Journal of Hydrology: Regional Studies 47 (2023) 101409

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

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

Comparison of gridded precipitation estimates for regional hydrological modeling in West and Central Africa

Christopher Kouakou [a,*], Jean-Emmanuel Paturel [b], Frédéric Satgé [c], Yves Tramblay [b], Dimitri Defrance [d], Nathalie Rouché [b]

[a] Ecole Doctorale Polytechnique (EDP), Institut National Polytechnique Félix Houphouët-Boigny (INP-HB), Yamoussoukro, Cote d'Ivoire
[b] HSM (Univ. Montpellier, CNRS, IRD), Montpellier, France
[c] ESPACE-DEV (Univ. Montpellier, IRD, Univ. Antilles, Univ. Guyane, Univ. Réunion), Montpellier, France
[d] The Climate Data Factory, Paris, France

ARTICLE INFO

Keywords:
Gridded precipitation estimates
GR2M
West Africa
Central Africa
Re-analysis
Observed data

ABSTRACT

Study region: Data-scarce basins located in West Africa and northern Central Africa.
Study focus: Multiple studies have shown that global gridded precipitation datasets could provide an alternative to the lack of observed data in Sub-Saharan Africa. This work evaluated 15 precipitation datasets based on satellite rainfall (ARC v.2, CHIRP v.2, CHIRPS v.2, PERSIANN-CDR, MSWEP v2.2 and TAMSAT v3), reanalysis (ERA5, JRA-55 Adj, MERRA-2 PRECTOT, MERRA-2 PRECTOTCORR, WFDEI-CRU and WFDEI-GPCC) and ground measurements (CPC v.1, CRU TS v.4.00 and GPCC v.7), as well as a regional estimation method, based on spatial proximity, for the parameters of a simple monthly water balance model, GR2M. The regional simulations of the GR2M model were evaluated based on a Kling-Gupta Efficiency score in a split-sample spatio-temporal validation scheme.
New hydrological insights for the region: The results show that among all the precipitation products, CHIRPS is the most effective for hydrological modeling in West and Central Africa at a monthly timestep. Also, among the top five products are WFDEI-CRU, CRU, WFDEI-GPCC and GPCC. Overall, regional hydrological modeling is more effective for basins smaller than 80,000 km^2. The method of regionalization by spatial proximity causes an overall drop in the ability of the various precipitation products to reproduce discharge, most notably with WFDEI-GPCC and GPCC. CHIRPS remains the best product in terms of KGE2 values in regionalization.

1. Introduction

1.1. Need for precipitation and streamflow data for West Africa

Water resources across the globe are under significant pressure due to rapid demographic increases, with strong demand for economic activities (agriculture, industry), leisure (depending on the standard of living), and quite simply subsistence. According to the United Nations World Report on the Development of Water Resources 2021, the consumption of "blue gold" will increase at a rate of approximately 1% per year until 2050, when it will be 20–30% above current levels (Burek et al., 2016; Wada et al., 2016).

* Corresponding author.
E-mail address: christopher.kouakou@inphb.ci (C. Kouakou).

https://doi.org/10.1016/j.ejrh.2023.101409
Received 8 November 2022; Received in revised form 28 April 2023; Accepted 1 May 2023
Available online 12 May 2023

On top of the demographic pressure, there is that of climate change. In West Africa and Central Africa, annual accumulations on 2011–2019 are stable relative to the 1981–2010 normal (State of the Climate in Africa 2019). However, the number of extreme events has been increasing in recent years. This results in the loss of human lives and the destruction of material goods, including crops in rural areas and infrastructure, which is already scarce and mainly concentrated in the few urban areas located in developing countries (World Meteorological Organization, 2021).

To control these risks, it is important to better understand the hazard and its impact. In other words, it is necessary to master rainfall information in the African space in order to face the challenge of forecasting the evolution of water resources and extreme events, the frequency of which has increased over the past two decades.

The task is all the tougher, however, because the observation networks (rainfall stations) are under-equipped, vandalized, sparse, and sometimes poorly maintained (Ekeu-wei et al., 2018; Ekeu-wei and Blackburn, 2018). According to the World Meteorological Organization, the current observation network in Africa is at one eighth of the ideal density, with only 22% of stations in 2019 meeting the requirements of the Global Climate Observing System (GCOS), versus 57% in 2011 (State of the Climate in Africa 2019). In addition to the problems mentioned, there is the problem of access to data managed by meteorological services, especially for the academic world, due to their high acquisition cost (Bodian et al., 2020; Panthou et al., 2014; Tramblay et al., 2021).

1.2. State of the art

There has been an increase in the design, evaluation, and improvement of gridded precipitation estimates (GPEs) covering various spatio-temporal windows, which appear as a response to the lack of data in certain countries (Satgé et al., 2021, 2020; Viana et al., 2021). These products fall under three categories based on the sources of information used (individually or in combination) to generate them: ground observation, outputs of the numerical model of the globe (or re-analyses) and estimates by satellite remote sensing.

The first category of GPEs is obtained by spatial interpolation of rainfall measurements from several rainfall stations using kriging methods, inverse distance weighting, and many other methods for interpolation between rain gauges (Harris et al., 2020; Schneider et al., 2014).

GPEs based on retrospective analyses, or re-analyses, which make up the second category, combine outputs of a numerical model with observations through data assimilation. Satellite observations mainly but also ground observation data are fed into Atmospheric General Circulation Models (AGCMs) for forecasts over the few hours preceding current weather or the chosen reference period (Gelaro et al., 2017). Estimates corrected with observational data are fed back into the AGCMs. This progressive data assimilation process produces time series fairly representative of the dynamics of certain climatic parameters such as precipitation, temperature, evapotranspiration, etc. (Reichle et al., 2017).

Finally, the third category of rainfall grid contains satellite precipitation estimates that are extracted from infrared and microwave recordings. Indeed, infrared recordings make it possible to deduce the temperature of the clouds and to estimate the quantity of precipitated water. On the other side, satellite microwave sensors allow precipitation estimation through measuring the microwave radiation emitted or reflected by the atmosphere and Earth's surface. In addition to the analysis of cloud dynamics, there are satellite estimates, again using the infrared and microwaves, of soil surface temperature and even soil humidity (Martens et al., 2017). Machine learning methods such as neural networks make it possible to take better advantage of this mass of data (Hsu et al., 1999, 1997; Sun et al., 2018). Furthermore, the ground-based GPEs and bias correction procedures based on physiographic indices ensure the design of more robust composite GPEs that take into account satellite estimates (Ashouri et al., 2016, 2015; Funk et al., 2015).

Many studies focus on the evaluation and improvement of these various products, mostly for Asia and Latin America, and increasingly for Africa. This is done through direct comparisons between the precipitation series (P) observed on the ground and the P series of the various estimation products. In 2012, seven P datasets were tested on the Volta basin: CMORPH, GPROF-v6, GSMaP-MVK v5, RFE-2.0, TMPA-v6, PERSIANN and ERA-Interim (Thiemig et al., 2013). Seven P datasets for Burkina Faso (ARC2, CHIRPS v.2, PERSIANN-CDR, RFE v.2, TAMSAT v.2, TMPA v.7 and TMPA RT v.7) were also evaluated on a daily, decadal, and monthly scale (Dembélé and Zwart, 2016). In 2017, TAMSAT v.3 was compared to the previous version (TAMSAT v.2) and to six P datasets (ARC 2, CHIRP v.2, CHIRPS v.2, CMORPH v.1, RFE and TMPA v.7) for several West African countries, specifically Nigeria, Niger, Uganda, Zambia and Mozambique (Maidment et al., 2017). On the scale of Ghana, the GPCC v.7, CHIRPS, CMORPH and TRMM 3B42 products were compared to series from 30 rainfall stations (Atiah et al., 2020). On a larger West African scale, two works evaluated four (TMPA v.7, Udel v. 3.1, CRU v.3, and ARC2) and twenty-three newer precipitation estimation products, respectively (Akinsanola et al., 2017; Satgé et al., 2020).

These GPEs are also used as forcing variables in hydrological and agroclimatic models. The simulated flow rates are compared to observations via various performance criteria. In 2013, nine estimation products covering Benin and Niger (CMORPH, EPSAT-SG, GPCP, GSMaP-MVK, GSMaPRT, RFE-2, TMPA-v6, TMPA-RT v6 and PERSIANN) were tested for hydrological applications with the Soil Conservation Service method and the GR4J model and for agrometeorological uses with the SARRA-H ("Système d'Analyse Régionale de Risques Agroclimatiques version Habillée") and EPIC (Environmental Policy Integrated Climate) crop models (Gosset et al., 2013; Ramarohetra et al., 2013). In 2017, the HBV light model was used to evaluate 10 precipitation datasets (CFSR, CHIRPS, CMORPH v.1 RAW and CRT, PERSIANN-CDR, RFE-2, TAMSAT v.2, TMPA v.7, TMPA-RT v.7 and GPCC) for six basins located in Burkina Faso, Nigeria and Ghana (Poméon et al., 2017). Likewise, TRMM, CMORPH, CPC-FEWS, PERSIANN and RFE2 were assessed on Senegal and Niger basins (Bodian et al., 2016; Casse et al., 2015; Stisen and Sandholt, 2010). More recently, in 2020, a study focused on the evaluation of 17 precipitation estimation datasets over the Volta basin using the fully distributed mesoscale Hydrologic Model (mHM) (Dembélé et al., 2020). Similarly, research was conducted on the use of gridded data as forcing data for the lumped hydrological model HMETS with a dataset consisting of GPCC, CPC Unified, CHIRPS, PERSIANN-CDR, MSWEP V1.0, JRA55, NCEP-CFSR,

ERA-I, and ERA5 series in the context of climate change in 850 African basins (Tarek et al., 2020).

It emerges from these two types of studies that products incorporating various sources of information, in particular based on ground measurements data, perform better than simple ground-based, raw satellite-based or raw re-analysis products. In addition, products derived entirely from satellite estimates or re-analyses show large biases with respect to local ground measurements timeseries. These deviations have a negative impact on the modeling of crops and flows in certain geographical contexts.

1.3. Objectives

This article aims to assess different types of gridded precipitation estimates products (ground-based, satellite-based, re-analysis and combined products) through rainfall-runoff modeling at a monthly time step and on a regional scale including basins from both West and Central Africa. Specifically, this involves an evaluation of the performance of a set of 15 precipitation estimation products at a monthly time step using a hydrological calibration-validation procedure. Our study contribution to the former ones relies on the combination of a large dataset with a wider regional scale. In fact, the previous studies on Africa mentioned above, apart from Tarek et al. (2020), focused on short number of isolated or transboundary basins either on West Africa or Central Africa but not both areas at once. This study includes also 9 additional GPEs to the products assessed by Tarek et al. (2020), among which a newer version of MSWEP V2.2.

Since ground data is often scarce in these regions, this evaluation is enriched with an analysis of the reliability and robustness of these products in the context of ungauged basins via a regionalization method.

2. Study area and data

2.1. Study area

Our study area is made up of nineteen main basins some of which cross national boundaries. The largest is the basin of the Niger River, which is 4200 km long and runs through Guinea, Mali, Niger, Burkina Faso, Benin, and finally Nigeria, where it flows into the Atlantic Ocean.

In this region straddling tropical and equatorial area, the seasons and the division into climatic categories are mainly governed by rain (Fig. 1). The rainfall regimes vary from a desert regime in the far north, in the Sahara zone, to more humid regimes towards the

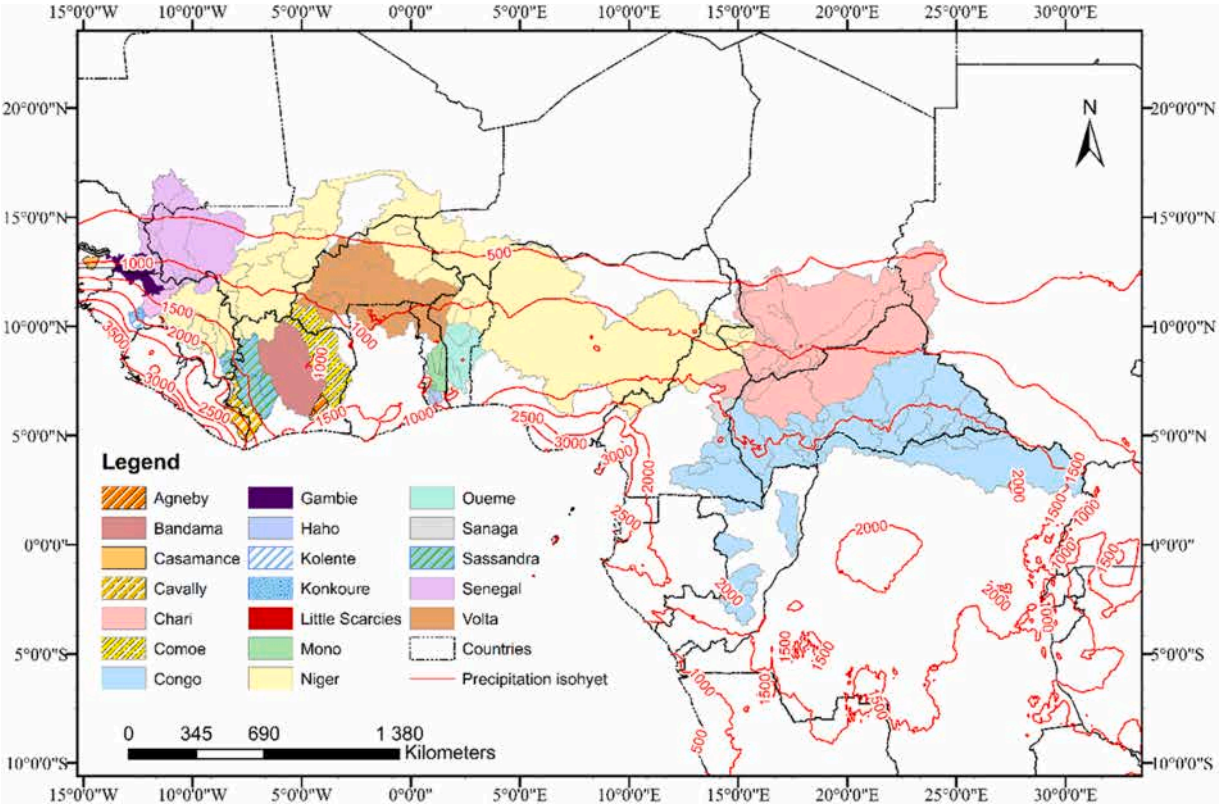

Fig. 1. Map of the study area: Annual rainfall (1970–2000) and main basins (Sources: (Fick and Hijmans, 2017); SIEREM database, HydroSciences Montpellier).

