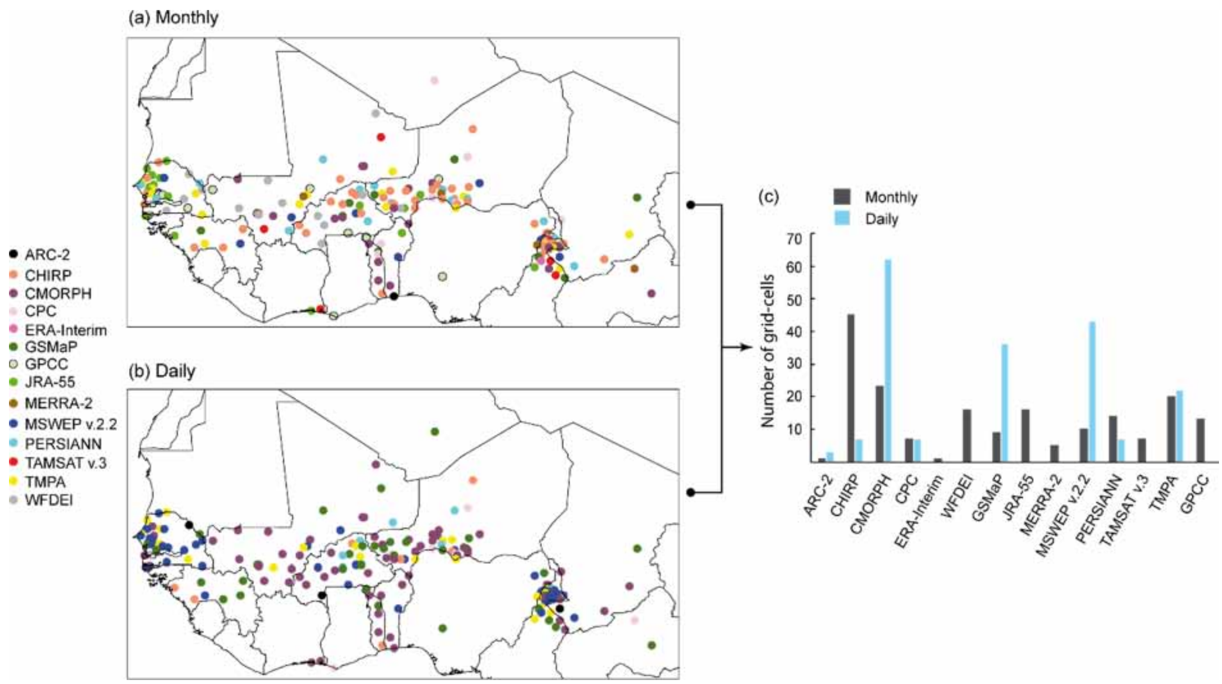

Fig. 15. Most efficient P-datasets at the grid-cell level. For simplification, the P-datasets were aggregated in main groups: GSMaP = GSMaP-RT + Adj v.6; TMPA = TMPA-RT + TMPA-Adj; JRA-55 = JRA-55 + JRA-55 Adj; PERSIANN = PERSIANN-RT + PERSIANN-Adj + PERSIANN-CDR.

interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abiodun, B.J., Adeyewa, Z.D., Ajayi, V.O., 2012. Modeling the impacts of reforestation on future climate in West Africa, 77–96, doi:10.1007/s00704-012-0614-1.

Adler, R.F., Huffman, G.J., Chang, A., Ferraro, F., Xie, P.-P., Janowiak, J., Rudolff, B., Schneider, U., Curtis, S., Bolvin, D., Gruber, A., Susskind, J., Arkin, P., Nelkin, E., 2003. The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979 – Present). J. Hydrometeorol. 1147–1167, 2003.

Adler, R.F., Gu, G., Huffman, G.J., 2012. Estimating climatological bias errors for the global precipitation climatology project (GPCP). J. Appl. Meteorol. Climatol. 84–99, 2012. https://doi.org/10.1175/JAMC-D-11-052.1.

Agutu, N.O., Awange, J.L., Zerihun, A., Ndehedehe, C.E., Kuhn, M., Fukuda, Y., 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., Ogunjobi, K.O., Ajayi, V.O., Adefisan, E.A., Omotosho, J.A., Sanogo, S., 2016. Comparison of five gridded precipitation products at climatological scales over West Africa Comparison of five gridded precipitation products at climatological scales over West Africa. Meteorol. Atmos. Phys. https://doi.org/10.1007/s00703-016-0493-6.

Anjum, M.N., Ding, Y., Shangguan, D., Tahir, A.A., Iqbal, M., Adnan, M., 2016. Comparison of two successive versions 6 and 7 of TMPA satellite precipitation products with rain gauge data over Swat Watershed, Hindukush Mountains, Pakistan. Atmos. Sci. Lett. 17 (4), 270–279. https://doi.org/10.1002/asl.654.

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

Ashouri, H., Hsu, K.L., Sorooshian, S., Braithwaite, D.K., Knapp, K.R., Cecil, L.D., Nelson, B.R., Prat, O.P., 2015. PERSIANN-CDR: daily precipitation climate data record from multisatellite observations for hydrological and climate studies. Bull. Am. Meteorol. Soc. 96 (1), 69–83. https://doi.org/10.1175/BAMS-D-13-00068.1.

Awange, J.L., Ferreira, V.G., Forootan, E., Andam-akorful, S. A. and Agutu, N. O., 2016. Uncertainties in remotely sensed precipitation, 323(April 2015), 303–323, doi:10.1002/joc.4346.

Beck, H.E., Pan, M., Roy, T., Weedon, G.P., Pappenberger, F., Dijk, A.I.J.M., Van, Huffman, G.J., Adler, R.F., Wood, E.F., 2019. Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS. pp. 207–224.

Beck, H.E., Wood, E.F., Pan, M., Fisher, C.K., Miralles, D.G., Van Dijk, A.I.J.M., McVicar, T.R., Adler, R.F., 2018. MSWEP V2 global 3-Hourly 0.1° precipitation methodology and quantitative assessment. Bull. Am. Meteorol. Soc. 473–500. https://doi.org/10.1175/BAMS-D-17-0138.1.

Becker, A., Finger, P., Rudolf, B., Schamm, K., Schneider, U., Ziese, M., Precipitation, G., Centre, C. and Wetterdienst, D., 2013. A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present, pp. 71–99, doi:10.5194/essd-5-71-2013.

Belabib, N., Zhao, F., Brocca, L., Huang, Y., Tan, Y., 2019. Near-real-time flood forecasting based on satellite precipitation products. Remote Sens. 11. https://doi.org/10.3390/rs11030252.

Bhuiyan, M.A.E., Nikolopoulos, E.I., Anagnostou, E.N., 2019. Machine learning-based blending of satellite and reanalysis precipitation datasets: a multi-regional tropical complex terrain evaluation. J. Hydrometeorol. 6806 (860), 1–42. https://doi.org/10.1175/JHM-D-19-0073.1.

Brunetti, M.T., Melillo, M., Peruccacci, S., Ciabatta, L., Brocca, L., 2018. Remote sensing of environment how far are we from the use of satellite rainfall products in landslide forecasting? Remote Sens. Environ. 210, 65–75. https://doi.org/10.1016/j.rse.2018.03.016.

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

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

Chen, M., Shi, W., Xie, P., Silva, V.B.S., Kousky, V.E., Higgins, R.W., Janowiak, J.E., 2008. Assessing objective techniques for gauge-based analyses of global daily precipitation. J. Geophys. Res. 113, 1–13. https://doi.org/10.1029/2007JD009132.

Collischonn, B., Collischonn, W., Tucci, C.E.M., 2008. Daily hydrological modeling in the Amazon basin using TRMM rainfall estimates. J. Hydrol. 360 (1–4), 207–216. https://doi.org/10.1016/j.jhydrol.2008.07.032.