Table 1
Characteristics and references of gridded precipitation estimates.

Acronym	Full name	Data Source	Temporal Coverage	Temporal Resolution	Spatial Coverage	Spatial Resolution	Latency	References
ARC-2	Africa Rainfall Climatology v.2	S,G	1983-present	Daily	Africa	0.1°	2 days	(Novella and Thiaw, 2013)
CHIRP v.2	Climate Hazards group InfraRed Precipitation v.2	S,R	1981-present	Daily	50°	0.05°	2 days	(Funk et al., 2015)
CHIRPS v.2	CHIRP with Stations v.2	S,R,G	1981-present	Daily	50°	0.05°	1 month	(Funk et al., 2015)
CPC v.1	Climate Prediction Center unified v.1	G	1979-present	Daily	Global	0.5°	1 day	(Chen et al., 2008; Xie et al., 2007)
CRU TS v.4.00	Climatic Research Unit Time-Series v.4.00	G	1901–2005	Monthly	Land/ Antarctica	0.5°	Irregular	(Harris et al., 2020)
ERA5	European Centre for Medium-Range Weather Forecasts ReAnalysis 5	R	1979-present	Hourly	Global	0.25°	3 months	(Hersbach et al., 2020)
GPCC v.7	Global Precipitation Climatology Center v.7	G	1901–2013	Monthly	Global	1°	Irregular	(Becker et al., 2013; Schneider et al., 2014)
JRA-55 Adj	Japanese 55-year ReAnalysis Adjusted	R,G	1959–2013	3 h	Global	0.56°	Stopped	(Iizumi et al., 2017)
MERRA-2 PRECTOT	Modern-Era Retrospective Analysis for Research and Applications 2 PRECipitation TOTal	R,S,G	1980-present	Hourly	Global	0.5°x 0.625°	2 Months	(Gelaro et al., 2017; Reichle et al., 2017)
MERRA-2 PRECTOTCORR	MERRA-2 PRECTOT bias CORRected	R,S,G	1980-present	Hourly	Global	0.5°x 0.625°	2 Months	(Gelaro et al., 2017; Reichle et al., 2017)
MSWEP v2.2	Multi-Source Weighted Ensemble Precipitation v.2.2	S,R,G	1979-present	3 h	Global	0.1°	Several months	(Beck et al., 2019)
PERSIANN-CDR	Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks - Climate Data Record	S,G	1983–2016	Daily	60°	0.25°	6 months	(Ashouri et al., 2015)
TAMSAT v.3	Tropical Applications of Meteorology using SATellite and ground-based observations v.3	S,G	1983-present	Daily	Africa	0.0375°	~1 week	(Maidment et al., 2017)
WFDEI-CRU	WATCH Forcing Data ERA Interim (WFDEI) corrected using Climatic Research Unit (CRU) dataset	R,G	1979–2016	3 h	Global	0.5°	Stopped	(Weedon et al., 2014)
WFDEI-GPCC	WATCH Forcing Data ERA Interim (WFDEI) corrected using Global Precipitation Climatology Center (GPCC) dataset	R,G	1979–2016	3 h	Global	0.5°	Stopped	(Weedon et al., 2014)

NB: The letters G, R and S respectively represent ground measurements, re-analyses and satellite estimates. There is an order of importance in the "Data Source" column: the first item listed is the preponderant source of rainfall information.

Atlantic Ocean and as we approach Central Africa (Cameroon, Congo). At the same time, annual accumulations vary from less than 200 mm/year to more than 5000 mm/year (Abiodun et al., 2012; Akinsanola et al., 2017).

2.2. Observed streamflow dataset

The flow data comes from the SIEREM database (http://www.hydrosciences.fr/SIEREM/), developed by researchers and technicians at HydroSciences Montpellier (Boyer et al., 2006). It consists of monthly flows observed at hydrometric stations (348 stations) distributed over the study area. These timeseries contain different rate of gaps with majority of end dates in the 1990 s and early 2000 s (Ouermi et al., 2019; Paturel et al., 2010).

2.3. Meteorological datasets

The rainfall dataset consists of GPEs based on satellite estimates (ARC-2, CHIRP v.2, CHIRPS v.2, PERSIANN-CDR, MSWEP v2.2, TAMSAT v3), re-analysis products (ERA5, JRA-55 Adj, MERRA-2 PRECTOT, MERRA-2 PRECTOTCORR, WFDEI-CRU, WFDEI-GPCC), and GPEs based on ground observations (CPC v.1, CRU TS v. 4.00 and GPCC v.7). The choice of the GPEs was led by the temporal coverage of SIEREM flow dataset. Table 1 summarizes the characteristics of these products by specifying the origins of the data used to create each precipitation grid and its spatial and temporal resolution. There is an order of importance in the “Data Source” column: the first item listed is the preponderant source of rainfall information. "R, S, G” means that the estimation product comes from re-analysis (R) reinforced by satellite estimates (S) and corrected with ground-based data (G). This column makes it possible to identify and compare the various types of GPEs selected for this study.

For hydrological modeling purposes, the Potential evapotranspiration data (PET) is taken from the CRU TS v.4.00 product (Harris et al., 2020) and is calculated by the FAO Penman–Monteith method (Allen et al., 1998).

For modelling process, both rainfall and evapotranspiration timeseries for each of the basins selected are extracted from weighted average of values of the pixels overlaying their areas. For instance, the weight of a pixel fully include in the basin is 1 while the weight is 0.5 for the pixel for which only mid-part is inside the basin. This extraction was made using scripts in R language on RStudio and Java scripts on Google Earth Engine. In cases where there is only one pixel that concerns the basin, the value of this pixel is directly taken as the average rainfall of the basin. This latter procedure is justified by the fact that GR2M lumped model uses catchment mean areal rainfall (Mouelhi et al., 2006) and the basins considered in this study are at least of medium size. Indeed, the areas of the selected basins ranged from 1 279 km^2 to 600,000 km^2.

3. Methodology

3.1. Hydrological model

This study uses the GR2M model (Mouelhi et al., 2006) integrated into the airGR package (Coron et al., 2017) developed in R by the french institute named “Institut national de recherche pour l’agriculture, l’alimentation et l’environnement–INRAE " (Fig. 2). This model was chosen because it has shown its robustness in many studies and is widely used by the scientific community in West and

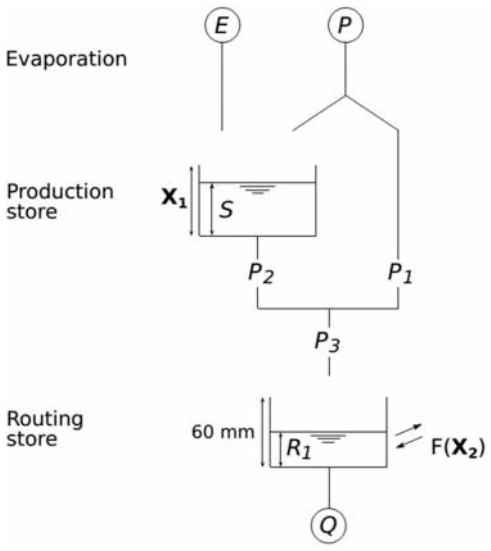

Fig. 2. Diagram of the Mouelhi GR2M model.
Source: Manuel airGR 2020.

Central Africa (Kouassi et al., 2013, 2012; Ouermi et al., 2019; Paturel et al., 2010; Sambou et al., 2020; Vamoryba et al., 2019).

3.2. GR2M model

GR2M is a monthly-timestep lumped model with two parameters X1 and X2 related to the production function and the transfer function, respectively. X1 is the storage capacity in mm, while X2 is the underground exchange coefficient. The input variables are the rainfall and the PET observed over the catchment area and expressed in mm. The output variable of the model is the simulated flow. For more details on the intermediate calculation stages of the production and transfer functions, see Mouelhi (2003).

3.3. Calibration-validation method

The calibration-validation method makes it possible to avoid a random choice for the division of the data series (Klemeš, 1986). Our approach is based on the split-sample test (SST) method, which stipulates that the data series should be split into two parts of equal length if the series is long enough (as was done in this paper). Initially, the first 50% of the data is used for calibration and the remaining 50% for validation, then the last 50% is used for calibration and the first 50% for validation. The model warm-up considers the year preceding the calibration sub-period.

The longest period of availability of both hydrometric data and precipitation estimates over this study area of interest is 1984–2005. With regards to the scarcity of ground data in our area of interest, 21 years appear as a quite long timeseries. Likewise, while 21 years may seem short in a climate perspective (with a climate normal computed on 30 years), several studies have shown that as little as 8 years could be sufficient to calibrate a hydrological model (Arsenault et al., 2018; Singh and Bárdossy, 2012). Other authors in the same region only used 10 years for calibration of hydrological models (for instance, Dembélé et al., 2020, Poméon et al., 2017). Then, to limit the impact of gaps in hydrometric series on the modelling, the selection of hydrometric stations was based on a maximum gap rate of 50% over the periods 1986–1995 and 1996–2005 (Table 2). Based on this gap threshold, 68 hydrometric stations were retained for the calibration-validation procedure. Using the GrandDam database (Tramblay et al., 2021), we made sure that the basins corresponding to the selected hydrometric stations were not or only slightly impacted by dams or human activities. Both 1985 and 1995 data are used for the model warm-up for respectively SST1 and SST2 periods. The spatial visualization of these stations and corresponding basins can be found in the Fig. 3. Fig. 4 shows temporality of the available data for the different hydrometric stations selected.

3.4. Model calibration criteria

KGE2 (according to its name in the airGR package), a modified formulation of the Kling and Gupta Efficiency criteria (KGE), was chosen as the criterion function for evaluating the performance of hydrological models (Kling et al., 2012).

$$\boldsymbol{KGE2} = \mathbf{1} - \sqrt{(\boldsymbol{r}-\mathbf{1})^2 + (\boldsymbol{\beta}-\mathbf{1})^2 + (\boldsymbol{\gamma}-\mathbf{1})^2} \tag{1}$$

with:

$$r = \frac{\sum_{i=1}^{N}(S_i - \mu_S)(O_i - \mu_O)}{\sqrt{\sum_{i=1}^{N}(S_i - \mu_S)^2}\sqrt{\sum_{i=1}^{N}(O_i - \mu_O)^2}}$$

$$\beta = \frac{\mu_S}{\mu_O}$$

$$\gamma = \frac{CV_S}{CV_O} = \frac{\sigma_S/\mu_S}{\sigma_O/\mu_O}$$

$\boldsymbol{S_i}$ stand for simulated flow for month i and $\boldsymbol{O_i}$ for observed flow for the same month i. $\boldsymbol{r}$ is the correlation coefficient between the simulated flows and the observed flows. $\boldsymbol{\beta}$ is the bias ratio between the Mean of the simulated flows μ_s and the Mean of the observed flows μ_o(in mm). $\boldsymbol{\gamma}$ is the variability ratio between the coefficients of variation. $\boldsymbol{CV_o}$ (resp. $\boldsymbol{CV_s}$) represents the coefficient of variation of the observed flow (resp. simulated flow). $\boldsymbol{\sigma_o}$(resp. $\boldsymbol{\sigma_s}$) represents the standard deviation of the observed flow (resp. simulated flow) in mm. The use of the variability ratio as a replacement for the standard deviation ratio initially included in the formulation of the KGE

Table 2
Pairs of calibration-validation periods.

Period	SST1	SST2
Calibration	1986–1995	1996–2005
Validation	1996–2005	1985–1995

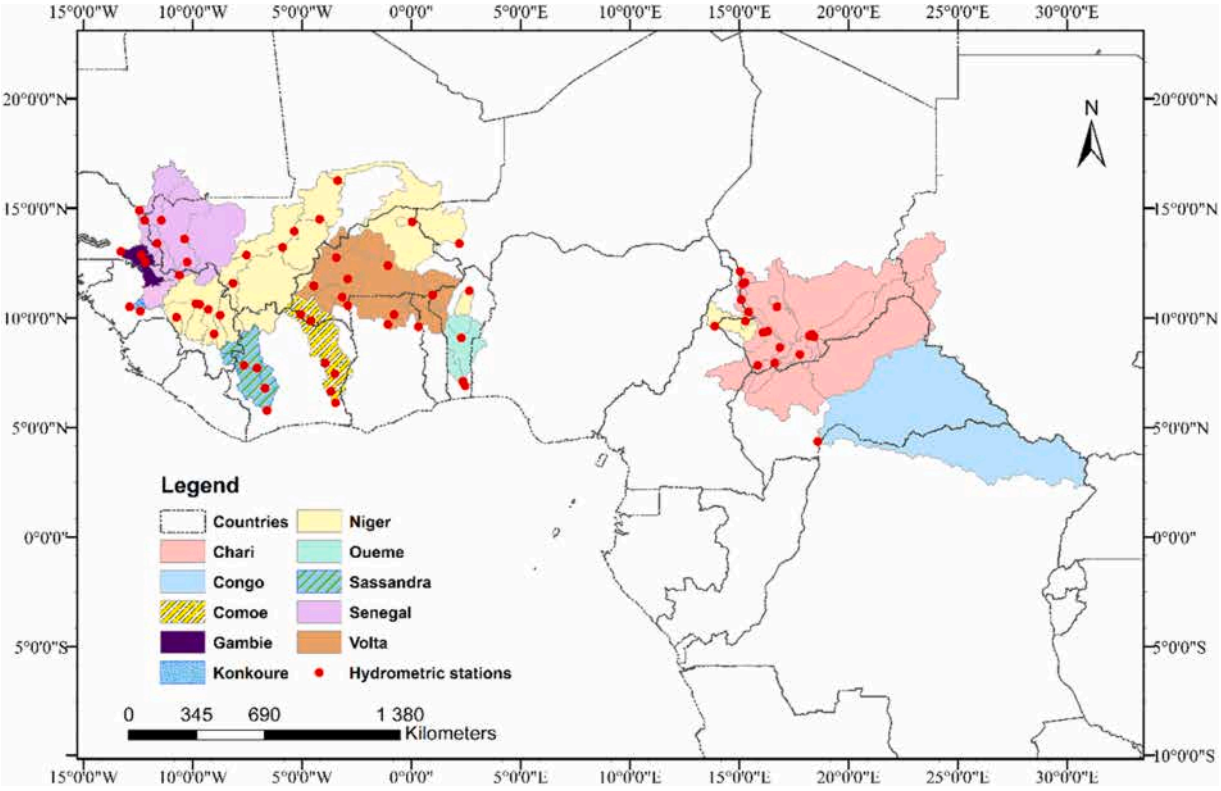

Fig. 3. Presentation of the 68 basins (ranged from 1 279 km^2 to 600,000 km^2) considered for calibration-validation procedure.

(Gupta et al., 2009) makes it possible to remove the cross-correlation between the factors β and γ.

The interpretation of KGE2 differs from that of NSE (Nash-Sutcliffe Efficiency). If $S_i = \mu_o$, NSE = 0, whereas KGE= −0.41. KGE2 must therefore be interpreted from the point of view of its components r, β and γ (Knoben et al., 2019). It makes possible to evaluate the capacity of a model to reproduce the temporal dynamics (parameter r) and the conservation of the duration of the flows (parameters β and γ). In addition, KGE2 tends towards the smallest of the triplet of factors (r, β, γ), thus allowing a hydrological interpretation of the results. For example, if KGE2 = 0.85, this means that the smallest of the factors (r, β *or* γ) is close to or slightly above 0.85.

The calculation of the KGE2 criterion was carried out using the square root of the flows in order to give the same weight to the different parts of the flood hydrograph: low water, high water and Mean water.

3.5. Model performance evaluation

First of all, the performance of the precipitation products was measured by the hydrological model performance over the calibration period. Three elements were then considered to judge the performance of the products. These are the median of the KGE2 values, the min-max distance (difference between the maximum value of KGE2 and the minimum value of KGE2 over all the catchments for a given rainfall product), and the interquartile distance (IQ = Q3 – Q1; Q1 = first quartile, Q3 = third quartile). Unlike KGE2, the best value of which must be close to 1, the min-max and interquartile distances must be as close as possible to 0 for better model performance.

Then, the validation performance was evaluated using the robustness ratio (Dakhlaoui et al., 2017) in addition to the previous elements via the following formula:

$$Robustness \quad ratio = \frac{KGE2_{val/receiver} - KGE2_{cal/receiiver}}{KGE2_{cal/receiver}} \tag{2}$$

where $KGE2_{val/receiver}$ designates the validation performance over the "receiver" (validation period) period using the parameters resulting from the calibration over the "donor" (calibration period) period, and $KGE2_{cal/receiver}$ is the calibration performance over the "receiver" period (Ouermi et al., 2019). The advantage of this formula in the context of our study area is to minimize the influence of the differences in monthly flow gaps between the calibration and validation periods on the performance of the models.

Since the ratios are mostly negative, the median and 3rd quartile absolute values of the ratios were retained for our analysis. The more the ratio tends towards 0, the more robust the model. Similarly, the more the median and 3rd quartile absolute values of the ratio tend towards 0, the more robust the model.

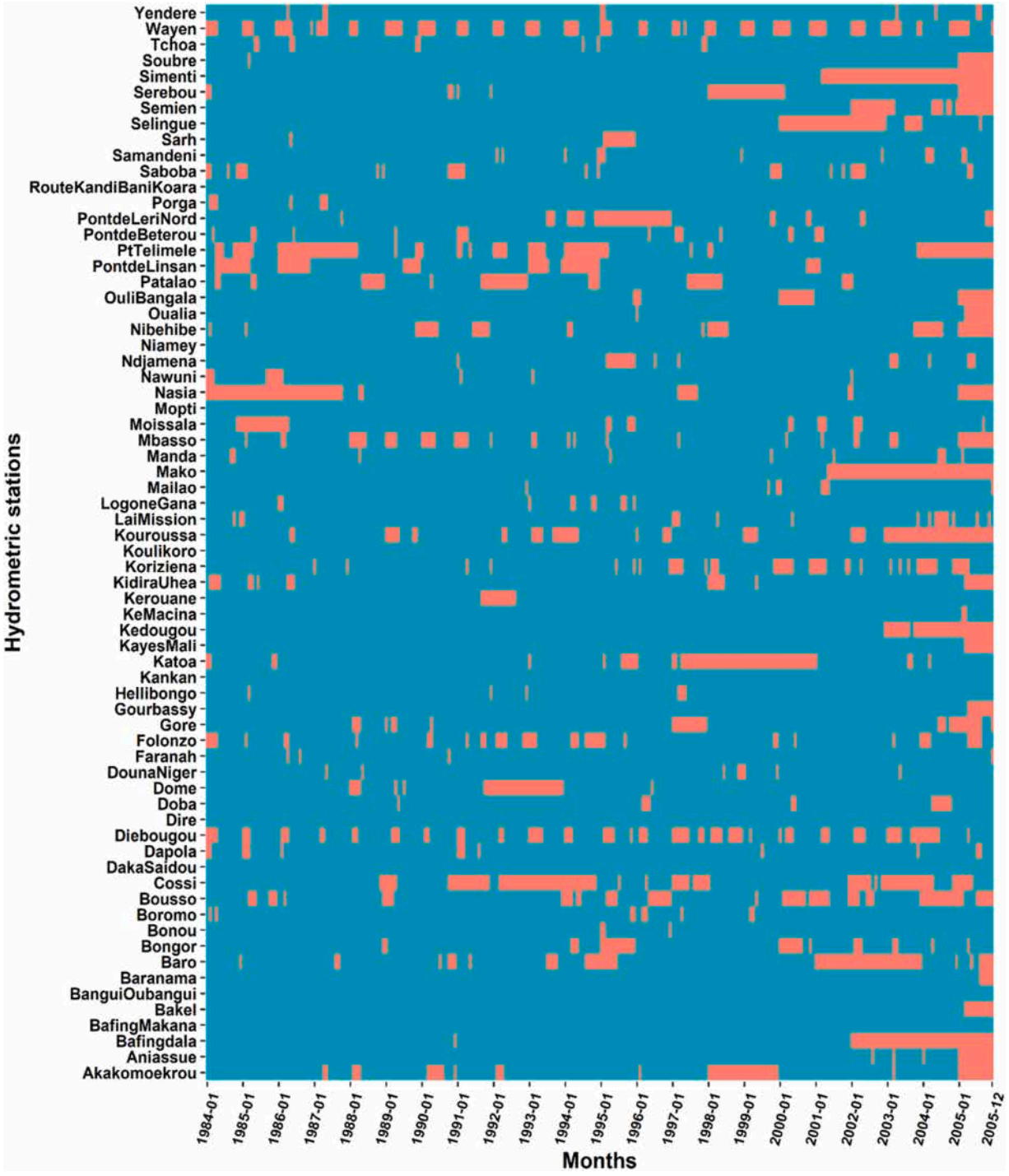

Fig. 4. Temporal availability of monthly flow data for the 68 hydrometric stations from 1984 to 01–2005–12. In blue the available flow and in red the missing data.

3.6. Regionalization (adaptation of Oudin's method)

Since hydroclimatic data is often very incomplete in West and Central Africa, regionalization techniques could be a good alternative for hydrological modeling on ungauged and data scarce basins (Siderius et al., 2018; Vamoryba et al., 2019). Our objective is to determine whether a regionalization method coupled with satellite products makes it possible to preserve the performance of hydrological modeling.

The regionalization method (Oudin et al., 2008) chosen is based on the selection of sets of parameters from the calibration on neighboring basins (geographical neighborhood or other), which are used in turn for the validation on the target basin. The final flow is the Mean of the various simulated flows. Based on the work of Defrance et al. (2014), we chose the three closest geographical neighbors

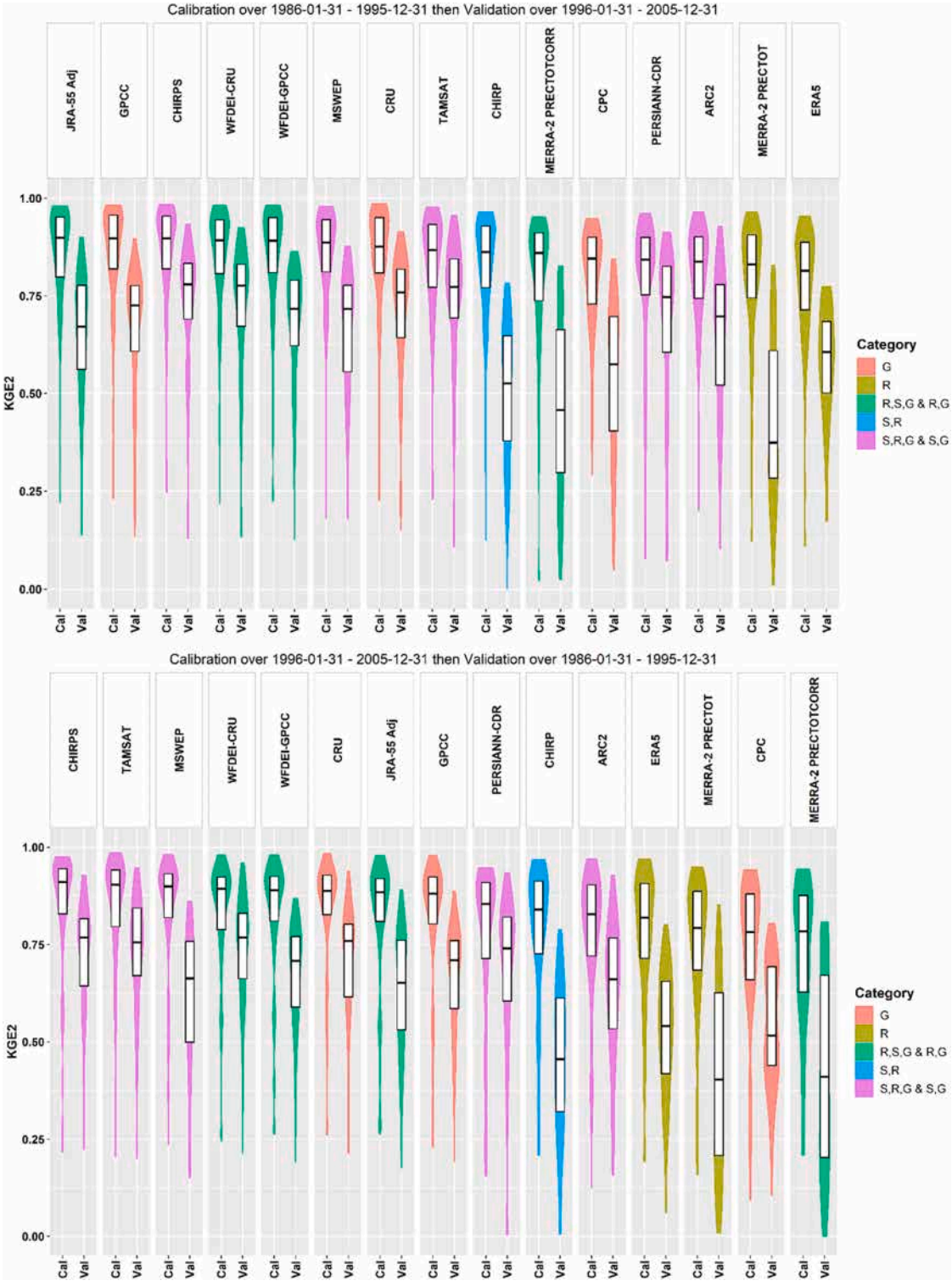

Fig. 5. Violin graphs summarizing KGE2 values obtained in calibration (Cal) and validation (Val) for the pair SST1 (top) and SST2 (bottom) – we fix the minimum of the criteria axis to 0 for a visibility question.

by considering the centroid of the basins.

In order to spatially assess the performance of the regionalization method, an index was calculated for each basin:

$$Variation = \frac{KGE2_{regionalisation} - KGE2_{local}}{KGE2_{local}} \quad (3)$$

4. Results

4.1. Overall performance of precipitation datasets

The calibration and validation of the GR2M model was carried out for each of the pairs of periods listed in Table 2.

Fig. 5 presents the "violins" of the values of the KGE2 score of the 15 rainfall products for the pairs of calibration/validation periods SST1 and SST2 (hereinafter respectively denoted pair SST1 and pair SST2) and the 68 stations selected. The rainfall products are arranged from left to right according to decreasing median value of KGE2 in calibration ("Cal" for calibration and "Val" for validation). The "violins" are read like a boxplot (min, max, median, and 1st and 3rd quartiles) but with additional information: the width of the "violins" is proportional to the number of values and thus transcribes the distribution of values.

To determine which is the best product, we decided to use various indicators related to KGE2, in calibration and validation, and the robustness criterion. In calibration, the median, min-max and interquartile distances (IQR) of the KGE2 values were considered. In validation, the median KGE2 and the median and 3rd quartile of the absolute values of the robustness ratio were used. We emphasize more on KGE2 value in calibration than its value in validation because we assume that if the good performance of the model is carried over from calibration to validation period, the model is likely to have a better temporal transferability for future application.