Cosgrove, W.J., Risberman, F.R. (Eds.), 2000. World Water Vision. Earthscan Publications Ltd, London.

De Paiva, R.C.D., Buarque, D.C., Collischonn, W., Bonnet, M.P., Frappart, F., Calmant, S., Bulhões Mendes, C.A., 2013. Large-scale hydrologic and hydrodynamic modeling of the Amazon River basin. Water Resour. Res. 49 (3), 1226–1243. https://doi.org/10.1002/wrcr.20067.

de Wit, A, Baruth, B, Boogaard, H, van Diepen, K, van Kraalingen, D, Micale, F, te Roller, J, Supit, I, van den Wijngaart, R, 2010. Using ERA-INTERIM for regional crop yield forecasting in Europe. Clim. Res. 44 (1), 41–53. https://doi.org/10.3354/cr00872.

Dembélé, M., Zwart, S.J., 2016. Evaluation and comparison of satellite-based rainfall products in Burkina Faso, West Africa. Int. J. Remote Sens. 37 (17), 3995–4014. https://doi.org/10.1080/01431161.2016.1207258.

Dinku, T., Ceccato, P., Grover-Kopec, E., Lemma, M., Connor, S.J., Ropelewski, C.F., 2007. Validation of satellite rainfall products over East Africa's complex topography. Int. J. Remote Sens. 28 (7), 1503–1526. https://doi.org/10.1080/01431160600954688.

Dorigo, W.A., Gruber, A., De Jeu, R.A.M., Wagner, W., Stacke, T., Loew, A., Albergel, C., Brocca, L., Chung, D., Parinussa, R.M., Kidd, R., 2015. Evaluation of the ESA CCI soil moisture product using ground-based observations. Remote Sens. Environ. 162, 380–395. https://doi.org/10.1016/j.rse.2014.07.023.

Doumounia, A., Gosset, M., Cazenave, F., Kacou, M., Zougmore, F., 2014. Rainfall monitoring based on microwave links from cellular telecommunication networks: first results from a West African test bed. Geophys. Res. Lett. 41 (16), 6015–6021. https://doi.org/10.1002/2014GL060724.

Ferraro, R.R., Smith, E.A., Berg, W., Huffman, G.J., 1998. A screening methodology for passive microwave precipitation retrieval algorithms. J. Atmos. Sci. 55 (9), 1583–1600. https://doi.org/10.1175/1520-0469(1998) 055<1583:ASMFPM>2.0.CO;2.

Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int. J. Climatol. 4315 (May), 4302–4315. https://doi.org/10.1002/joc.5086.

Fischer, E.M., Knutti, R., 2015. Anthropogenic contribution to global occurrence of heavy-precipitation and high-temperature extremes. Nat. Clim. Chang. 5 (April). https://doi.org/10.1038/NCLIMATE2617.

Funk, C., Verdin, A., Michaelsen, J., Peterson, P., Pedreros, P., Husak, G., 2015. A global satellite-assisted precipitation climatology. Earth Syst. Sci. Data 7, 275–287. https://doi.org/10.5194/essd-7-275-2015.

Gao, Z., Long, D., Tang, G., Zeng, C., Huang, J., Hong, Y., 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.

Gebregiorgis, A.S., Hossain, F., 2013. Understanding the dependence of satellite rainfall uncertainty on topography and climate for hydrologic model simulation. IEEE Trans. Geosci. Remote Sens. 51 (1), 704–718. https://doi.org/10.1109/TGRS.2012.2196282.

Gelaro, R., McCarty, W., Suarez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., Da Silva, A.M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J.E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S.D., Sienkiewicz, M., Zhao, B., 2017. The Modern-Era Retrospective Analysis for Research and Applications 2017 (2), 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1.

Giorgi, F., Coppola, E., Raffaele, F., 2018,. Threatening levels of cumulative stress due to hydroclimatic extremes in the 21st century. NPJ Clim. Atmos. Sci. https://doi.org/10.1038/s41612-018-0028-6.