Fig. 5 illustrates the interest of a determination based on criteria other than just the median value of KGE2 in calibration or in validation. A product that is good in calibration is not necessarily good in validation (for example CHIRP, MERRA-2 (PRECTOT and PRECTOTCORR), ERA5 and CPC) and vice versa. Also, the difference between the median values of KGE2 in calibration (Cal) and validation (Val) for these GPEs is very important. This demonstrates the interest of taking robustness into account (Eq. (2)). In addition, the IQR and min-max of the KGE2 values make it possible to determine whether the modeling performance of the GPEs products carries over to a variety of catchment areas and incidentally to various geographical and climatic zones.

Fig. 5 also shows that the shape of the violins is not symmetrical in calibration, just like the box: there are more high values than low values. This implies that for most of the basins selected, the GR2M model integrates the various rainfall products well, even if these products are very different from each other. In validation, the shape of the violins is more symmetrical.

We then implemented a ranking grid that assigns a number of points according to the rank obtained: the better the rank of a product, the more points the product is given (150 points for the best product, 10 points for the worst). For example, Table 3 indicates the number of points allocated to rainfall products according to the rank of the median value of KGE2 obtained in calibration for the pair SST1.

When comparing the pairs SST1 and SST2, the ranking of the rainfall products is not the same, but we still generally find the same products in the first half and in the second half of the results. There is also a diversity of performances for the same type of product: some are good, others are not. On the other hand, the GPEs resulting from re-analyses alone are poor for both pairs of periods.

Table 4 summarizes the scores of each rainfall product according to the 6 criteria chosen in our approach. The following lessons can be learned:

- The best calibration product is CHIRPS, which combines ground, re-analysis and satellite data, with a total of 770/900. It is ranked ahead of products like CRU (680/900) and GPCC (670/900), which are based on ground observations alone.
- The best product in validation is WFDEI-CRU (combination of re-analysis and ground data), with a total of 820/900, ahead of TAMSAT (810/900) and PERSIANN-CDR (800/900), which are combinations of satellite and ground data. CHIRPS (770/900) and CRU (700/900) are ranked 4th and 5th in validation.
- The worst products are MERRA, ERA 5 (re-analysis and re-analysis/ground data) and CPC (ground data).
- CHIRP (satellite and re-analysis), unlike the version corrected with ground data (CHIRPS), has poor results. Only 4 other products are less effective.
- Products based only on re-analyses or only on satellite outputs (or composites of the two) are not very effective (CHIRP, ERA5, MERRA).

Overall, the best products are either composites of ground/re-analysis/satellite data or based on ground data only, which shows the importance of ground data. The poor performance of the CPC product may be linked to the low density of the stations that were used to generate it for our study area.

Since the robustness ratio reflects a deviation from KGE2 in calibration, it is essential that the value of KGE2 in calibration be good.

Table 3
Precipitation estimate grid performance rating: example for the pair SST1 (calibration on 1986–1995 and validation on 1996–2005).

Grid	JRA55 adj	GPCC	CHIRPS	…….	ARC2	MERRA-2 PRECTOT	ERA 5
Rating	150	140	130	……	30	20	10

Table 4
Ranking of gridded precipitation estimates by performance criterion in calibration and validation (KGE2 in calibration and validation, ratio (Dakhlaoui) in validation). In **red** observation products; in **dark blue** satellite products corrected with ground data and re-analyses; in **light blue** satellite products associated with re-analyses; in **green** re-analyses corrected with ground data; and in black italics the products of re-analyses alone. The maximum possible score and the score obtained by each rainfall product for each of the criteria are indicated in parentheses in the title cells and in the cells of the table, respectively.

Rank	CALIBRATION			VALIDATION		
	Median $KGE2_{Cal}$ (max 300)	Min-Max $KGE2_{Cal}$ (max 300)	IQ $KGE2_{Cal}$ (max 300)	Median (| Ratio |) (max 300)	Q3 (| Ratio |) (max 300)	Median $KGE2_{Val}$ (max 300)
1	**CHIRPS** (280)	**WFDEI-GPCC** (260)	**MSWEP** (280)	**PERSIANN-CDR** (300)	**PERSIANN-CDR** (300)	**CHIRPS** (300)
2	**JRA-55 Adj** (240)	**JRA-55 Adj** (250)	**CHIRPS** (250)	**TAMSAT** (280)	**TAMSAT** (280)	**WFDEI-CRU** (280)
3	**WFDEI-CRU** (240)	**CHIRPS** (240)	**CRU** (250)	**WFDEI-CRU** (260)	**WFDEI-CRU** (280)	**CRU** (250)
4	**MSWEP** (230)	**CRU** (240)	**WFDEI-GPCC** (230)	**CHIRPS** (240)	**CHIRPS** (230)	**TAMSAT** (250)
5	**GPCC** (220)	**GPCC** (230)	**GPCC** (220)	**CRU** (220)	**CRU** (230)	**PERSIANN-CDR** (220)
6	**TAMSAT** (220)	**TAMSAT** (200)	**JRA-55 Adj** (220)	**WFDEI-GPCC** (200)	**WFDEI-GPCC** (200)	**GPCC** (200)
7	**WFDEI-GPCC** (220)	**WFDEI-CRU** (200)	**WFDEI-CRU** (220)	**ARC2** (180)	**GPCC** (180)	**WFDEI-GPCC** (180)
8	**CRU** (190)	**MSWEP** (180)	**ARC2** (140)	**GPCC** (160)	**ARC2** (160)	**MSWEP** (160)
9	**CHIRP** (130)	**CHIRP** (140)	**PERSIANN-CDR** (130)	**JRA-55 Adj** (140)	**JRA-55 Adj** (130)	**ARC2** (140)
10	**PERSIANN-CDR** (110)	**ARC2** (120)	**CHIRP** (120)	**MSWEP** (120)	**MSWEP** (130)	**JRA-55 Adj** (120)
11	**ARC2** (80)	*ERA5* (110)	**TAMSAT** (120)	**CPC** (90)	*ERA5* (100)	*ERA5* (100)
12	**CPC** (70)	*MERRA-2 PRECTOT* (100)	*ERA5* (80)	*ERA5* (90)	**CPC** (80)	**CPC** (80)
13	**MERRA-2 PRECTOTCORR** (70)	**PERSIANN-CDR** (70)	*MERRA-2 PRECTOT* (80)	**CHIRP** (60)	**CHIRP** (60)	**CHIRP** (60)
14	*ERA5* (50)	**CPC** (30)	**CPC** (30)	**MERRA-2 PRECTOTCORR** (40)	**MERRA-2 PRECTOTCORR** (40)	**MERRA-2 PRECTOTCORR** (40)
15	*MERRA-2 PRECTOT* (50)	**MERRA-2 PRECTOTCORR** (30)	**MERRA-2 PRECTOTCORR** (30)	*MERRA-2 PRECTOT* (20)	*MERRA-2 PRECTOT* (20)	*MERRA-2 PRECTOT* (20)

To take this into account, we have established a "definitive ranking" based on a mean between the performance in calibration and that in validation, reduced to a score out of 20 (Table 5).

Table 5 confirms the observations made for Table 4:

- CHIRPS (Climate Hazards group InfraRed Precipitation with Stations) is the best product, followed by WFDEI-CRU and CRU.
- There is a big difference between the CHIRPS product and the raw CHIRP version without ground data.
- Products based on ground data alone (except CPC) are always well ranked.
- Among the satellite/ground/re-analysis composite products, CHIRPS clearly stands out from the other 4, in particular MSWEP and ARC2.
- Of the two products dedicated to Africa, TAMSAT and ARC2, only the first is up to expectations, with a good position of 4th out of 15 versus 10th place for ARC2.

4.2. *Spatial performance of precipitation datasets*

Spatial analysis shows whether satellite products systematically lead to poor/good model performance in a given geographical area. It also makes it possible to determine the influence of the size of the basins on the performance of the GPEs associated with the GR2M model. Our analysis is based on a subdivision of our study area into two subsets, West Africa (WA) and northern Central Africa (CA).

Fig. 6 presents the mean values of KGE2 in validation for the two pairs SST1 and SST2. We have chosen the threshold of 0.6 as the minimum acceptable value (Kling et al., 2012). This figure exhibits a balanced performance of the GR2M model for a given rainfall product, among the 7 best gridded precipitation estimates (Table 5), between WA and CA areas. In fact, colors observed on WA are quite the same over CA. This balanced result is not observed with poorly ranked rainfall products such as MERRA2 (PRECTOT and PRECTOTCORR), ERA5, and CPC. These latter products performances seem worst over CA as colors are fully red and dark red apart from CPC (KGE2 < 0.7) while on WA yellow and blue colors can be observed on few basins (KGE2 ≥ 0.7). However, singularities appear for certain basins for which almost all the rainfall products lead to poor model performance (KGE2 <0.6). There can be many reasons for this, and a more detailed analysis of the data and results provides some explanations: the data is too incomplete, or the Mouelhi GR2M model cannot be adapted to the basins because the hydrological regime of these basins is of a complex nature (often the case of large basins that cover multiple rainfall regimes) or because the basin is heavily anthropized. These basins in question are Cossi and Baranama, medium sized basins on the Niger, and Niamey, a bigger basin located downstream of the Inner Niger Delta (IND), Logone Gana on the Chari, and Soubré on the Sassandra, located downstream of the Buyo hydroelectric dam in Côte d'Ivoire.

Fig. 7 provides a comparative summary of the minimum, maximum, quartiles and Means of KGE2 obtained over West Africa (WA) and northern Central Africa (CA) for the pairs SST1 and SST2. This figure is subdivided into 8 blocks on the basis of 3 pieces of information: the study area (WA or CA), the calibration/validation period pairs (SST1 and SST2), and the processing phase (calibration or validation). The results presented in Fig. 7 are summarized in Table 6.

Table 5
Overall ranking of gridded precipitation estimates. In **red** observation products; in **dark blue** satellite products corrected with ground data and re-analyses; in **light blue** satellite products associated with re-analyses; in **green** re-analyses corrected with ground data; and in black italics the products of re-analyses alone.

Rank	PRODUCT	PERFORMANCE (Cal +Val) / 2 (max = 900)	Mean/20
1	**CHIRPS**	770	**17.11**
2	**WFDEI-CRU**	740	**16.44**
3	**CRU**	690	**15.33**
4	**TAMSAT**	675	**15.00**
5	**WFDEI-GPCC**	645	**14.33**
6	**GPCC**	605	**13.44**
7	**PERSIANN-CDR**	555	**12.33**
8	**JRA-55 Adj**	550	**12.22**
9	**MSWEP**	550	**12.22**
10	**ARC2**	410	**9.11**
11	**CHIRP**	285	**6.33**
12	*ERA5*	265	**5.89**
13	**CPC**	190	**4.22**
14	*MERRA-2 PRECTOT*	145	**3.22**
15	**MERRA-2 PRECTOTCORR**	125	**2.78**

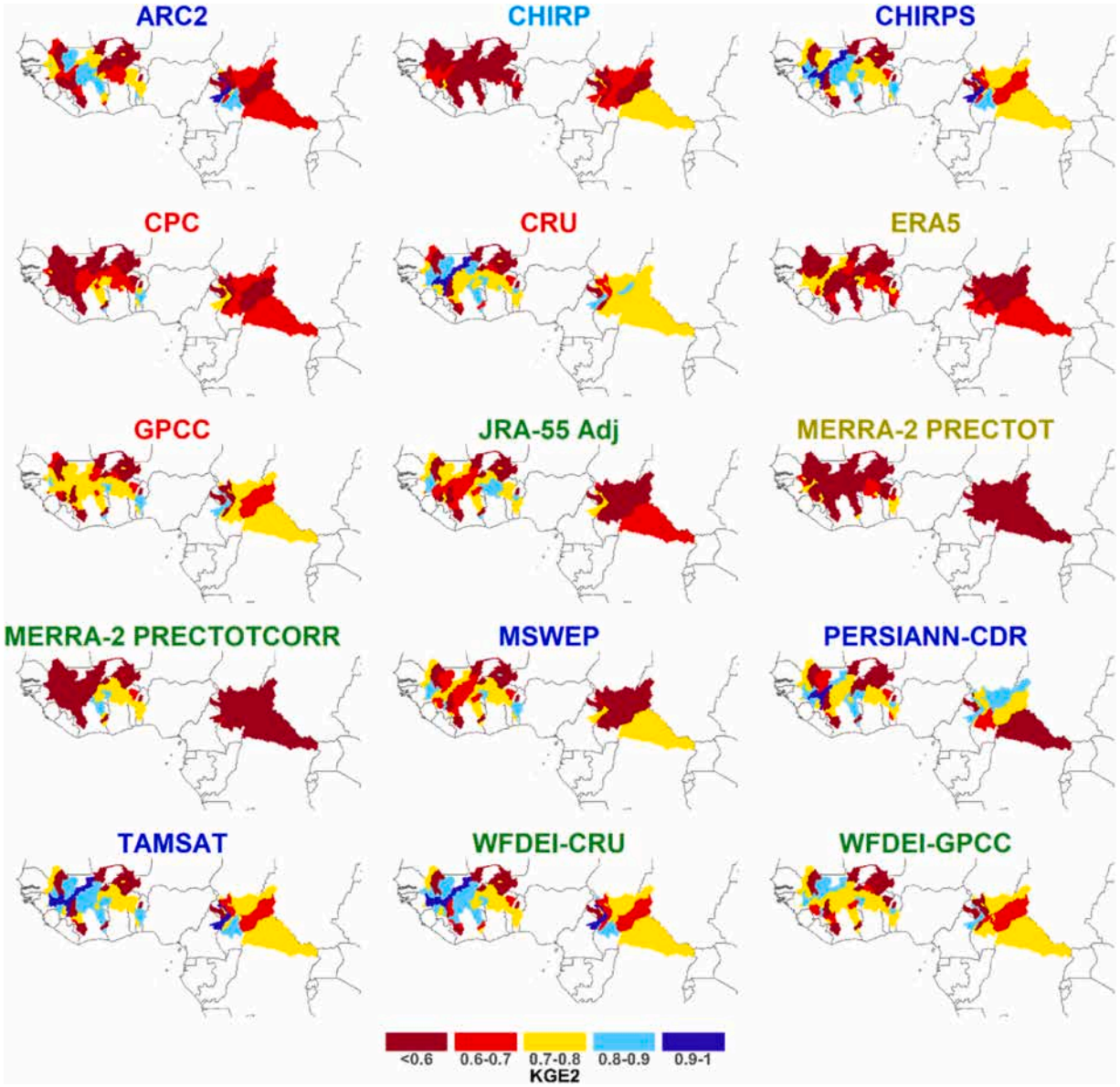

Fig. 6. Spatial representation of the average between the KGE2 values obtained in validation on SST1 and on SST2. In **red** observation products; in **dark blue** satellite products corrected with ground data and re-analyses; in **light blue** satellite products associated with re-analyses; in **green** re-analyses corrected with data on the ground; and in **green gold** re-analysis products alone.