Gosset, M., Viarre, J., Quantin, G., Alcoba, M., 2013. Evaluation of several rainfall products used for hydrological applications over West Africa using two high-resolution gauge networks. Q. J. R. Meteorol. Soc. 139 (673), 923–940. https://doi.org/10.1002/qj.2130.

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

Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. J. Hydrol. 377 (1–2), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003.

Hsu, K., Gao, X., Sorooshian, S., Gupta, H.V., 1997. Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks. J. Appl. Meteorol. 36 (9), 1176–1190. https://doi.org/10.1175/1520-0450(1997) 036<1176:PEFRSI>2.0.CO;2.

Huffman, G.J., Adler, R.F., Bolvin, D.T., Nelkin, E.J., 2010. The TRMM Multi-satellite Precipitation Analysis (TMPA). Satell. Rainfall Appl. Surf. Hydrol. 3–22. https://doi.org/10.1007/978-90-481-2915-7_1.

Huffman, G., Bolvin, D., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E. J., Sorroshian, S., Tan, J. and Xie, P.: Algorithm Theoretical Basis Document (ATBD) NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG)., 2018.

Hussain, Y., Satgé, F., Hussain, M.B., Martinez-Caravajal, H., Bonnet, M.-P., Cardenas-Soto, M., Llacer Roig, H., Akhter, G., 2017. Performance of CMORPH, TMPA and PERSIANN rainfall datasets over plain, mountainous and glacial regions of Pakistan. Theor. Appl. Climatol. https://doi.org/10.1007/s00704-016-2027-z.

Izumi, T., Takikawa, H., Hirabayashi, Y., Hanasaki, N., Nishimori, M., 2017. Contributions of different bias-correction methods and reference meteorological forcing data sets to uncertainty in projected temperature and precipitation extremes. J. Geophys. Res. Atmos. 122 (15), 7800–7819. https://doi.org/10.1002/2017JD026613.

Joyce, R.J., Janowiak, J.E., Arkin, P.A., Xie, P., 2004. CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution. J. Hydrometeorol. 5 (3), 487–503. https://doi.org/10.1175/1525-7541(2004) 005<0487:CAMTPG>2.0.CO;2.

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

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K., Takahashi, K., 2015. The JRA-55 Reanalysis : General Specifications and Basic Characteristics. J. Meteorol. Soc. Japan 93 (1), 5–48. https://doi.org/10.2151/jmsj.2015-001.

Lebel, T., Taupin, J.D., D'Amato, N., 1997. Rainfall monitoring during HAPEX-Sahel. 1. General rainfall conditions and climatology. J. Hydrol. 189, 74–96.

Levizzani, V., Amorati, R., Meneguzzo, F., 2002. A review of satellite-based rainfall estimation. Methods.

Li, J., Heap, A.: Spatial Interpolation Methods: A Review for Environmental Scientists. [online] Available from: http://www.ga.gov.au/corporate_data/68229/Rec2008_023.pdf. (Accessed 4 November 2014), 2008.

Ma, Y., Yang, Y., Han, Z., Tang, G., Maguire, L., Chu, Z., 2018. Comprehensive evaluation of ensemble multi-satellite precipitation dataset using the dynamic bayesian model averaging scheme over the Tibetan plateau. J. Hydrol. 556, 634–644. https://doi.org/10.1016/j.jhydrol.2017.11.050.

Maggioni, V., Massari, C., 2018. On the performance of satellite precipitation products in riverine flood modeling: a review. J. Hydrol. 558, 214–224. https://doi.org/10.1016/j.jhydrol.2018.01.039.

Maggioni, V., Meyers, P.C., Robinson, M.D., 2016. A review of merged high-resolution satellite precipitation product accuracy during the tropical rainfall measuring mission (TRMM) era. J. Hydrometeorol. 17 (4), 1101–1117. https://doi.org/10.1175/JHM-D-15-0190.1.