For the entire set of GPEs, comparisons of KGE2 values shows slight differences depending on statistical specifications analyzed: min, max, 1st and 3rd quartiles and median. Indeed, there is globally a slight drop of mean values of KGE2 in calibration from SST1 to SST2 for both areas WA (median of the means going from 0.86 to 0.84 for the top 7; from 0.81 to 0.78 for the others) and CA (median of the means going from 0.81 to 0.78 for the top 7; range of means from 0.75 to 079–0.70–077 for the 8 others GPEs). This drop is also observed in validation on both WA (from range of means 0.69–0.74–0.68–0.74 for top 7 GPEs; from median of means 0.56–0.54 for the 8 remaining) and CA (from 0.54 to 0.48 for the 8 bottom GPEs). It is only in validation phase for the 7 best products that KGE2 mean values are slightly increasing (from median 0.66–0.68) from SST1 to SST2 over CA. Such overall drop from SST1 to SST2 is also observed for median and 3rd quartiles with little exceptions. Min and max values are more mixed. Besides, KGE2 values in calibration and validation are globally higher on WA than in CA. This fact is stressed by the important number of values in bold for WA related columns in Table 6. It finally appears that there are more differences between CA and WA than between the SSTs.

Fig. 8 presents the values of the quartiles of KGE2 according to the size of the basins in the SST period pairs. Table 7 is a summary of the results presented by Fig. 8.

The performance of the GR2M models in terms of quartiles according to the size of the basins resembles a flat bell curve: the best

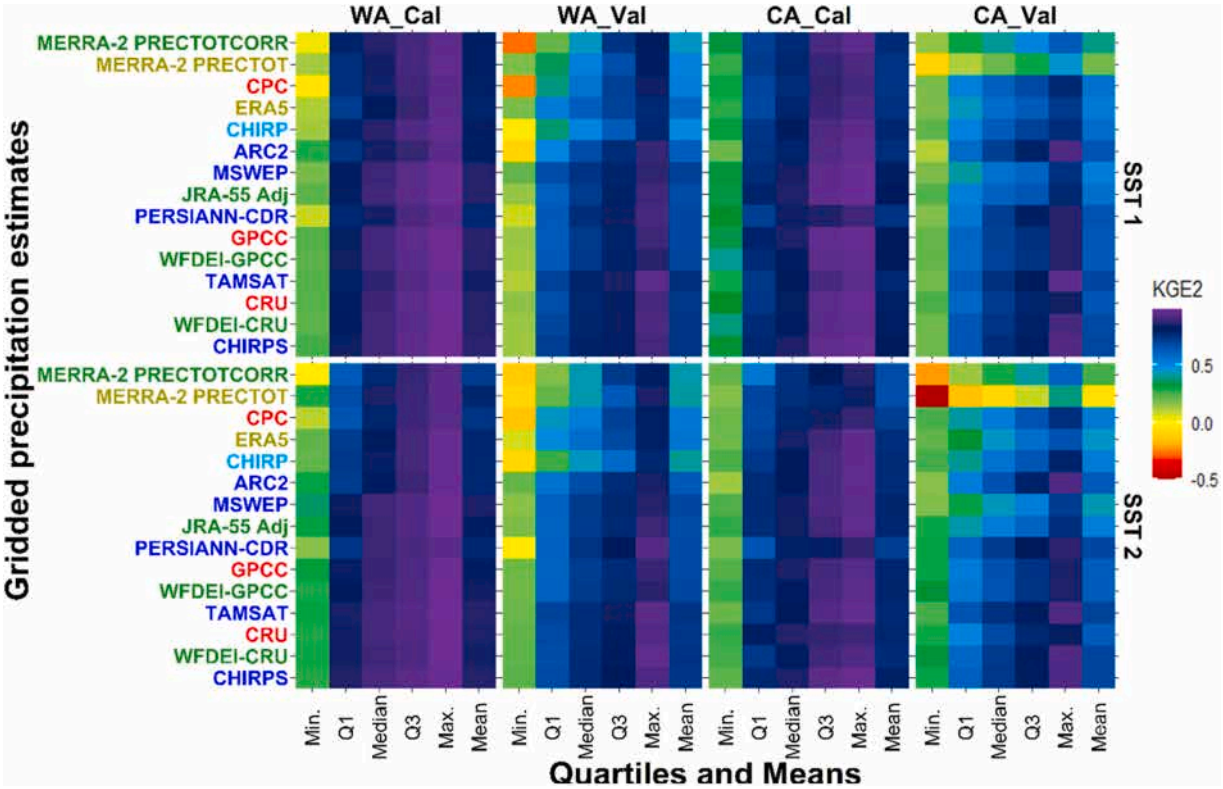

Fig. 7. min/max, quartiles, and averages values of KGE2 (on the abscissa) obtained over West Africa and Central Africa for the various rainfall products (ordinate) and for SST's. GPEs are ordered according to the ranking in Table 5, from bottom to the top.

Table 6
Median (and range) of different characteristics of KGE2: min, max, quartiles (Q1, median and Q3), and mean values for SST1 and SST2 period pairs. In bold case and for each line, the max values among median of KGE2 characteristics for different periods of calibration and validation.

7 best products					
	Characteristic	**CA_Cal, N = 7**	**CA_Val, N = 7**	**WA_Cal, N = 7**	**WA_Val, N = 7**
SST1	**Min.**	**0.32** (0.32, 0.36)	**0.20** (0.19, 0.21)	0.23 (0.22, 0.23)	0.13 (0.12, 0.13)
	Q1	0.77 (0.74, 0.78)	0.60 (0.59, 0.64)	**0.83** (0.81, 0.84)	**0.67** (0.63, 0.70)
	Median	0.84 (0.83, 0.85)	0.72 (0.71, 0.75)	**0.90** (0.89, 0.90)	**0.78** (0.75, 0.79)
	Q3	**0.95** (0.93, 0.96)	0.80 (0.77, 0.81)	**0.95** (0.94, 0.95)	**0.84** (0.81, 0.84)
	Max.	0.97 (0.96, 0.97)	0.87 (0.86, 0.92)	**0.98** (0.98, 0.98)	**0.91** (0.90, 0.92)
	Mean	0.81 (0.79, 0.82)	0.66 (0.64, 0.68)	**0.86** (0.85, 0.86)	**0.73** (0.69, 0.74)
SST2	**Min.**	0.23 (0.21, 0.25)	**0.29** (0.28, 0.31)	**0.27** (0.27, 0.28)	0.20 (0.19, 0.21)
	Q1	0.76 (0.73, 0.77)	0.59 (0.53, 0.60)	**0.84** (0.83, 0.85)	**0.67** (0.61, 0.68)
	Median	0.83 (0.81, 0.85)	0.72 (0.67, 0.73)	**0.90** (0.89, 0.91)	**0.76** (0.74, 0.77)
	Q3	0.91 (0.89, 0.92)	0.81 (0.76, 0.82)	**0.93** (0.93, 0.94)	**0.82** (0.79, 0.82)
	Max.	0.94 (0.91, 0.96)	0.87 (0.85, 0.93)	**0.98** (0.98, 0.98)	**0.94** (0.91, 0.94)
	Mean	0.78 (0.77, 0.79)	0.68 (0.63, 0.70)	**0.84** (0.84, 0.86)	**0.73** (0.68, 0.74)
8 remaining products					
	Characteristic	**CA_Cal, N = 8**	**CA_Val, N = 8**	**WA_Cal, N = 8**	**WA_Val, N = 8**
SST1	**Min.**	**0.30** (0.26, 0.33)	**0.17** (0.13, 0.19)	0.12 (0.09, 0.19)	0.07 (−0.13, 0.17)
	Q1	0.73 (0.71, 0.75)	**0.47** (0.38, 0.50)	**0.77** (0.75, 0.80)	0.43 (0.36, 0.54)
	Median	0.79 (0.78, 0.82)	**0.59** (0.52, 0.62)	**0.85** (0.84, 0.87)	0.58 (0.50, 0.69)
	Q3	**0.92** (0.89, 0.94)	0.64 (0.57, 0.69)	0.91 (0.90, 0.93)	**0.72** (0.69, 0.77)
	Max.	0.95 (0.92, 0.97)	0.76 (0.70, 0.78)	**0.96** (0.95, 0.97)	**0.84** (0.82, 0.88)
	Mean	0.76 (0.75, 0.79)	0.54 (0.48, 0.56)	**0.81** (0.80, 0.84)	**0.56** (0.50, 0.64)
SST2	**Min.**	0.20 (0.18, 0.24)	**0.19** (0.06, 0.24)	**0.25** (0.18, 0.29)	-0.01 (−0.10, 0.16)
	Q1	**0.72** (0.69, 0.76)	0.36 (0.25, 0.41)	**0.72** (0.67, 0.75)	**0.46** (0.24, 0.56)
	Median	0.82 (0.78, 0.84)	0.48 (0.40, 0.53)	**0.83** (0.81, 0.86)	**0.54** (0.44, 0.68)
	Q3	0.90 (0.84, 0.92)	0.58 (0.48, 0.65)	**0.90** (0.89, 0.92)	**0.71** (0.67, 0.77)
	Max.	0.95 (0.88, 0.96)	0.74 (0.64, 0.76)	**0.97** (0.95, 0.97)	**0.83** (0.80, 0.85)
	Mean	0.76 (0.70, 0.77)	0.48 (0.38, 0.52)	**0.78** (0.77, 0.80)	**0.54** (0.42, 0.64)

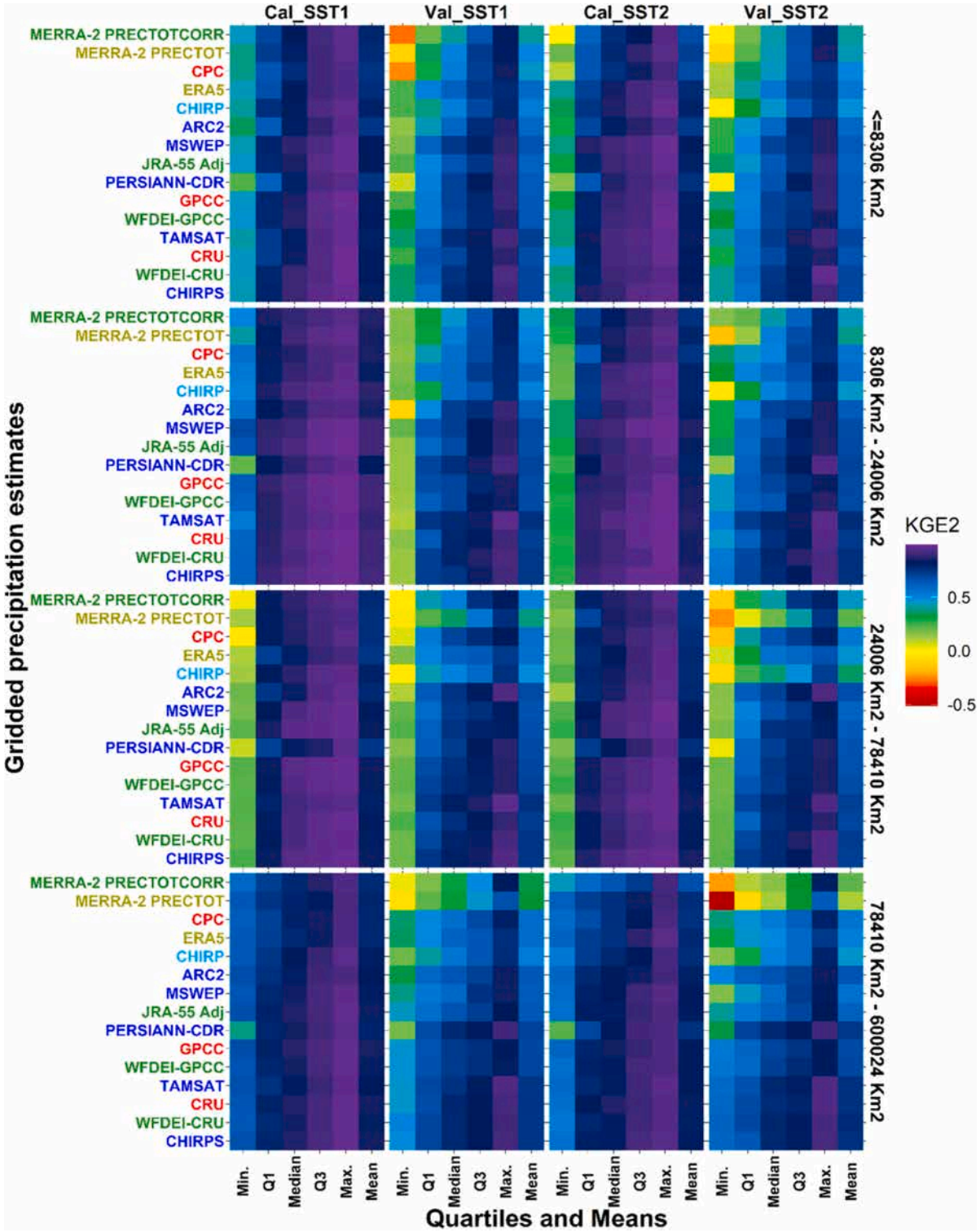

Fig. 8. min/max, quartiles, and means values of KGE2 according to the size of the catchment areas (divided according to surface area quartiles, same number of basins per category) and for different SST's. GPEs are ordered according to the ranking in Table 5, from bottom to the top.