Maidment, R.I., Grimes, D., Allan, R.P., Tarnavsky, E., Stringer, M., Hewison, T., Roebeling, R., Black, E., 2014. The 30 year TAMSAT African rainfall climatology and time series (TARCAT) data set. J. Geophys. Res. [Atmos.] 10619–10644. https://doi.org/10.1002/2014JD021927.Received.

Maidment, R.I., Grimes, D., Black, E., Tarnavsky, E., Young, M., Greatrex, H., Allan, R.P., Stein, T., Nkonde, E., Senkunda, S., Misael, E., Alcántara, U., 2017. Data Descriptor: a new, long-term daily satellite-based rainfall dataset for operational monitoring in Africa. Nat. Publ. Gr. 4, 1–19. https://doi.org/10.1038/sdata.2017.63.

Massari, C., Crow, W., Brocca, L., 2017. An assessment of the accuracy of global rainfall estimates without ground-based observations. Hydrol. Earth Syst. Sci. Discuss. 1–24. https://doi.org/10.5194/hess-2017-163, 2017.

Messer, H., Zinevich, A., Alpert, P., 2006. Environmental monitoring by wireless communication networks. 312(5774). Science 80, 713. https://doi.org/10.1126/science.1120034.

Muhammad, W., Yang, H., Lei, H., Muhammad, A., Yang, D., 2018. Improving the regional applicability of satellite precipitation products by ensemble algorithm. Remote Sens. 10 (4), 1–19. https://doi.org/10.3390/rs10040577.

Nikolopoulos, E.I., Anagnostou, A.N., Borga, M., 2013. Using high-resolution satellite rainfall products to simulate a major flash flood event in Northern Italy. J. Hydrometeorol. 14, 171–185. https://doi.org/10.1175/JHM-D-12-09.1.

Novella, N.S., Thiaw, W.M., 2012. African rainfall climatology version 2 for famine early warning systems. J. Appl. Meteorol. Climatol. 52 (1996), 588–606. https://doi.org/10.1175/JAMC-D-11-0238.1.

Overeem, A., Leijnse, H., Uijlenhoet, R., 2011. Measuring urban rainfall using microwave links from commercial cellular communication networks. Water Resour. Res. 47 (12), 1–16. https://doi.org/10.1029/2010WR010350.

Poméon, T., Jackisch, D., Diekkrüger, B., 2017. J. Hydrol. 547, 222–235. https://doi.org/10.1016/j.jhydrol.2017.01.055.

Prakash, S., Mitra, A.K., Momin, I.M., 2014b. Comparison of TMPA-3B42 Versions 6 and 7 Precipitation Products with Gauge-Based Data over India for the Southwest Monsoon Period. pp. 346–362, doi:10.1175/JHM-D-14-0024.1.

Prakash, S., Sathiyamoorthy, V., Mahesh, C., Gairola, R.M., 2014a. An evaluation of high-resolution multisatellite rainfall products over the Indian monsoon region. Int. J. Remote Sens. 35 (9), 3018–3035. https://doi.org/10.1080/01431161.2014.894661.

Rahman, K.U., Shang, S., Shahid, M., 2018. Developing an ensemble precipitation algorithm from satellite products and its topographical and seasonal evaluations over Pakistan. Remote Sens. https://doi.org/10.3390/rs10111835.

Rahmawati, N., Lubczynski, M.W., 2017. Validation of satellite daily rainfall estimates in complex terrain of Bali Island, Indonesia. Theor. Appl. Climatol. 1–20. https://doi.org/10.1007/s00704-017-2290-7.

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

Reichle, R.H., Liu, Q., Koster, R.D., Draper, C.S., Mahanama, S.P.P., Partyka, G.S., 2017. Land Surface Precipitation in MERRA-2. J. Clim. 2, 1643–1664. https://doi.org/10.1175/JCLI-D-16-0570.1.

Ringard, J., Becker, M., Seyler, F., Linguet, L., 2015. Temporal and spatial assessment of four satellite rainfall estimates over French Guiana and north Brazil. Remote Sens. 7 (12), 16441–16459. https://doi.org/10.3390/rs71215831.