Table 7
Median, interquartile range, mean (standard deviation), and min-max of KGE2 according to basin area classes.

	Characteristic	≤ 8310 KM2 N = 510	8310–24000 KM2 N = 510	24010–78410 KM2 N = 510	≥ 78410 KM2 N = 510
Cal	Median	0.84	0.89	0.88	0.82
	IQR	0.70–0.92	0.84–0.94	0.79–0.93	0.74–0.89
	Mean (SD)	0.79 (0.18)	0.85 (0.16)	0.80 (0.22)	0.80 (0.11)
	Range	-0.01–0.98	0.21–0.99	-0.04–0.98	0.23–0.97
Val	Median	0.65	0.69	0.7	0.65
	IQR	0.43–0.76	0.53–0.79	0.48–0.78	0.49–0.76
	Mean (SD)	0.59 (0.22)	0.63 (0.21)	0.61 (0.24)	0.60 (0.22)
	Range	-0.28–0.96	-0.16–0.95	-0.24–0.96	-0.51–0.93

N = (68 stations /4)* (15 GPEs)* (2 SST periods) = 17 * 15 * 2

values are obtained for the 2 intermediate classes of surface area of the basins studied. The products ranked in our top 7 for overall performance (Table 5) still have better results than the remaining 8.

4.3. Regionalization (adaptation of Oudin's method)

Fig. 9 presents violin graphs of the KGE2 values obtained by the regionalization procedure (Reg2) (Oudin et al., 2008) in comparison to an in-situ calibration (Local). As expected, the regionalization method generally leads to a loss of performance, while preserving the median KGE2 above 0.5 for the majority of satellite products. Increases in performance can be seen when it comes to the lower-rated rainfall products. However, the performance criteria remain low, much lower than those obtained with the best rainfall products.

The five best rainfall products for the pairs SST1 and SST2 are four mixed products with a main satellite component (TAMSAT, CHIRPS, ARC2 and PERSIANN-CDR) and a WFDEI-CRU corrected re-analysis product. Among these five best products, three were in the previous top five (CHIRPS, WFDEI-CRU and TAMSAT) and two were lower in the ranking (PERSIANN-CDR and ARC2).

The summary presented by Table 8 is an adaptation of the performance assessment method proposed above (Table 4). It shows, according to the criterion of median KGE2, a relative conservation of the ranking in terms of overall performance of the precipitation products. It should be noted, however, that the product ARC2, which was ranked 10th in Table 5, now went up 6 places in the ranking (Table 9). PERSIANN-CDR and TAMSAT only went up two and three places respectively. In addition, the min-max and interquartile distance criteria confirm that the regionalization presents a sharp drop in performance in certain basins, as shown in Fig. 8. This is even the case for well-ranked products such as CRU and GPCC, which are now in 7th and 9th place (Table 9).

Fig. 10 shows the spatial distribution of the mean rate of variation between the KGE2s in local validation and the KGE2s resulting from the regionalization procedure for the SST1 and SST2 pairs. The results are globally balanced between WA and CA. Also, there is bluer (variation > 0.1) for the GPEs ranked lower in Table 5. This reflects a positive impact of regionalization on the performance of these products. The result is reversed for the good GPEs based only on ground observations (CRU and GPCC), with an increase in the interquartile distance (Q1-Q3) of the KGE2 values (Table 8) due to a drop in the KGE2 value of more than 20% for the majority of basins (Fig. 10; variation < -0.2).

5. Discussion

5.1. Importance of bias correction and the quality of data source in GPEs performances

The mixed product CHIRPS (satellite, re-analysis and ground data) appears to be the best product overall in our study area when using the GR2M model, Mouelhi version. It is followed by products based on ground observations and products based on re-analysis and satellite coupled with ground observations. It is therefore noted that the integration of ground observation data strongly influences the ability of a rainfall product to reproduce good validation with the GR2M model. The comparison of the CHIRPS product to its raw CHIRP version, free of ground data, is explicit from this point of view. The differences between the "raw" version and the version "adjusted using ground observations" are more nuanced with the re-analysis products (MERRA-2 PRECTOTCORR and MERRA-2 PRECTOT). This result highlights the ineffectiveness of the bias correction method of the corrected version MERRA-2 PRECTOTCORR on our study area. It is also important to note that no raw re-analysis or satellite product could obtain an overall score greater than or equal to 12 in our ranking (Table 5).

"Mixed" products seem more appropriate for good modeling of the rainfall-runoff relationship. However, the density of the climate stations and the quality of the raw satellite data are sources of sometimes significant differences between satellite-based GPEs even in the top five P datasets. Furthermore, specificities of algorithms and observed rainfall datasets for algorithms calibration or bias correction of output timeseries amplify these differences among the GPEs. Despite, both TAMSAT and CHIRPS approaches are mainly based on Cold Cloud Duration (CCD) rainfall estimates from satellite thermal infrared recording, CHIRPS has a better performance. Especially, TAMSAT relies on Meteosat TIR imagery provided by The European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) while CHIRPS relies on two geosynchronous TIR archives, the 1981–2008 Globally Gridded Satellite (GriSat) archive produced by NOAA's National Climate Data Center and the 2000-present NOAA Climate Prediction Center dataset (CPC TIR). However the asset of CHIRPS is that it also includes Tropical Rainfall Measuring Mission 2B31 microwave precipitation estimates,

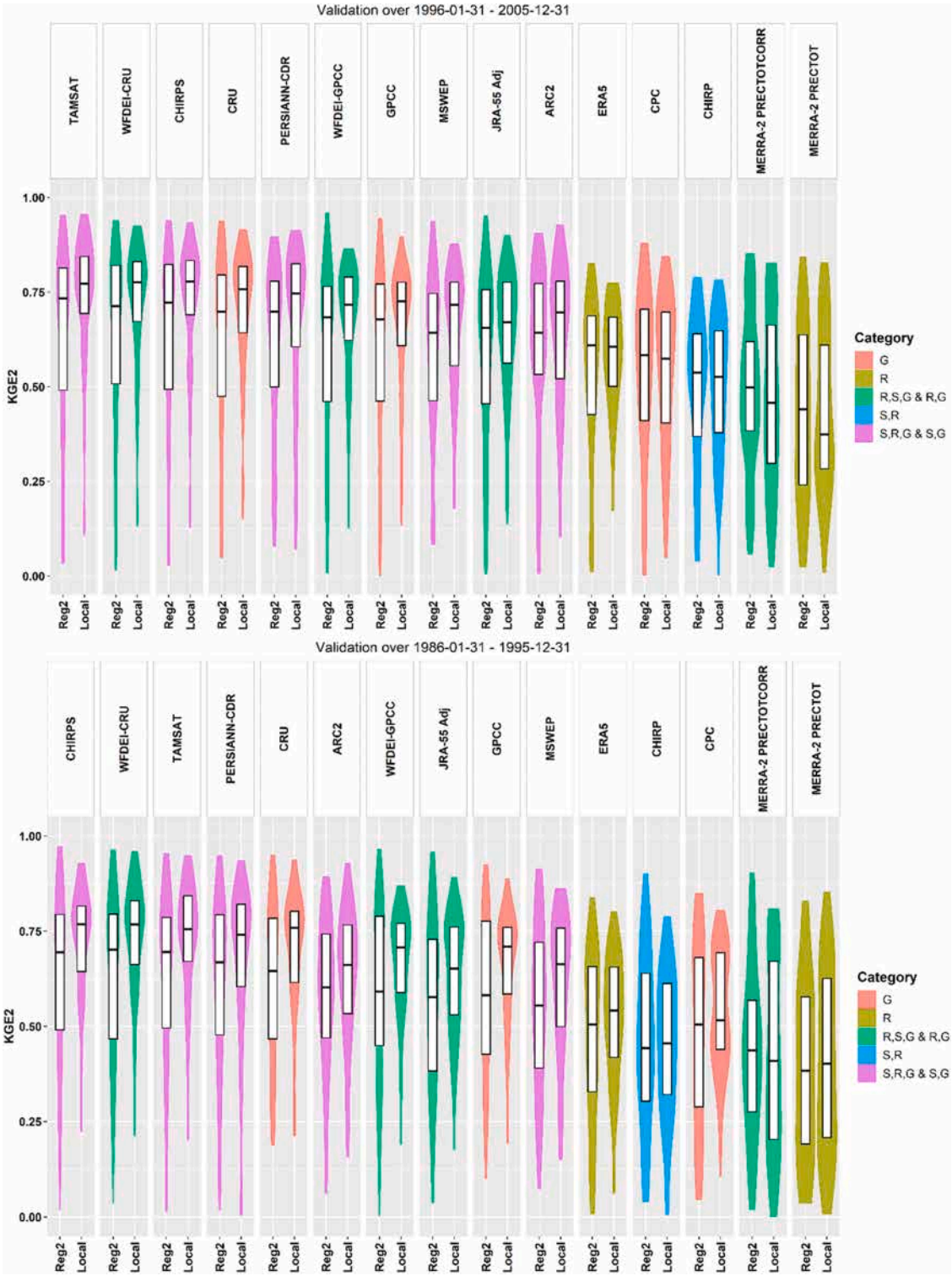

Fig. 9. Violin graphs of KGE2 values obtained in validation by the regionalization method according to the validation period.

CMORPH microwave-plus-infrared based precipitation estimates, monthly mean geostationary infrared brightness temperatures, and land surface temperature estimates in addition to a wider meteorological database (Funk et al., 2015; Maidment et al., 2017). Likewise, WFDEI-CRU and WFDEI-GPCC rely on WATCH Forcing Data methodology applied to the reanalysis ERA-Interim data. Nevertheless, CRU data have a higher spatial resolution (0.5°) compared to GPCC (1°) linked with a better density of ground stations providing

Table 8
Ranking of gridded precipitation estimates by performance criterion in Regionalization.

Rank	REGIONALIZATION			VARIATION	
	Median $KGE2_{Reg}$ (max 300)	Min-Max $KGE2_{Reg}$ (max 300)	IQ $KGE2_{Reg}$ (max 300)	Median (|Ratio|) (max 300)	Q3 (|Ratio|) (max 300)
1	CHIRPS (280)	*MERRA-2 PRECTOT* (260)	CHIRPS (240)	TAMSAT (290)	TAMSAT (280)
2	TAMSAT (280)	WFDEI-CRU (250)	MSWEP (230)	PERSIANN-CDR (280)	ARC2 (270)
3	WFDEI-CRU (240)	ARC2 (220)	JRA-55 Adj (210)	WFDEI-CRU (270)	CHIRPS (260)
4	CRU (230)	TAMSAT (220)	TAMSAT (210)	ARC2 (210)	WFDEI-CRU (230)
5	PERSIANN-CDR (230)	GPCC (200)	WFDEI-CRU (210)	CRU (200)	PERSIANN-CDR (220)
6	WFDEI-GPCC (190)	CRU (190)	*ERA5* (200)	WFDEI-GPCC (200)	MSWEP (190)
7	ARC2 (160)	MSWEP (160)	CHIRP (180)	CHIRPS (190)	WFDEI-GPCC (190)
8	GPCC (160)	PERSIANN-CDR (160)	WFDEI-GPCC (180)	GPCC (170)	CRU (160)
9	JRA-55 Adj (150)	WFDEI-GPCC (150)	GPCC (170)	MSWEP (160)	JRA-55 Adj (150)
10	MSWEP (140)	CHIRP (130)	ARC2 (160)	JRA-55 Adj (130)	*ERA5* (130)
11	*ERA5* (100)	CHIRPS (130)	CRU (120)	*ERA5* (100)	GPCC (120)
12	CHIRP (70)	*ERA5* (130)	PERSIANN-CDR (120)	CHIRP (70)	CHIRP (70)
13	CPC (70)	JRA-55 Adj (80)	MERRA-2 PRECTOT (80)	CPC (70)	CPC (70)
14	MERRA-2 PRECTOTCORR (40)	CPC (60)	MERRA-2 PRECTOTCORR (60)	MERRA-2 PRECTOTCORR (40)	*MERRA-2 PRECTOT* (30)
15	*MERRA-2 PRECTOT* (20)	MERRA-2 PRECTOTCORR (60)	CPC (30)	*MERRA-2 PRECTOT* (20)	MERRA-2 PRECTOTCORR (30)

observed rainfall (Weedon et al., 2014). As a result, WFDEI-CRU (resp. CRU) shows better performances for rainfall-runoff modelling than WFDEI-GPCC (resp. GPCC) on the basins considered in this study.

The ARC2 product, developed especially for the African continent, does not perform well in hydrological modeling. The TAMSAT product, also developed for the African continent, did better: 4th place in our overall ranking. If both products are based on EUMETSAT IR database, TAMSAT includes 30 min (prior to July 2006) and 15 min (till July 2006) TIR imagery unlike ARC2 which includes 3-hourly IR data. Algorithms employed are also different (Novella and Thiaw, 2013).

Among the ground observation products, CPC performs much worse than CRU and GPCC. This is probably related to the spatial density of the stations over West and Central Africa which are involved in the design of these ground observation data based GPEs.