Saeed, F., Bethke, I., Fischer, E., Legutke, S., Shiogama, H., Stone, D.A., Schleussner, C.-P., 2018. Robust changes in tropical rainy season length at OPEN ACCESS Robust changes in tropical rainy season length at 1. 5 ∘ C. Environ. Res. Lett. 13. https://doi.org/10.1088/1748-9326/aab797.

Salles, L., Satgé, F., Roig, H., Almeida, T., Olivetti, D., Ferreira, W., 2019. Seasonal effect on spatial and temporal consistency of the new GPM-Based IMERG-v5 and GSMaP-v7 Satellite Precipitation Estimates in Brazil's Central Plateau Region. Water 11. https://doi.org/10.3390/w11040668.

Satgé, F., Bonnet, M.-P., Gosset, M., Molina, J., Hernan, W., Yuque Lima, Pillco Zolá, R., Timouk, F., Garnier, J., 2016. Assessment of satellite rainfall products over the Andean plateau. Atmos. Res. 167, 1–14. https://doi.org/10.1016/j.atmosres.2015.07.012.

Satgé, F., Espinoza, R., Zolá, R., Roig, H., Timouk, F., Molina, J., Garnier, J., Calmant, S., Seyler, F., Bonnet, M.-P., 2017a. Role of climate variability and human activity on Poopó Lake droughts between 1990 and 2015 assessed using remote sensing Data. Remote Sens. 9 (3), 218. https://doi.org/10.3390/rs9030218.

Satgé, F., Xavier, A., Zolá, R., Hussain, Y., Timouk, F., Garnier, J., Bonnet, M.-P., 2017b. Comparative assessments of the latest GPM mission's spatially enhanced satellite rainfall products over the main Bolivian watersheds. Remote Sens. 9 (4), 369. https://doi.org/10.3390/rs9040369.

Satgé, F., Ruelland, D., Bonnet, M., Molina, J., Pillco, R., Hydrosciences, U.M.R., Bataillon, P.E., Cedex, M., Hydrosciences, U.M.R., Bataillon, P.E., Cedex, M., 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.

Schneider, U., Becker, A., Finger, P., Meyer-christoffer, A., Ziese, M., Rudolf, B., 2014. GPCC ' s new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle, , 15–40, doi:10.1007/s00704-013-0860-x.

Sorooshian, S., Hsu, K.-L., Gao, X., Gupta, H.V., Imam, B., Braithwaite, D., 2000.

Evaluation of PERSIANN System Satellite-Based Estimates of Tropical Rainfall. Bull. Am. Meteorol. Soc. 81 (9), 2035–2046. https://doi.org/10.1175/1520-0477(2000)081<2035:EOPSSE>2.3.CO;2.

Su, J., Lü, H., Wang, J., Sadeghi, A.M., Zhu, Y., 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 (11), 1–19. https://doi.org/10.3390/rs9111176.

Sultan, B., Roudier, P., Quiron, P., Alhassane, A., Muller, B., Dingkuhn, M., Ciais, P., Guimberteau, M., Traore, S., Baron, C., 2013. Assessing climate change impacts on sorghum and millet yields in the Sudanian and Sahelian savannas of West Africa. Environ. Res. Lett. 8. https://doi.org/10.1088/1748-9326/8/1/014040.

Sun, W., Ma, J., Yang, G., 2018b. Statistical and hydrological evaluations of multi-satellite precipitation products over Fujiang River Basin in Humid Southeast China. Remote Sens. 10. https://doi.org/10.3390/rs10121898.

Sun, Q., Miao, C., Duan, Q., Kong, D., 2015. Would the ' real ' observed dataset stand up ? A critical examination of eight observed gridded climate datasets for China. Environ. Res. Lett. 9. https://doi.org/10.1088/1748-9326/9/1/015001.

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorroshian, S., Hsu, K.-L., 2018a. A review of global precipitation data sets: data, sources, estimation, and intercomparisons. Rev. Geophys. 56, 79–107. https://doi.org/10.1002/2017RG000574.