This study focused on "old" products due to the availability of the data (before the year 2000). New emerging products of the same type as CHIRP (raw satellite estimates not corrected with ground data) like GPM IMERG (Early and Late Run) and GSMaP (Real-Time) seem to show less differences with respect to their corrected versions than observed between CHIRP and CHIRPS (the corrected version). For instance, Dembele et al. (2020) obtain results similar to ours with the same precipitation estimates products but with a different hydrological model (mHM) operating with a daily timestep over the Volta Basin in Africa. On the other hand, Jiang and

Table 9
Overall ranking of gridded precipitation estimates in regionalization.

Rank	PRODUCT	PERFORMANCE (Cal +Val) (max = 1500)	Mean/20
1	TAMSAT	1280	**17.07**
2	WFDEI-CRU	1240	**16.53**
3	CHIRPS	1100	**14.67**
4	ARC2	1020	**13.60**
5	PERSIANN-CDR	1010	**13.47**
6	WFDEI-GPCC	910	**12.13**
7	CRU	900	**12.00**
8	MSWEP	880	**11.73**
9	GPCC	820	**10.93**
10	JRA-55 Adj	720	**9.60**
11	ERA5	660	**8.80**
12	CHIRP	520	**6.93**
13	MERRA-2 PRECTOT	410	**5.47**
14	CPC	300	**4.00**
15	MERRA-2 PRECTOTCORR	230	**3.07**

Bauer-Gottwein, using the HBV model with a daily timestep on basins in China, show that the performance of the IMERG-E (raw) and IMERG-F (corrected) versions is similar (Jiang and Bauer-Gottwein, 2019). They also show that IMERG-E has advantages over TRMM 3B42, even though the latter is adjusted with GPCC data. The quality of GPM IMERG products can be attributed to the coupling of various satellite sources and sensors (microwave and infrared) with fine temporal resolution precipitation data from synoptic stations. GSMaP-RT has also showed a better performance to estimate rainfall at daily timestep over West Africa when compared to its gauge corrected version (Satgé et al., 2020).

5.2. Spatial and temporal consistency of performance of rainfall estimation products

Following the overall analysis of the performance of the GR2M model on both SST period pairs and the associated basins with the gridded precipitation estimates as forcing data, a study of the spatial performance was carried out. The first objective was to see the effectiveness of the estimation products in the two major geographical areas of West Africa, with a greater number of basins, and Central Africa, which is less represented. It appears, while taking into account the asymmetry in terms of number, that the overall performance of GPEs is better on WA with higher values compared to CA (Fig. 6). This fact is shown in Fig. 6 by a greater density of blue coloring over WA. Also, KGE2 values are globally greater in SST1 than in SST2. Nevertheless, slight variations are observed depending on the quartile and the GPE considered. Fig. 7 shows an increase in KGE2 min and max values in calibration and validation on the SST2 period pair for the WA zone compared to SST1 period pair for the top seven P datasets. That is more mitigated for CA zone.

In order to better understand the impact of the choice of the study period on the performance of the GR2M model, a study of the statistical characteristics of the precipitation data was undertaken but not presented here. The CRU rainfall product was used as reference data for the study of the stationary or non-stationary nature of the time series, as well as for the detection of possible breaks in the mean. This study shows that 82% of the series are non-stationary in WA versus 43% in CA. Breaks are detected in the non-stationary series around 1990–1994. The year of rupture is generally 1993, and there are more ruptures in the WA zone (56% for WA, 47% for CA). The post 1993 period is wetter than 1980–1993 period. Unlike WA, the CA zone is under the influence of the equatorial regime, resulting in a mixed hydrological regime (2 maxima and 2 minima per year). Thus, mean biases of the top 7 GPEs over Sahel region in WA evaluated by Satgé et al. (2020) in the range 0.85–1.15 are better compensated by GR2M when calibrating the model on 1996–2005 and validating on 1986–1995.

The analysis of Fig. 8 highlights an influence of the surface area of the basins on the performance of the GR2M model, characterized by a slight drop in performance below 8300 km^2 and above 80,000 km^2 for both SST period pairs. We can conclude that the influence of the catchment area is not a major factor in the performance of the GR2M model, unlike the choice of the calibration-validation period. The latter criterion leads to greater variations in performance than the former.

5.3. Regionalization efficiency

The application of a regionalization method leads to a different ranking of the P datasets than that observed previously. The performance of satellite products like ARC2, PERSIANN-CDR, TAMSAT and MSWEP improves. This observation testifies to the fact that the regionalization reduces the impact of the biases in these precipitation series on flow simulation. These biases are often linked to tendencies of GPEs to overestimate or underestimate rainfall amount in certain localities (Satgé et al., 2020).

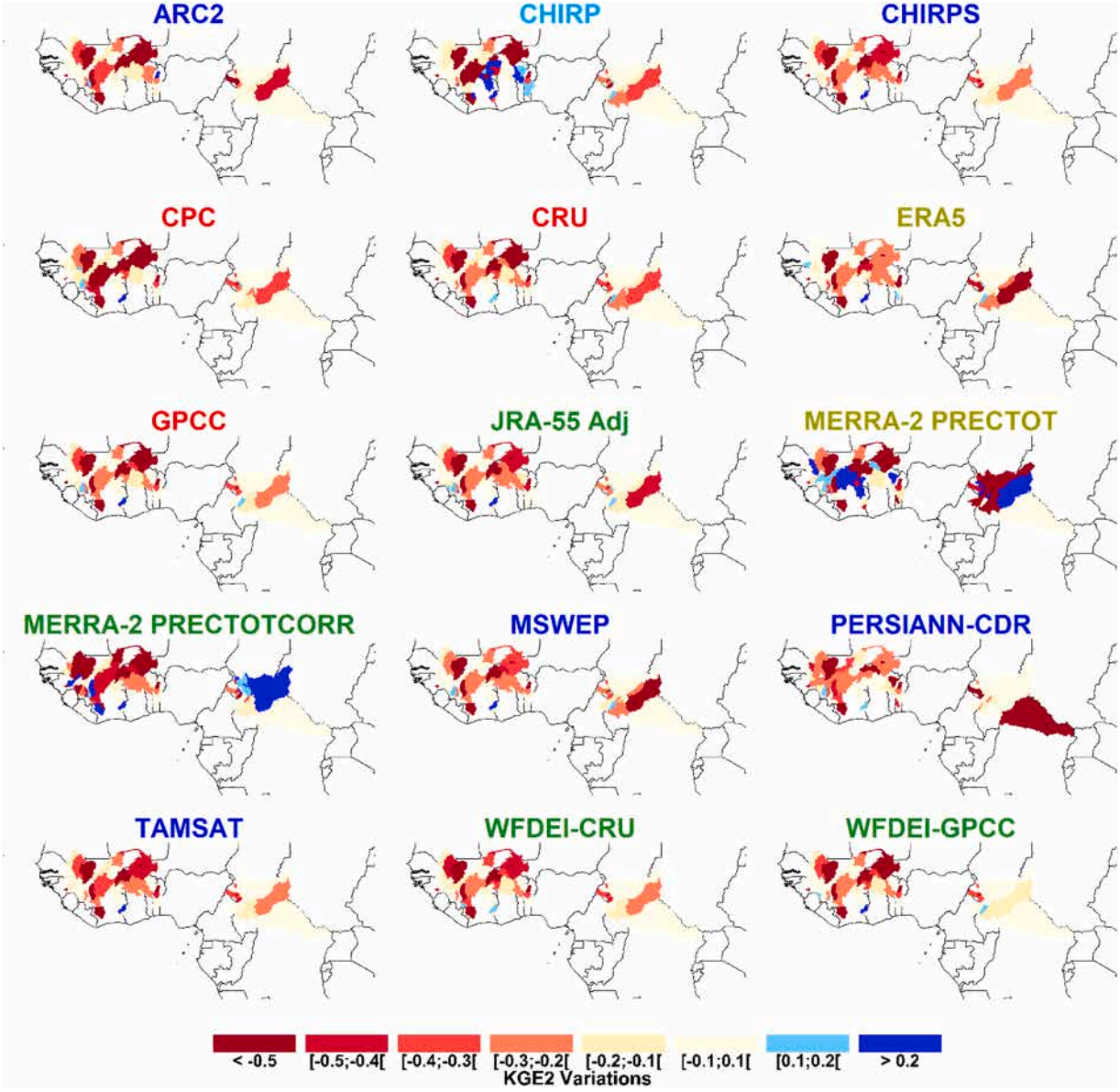

Fig. 10. Spatial representation of the rate of variation between the KGE2 obtained in validation with local parameters and those resulting from regionalization. If < 0, Regionalization $<$ local calibration, and vice versa. In **red** observation products; in **dark blue** satellite products corrected with ground data and re-analyses; in **light blue** satellite products associated with re-analyses; in **green** re-analyses corrected with data on the ground; and in **green gold** re-analysis products alone.

In addition, the results in the Inner Niger Delta are improved following the calibration on the 1996–2005 wet period and the validation on the 1986–1995 dry period for the majority of the precipitation estimation products. The cause could be the strong influence of precipitation and runoff upstream on the hydrological conditions of the IND, which is accentuated in the wet period. Indeed, since rainfall is low there ($<$ 500 mm/year), it has a very slight or delayed influence on runoff in the region.

Although the decrease in robustness of the observation products is in agreement with previous work on regionalization, the CRU product is better than the GPCC product. The reason for this difference must lie in the larger mesh of the observation stations involved in the design of CRU TS v4.00 (Harris et al., 2020) in our study area.

In our work, we have limited ourselves to the use of an adaptation of the method of spatial proximity of basin centroids according to Oudin et al. (2008). The use of other methods based on krige and multicriteria similarity among basins could improve the positive effect of regionalization on rainfall-runoff model performance (Goswami et al., 2007).

6. Conclusion

Our work made it possible to evaluate a set of 15 P datasets: three based only on ground observation data (CPC v.1, CRU TS v.4.00 and GPCC v.7); six based on reanalysis (ERA5, JRA-55 Adj, MERRA-2 PRECTOT, MERRA-2 PRECTOTCORR, WFDEI-CRU, WFDEI-GPCC); and six based on satellite data (ARC v.2, CHIRP v.2, CHIRPS v. 2, PERSIANN-CDR, MSWEP v2.2, TAMSAT v3.0).

The approach used is based on the outputs of the GR2M hydrological model, fed with GPEs as forcing data, as well as on a regionalization according to the near-neighbor spatial method.

The results prove the usefulness of ground observation data insofar as the best P datasets are those corrected with ground data, such as GPCC, GPCP and CRU. Of course, the GPCC and CRU observation based GPEs performed well on the majority of the basins. As for the poor performance of the CPC observation product, it highlights the need for a good mesh of rainfall stations for the constitution of gridded precipitation estimates. The blended product CHIRPS combines ground observations from great international and national meteorological databases with re-analysis timeseries and satellite estimates. Furthermore, from a spatial point of view, the CHIRPS results are balanced over CA and WA.

Regionalization by the near-neighbor method (Defrance et al., 2014; Oudin et al., 2008) shows that some previously poorly ranked products such as TAMSAT, PERSIANN-CDR and ARC2 become usable. More complex regionalization methods could further improve model outputs for a larger number of GPEs.

The period of availability of both hydrometric data and P datasets led us to limit ourselves to 15 P datasets over the period 1985–2005. It would be advantageous, however, to look into the use of more recent products, which are promising according to the recent bibliography (in certain regions of the world, particularly in Asia), such as CMORPH, GSMaP and GPM IMERG for hydrological modeling in the WA and CA zone. Even in their "raw" version (without ground observation), these products manage to reproduce what is observed on the ground well.

CRediT authorship contribution statement

Christopher Kouakou: Developed the methodology, Set up the data sets and models, Performed the numerical processing, Wrote the bulk of the paper. **Jean-Emmanuel Paturel**: Contributed to the setup of the methodology, datasets and models through discussions with the lead author, Contributed to the discussion of the results and the writing of the paper. **Frédéric Satgé**: Contributed to the setup of the methodology and the datasets through discussions with the lead author, Contributed to the discussion of the results and the writing of the paper. **Yves Tramblay**: Contributed to the setup of the methodology through discussions with the lead author, Contributed to the discussion of the results and the writing of the paper. **Dimtri Defrance**: Contributed to the discussion of the results and the writing the paper. **Nathalie Rouché**: Contributed to the setup of the datasets through discussions with the lead author.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

We would like to thank all the agents of the national services who measure the hydroclimatic data. We would also like to thank HydroSciences Montpellier, which provided the data through their database SIEREM, and Nathalie Rouché, who carried out critical and investigative analyses on this streamflow data. We would like to acknowledge Frédéric Satgé for preprocessing and providing ready-to-use gridded precipitation data, and GDRI-Sud and ActNAO for funding the research. Finally, we would like to thank all the co-authors for their valuable contributions to this manuscript.

References

Abiodun, B.J., Adeyewa, Z.D., Oguntunde, P.G., Salami, A.T., Ajayi, V.O., 2012. Modeling the impacts of reforestation on future climate in West Africa. Theor. Appl. Clim. 110, 77–96. https://doi.org/10.1007/s00704-012-0614-1.

Akinsanola, A.A., Ogunjobi, K.O., Ajayi, V.O., Adefisan, E.A., Omotosho, J.A., Sanogo, S., 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.

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop Evapotranspiration - Guidelines for Computing Crop Water Requirements - FAO Irrigation and drainage. FAO,, Rome, Italy.

Arsenault, R., Brissette, F., Martel, J.-L., 2018. The hazards of split-sample validation in hydrological model calibration. J. Hydrol. 566, 346–362. https://doi.org/10.1016/j.jhydrol.2018.09.027.

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, 69–83. https://doi.org/10.1175/BAMS-D-13-00068.1.

Ashouri, H., Nguyen, P., Thorstensen, A., Hsu, K., Sorooshian, S., Braithwaite, D., 2016. Assessing the Efficacy of High-Resolution Satellite-Based PERSIANN-CDR Precipitation Product in Simulating Streamflow. J. Hydrometeorol. 17, 2061–2076. https://doi.org/10.1175/JHM-D-15-0192.1.