Tang, G., Ma, Y., Long, D., Zhong, L., Hong, Y., 2016. Evaluation of GPM Day-1 IMERG and TMPA Version-7 legacy products over Mainland China at multiple spatio-temporal scales. J. Hydrol. 533, 152–167. https://doi.org/10.1016/j.jhydrol.2015.12.008.

Tang, G., Behrangi, A., Long, D., Li, C., Hong, Y., 2018. Accounting for spatiotemporal errors of gauges: a critical step to evaluate gridded precipitation products. J. Hydrol. 559, 294–306. https://doi.org/10.1016/j.jhydrol.2018.02.057.

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

Thiemig, V., Rojas, R., Zambrano-Bigiarini, M., Levizzani, V., De Roo, A., 2012. Validation of satellite-based precipitation products over sparsely gauged African river basins. J. Hydrometeorol. 13 (6), 1760–1783. https://doi.org/10.1175/JHM-D-12-032.1.

Tian, Y., Peters-Lidard, C.D., Eylander, J.B., Joyce, R.J., Huffman, G.J., Adler, R.F., Hsu, K.-L., Turk, F.J., Garcia, M., Zeng, J., 2009. Component analysis of errors in satellite-based precipitation estimates. J. Geophys. Res. 114 (D24), D24101. https://doi.org/10.1029/2009JD011949.

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

Ushio, T., Sasashige, K., Kubota, T., Shige, S., Okamoto, K., Aonashi, K., Inoue, T., Takahashi, N., Iguchi, T., Kachi, M., Oki, R., Morimoto, T., Kawasaki, Z.-I., 2009. A Kalman filter approach to the global satellite mapping of precipitation (GSMaP) from combined passive microwave and infrared radiometric data. J. Meteorol. Soc. Japan 87A, 137–151. https://doi.org/10.2151/jmsj.87A.137.

Xie, P., Yatagai, A., Chen, M., Hayasaka, T., Fukushima, Y., Liu, C., Yang, S., 2007. A gauge-based analysis of daily precipitation over East Asia. J. Hydrometeorol. 607–626. https://doi.org/10.1175/JHM583.1.

Yamamoto, M.K., Shige, S., 2014. Implementation of an orographic/nonorographic rainfall classification scheme in the GSMaP algorithm for microwave radiometers. Atmos. Res. 163, 36–47. https://doi.org/10.1016/j.atmosres.2014.07.024.

Yuan, F., Zhang, L., Min, K., Soe, W., Ren, L., Zhao, C., 2019. Applications of TRMM- and GPM-Era multiple- satellite precipitation products for flood simulations at sub-daily scales in a sparsely gauged watershed in Myanmar. Remote Sens. https://doi.org/10.3390/rs11020140.

Zambrano-Bigiarini, M., Nauditt, A., Birkel, C., Verbist, K., Ribbe, L., 2017. Temporal and spatial evaluation of satellite-based rainfall estimates across the complex topographical and climatic gradients of Chile. Hydrol. Earth Syst. Sci. 21 (2), 1295–1320. https://doi.org/10.5194/hess-21-1295-2017.

Zeng, Q.Y., Wang, L., Chen, Z., Wang, H., Zhu, B., 2018. Li Inter-comparison and evaluation of remote sensing precipitation products over China from 2005 to 2013. Remote Sens. 10 (2). https://doi.org/10.3390/rs10020168.

Zhang, Y., Li, Y., Ji, X., Luo, X., Li, X., 2018. Evaluation and hydrologic validation of three satellite-based precipitation products in the upper catchment of the Red River Basin, China. Remote Sens. 10, 1–22. https://doi.org/10.3390/rs10121881.

Zinevich, A., Alpert, P., Messer, H., 2008. Estimation of rainfall fields using commercial microwave communication networks of variable density. Adv. Water Resour. 31 (11), 1470–1480. https://doi.org/10.1016/j.advwatres.2008.03.003.