Atiah, W.A., Tsidu, G.M., Amekudzi, L.K., 2020. Investigating the merits of gauge and satellite rainfall data at local scales in Ghana, West Africa. Weather Clim. Extrem. 30, 100292 https://doi.org/10.1016/j.wace.2020.100292.
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., 2019. MSWEP V2 Global 3-Hourly 0.1° precipitation: methodology and quantitative assessment. Bull. Am. Meteorol. Soc. 100, 473–500. https://doi.org/10.1175/BAMS-D-17-0138.1.
Becker, A., Finger, P., Meyer-Christo, A., 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 29.
Bodian, A., Dezetter, A., Deme, A., Diop, L., 2016. Hydrological Evaluation of TRMM Rainfall over the Upper Senegal River Basin. Hydrology 3, 15. https://doi.org/10.3390/hydrology3020015.
Bodian, A., Diop, L., Panthou, G., Dacosta, H., Deme, A., Dezetter, A., Ndiaye, P.M., Diouf, I., Vischel, T., 2020. Recent Trend in Hydroclimatic Conditions in the Senegal River Basin. Water 12, 436. https://doi.org/10.3390/w12020436.
Boyer, J.F., Dieulin, C., Rouche, N., Cres, A., Servat, E., Paturel, J.E., Mahé, G., 2006. SIEREM: an environmental information system for water resources. IAHS 308.
Burek, P., Langan, S., Cosgrove, W., Fischer, G., Kahil, T., Magnusziewski, P., Satoh, Y., Tramberend, S., Wada, Y., Wiberg, D., 2016, The Water Futures and Solutions Initiative of IIASA.
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., 6th Workshop Int. Precip. Work. Group 163, 162–176. https://doi.org/10.1016/j.atmosres.2015.01.010.
Chen, J., Del Genio, A.D., Carlson, B.E., Bosilovich, M.G., 2008. The Spatiotemporal Structure of Twentieth-Century Climate Variations in Observations and Reanalyses. Part I: Long-Term Trend. J. Clim. 21, 2611–2633. https://doi.org/10.1175/2007JCLI2011.1.
Coron, L., Thirel, G., Delaigue, O., Perrin, C., Andréassian, V., 2017. The suite of lumped GR hydrological models in an R package. Environ. Model. Softw. 94, 166–171. https://doi.org/10.1016/j.envsoft.2017.05.002.
Dakhlaoui, H., Ruelland, D., Tramblay, Y., Bargaoui, Z., 2017. Evaluating the robustness of conceptual rainfall-runoff models under climate variability in northern Tunisia. J. Hydrol. 550, 201–217. https://doi.org/10.1016/j.jhydrol.2017.04.032.
Defrance, D., Javelle, P., Organde, D., Ecrepont, S., Andréassian, V., Arnaud, P., 2014. Using damage reports to assess different versions of a hydrological early warning system (preprint). Hydrometeorol. /Model. Approaches. https://doi.org/10.5194/hessd-11-4365-2014.
Dembélé, M., Zwart, S.J., 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.
Dembélé, M., Schaefli, B., van de Giesen, N., Mariéthoz, G., 2020. Suitability of 17 rainfall and temperature gridded datasets for largescalehydrological modelling in West Africa (preprint). Hydrometeorol. /Remote Sens. GIS. https://doi.org/10.5194/hess-2020-68.
Ekeu-wei, I., Blackburn, G., Pedruco, P., 2018. Infilling Missing Data in Hydrology: Solutions Using Satellite Radar Altimetry and Multiple Imputation for Data-Sparse Regions. Water 10, 1483. https://doi.org/10.3390/w10101483.
Ekeu-wei, I.T., Blackburn, G.A., 2018. Applications of open-access remotely sensed data for flood modelling and mapping in developing regions. Hydrology 5, 39. https://doi.org/10.3390/hydrology5030039.
Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: new 1–km spatial resolution climate surfaces for global land areas. Int. J. Clim. 37, 4302–4315. https://doi.org/10.1002/joc.5086.
Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., Michaelsen, J., 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.
Gelaro, R., McCarty, W., Suárez, 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, Version 2 (MERRA-2). J. Clim. 30, 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1.
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, 923–940. https://doi.org/10.1002/qj.2130.
Goswami, M., O'Connor, K.M., Bhattarai, K.P., 2007. Development of regionalisation procedures using a multi-model approach for flow simulation in an ungauged catchment. J. Hydrol. 333, 517–531. https://doi.org/10.1016/j.jhydrol.2006.09.018.
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, 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003.
Harris, I., Osborn, T.J., Jones, P., Lister, D., 2020. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. Sci. Data 7, 109. https://doi.org/10.1038/s41597-020-0453-3.
Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J., 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. 146, 1999–2049. https://doi.org/10.1002/qj.3803.
Hsu, K., Gao, X., Sorooshian, S., Gupta, H.V., 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.
Hsu, K., Gupta, H.V., Gao, X., Sorooshian, S., 1999. Estimation of physical variables from multichannel remotely sensed imagery using a neural network: Application to rainfall estimation. Water Resour. Res. 35, 1605–1618. https://doi.org/10.1029/1999WR900032.
Iizumi, 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, 7800–7819. https://doi.org/10.1002/2017JD026613.
Jiang, L., Bauer-Gottwein, P., 2019. How do GPM IMERG precipitation estimates perform as hydrological model forcing? Evaluation for 300 catchments across Mainland China. J. Hydrol. 572, 486–500. https://doi.org/10.1016/j.jhydrol.2019.03.042.
Klemeš, V., 1986. Operational testing of hydrological simulation models. Hydrol. Sci. J. 31, 13–24. https://doi.org/10.1080/02626668609491024.
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.
Knoben, W.J.M., Freer, J.E., Woods, R.A., 2019. Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. Hydrol. Earth Syst. Sci. 23, 4323–4331. https://doi.org/10.5194/hess-23-4323-2019.
Kouassi, A.M., Bi, T.M.N., Kouamé, K.F., Kouamé, K.A., Okaingni, J.-C., Biemi, J., 2012. Application de la méthode des simulations croisées à l'analyse de tendances dans la relation pluie-débit à partir du modèle GR2M: cas du bassin versant du N'zi-Bandama (Côte d'Ivoire). Comptes Rendus Geosci. 344, 288–296. https://doi.org/10.1016/j.crte.2012.02.003.
Kouassi, A.M., Koffi, Y.B., Kouamé, K.F., Lasm, T., 2013, Application d'un modèle conceptuel et d'un modèle de réseaux de neurones artificiels à la simulation des débits annuels dans le bassin versant du N'zi-Bandama (Côte d'Ivoire) 13.
Maidment, R.I., Grimes, D., Black, E., Tarnavsky, E., Young, M., Greatrex, H., Allan, R.P., Stein, T., Nkonde, E., Senkunda, S., Alcántara, E.M.U., 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.
Martens, B., Miralles, D.G., Lievens, H., van der Schalie, R., de Jeu, R.A.M., Fernández-Prieto, D., Beck, H.E., Dorigo, W.A., Verhoest, N.E.C., 2017. GLEAM v3: satellite-based land evaporation and root-zone soil moisture. Geosci. Model Dev. 10, 1903–1925. https://doi.org/10.5194/gmd-10-1903-2017.
Mouelhi, S., Michel, C., Perrin, C., Andréassian, V., 2006. Stepwise development of a two-parameter monthly water balance model. J. Hydrol. 318, 200–214. https://doi.org/10.1016/j.jhydrol.2005.06.014.
Novella, N.S., Thiaw, W.M., 2013. African Rainfall Climatology Version 2 for Famine Early Warning Systems. J. Appl. Meteorol. Climatol. 52, 588–606. https://doi.org/10.1175/JAMC-D-11-0238.1.

Oudin, L., Andréassian, V., Perrin, C., Michel, C., Le Moine, N., 2008. Spatial proximity, physical similarity, regression and ungaged catchments: A comparison of regionalization approaches based on 913 French catchments: A COMPARISON OF REGIONALIZATION APPROACHES ON 913 CATCHMENTS. Water Resour. Res 44. https://doi.org/10.1029/2007WR006240.

Ouermi, K.S., Paturel, J.-E., Adounpke, J., Lawin, A.E., Goula, B.T.A., Amoussou, E., 2019. Comparison of hydrological models for use in climate change studies: A test on 241 catchments in West and Central Africa. Comptes Rendus Geosci. 351, 477–486. https://doi.org/10.1016/j.crte.2019.08.001.

Panthou, G., Vischel, T., Lebel, T., 2014. Recent trends in the regime of extreme rainfall in the Central Sahel. Int. J. Climatol. 34, 3998–4006. https://doi.org/10.1002/joc.3984.

Paturel, J.-E., Boubacar, I., L'aour-Cres, A., Mahe, G., 2010. Note de recherche: Grilles mensuelles de pluie en Afrique de l'Ouest et Centrale. rseau 23, 325–333. https://doi.org/10.7202/045095ar.

Poméon, T., Jackisch, D., Diekkrüger, B., 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.

Ramarohetra, J., Sultan, B., Baron, C., Gaiser, T., Gosset, M., 2013. 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., Draper, C.S., Liu, Q., Girotto, M., Mahanama, S.P.P., Koster, R.D., De Lannoy, G.J.M., 2017. Assessment of MERRA-2 Land Surface Hydrology Estimates. J. Clim. 30, 2937–2960. https://doi.org/10.1175/JCLI-D-16-0720.1.

Sambou, S., Dacosta, H., Diouf, R.N., Diouf, I., Kane, A., 2020, Hydropluviometric variability in non-Sahelian West Africa: case of the Koliba/Corubal River Basin (Guinea and Guinea-Bissau). Proc. IAHS 383, 171–183. https://doi.org/10.5194/piahs-383–171-2020.

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

Satgé, F., Pillot, B., Roig, H., Bonnet, M.-P., 2021. Are gridded precipitation datasets a good option for streamflow simulation across the Juruá river basin, Amazon? J. Hydrol. 602, 126773 https://doi.org/10.1016/j.jhydrol.2021.126773.

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. Theor. Appl. Clim. 115, 15–40. https://doi.org/10.1007/s00704-013-0860-x.

Siderius, C., Biemans, H., Kashaigili, J.J., Conway, D., 2018. Going local: Evaluating and regionalizing a global hydrological model's simulation of river flows in a medium-sized East African basin. J. Hydrol.: Reg. Stud. 19, 349–364. https://doi.org/10.1016/j.ejrh.2018.10.007.

Singh, S.K., Bárdossy, A., 2012. Calibration of hydrological models on hydrologically unusual events. Adv. Water Resour. 38, 81–91. https://doi.org/10.1016/j.advwatres.2011.12.006.

Stisen, S., Sandholt, I., 2010. Evaluation of remote-sensing-based rainfall products through predictive capability in hydrological runoff modelling. Hydrol. Process. 24, 879–891. https://doi.org/10.1002/hyp.7529.

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., Hsu, K., 2018. A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons. Rev. Geophys. 56, 79–107. https://doi.org/10.1002/2017RG000574.

Tarek, M., Brissette, F.P., Arsenault, R., 2020. Comparison of gridded datasetsfor the simulation of streamflowin Africa. IOP Conf. Ser.: Mater. Sci. Eng. 974, 012001 https://doi.org/10.1088/1757-899X/974/1/012001.

Thiemig, V., Rojas, R., Zambrano-Bigiarini, M., De Roo, A., 2013. Hydrological evaluation of satellite-based rainfall estimates over the Volta and Baro-Akobo Basin. J. Hydrol. 499, 324–338. https://doi.org/10.1016/j.jhydrol.2013.07.012.

Tramblay, Y., Rouché, N., Paturel, J.-E., Mahé, G., Boyer, J.-F., Amoussou, E., Bodian, A., Dacosta, H., Dakhlaoui, H., Dezetter, A., Hughes, D., Hanich, L., Peugeot, C., Tshimanga, R., Lachassagne, P., 2021. ADHI: the African Database of Hydrometric Indices (1950–2018). Earth Syst. Sci. Data 13, 1547–1560. https://doi.org/10.5194/essd-13-1547-2021.

Vamoryba, F., Amidou, D., Emile, S.G., Albert, G.B.T., Issiaka, S., 2019, Test of simulation of discharges from an ungauged basin in the South-West Côte d'Ivoire in equatorial climate 10, 5.

Viana, J.F., de, S., Montenegro, S.M.G.L., da Silva, B.B., da Silva, R.M., Srinivasan, R., Santos, C.A.G., Araujo, D.C., dos, S., Tavares, C.G., 2021. Evaluation of gridded meteorological datasets and their potential hydrological application to a humid area with scarce data for Pirapama River basin, northeastern Brazil. Theor. Appl. Clim. 145, 393–410. https://doi.org/10.1007/s00704-021-03628-7.

Wada, Y., Flörke, M., Hanasaki, N., Eisner, S., Fischer, G., Tramberend, S., Satoh, Y., van Vliet, M.T.H., Yillia, P., Ringler, C., Burek, P., Wiberg, D., 2016. Modeling global water use for the 21st century: the Water Futures and Solutions (WFaS) initiative and its approaches. Geosci. Model Dev. 9, 175–222. https://doi.org/10.5194/gmd-9-175-2016.

Weedon, G.P., Balsamo, G., Bellouin, N., Gomes, S., Best, M.J., Viterbo, P., 2014. The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. Water Resour. Res. 50, 7505–7514. https://doi.org/10.1002/2014WR015638.

World Meteorological Organization, W., 2021, State of the global climate 2020.

Xie, P., Chen, M., Yang, S., Yatagai, A., Hayasaka, T., Fukushima, Y., Liu, C., 2007. A gauge-based analysis of daily precipitation over East Asia. J. Hydrometeorol. 8, 607–626. https://doi.org/10.1175/JHM583.1.